

Peters, Jan Cornelius

Working Paper

Quantifying the effect of labor market size on learning externalities

Economics Working Paper, No. 2016-11

Provided in Cooperation with:

Christian-Albrechts-University of Kiel, Department of Economics

Suggested Citation: Peters, Jan Cornelius (2016) : Quantifying the effect of labor market size on learning externalities, Economics Working Paper, No. 2016-11, Kiel University, Department of Economics, Kiel

This Version is available at:

<https://hdl.handle.net/10419/144741>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

quantifying the effect of labor market size on learning externalities

by Jan Cornelius Peters



Quantifying the effect of labor market size on learning externalities

Jan Cornelius Peters*

Cornelius.Peters2@iab.de
+49/431/3395-3925

IAB Northern Germany, Regional Research Network of the Institute for Employment Research,
Projensdorfer Str. 82, 24106 Kiel, Germany

August 1, 2016

Abstract

This paper provides empirical evidence that individual labor productivity significantly depends on the size of the local labor market in which a worker previously acquired work experience. The analysis uses German micro data from the Institute for Employment Research (IAB) on transitions to employment within the period 2005 to 2011 and individual employment biographies from 1975 onwards. Analyzing the wages associated with the newly established employment relationships, suggests that dynamic agglomeration economies in general, and learning externalities in particular, play an important role in explaining individual labor productivity. Workers receive a significantly higher wage after acquiring experience in urban than in non-urban labor markets. Doubling local employment in all labor markets where experience was acquired, increases the productivity of a worker with two years of work experience by more than 0.7 percent. After 10 years of experience the corresponding gain amounts to about three percent, after 30 years to about four to five percent. A key factor seems to be an above average share of high-skilled labor within large urban labor markets which is supposed to enhance local learning opportunities.

JEL classification: R10, R23, J31

Keywords: Agglomeration economies, Human capital externalities, Learning, Urban wage growth premium, Transition to employment

*I gratefully thank Anette Haas, Silke Hamann, Henning Meier, Malte Reichelt, Anja Rossen and especially Johannes Bröcker and Annekatrin Niebuhr for their very helpful remarks and suggestions. Furthermore, I have benefited from fruitful discussions with seminar participants at the IEB/UEA Summer School in Urban Economics 2016, the Institute for Employment Research (IAB), Kiel University (CAU) and the Kiel Institute for the World Economy (IfW). The usual disclaimer applies.

1 Introduction

There is extensive empirical evidence that wages in large local labor markets are significantly higher than wages in smaller ones. Figure 1 illustrates disparities in regional wages for Germany. The difference in the regional wage level of the largest labor market region and the smallest amounts to almost 40 percent. Similar significant disparities are observed for other countries; see, e.g., Glaeser and Maré (2001) for the United States and Combes et al. (2008) for France. This raises the question, why firms in agglomerations¹ pay high wages and do not relocate to regions in which labor is cheaper.²

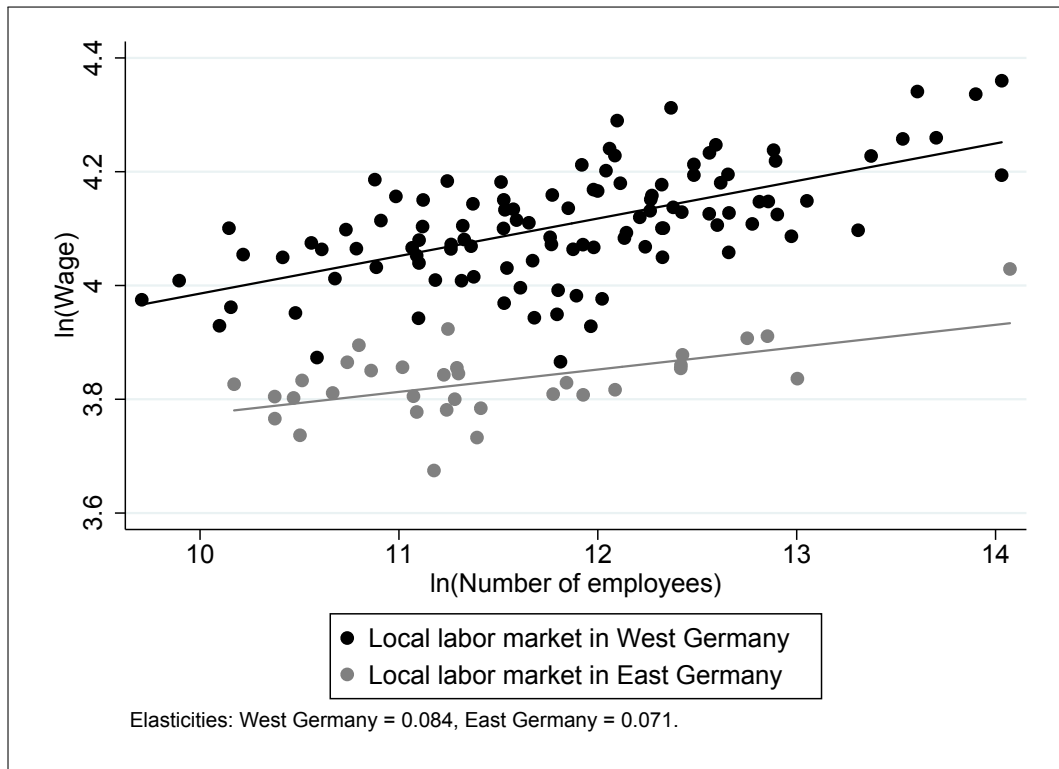


Figure 1: Local labor market size and regional wages.

Note: The figure refers to average regional wages that are paid at the beginning of new employment relationships starting in 2011 and the size of regional labor markets in terms of number of employees subject to social security contributions.

Glaeser and Maré (2001: 317) note, “The best explanation [...] is that these higher wages are compensated for by higher productivity.” They distinguish three possible reasons as to why workers might be more productive in urban than in non-urban labor markets. Most of the underlying theoretical arguments were already discussed in the 19th century by List (1838), Roscher (1878), and Marshall (1890). Nevertheless, it is still an ongoing issue as understanding the mechanisms behind this phenomenon refers to the fundamental question of urban economics about why cities exist (Glaeser/Maré, 2001).

The first explanation refers to the sorting of more able workers into cities, arguing that higher wages in cities reflect higher abilities of its workers. Regarding the importance of sorting, there still is an ongoing debate. While Combes et al. (2008) conclude ‘sorting matters’, De la Roca and Puga (2016) find that it is less important

¹ I use terms like ‘agglomeration’, ‘city’, ‘urban region’, ‘large local labor market’ interchangeably to improve readability.

² Explanations why workers do not fully concentrate in cities where wages are higher refer to higher costs of living and urban disamenities (Glaeser/Maré, 2001).

once dynamic agglomeration economies (see below) are considered. D’Costa and Overman (2014) in turn argue that sorting plays a role in explaining why wages in agglomerations grow faster.

The second explanation is static agglomeration benefits based on sharing, matching, and learning. Duranton and Puga (2004) discuss the different mechanisms in detail. Sharing refers to, among other things, benefits of sharing gains from variety and individual specialization. Matching advantages arise if the chances of matching and the quality of successful matches between job seekers and vacancies increase with the size of the local labor market. An example of learning effects that result in a static agglomeration benefit is where firms learn more from other firms that are located nearby and thereby improve their own production process. Static agglomeration benefits are supposed to generate an immediate wage increase if workers move from a non-urban to an urban environment, whereas leaving a city leads to an immediate wage decline (Glaeser/Maré, 2001). There is ample evidence that static agglomeration benefits exist. Recent studies find that doubling employment density leads to a wage increase by about one to two percent (e.g., Combes et al., 2008, De la Roca/Puga, 2016). Combes and Gobillon (2015) provide a comprehensive overview.

The third explanation is dynamic agglomeration benefits, which are the main interest of this paper. They are supposed to enhance individual wage growth in cities. There are primarily two mechanisms discussed in this context which were already mentioned by List (1838). He recognized, *inter alia*, that concentration of economic activity enhances the individual opportunities for workers to acquire new skills and that accessibility helps to find new jobs. In the recent discussion, the mechanisms are labeled learning and coordination, also referred to as dynamic matching. Similar to List’s considerations, the coordination hypothesis suggests that urban density makes it “easier for workers to find the best jobs for themselves” (Glaeser/Maré, 2001: 322). The learning hypothesis proposes that workers accumulate more human capital by working in an agglomerated urban environment than by working elsewhere and that these workers can make use of this knowledge even if they migrate to another labor market. The objective of this paper is to analyze the importance of this channel.

If working in a large labor market fosters individual accumulation of knowledge and thereby increases the human capital endowment of a worker, the marginal value of work experience should be higher, the larger the labor markets have been in which experience was acquired. Pioneered by the work of De la Roca and Puga (2016), recent findings support this hypothesis (see also Andersson et al., 2014, D’Costa/Overman, 2014, Matano/Naticchioni, 2016). These studies show that the value of work experience which was acquired in the largest cities of a country has a significant higher value than experience acquired elsewhere and that this manifests in higher wages. It raises the question of how the value of experience exactly depends on the size of the local labor markets in which it was acquired. To provide such insights, is the main contribution of this paper. In contrast to previous papers, I explicitly estimate the *elasticity* of wage with respect to the size of the local labor market in which work experience was acquired, employing a micro-econometric framework described by Combes and Gobillon (2015).

Using administrative data from the Institute for Employment Research (IAB) on individual employment biographies, I am able to follow a worker’s employment history back to 1975 and observe the size of all regional labor markets in Germany in which he or she acquired work experience. This information is used to estimate the extent to which wages depend on the size of the labor market where previous work experience was acquired. The wage information refers to individual new employment relationships in Germany starting between 2005 and 2011. These wages contain important information as to how firms value work experience when they hire a worker depending on the size of the labor market in which it was acquired. By considering the wages associated with new employment relationships, I reduce the risk of an omitted variable bias since other factors which

are unobserved by the econometrician gain importance for productivity as tenure increases, like professional development training offered by the firm (Hamann et al., 2016).

The identified dynamic agglomeration benefit should be strongly related to learning externalities. I control for sorting of more able workers into large labor markets and static agglomeration benefits by means of individual as well as region fixed effects. Furthermore, I include the number of previous employers in order to control for dynamic matching according to the coordination hypothesis. Time varying individual characteristics and the individual labor market biographies of the workers are considered, as well as time varying characteristics of the local industry.

A further contribution of this paper is that it explicitly takes into account that large local labor markets are typically not only characterized by a large workforce but also by an above average share of high-skilled labor. Presuming that those workers one learns from the most are the skilled ones, it is not clear-cut whether dynamic agglomeration gains in general, and learning benefits in particular, stem from the large local workforce or from the high share of high-skilled labor. This paper provides corresponding evidence.

The remainder of the paper is structured as follows. In section 2 I review the corresponding literature with a focus on learning externalities. In section 3 I describe the methodology of my empirical analysis and in section 4 the data set. In section 5 I discuss the results of my empirical analysis and finally, in section 6 I set out my conclusions.

2 Related literature

Puga (2010) as well as Combes and Gobillon (2015), who review the literature on the magnitude and causes of agglomeration economies, conclude that most empirical studies analyze the overall impact of local characteristics on some local outcome rather than the importance of the different mechanisms behind agglomeration economies. Furthermore, they note, the role of learning in cities in particular is still an open question and further work is needed to fully understand the impact of this underlying mechanism.

In his paper 'Learning in Cities' Glaeser (1999: 255) state that "the primary informational role of cities may not be in creating cutting edge technologies, but rather in creating learning opportunities for everyday people". He argues that agglomeration fosters individual learning opportunities since the latter increase with the rate of new contacts between individuals and this is highest in a dense urban environment.

Although it is still an open question how to provide micro-foundations for the assumption that the learning function increases with the local number of skilled workers one can learn from (Duranton/Puga, 2004), there is empirical evidence that wages in large cities grow faster than elsewhere and one mechanism behind this phenomenon may be learning externalities. Glaeser and Maré (2001) were the first who empirically analyzed why density increases individual productivity, addressing sorting of more productive workers into cities by means of individual fixed effects. To disentangle static and dynamic agglomeration benefits, Glaeser and Maré (2001) analyze wages of migrants within the United States, where they find evidence of both types of agglomeration benefits. Rural-to-urban migrants receive quickly a wage increase which points to static agglomeration benefits. In addition, the urban-to-rural wage premium increases after migration with the amount of time spent in cities which points to dynamic agglomeration benefits. Furthermore, the results support the hypothesis that dynamic agglomeration benefits are at least partly based on learning effects. Wages decline only little when a worker migrates from a metropolitan to a rural area.

As emphasized by Wheeler (2006) and Yankow (2006) dynamic agglomeration benefits may not only arise due to learning effects but also due to improved firm-worker matches over time. It relates to benefits of searching jobs in urban rather than in non-urban labor markets. In urban areas the number of job openings is larger and search costs are lower. According to this 'coordination hypothesis', wage growth in cities should be related to job changes rather than to faster wage growth on a job. To differentiate between learning and dynamic matching effects, both authors suggest to distinguish between within-job and between-job wage growth. If within-job wage growth is larger in urban than in non-urban labor markets, this supports the learning hypothesis. Evidence of faster between-job wage growth in big cities would support the idea of improved dynamic matching.

Similarly to Glaeser and Maré (2001), Wheeler (2006) as well as Yankow (2006) find that an important part of the urban wage premium is a wage growth effect. Like Glaeser and Maré (2001), both analyze wage effects based on migrants between rural and urban regions (and vice versa). Yankow (2006) focuses on the role of between-job wage growth as a driver of the urban wage growth premium. He finds that the wage gain per job change does not significantly differ between urban and non-urban regions. However, he observes that the number of job changes is larger in big cities. This suggests that parts of the dynamic agglomeration benefit are based on more frequent job changes in cities. The role of learning effects is not explicitly analyzed by Yankow (2006).

In contrast, Wheeler (2006) considers both: between-job and within-job wage growth. He finds evidence in favor of the coordination hypothesis, as labor market density has a positive impact on between-job wage growth. However, this result is not robust when individual fixed effects are included in the analysis. With regard to within-job wage growth, Wheeler (2006) finds no significant impact of the local labor market, using three different measures of local market scale. D'Costa and Overman (2014), who control for unobserved individual characteristics that impact on wage growth, obtain a similar result for Great Britain, when they compare wage growth of workers who move from rural to urban areas with workers who stay in rural areas. However, they observe for rural workers who "worked in a city at some point" (D'Costa/Overman, 2014: 174) a faster wage growth compared to rural workers with no city experience. Once city experience is taken into account, they find that workers who have either current or past experience in an urban environment benefit from faster wage growth than workers who have never had any urban experience. Furthermore, their results indicate that the wage growth of rural workers with past urban experience are based on within-job wage growth. D'Costa and Overman (2014) interpret this finding as evidence in favor of the learning hypothesis. Individuals with past urban experience are able to make use of their accumulated knowledge even after relocating to rural areas.

Lehmer and Möller (2010) examine, among other questions, whether faster wage growth in urban regions is tied to the firm level or whether there is evidence of externalities provided by the urban environment. They analyze the wage effects of mobility between firms as well as between regions for Germany, distinguishing between small and large firms as well as between rural and urban regions. Their findings suggest that wage gains of moving from small to large firms increase in the years after migration, pointing to faster knowledge accumulation within large firms. However, even controlling for this effect, Lehmer and Möller (2010) provide evidence that wages in urban areas grow at a higher rate than wages in rural areas. This supports the view that the urban environment fosters the individual accumulation of knowledge as well.

"Currently one of the most complete studies on the topic" (Combes/Gobillon, 2015: 263) is the one by De la Roca and Puga (2016).³ As with previous authors, they control for unobserved individual characteristics that

³ Combes and Gobillon (2015) refer to an earlier version of De la Roca/Puga (2016).

impact on the individual wage level and distinguish static and dynamic effects. De la Roca and Puga (2016) provide empirical evidence that strongly suggest that dynamic benefits depend on learning effects. In order to identify dynamic agglomeration economies, they do not analyze individual wage growth after migrating from rural to urban areas like the studies discussed so far, but instead analyze the value of individual work experience, depending on the size of the city in which it was acquired. They distinguish experience that was acquired in three groups of Spanish cities. Their results show that experience acquired in the largest Spanish cities has the highest value. In addition, experience acquired in the next largest cities is also more valuable than experience acquired in the rest of the country. It suggests that a worker accumulates in a large city per unit of experience more valuable human capital than in a small city. Similar results for Italy are obtained by Matano and Naticchioni (2016) focusing on young workers, and by Hamann et al. (2016) for Germany when analyzing transitions to employment. The latter find that employers value work experience higher when they hire a new employee if experience was acquired in agglomerations rather than in less dense local labor markets. This manifests in higher wages.

De la Roca and Puga (2016) also provide evidence for the learning hypothesis as a driver of the urban wage growth premium. They find that the additional value of experience acquired in the biggest Spanish cities does not significantly change when a worker leaves the largest cities to work somewhere else.

Some authors also provide evidence of heterogeneity in dynamic agglomeration benefits, depending on individual characteristics. The results by Glaeser and Maré (2001) suggest that the urban wage growth premium is larger for younger than for older workers. Results obtained by D'Costa and Overman (2014) support this finding.

De la Roca and Puga (2016) observe that particularly workers with a high 'innate ability', measured by individual fixed effects, benefit from working in large cities through learning effects. In contrast, Matano and Naticchioni (2016) find for Italy that workers at the bottom of the wage distribution experience a stronger wage increase in the years after a move from a province of low to a province of high density than workers at the top of the wage distribution. For the latter, they also observe a significant static wage premium. Matano and Naticchioni (2016) also distinguish benefits of firm tenure (as a proxy for within-job wage growth) and job changes. Their results indicate that workers at the bottom of the wage distribution benefit from tenure, pointing to learning effects. For workers at the top of the wage distribution tenure is according to Matano and Naticchioni (2016) less important. This group benefits rather from job changes, indicating that highly dense regions promote the efficiency of matches between firms and workers.

Gould (2007) and Andersson et al. (2014) also provide evidence of heterogeneous learning effects. Gould (2007) applies a structural model that accounts for the endogeneity of location choices, and finds for the United States that work experience in a city significantly increases the wage of white-collar workers, even after relocation to the rural area. For blue-collar workers he does not find such an effect. Andersson et al. (2014) distinguish workers according to tasks. They find that workers with a high fraction of non-routine tasks receive a higher wage after a move from Sweden's three largest metropolitan regions to the rest of the country than workers that previously did not work in one of the large local labor markets. For workers with a low fraction of non-routine tasks they do not observe corresponding wage differentials.

The literature discussed so far focuses on externalities arising from the spatial concentration of economic activity and employment. However, a related strand of literature exists that discusses the role of local learning externalities as well (Heuermann et al., 2010). This strand analyzes the external effects of localized human

capital. It focuses on the composition of the local workforce with respect to skills, rather than on labor market size. Comprehensive overviews are given by Moretti (2004b) and Duranton (2007).

To analyze human capital externalities, typically some indicator for productivity is regressed on the share of (high-)skilled workers in local employment. As summarized by Heuermann et al. (2010), one underlying mechanism on why local human capital impacts on wages is that firms invest more in physical capital when they are located in a human capital intensive region. This increases also the productivity of low-skilled workers through working with a larger stock of physical capital.⁴ Another reason why localized human capital is supposed to impact on individual productivity is learning externalities. Already List (1838) noted, the interaction of skilled and unskilled workers presumably leads to imitations by the latter. Furthermore, skilled workers are expected to benefit from each other because meetings between skilled workers with different ideas might generate new ideas and knowledge (Jovanovic/Rob, 1989).

Empirical evidence of productivity effects of local human capital is provided by, e.g., Moretti (2004c). He finds that the value added of manufacturing plants grows faster in those cities in the United States in which the share of college graduates increases at a higher rate than in other cities. However, most frequently human capital externalities are analyzed by comparing wages of workers across cities with differing levels of human capital (Heuermann et al., 2010), among others Moretti (2004a) for the United States and Heuermann (2011) as well as Braakmann (2009) for Germany.

One difficulty when analyzing the impact of an increasing local share of skilled labor on individual productivity is the identification of the externality effect. As stressed by Moretti (2004a) and Ciccone and Peri (2006) and reviewed by Combes and Gobillon (2015) not only human capital externalities are at work, but individual wages also depend on the composition of the local workforce due to the imperfect substitutability between workers of different skill levels. If this is not taken into account, only a composite effect is identified.

Furthermore, Combes and Gobillon (2015) note that most studies focusing on localized human capital externalities neglect that the degree of agglomeration, and the regional share of high-skilled labor, are typically correlated. Most of these studies only consider the share of high-skilled labor and no other regional characteristic in their empirical analysis. Therefore, the human capital variable likely also captures the effect of other regional characteristics such as labor market density (Combes/Gobillon, 2015). An exception is, e.g., Heuermann (2011) who controls for the density of the local labor market.

Vice versa, if only density is considered, as typically done by the literature on the urban wage premium discussed above, the identified effect might also capture the effect of the high share of high-skilled labor. However, Heuermann et al. (2010) as well as Combes and Gobillon (2015) argue that the latter effect may be interpreted as one *channel* of agglomeration economies. It may reinforce local learning opportunities in large labor markets.

Taken together, the existing literature on dynamic agglomeration economies predominantly analyzes whether workers migrating from non-urban to urban areas experience thereafter a faster wage growth than workers who remain in non-urban areas (Glaeser/Maré, 2001, Yankow, 2006, Lehmer/Möller, 2010, D’Costa/Overman, 2014, Matano/Naticchioni, 2016). Even though the results are heterogeneous, the overall impression is that

⁴ The argument for the larger investment in physical capital is that physical and human capital are complementary in the production process and that firms anticipate that the contemporary aggregate level of human capital determines future educational attainments (Heuermann et al., 2010).

wage growth in urban areas is in fact stronger than in non-urban areas. In addition, there is also evidence that workers who migrate from urban to non-urban areas thereafter still benefit from their work experience acquired in the urban labor market. This is typically interpreted as support for the learning hypothesis. However, the magnitude of dynamic gains, depending on the size of the local labor market in which experience was acquired, has not been analyzed by these studies.

Inspired by the work of De la Roca and Puga (2016), a few studies recently estimated the value of work experience by cities. They distinguish two or three classes of large cities and compare the value of work experience acquired there to the value of work experience acquired in the rest of the country. In line with the learning hypothesis they find a positive relationship between the value of city-class experience and the size of the cities within the class. This indicates that a worker accumulates more human capital per unit of experience the larger the local labor market is in which experience is acquired. However, general conclusions with regard to the impact of city size on learning benefits are difficult as these studies distinguish only very few cities. Furthermore, the underlying mechanisms remain unclear, i.e., which characteristics of urban labor markets are important for the emergence of learning externalities. Is it a scale effect, meaning the large number of workers who provide different ideas and knowledge or, as suggested by the literature on human capital externalities, the typically high share of high-skilled labor in urban labor markets?

Distinguishing experience acquired in 141 German labor market regions in the period 1975-2011, this paper explicitly estimates the impact of labor market size on dynamic agglomeration benefits. I also provide insights with regard to the importance of labor market size relative to the local share of high-skilled labor. The identified effects are supposed to be strongly related to learning externalities. I focus on wages associated with transitions to new employment relationships that indicate how firms value previous work experience when they hire a worker. Furthermore, I control for various other channels of agglomeration economies and show that the value of experience predominately is determined by the size of the labor market in which experience was acquired and portable to other regions.

3 Methodology

3.1 Empirical model

In order to quantify the benefit of acquiring work experience in large local labor markets, I analyze the wages associated with new employment relationships after transitions to employment. These wages contain important information on how firms value previously acquired work experience when they hire a new employee. As tenure increases, other factors which are unobserved by the econometrician will gain importance for productivity, like on-the-job and professional development training offered by the firm (Hamann et al., 2016). Hence, considering the wage of newly established employment relationships reduces the risk of an omitted variable bias.

In my analysis, I make use of a micro-econometric framework described by Combes and Gobillon (2015). Suppose a worker i is hired by an establishment at date t and the logarithm of the wage which the worker receives, $w_{i,t}$, is given by equation (1); compare Combes and Gobillon (2015: eq. 5.20).

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times v_{r(i,\tau),\tau,r(i,t)} + \mathbf{x}'_{i,t} \boldsymbol{\beta} + \varepsilon_{i,t} \quad (1)$$

u_i denotes an individual fixed effect and $\mu_{r,y(t)}$ a fixed effect for local labor market r , i.e., the labor market individual i starts to work in at date t . These fixed effects are allowed to vary across years y . $O_{i,\tau}$ is a dummy variable taking the value 1 if individual i was working in the past at date τ and 0 else. $v_{s,\tau,r}$ is a time varying fixed effect for working at date τ in region s and at date t in region r .⁵ $\mathbf{x}_{i,t}$ is a vector of time-varying individual characteristics with parameter vector β and $\varepsilon_{i,t}$ is an error term. In the empirical analysis subscript t refers to days between January 1, 2005 and December 31, 2011 and $\tau = 1$ to January 1, 1975.

The worker fixed effect captures all unobserved time-invariant characteristics of a worker that determine the individual wage. The region-year fixed effect refers to the region specific productivity level that is based on static local effects which may vary over time. Dynamic local effects are captured by $I(O_{i,\tau} = 1) \times v_{s,\tau,r}$. Given that individual i was working at date τ which is indicated by $O_{i,\tau}$, the parameter $v_{s,\tau,r}$ can be interpreted as the weight of the work experience that individual i acquired at date τ . This marginal value of experience depends on the labor market and date of acquirement, indicated by s and τ , respectively, and the labor market in which the experience was used, r . $\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times v_{s,\tau,r}$ captures the wage premium that individual i receives at date t due to his or her entire work experience acquired until date $t - 1$.⁶

Imagine, I ignore that the value of experience depends on the time and the location of its acquirement and on where it is used, then, $v_{s,\tau,r} = v$ and (1) reduces to a frequently applied Mincer type wage equation:

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + vE_{i,t} + \mathbf{x}_{i,t}'\beta + \varepsilon_{i,t}, \quad (2)$$

where $E_{i,t}$ denotes the total amount of work experience which is given by the number of days of employment until date $t - 1$, i.e., $E_{i,t} = \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1)$. In both equations, (1) and (2), v refers to the marginal value of work experience. In (1) it is allowed to vary across labor markets, in (2) it is not.

In order to identify dynamic agglomeration economies based on (1), Combes and Gobillon (2015) suggest a two-step procedure. The first step requires to estimate $v_{s,\tau,r}$ for all combinations of s , t and r , in addition to all other parameters of equation (1). In a second step, the estimated parameters $\hat{v}_{s,\tau,r}$ are regressed on local characteristics such as labor market size to quantify its impact on the marginal value of experience. A severe problem of this approach is that the number of parameters which have to be estimated in the first step amounts to the square of the number of labor market regions times the number of considered dates (Combes/Gobillon, 2015). Moreover, it is not possible to take into account that the marginal return to experience acquired in labor market s at date τ may vary across individuals depending on the amount of previously acquired experience (see below).

De la Roca and Puga (2016) estimate a special case of equation (1) where they impose restrictions on $v_{s,\tau,r}$ such that the number of parameters reduces. Using Spanish data, they distinguish experience only according to three classes of cities in which it was acquired: in Madrid/Barcelona, Valencia/Sevilla/Zaragoza, or elsewhere in Spain. Furthermore, they assume that the value of experience acquired in one of the classes does not depend on the day when it was acquired. In this case, equation (1) simplifies to (3); compare De la Roca and Puga

⁵ It is important to distinguish the subscripts s and r . $s = r(i, \tau)$ denotes the labor market in which worker i acquired experience at date τ . $r = r(i, t)$ indicates in which local labor market individual i starts to work at date t .

⁶ Because information on self-employment is not available, I only consider spells of employment when measuring work experience. I assume that knowledge based on work experience does not depreciate over time; compare, e.g., De la Roca and Puga (2016), Matano and Naticchioni (2016).

(2016: eq. 1).

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + \sum_{s=1}^3 v_{s,r(i,t)} E_{i,s,t} + \mathbf{x}_{i,t}' \boldsymbol{\beta} + \varepsilon_{i,t} \quad (3)$$

$E_{i,s,t}$ denotes the amount of experience that was acquired until $t - 1$ in, respectively, Madrid or Barcelona, Valencia, Sevilla, or Zaragoza, or elsewhere in Spain. This approach enables to compare the marginal value of experience acquired in different groups of cities to assess the magnitude of dynamic agglomeration benefits. However, to obtain general results with regard to the impact of labor market size on dynamic gains is difficult. The two-step procedure suggested by Combes and Gobillon (2015) requires in the first step not only few parameters $v_{s,\tau,r}$ to be estimated, but many.

I adopt the strategy by De la Roca and Puga (2016) as a point of departure and estimate a model such as (3) where I distinguish only between two groups of German labor markets. Thereafter, I apply an alternative estimation strategy that is also discussed by Combes and Gobillon (2015) in order to obtain more general results. I express $v_{s,\tau,r}$ in (1) by a function of labor market size, equivalent to the second step of the two-step procedure and estimate to which extent wages associated with new employment relationships depend on the size of the labor market in which experience was acquired. At this point, I assume that the value of experience is only determined by the size of the labor market where experience was acquired and does not vary depending on where it is used, i.e., $v_{s,\tau,r} = v_{s,\tau}$. This assumption is relaxed at the end of my analysis. Furthermore, I follow Combes et al. (2008), Andersson et al. (2014), De la Roca and Puga (2016), and others, and assume that there is a log-linear relationship between net-agglomeration gains and labor market size denoted by $Z_{s,\tau}$. More precisely, I assume that $v_{s,\tau}$ is given by equation (4). It may be interpreted as a learning function capturing externalities of acquiring work experience in large local labor markets.⁷

$$v_{s,\tau} = \begin{cases} \delta \ln \left(\frac{Z_{s,\tau}}{\underline{Z}} \right) & \text{if } Z_{s,\tau} > \underline{Z} \\ 0 & \text{else} \end{cases} \quad (4)$$

In accordance with Duranton and Puga (2004), marginal learning benefits with regard to labor market size are assumed to be positive but decreasing. $\delta \ln \left(\frac{Z_{s,\tau}}{\underline{Z}} \right)$ denotes the value of the additional human capital that a worker accumulates by working at date τ in a labor market with size $Z_{s,\tau}$, given the labor market size exceeds a certain (unknown) threshold \underline{Z} . If the local labor market is smaller than \underline{Z} , a worker does not acquire any knowledge he or she can make use of in the future.⁸ In order to obtain the baseline model of my empirical analysis, I rewrite the learning function such that it is given by:

$$v_{s,\tau} = \begin{cases} \gamma + \delta \ln Z_{s,\tau} & \text{if } Z_{s,\tau} > \underline{Z} \\ 0 & \text{else,} \end{cases} \quad (5)$$

with $\gamma = -\delta \ln \underline{Z}$. Given that $Z_{s,\tau}$ is always larger than \underline{Z} , inserting (5) into (1) leads to:

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + \gamma E_{i,t} + \delta \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau} + \mathbf{x}_{i,t}' \boldsymbol{\beta} + \varepsilon_{i,t}, \quad (6)$$

with $z_{s,\tau} = \ln Z_{s,\tau}$ and $E_{i,t}$ being individual work experience measured by the number of days a worker was

⁷ Combes and Gobillon (2015) state that external effects do not need to depend only on the size, but also on other regional characteristics of labor market s . Later in my paper, I also take into account the skill structure of the local labor force. Thereby, I also address that Duranton and Puga (2004) argue that the learning function depends on the local amount of *skilled* labor.

⁸ I gratefully thank Johannes Bröcker for the suggestion to introduce \underline{Z} in the learning function.

employed before date t (compare equation (2)). After estimating equation (6) it has to be verified whether all local labor markets are indeed larger than \underline{Z} which is given by $\exp(-\hat{\gamma}/\hat{\delta})$. If some local labor markets are smaller than \underline{Z} , the size of the respective labor market has to be set to \underline{Z} (leaving $E_{i,t}$ unchanged) and an iterative procedure has to be applied in order to obtain the solution for equation (4).

$\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau}$ in equation (6) may be interpreted as the total number of local learning opportunities of individual i until date $t - 1$ depending on the size of the labor markets in which experience was acquired. The corresponding parameter δ denotes the elasticity of wage with respect to the size of the labor market in which individual i was working at one day in the past. Accordingly, the productivity at date t as reflected in wage would be $\delta E_{i,t}$ percent higher if all labor markets individual i was working in before would have been on average one percent larger. For interpretation it is important to note that the partial derivative of $w_{i,t}$ with respect to $v_{s,\tau}$ is unity, given that individual i was employed at date τ (see equation 1). Therefore, the estimated *elasticity of wage* with respect to the size of the labor market in which experience was acquired can also be interpreted as the *elasticity of the marginal value of experience* with respect to labor market size (compare equation 4).

So far, I implicitly assumed that the amount of *already accumulated human capital* does not impact on the return of acquiring an additional day of experience with respect to the individual knowledge stock. Results by De la Roca and Puga (2016) suggest that this assumption is too restrictive. They show that the marginal value of experience decreases in the number of days an individual worked in the past and, especially, if the employers were located in a large labor market. Taking this into account, I augment the learning function such that the value of work experience acquired by working at date τ , $I(O_{i,\tau} = 1) \times v_{s,\tau,r}$ in equation (1), is given by equation (7).⁹ θ denotes the rate by which the marginal return to experience decreases in the amount of human capital that individual i accumulated before date τ , $Q_{i,\tau}$.¹⁰

$$I(O_{i,\tau} = 1) \times v_{s,\tau,i} = \begin{cases} \delta \ln\left(\frac{Z_{s,\tau}}{\underline{Z}}\right) - \theta Q_{i,\tau} & \text{if } Z_{s,\tau} > \underline{Z} \text{ and } O_{i,\tau} = 1 \\ 0 & \text{else} \end{cases} \quad (7)$$

with $Q_{i,1} = 0$ and

$$Q_{i,\tau} = \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times v_{r(i,k),k,i} \quad (8)$$

for $\tau > 1$.

Since $I(O_{i,\tau} = 1) \times v_{s,\tau,i}$ is defined as $Q_{i,\tau+1} - Q_{i,\tau}$, equation (7) corresponds to the first order difference equation given by (9).¹¹ Its solution is given by equation (10), where $\tilde{E}_{i,\tau}$ denotes the number of days of employment before date τ in labor markets *larger than* \underline{Z} . For details see appendix A.

$$Q_{i,\tau+1} = \begin{cases} \delta \ln\left(\frac{Z_{s,\tau}}{\underline{Z}}\right) + (1 - \theta)Q_{i,\tau} & \text{if } Z_{s,\tau} > \underline{Z} \text{ and } O_{i,\tau} = 1 \\ Q_{i,\tau} & \text{else} \end{cases} \quad (9)$$

⁹ In equation (7), I still assume that the value of previously acquired work experience is the same in all labor markets.

¹⁰ θ is supposed to be in the interval $(0; 1)$. The marginal value of experience acquired in labor market s at date τ now varies across individuals. Therefore, I add the subscript i .

¹¹ I gratefully thank Johannes Bröcker for pointing this out and to give me the decisive hint how to obtain the solution of (9) referring to Sydsæter et al.

$$Q_{i,\tau} = \delta \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times \max \left\{ \ln \left(\frac{Z_{r(i,k),k}}{\underline{Z}} \right); 0 \right\} \times (1 - \theta)^{\tilde{E}_{i,\tau} - \tilde{E}_{i,k+1}} \quad (10)$$

Hence, the value of human capital acquired before date τ , $Q_{i,\tau}$, is proportional to the weighted sum of the size of all labor markets in which previous work experience was acquired. For $k = \tau - 1$, the weight $(1 - \theta)^{\tilde{E}_{i,\tau} - \tilde{E}_{i,k+1}}$ equals one, meaning that the size of the labor market in which a worker acquired the last day of work experience is not discounted. In contrast, the size of the labor market where the first day of experience was acquired has the lowest weight, $(1 - \theta)^{\tilde{E}_{i,\tau} - 1}$, addressing that the human capital acquired at this day, diminishes the amount of human capital that is additionally acquired at any day in the future.

In order to obtain estimates for θ , δ , and \underline{Z} based on the augmented learning function given by equation (7), I again replace $\ln \underline{Z}$ by $-\gamma/\delta$ such that equation (10) can be written as:

$$\begin{aligned} Q_{i,\tau} = & \gamma \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times (1 - \theta)^{\tilde{E}_{i,\tau} - \tilde{E}_{i,k+1}} + \\ & + \delta \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times \max \{ \ln Z_{r(i,k),k}; \ln \underline{Z} \} \times (1 - \theta)^{\tilde{E}_{i,\tau} - \tilde{E}_{i,k+1}}. \end{aligned} \quad (11)$$

Replacing the term $\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times v_{r(i,\tau),\tau,r(i,t)}$ in equation (1) by this expression yields equation (12).¹² Recall that both, $\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times v_{r(i,\tau),\tau,r(i,t)}$ as well as $Q_{i,t}$, denote the value of experience acquired before date t .

$$\begin{aligned} w_{i,t} = & u_i + \mu_{r(i,t),y(t)} + \gamma \sum_{k=1}^{t-1} I(O_{i,k} = 1) \times (1 - \theta)^{E_{i,t} - E_{i,k+1}} + \\ & + \delta \sum_{k=1}^{t-1} I(O_{i,k} = 1) \times z_{r(i,k),k} \times (1 - \theta)^{E_{i,t} - E_{i,k+1}} + \mathbf{x}_{i,t}' \boldsymbol{\beta} + \varepsilon_{i,t} \end{aligned} \quad (12)$$

Equation (12) takes into account, that workers accumulate less human capital per day, the more human capital he or she has previously acquired. This depends on the number of previous days of employment and the size of the labor markets in which previous employers were located. The crucial parameter is θ . Since equation (12) is highly non-linear in θ , I run several regressions with different values for θ . Once a value is imposed, the linear fixed effects estimator can be used to estimate the remaining parameters. I check afterwards which specification minimizes the residual sum of squares. For $\theta = 0$, equation (12) simplifies to the baseline model given by equation (6).

After estimating equations (6) and (12) which are based on the assumption that the value of experience is the same in all labor markets, I let the parameters of the learning function vary across different groups of local labor markets which are defined according to their size. This provides evidence as to whether the value of experience depends only on the size of the labor market in which experience was acquired or also on the size of the labor market in which it is used.

¹² Again, I assume that all labor markets are larger than threshold \underline{Z} which also implies that $\tilde{E}_{i,t} = E_{i,t}$. This assumption has to be verified after an estimate for \underline{Z} is obtained.

3.2 Econometric issues

When analyzing the impact of labor market size on wages, endogeneity has to be taken into account. A detailed discussion is provided by Combes et al. (2011). The risk of biased estimates due to omitted variables should be reduced by the setup of my empirical analysis. I control for all time-invariant characteristics of the worker by means of individual fixed effects as well as for time-varying characteristics like educational degree and pre-employment status. Region-year fixed effects control for all time-variant and time-invariant characteristics of the local labor market that lead to static regional wage differentials. In addition, I consider observable characteristics of the local industry, skill specific local unemployment rates as well as industry fixed effects. The latter capture all time-invariant industry specific factors that have an impact on the wage.

Also the risk of reverse causality should be of minor concern here. The pivotal explanatory variable is the size of the labor markets in which an individual acquired experience *before* date t , not the size of the labor market at the date at which the analyzed wage is paid. Of course, the size of the labor market in which experience has been acquired and where it is used are likely significantly correlated. However, the included region-year fixed effects control for all characteristics of the region in which individual i works at date t , including its size.

As noted by Combes and Gobillon (2015), a second econometric issue that has to be discussed refers to the computation of standard errors. Moulton (1990) shows that it is important to account for cross-sectional correlation in the error terms if a model explains individual outcomes by characteristics of the regional environment. As matrix $\mathbf{X}_{i,t}$ also contains information of the local industry, the appropriate solution would be to cluster the standard errors at the local industry level. However, this is not possible as the model includes individual fixed effects and because workers are mobile between regions and industries. The standard errors that I report are clustered at individual level. They are robust with regard to heteroskedasticity and serial correlation in the error terms (Wooldridge, 2013).¹³

4 Data

4.1 Individual data set

In order to quantify the impact of labor market size on the value of work experience, I analyze wages of 525,265 new employment relationships in Germany within the period 2005 to 2011. The new employment relationships are identified using detailed information on individual labor market biographies. The latter also enable to observe where and when work experience was acquired as information on the date and location of previous spells of employment is available.

The information on labor market biographies is taken from the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB). Among other information, the IEB contains very reliable micro data on employment which comes from the integrated notification procedure for health, pension and unemployment insurance.¹⁴ My data set comprises a 5 percent random sample of all employees with at least one notification

¹³ I obtain very similar standard errors if I compute them as suggested by Driscoll and Kraay (1998). Those standard errors are robust to general forms of cross-sectional and serial dependence in the error terms (see also Hoechle, 2007). The additional results are available upon request.

¹⁴ For a more detailed description of the IEB see vom Berge et al. (2013).

to social security between 2005 and 2011.¹⁵

I exclude all individuals for which it is not possible to observe the full employment biography. The setup of my analysis requires information on all locations where previous work experience was acquired.¹⁶ Among others, I exclude all individuals born before 1960 because the IEB contains only information on employment from 1975 onwards. A detailed description is provided in appendix B. For the remaining sample of workers my data set captures all spells of employment subject to social security contributions. I use them to identify transitions to employment within the period between 2005 and 2011 focusing on new full-time employment outside the public sector and outside the temporary employment industry. For the new employment relationships I observe the corresponding gross daily wage as well as further information on the new job (e.g., kind of occupation) and the worker (e.g., age, educational attainment and sex). The wage information in the IEB is right censored, as firms report earnings only up to the upper limit for social security contributions. Therefore, I partly impute the wages, see appendix B.

Using information on all previous spells of employment, I compute the individual labor market experience at the date at which the new employment relationship starts. Moreover, I generate important control variables that provide information on the recent labor market biography with respect to the pre-employment status, length of non-employment before the transition to employment, unemployment benefits, and participation in measures of active labor market policy. The information is also taken from the IEB. Detailed information on all variables that I use is provided by table B.1 in appendix B. Summary statistics can be found in table C.1 in appendix C.

Descriptive statistics indicate that a large share of the considered new employment relationships refers to rather young workers with few years of labor market experience. One likely explanation is that workers change jobs more frequently when they are young to find the job they like most as noted by Wheeler (2008). The mean work experience in my data set amounts to about 9.4 years. The analyses by Glaeser and Maré (2001), Wheeler (2006), Yankow (2006) as well as Gould (2007) are based on a data set with a similar sample mean.¹⁷

The establishment identifier in the IEB is used to identify the number of different establishments an individual was working in before the considered new employment relationship. I use this information as a control variable addressing that frequent job changes are discussed to improve the matching between firms and workers.¹⁸ Additionally, the establishment identifier allows me to add important information on the establishment to the individual data set, e.g., location, industry, establishment size as well as skill and age structure of the staff. The data is taken from the Establishment History Panel (BHP).

I also merge information on the local industry as well as on skill specific local unemployment rates. I compute corresponding indicators based on data taken from the (un-)employment statistics of the Federal Employment Agency (FEA).

4.2 The size of local labor markets

Di Addario and Patacchini (2008) show for Italy that the effect of local density on wages rapidly declines in

¹⁵ Hamann et al. (2016) use the described data set to analyze the impact of labor market density on the quality of matches between job seekers and vacancies.

¹⁶ The IEB does not contain information on the self-employed and civil servants. Therefore, my analysis is based on information on employment subject to social security contributions only.

¹⁷ Gould (2007) uses the same data set as the other authors, but does not provide summary statistics in his paper.

¹⁸ Different units of one firm that are located in different municipalities are considered as independent establishments. It is not possible to identify whether different establishments belong to the same firm.

distance. Only for population mass within a distance up to 12 kilometers, they find a significant impact on wages. It indicates that agglomeration benefits depend on the immediate environment.

Taking this into account and addressing that learning externalities are thought to crucially depend on interaction between individuals, I choose labor market regions as spatial units for my analysis. They are supposed to represent integrated local labor markets and have on average a radius of about 27 km (see table 1). Their definition is taken from Kosfeld and Werner (2012), who define 141 regional labor market regions employing the factor analysis to German commuter structure between NUTS-3 regions. Because the regions contain NUTS-3 regions connected through intense commuting, I assume that workers exchange knowledge exclusively within these regions and that their size in terms of employment is an appropriate measure for local learning opportunities.

One critical issue is that commuting patterns change over time, wherefore the chosen definition of labor market regions might be an inappropriate approximation of local labor markets decades ago. However, when estimating equation (6) which is based on the simplified learning function, the within transformation that is applied to eliminate the individual fixed effect wipes out all information on experience that was acquired before the date at which the first analyzed new employment relationship of worker i started. Hence, in the baseline specification $\hat{\delta}$ and $\hat{\gamma}$ are only based on information that refers to the period 2005 to 2011. Estimating equation (12) which is based on the augmented learning function, requires information on all previous employment spells of a worker. In this case, the size of the labor market in which experience was acquired is weighted lower the more additional experience was acquired thereafter; see also equation (10). Nevertheless, as a robustness check, I repeat my estimations focussing on those workers who only acquired work experience in 1995 or later, i.e., at least five years after reunification of East and West Germany. Overall, the results based on the full sample are confirmed as discussed in the next section.

Table 1: Size of German labor market regions

	Local labor market size in terms of. . .		
	radius in km [†]	thousand employees [‡]	share of high-skilled labor [§]
Minimum	10.3	13.8	2.3
Median	26.0	110.8	6.0
Mean	27.0	179.1	7.0
Maximum	51.8	1206.5	20.0

N=141.

[†] Under the assumption that the regions are circular. [‡] Number of employees subject to social security contributions. [§] Share of workers with a university degree.

Note: Definition of the regions according to Kosfeld and Werner (2012). The statistics are averages for the considered period, i.e., for West Germany 1975-2011, for East Germany 1993-2011.

Local learning externalities should crucially depend on the amount of localized knowledge and ideas. I approximate this amount by the number of employees working in a labor market. The data is taken from the Employment History of the IAB and refers to annual regional employment at June 30th. The size of the labor market regions in terms of employment varies between 14 thousand and 1,2 million employees (region specific averages, see table 1).

After identifying the effect of total regional employment on the value of experience as reflected in higher future wages, I analyze to which extent the dynamic agglomeration benefit may be explained by an above average share of high-skilled labor within large urbanized labor markets. I define this as the local share of workers with

a university degree which varies between two percent and 20 percent (see table 1). The correlation between total regional employment and the local share of high-skilled workers is 0.38.

4.3 The mobility of labor

The mobility of labor is an important aspect as it determines to which extent learning benefits manifest in regional wage disparities. Considering a certain degree of learning externalities, wage disparities between small and large local labor markets arise if workers are immobile between local labor markets that differ in size since then, individuals who work in large local labor markets accumulated in the past, *ceteris paribus*, on average more human capital than individuals working the same number of days in small local labor markets.

Descriptive statistics show that a worker acquired work experience in either small or large local labor markets. Mobility between labor markets that significantly differ in size is rather low. Table 2 provides information on the ratio of the size of the largest labor market and the size of the smallest labor market in which a worker acquired experience before one of the considered new employment relationships. In more than 60 percent of the observations the largest labor market was less than twice as large as the smallest labor market, and in less than ten percent more than ten times as large.¹⁹

Table 2: The size of the largest labor market in which a worker acquired experience relative to size of the smallest labor market

Percentile	Size of largest labor market Size of smallest labor market
5	1.020
10	1.037
20	1.059
30	1.085
40	1.145
50	1.268
60	1.806
70	3.055
80	4.879
90	9.004
95	14.181

N=525,265.

Note: For each considered new employment relationship I identified the largest and the smallest labor market in which the respective worker previously acquired experience. The size of the local labor markets is measured in terms of number of employees and varies across years.

Table 3 compares the size of the labor markets where experience was acquired and where it is used distinguishing five categories of labor market regions according to their size. In 61 percent of all considered new employment relationships the new employer is located in a region that belongs to the same category as the average region in which experience was acquired (main diagonal), in 25 percent the current region is larger than previous ones (upper triangle) and in 14 percent smaller (lower triangle).

Workers are not only rather immobile between local labor markets that belong to different categories as defined in table 3, but also between labor markets that are of a comparable size. The sample mean of total work experience is 9.4 years (see table C.1). On average, 59 percent of this previous work experience (5.7 years)

¹⁹ Recall that the largest labor market is more than ten times as large as the median labor market (see table 1).

refer to employment in the same labor market in which the new employer is located. The median amounts to 79 percent. 36 percent of the workers acquired all their work experience in that particular region. On the opposite, 25 percent of the workers were never previously employed in the region in which the new employer is located.

Taken together, workers tend to continue to work in the labor market in which they acquired experience, or in a region of a comparable size. Therefore, if the individual accumulation of knowledge is significantly larger in large than in small local labor markets, as proposed by the learning hypothesis, it likely results in wage disparities between small and large regions.

In order to test the hypothesis that the value of work experience depends not only on the size of the labor market in which experience was acquired, but also on the size of the labor market in which it is used, it requires that at least some workers start to work in a region that is of a different size to the average region in which individual experience was acquired. Despite the rather low degree of mobility, there presumably is sufficient variation to test this hypothesis. Corresponding results are reported at the end of the next section. Before this, the next section discusses the results for the control variables and the impact of the size of the labor market in which experience was acquired on experience's value and future productivity as reflected in wages.

Table 3: The size of labor market in which experience was acquired and in which it is used

Share in %		Size of labor market in which experience is used in thousand employees					Total
		≤ 125	125-250	250-500	500-1000	> 1000	
Average size of	≤ 125	11.88	3.74	1.34	0.57	1.14	18.68
labor market in which	125-250	3.07	16.91	3.92	1.37	2.09	27.35
a worker acquired	250-500	1.72	3.53	16.94	3.00	2.70	27.89
experience in	500-1000	0.79	1.61	1.86	8.72	5.42	18.39
thousand employees	> 1000	0.37	0.42	0.25	0.21	6.45	7.69
Total		17.82	26.20	24.31	13.88	17.80	100.00

N=525,265.

Note: The average size of the labor market in which a worker acquired experience denotes the size of the different labor markets where the worker acquired experience weighted by the length of employment in the respective labor market.

5 Results

5.1 Control variables

Table 4 summarizes the regression results for the control variables that I use in my analysis. Column 1 refers to ordinary least squares estimation (OLS). Since the available information on educational degrees are only imperfect measures of skills, OLS results are likely biased. To address that workers might sort on unobserved abilities into large labor markets, I include individual fixed effects (FE, specification (2)) as introduced by Glaeser and Maré (2001). This leads to a reduction of my sample, since 188,979 observations refer to workers, for whom I only observe one new employment relationship between 2005 and 2011. The fixed effects estimation is therefore based on 336,286 observations for 126,862 employees.

The comparison of OLS and FE results shows that the OLS estimates are in most cases biased upwards. However, the sign of the estimated coefficients is in both models almost always the same and as expected. The larger the highest educational degree of a worker, the larger is the wage at the beginning of the considered

new employment relationship. For example, workers with a university degree receive a 26 percent ($e^{0.233} - 1$) higher wage than workers with a secondary/intermediate school leaving certificate and a completed vocational training.

Since the wage rate that I use as a dependent variable refers to the average wage rate that is paid until December 31 of the year in which the employment relationship starts (see appendix B), I include the length of the considered spell of employment, measured in months. It confirms that tenure affects the wage rate positively, the monthly increase amounts to 0.8 percent.

As results by Yankow (2006) suggest, dynamic agglomeration economies are based not only on learning effects, but also on a higher quantity of matches between workers and firms. Therefore, I include the number of previous employers as control variable, meaning the number of different establishments a worker was working in until the considered new spell of employment starts. If mobility between establishments enhances the quality of matches, the number of previous employers has a positive impact on wages. The empirical results confirm this hypothesis. I do find the expected positive impact when controlling for unobserved individual characteristics. The corresponding elasticity amounts to about 0.06.

With respect to the pre-employment status of a worker, the results show that the longer the period of non-employment before the considered new spell of employment, the lower is the corresponding wage. Non-employment of more than one year results in a wage loss of about four percent. Following Mincer and Ofek (1982), a reasonable explanation is that human capital depreciates during periods of non-employment. However, it is worth noting that Mincer and Ofek (1982) also provide evidence that eroded human capital is restored after an individual returns to work. This is important as I assume that the value of work experience does not depreciate over time even if workers were at some point in the past out of work.

In addition, the estimation results indicate that a worker receives a 2.5 percent lower wage if he or she was registered by the Federal Employment Agency as a job seeker before the considered transition to employment than if he or she was not. This points to a selection effect. If workers have good labor market expectations they register relatively seldom as a job seeker. A similar explanation likely holds for the negative effect of participation in measures of active labor market policies. Furthermore, workers who received unemployment benefit (ALG I) before the considered transition to employment have a 0.8 percent higher productivity as reflected in wage than workers who received no public assistant benefit or unemployment assistance. Again, this likely is related to proximity to the labor market. Unemployment benefit in Germany is only paid within the first year after the end of an employment spell of at least one year (with exceptions). Thereafter, either no public assistant benefit or unemployment assistance is paid, depending on the wealth of the household.

In order to address heterogeneity in firm productivity, I include establishment size and information on the firm's workforce composition with respect to the skill level and the age of the workers. Furthermore, time-invariant heterogeneity across firms belonging to different industries is captured by industry fixed effects. The results confirm that large firms and firms with a more skilled labor workforce are more productive than others and pay higher wages. Doubling an establishment workforce comes along with an about one percent higher productivity and a ten percentage point increase in the share of high-skilled workers with a 0.6 percent higher productivity. The age structure of a firm's workforce is correlated with individual productivity as well. A shift from younger towards middle aged or older workers comes along with higher wages. A reasonable explanation for this are complementarities between differently aged workers. More than 50 percent of the analyzed wages refer to young workers. Therefore, the positive coefficients may be explained by a high productivity of young workers

Table 4: Results for control variables

Dependent variable: logarithmic wages in new employment relationships				
	OLS		FE	
	(1)		(2)	
Education:				
Secondary / intermediate school leaving certificate				
... without completed vocational training	−0.131***	(0.014)	−0.023	(0.027)
... with completed vocational training		<i>Reference</i>		
Upper secondary school leaving certificate				
... without completed vocational training	−0.019	(0.015)	−0.084***	(0.029)
... with completed vocational training	0.095***	(0.002)	0.012**	(0.005)
Completion of a university of applied sciences	0.246***	(0.009)	0.167***	(0.016)
College / university degree	0.373***	(0.009)	0.233***	(0.016)
Female worker	−0.163***	(0.001)		
Tenure	0.012***	(0.002)	0.008***	(0.002)
ln(Number of previous employers)	−0.015***	(0.001)	0.058***	(0.002)
Length of non-employment				
0-28 days (job-to-job transition)		<i>Reference</i>		
28-92 days	−0.053***	(0.001)	−0.030***	(0.002)
93 days - 1 year	−0.073***	(0.001)	−0.034***	(0.002)
> 1 year	−0.088***	(0.001)	−0.043***	(0.002)
Pre-employment status				
Not registered as job seeker		<i>Reference</i>		
Unemployed and registered as a job seeker	−0.063***	(0.002)	−0.025***	(0.002)
Not unemployed, but registered as a job seeker	−0.071***	(0.001)	−0.026***	(0.002)
Participation in measures of active labor market policy	−0.024***	(0.002)	−0.012***	(0.002)
Public assistance benefits				
No benefit		<i>Reference</i>		
Unemployment benefit (ALG I)	0.012***	(0.002)	0.008***	(0.002)
Unemployment assistance (ALG II, ALHI)	−0.030***	(0.002)	−0.001	(0.002)
ln(Number of workers within the establishment)	0.033***	(0.000)	0.017***	(0.000)
Share of high-skilled workers in establishment	0.173***	(0.003)	0.057***	(0.004)
Share of low-skilled workers in establishment	−0.059***	(0.002)	−0.025***	(0.003)
Share of middle aged workers in establishment	0.131***	(0.003)	0.076***	(0.003)
Share of older workers in establishment	0.099***	(0.003)	0.072***	(0.004)
ln(Employment share of the industry within the region)	0.008***	(0.001)	0.001	(0.001)
ln((Herfindahl index based on local industry shares) ^{−1})	−0.003	(0.009)	−0.020	(0.014)
Share of high-skilled workers in local industry	0.160***	(0.010)	0.086***	(0.012)
Share of low-skilled workers in local industry	−0.022***	(0.009)	−0.008	(0.011)
ln(Local unemployment rate among high-skilled labor)	−0.002	(0.006)	−0.083***	(0.010)
ln(Local unemployment rate among skilled labor)	0.011*	(0.006)	−0.023**	(0.009)
ln(Local unemployment rate among low-skilled labor)	0.033***	(0.007)	−0.013	(0.012)
Constant	3.702***	(0.035)	3.684***	(0.052)
Observations	525,265		336,286	
OLS: R ² , FE: within R ²	0.616		0.183	
Worker fixed effects	No		Yes	

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include industry, occupation, and region-year fixed effects as well as total work experience and the pivotal variable $\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau}$, see equation (6) and column (1) and (2) of table 6.

if their share is low. This interpretation is inline with results obtained by, e.g., Garloff and Roth (2016). They show that productivity of young workers is higher, the lower their share is in the local workforce.²⁰

The agglomeration economies literature points out that the local industry structure also determines productivity. As for example formally shown by Duranton and Puga (2004), localization economies generate advantages to urban specialization if agglomeration causes congestion costs. Therefore, I control for localization economies by using the local industry share. To address that industrial diversity might also be beneficial due to urbanization externalities as suggested by Jacobs (1969) and formally shown by Duranton and Puga (2001), I follow Combes et al. (2004) and consider the inverse of a Herfindahl index based on the industry shares within local employment. If all industries have an equal share in the local industry, the inverse of a Herfindahl index corresponds to the number of locally operating industries. When industries have unequal shares, it indicates the ‘equivalent’ number of industries, i.e., the number of industries that would result in the same degree of industrial concentration, given equal industry shares.²¹ The results of the fixed effect model suggest that neither the share of the own industry nor the diversity of the industry structure in the local labor market affects individual productivity.

As discussed in the literature review, another strand of literature focuses on the impact of the amount of localized human capital on wages referring to human capital externalities (see, e.g., Moretti, 2004a, Ciccone/Peri, 2006). Parts of these effects are captured by the region fixed effects included in the model. In addition, I also consider the industry specific regional share of high and low-skilled labor in my analysis. The empirical results point to a positive effect of the local industry specific share of high-skilled labor. A ten percentage points increase in this share is associated with a 0.9 percent higher productivity. However, for the interpretation of this contemporaneous effect it is important to recall that only the composite of an externality effect and a substitution effect is identified. To identify learning benefits, which also might depend on human capital externalities, this paper does not focus on the analysis of contemporaneous effects, but on benefits of having *previously* worked in an urbanized labor market with a large number of workers and a high share of high-skilled labor. The corresponding results are discussed in the next sub-section.

Finally, I control for skill-specific unemployment rates to address that the literature on the wage curve provides robust empirical evidence of a negative relationship between wages and unemployment (e.g., Blanchflower/Oswald, 1990). I do obtain negative elasticities, although the coefficient of the regional unemployment rate among low skilled labor is not statistically different from zero.

5.2 Main result - The benefit of acquiring experience in large labor markets

Table 5 summarizes results that I obtain adopting the empirical strategy suggested by De la Roca and Puga (2016) to analyze dynamic agglomeration benefits that is described by equation (3). In addition to the variables reported in table 5 and the control variables reported in table 4, all models contain industry, occupation, and region-year fixed effects. As with De la Roca and Puga (2016) for Spain and Matano and Naticchioni (2016) for Italy, I focus on the benefit of acquiring experience in Germany’s largest local labor markets: Berlin, Hamburg

²⁰ I also estimate specifications without firm variables as they might cause a simultaneity bias in the estimations (see discussion by Combes/Gobillon, 2015). The results with regard to my pivotal explanatory variables change only marginally. The results are available from the author upon request.

²¹ As suggested by Combes and Gobillon (2015), I remove the own industry from the computation of the index which eases interpretation since the share of the own industry within the local economy already measures local specialization.

and Munich.²² I also include the amount of overall experience in the analysis so that the estimated effect of experience acquired in Berlin, Hamburg or Munich refers to the difference in the value of experience acquired there and experience acquired elsewhere in Germany.

Table 5: The value of labor market specific work experience

Dependent variable: logarithmic wages in new employment relationships			
	(1)	(2)	(3)
Experience Berlin / Hamburg / Munich	0.003*** (0.000)	0.006*** (0.002)	0.013*** (0.002)
Experience Berlin / Hamburg / Munich \times experience [†]			-0.005*** (0.000)
Experience	0.016*** (0.000)	0.033*** (0.001)	0.065*** (0.001)
(Experience) ²			-0.001*** (0.000)
Experience Berlin / Hamburg / Munich, now elsewhere [†]	0.060*** (0.005)	0.007 (0.006)	0.020 (0.016)
Experience Berlin / Hamburg / Munich \times experience, now elsewhere [†]			0.001 (0.001)
Experience outside Berlin / Hamburg / Munich , now in 3 largest	0.007*** (0.000)	0.001** (0.001)	0.007*** (0.002)
Experience outside Berlin / Hamburg / Munich \times experience, now in 3 largest [†]			-0.003*** (0.001)
New employment relationships	525,265	336,286	336,286
OLS: R ² , FE: within R ²	0.613	0.180	0.191
Worker fixed effects	No	Yes	Yes

[†] Coefficient and standard error multiplied by 10.

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. Work experience measured on a daily bases and expressed in years. All models include control variables as well as year, industry, occupation, and region fixed effects (see table 4).

Similar to De la Roca and Puga (2016) as well as to Matano and Naticchioni (2016), I find that one year of experience acquired in the largest German local labor markets has a significantly higher value than experience acquired in the rest of the country. The inclusion of individual fixed effects in specification (2) let the value of experience increase, indicating a negative correlation of unobserved abilities and work experience.²³ The magnitude of the agglomeration benefit is discussed below. In order to test whether the value of experience depends on where it is used, I consider similar to De la Roca and Puga (2016) interaction effects between the experience variables and an indicator for moving to the respective other group of labor markets. The corresponding results of the fixed effects model in column (2) suggest that the value of experience acquired in the largest labor markets does not change when transferring it to smaller labor markets. The value of experience acquired outside the large labor markets increases slightly if it is used in Berlin, Hamburg or Munich.

Following De la Roca and Puga (2016), specification (3) additionally contains the square of experience to let the marginal value of experience decay with more experience. Furthermore, interaction effects of experience acquired in the largest labor markets and overall experience are included. They allow for heterogeneous effects for less and more experienced workers (De la Roca/Puga, 2016). Qualitatively, I obtain the same results as De la Roca and Puga (2016): (i) Experience acquired in the largest local labor markets has a significantly higher value than experience acquired elsewhere. (ii) The marginal gain of working in one of the largest labor

²² The local labor market regions of Berlin, Hamburg and Munich are considered as one group. I also estimated a specification where I distinguish experience acquired in the three labor markets. When controlling for unobservable abilities of the workers by means of individual fixed effects, I found no significant differences between the value of experience acquired in the three largest German cities.

²³ This effect also shows up in the results reported by De la Roca and Puga (2016: table 1)

markets is higher for individuals with less work experience than for more experienced workers. This is particularly true for workers who previously worked elsewhere in the country. A reasonable explanation is that, *ceteris paribus*, the latter workers so far accumulated less human capital than individuals who worked in the largest labor markets. (iii) The value of experience acquired in the largest labor markets is highly portable to smaller labor markets which strongly supports the hypothesis that learning externalities play an important role. (iv) Experience acquired in the rest of the country has a higher value if it is used in the largest local labor markets than in the rest of the country. However, the location where experience was acquired has a stronger impact on its value than the location where it is used. The first year of experience acquired in Berlin, Hamburg or Munich increases wages by about 1.3 percent ($e^{0.01338-0.00047} - 1$) relative to having worked elsewhere and independently of the new job location. In comparison, the value of the first year of experience acquired outside the largest local labor markets increases by less than 0.7 percent if the worker moves to Berlin, Hamburg or Munich. Qualitatively, the results are the same as obtained by De la Roca and Puga (2016) for Spain. Quantitatively, the identified agglomeration benefit for the largest German agglomerations is somewhat smaller than the dynamic agglomeration gains obtained by De la Roca and Puga (2016) for Madrid and Barcelona and by Matano and Naticchioni (2016) for Rome and Milan. Their results indicate that the value of the first year work experience acquired in the largest cities of the considered country has, respectively, a three percent and two percent higher value than the first year acquired elsewhere.²⁴

The results reported in table 5 and by other scholars raise the question of how the value of experience exactly depends on the size of the local labor market in which it was acquired. To provide corresponding insights, I estimate the parameters of equations (6) and (12). The estimates of the pivotal parameters are summarized in table 6. Again, I observe that the value of experience is biased downwards if I do not control for unobserved characteristics by means of individual fixed effects.

Column (1) and (2) of table 6 contain results for equation (6) referring to the simplified learning function given by equation (4). It is assumed that the marginal value of experience does not depend on the amount of previous work experience and already accumulated human capital. The findings support the hypothesis that labor market size fosters the individual accumulation of knowledge. Labor market size positively impacts on the value of the acquired experience which reflects in higher future wages. According to the specification with individual fixed effects, the elasticity of wage with respect to the size of the labor market in which experience was acquired at one day in the past, denoted by δ , amounts to 0.116×10^{-4} . This effect may also be interpreted as the elasticity of the marginal value of work experience with respect to labor market size (see equation (4)).

Based on the estimated parameters $\hat{\delta}$ and $\hat{\gamma}$ (not reported), I compute \hat{Z} , the estimate of the threshold that indicates beyond which labor market size learning externalities arise (see equation (4)). The results strongly suggest that workers accumulate human capital even if they work in the smallest German local labor markets with a local workforce of about 15,000 employees. The estimated critical value amounts to about 100 employees.

The results summarized in column (1) and (2) are based on the assumption that the marginal return to work experience does not depend on the amount of previously acquired work experience and knowledge. However, the results reported in table 5 indicate that this assumption is too restrictive. Therefore, column (3) of table 6 contains results for equation (12) which is based on the augmented learning function given by equation (7). The significant differences in the parameter estimates between column (2) and (3) confirm that it is important

²⁴ The estimated earnings premia are only to a limited extent comparable across countries as the largest local labor markets within the different countries and also the respective reference, i.e., the country specific ‘average’ local labor market differ in size.

Table 6: Estimates of the parameters of the learning function

	(1)	(2)	(3)	(4)
$\hat{\delta}^\dagger$	0.059*** (0.002)	0.116*** (0.016)	0.184*** (0.020)	0.199*** (0.025)
\hat{z}	140.101 (33.936)	97.489 (111.521)	0.632 (0.901)	1.136 (1.832)
θ^\dagger	0.000	0.000	2.358	3.138
New employment relationships	525,265	336,286	336,286	214,319
OLS: R^2 , FE: within R^2	0.616	0.183	0.195	0.260
Worker fixed effects	No	Yes	Yes	Yes

[†] Coefficients and standard errors multiplied by 10,000.

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. \hat{z} calculated based on delta method and $\hat{\delta}$ and $\hat{\gamma}$ (not reported). Column (1) and (2) refer to the baseline model given by equation (6) which is based on the simplified learning function (equation (4)). Column (3) and (4) report results for equation (12). It is based on the augmented learning function (equation (7)). The results in column (3) are obtained using the full sample. The specification reported in column (4) is restricted to workers who acquired work experience only 1995 or later. All specifications include control variables as well as industry, occupation, and region-year fixed effects (see table 4).

to address that a worker acquires less additional knowledge by working the more work experience and human capital he or she previously acquired, i.e. the larger the individual stock of knowledge. This effect is captured in the augmented learning function by θ . Equation (12) which I estimate in order to determine the parameters of the augmented learning function is highly non-linear in this parameter. To overcome this problem, I chose different values for θ and estimated equation (12) several times. Once a value is imposed, a linear fixed-effects estimator can be applied to estimate the other parameters. Column (3) contains the results that I obtain when setting θ equal to 2.358×10^{-4} . This value should at least be close to the ‘true’ value as this specification results in a smaller residual sum of squares than all other specifications that I estimated.²⁵ Suppose a worker was employed at each day within the last 365 days, then a θ of 2.358×10^{-4} means that the human capital which was acquired at date $t - 365$ is weighted by factor 0.918 ($= (1 - 0.0002358)^{364}$). The weight is smaller than unity since the knowledge that was acquired at date $t - 365$ diminished the marginal gain in human capital at each other day of the last year. In contrast, the weight of the knowledge that was acquired at day $t - 1$ is unity.

With respect to δ , I now obtain an elasticity of 0.184×10^{-4} , and \hat{z} is about one. Both estimates are confirmed by the results reported in column (4). The latter are based on a reduced sample where I restrict the analysis to individuals who acquired experience not before 1995. This robustness check takes into account that the chosen definition of labor market regions might be an inappropriate approximation for the shape of local labor markets decades ago which would result in a measurement error in the pivotal explanatory variable, i.e., the size of the labor markets in which experience was acquired. However, it does not seem to be a severe problem here as the results reported in column (3) and (4) are very similar, although a different value of θ minimizes the residual sum of squares. Using the reduced sample, I obtain the smallest residual sum of squares if I set θ equal to 3.138×10^{-4} (see table C.3), meaning that human capital acquired, e.g., at date $t - 365$ is weighted by 0.892, given that a worker was employed each day within the last year. Furthermore, the reduction of the sample let the fit of the model increase. Now it explains more than one quarter of the within variation.

The nature of dynamic agglomeration effects in general, and learning externalities in particular, is that benefits

²⁵ The results that I obtain for other values of θ are reported in table C.2. Starting with $\theta = 1 \times 10^{-4}$ (not reported in table C.2), I gradually increased / decreased θ by 10 percent. I also estimated a specification with θ close to zero. The result corresponds to the result reported in column (2) of table 6 (compare with specification 22 in table C.2).

of working in large labor markets accumulate over time. This is shown by figure 2. Based on the regression results reported in columns (3) and (4) of table 6, it illustrates for different levels of experience the elasticity of wage with respect to the size of the labor market in which previous experience was acquired. Consider for example a worker with two years of work experience at date t . The corresponding elasticity is given by about 0.01. In other words, doubling the size of all labor markets in which the two years of experience were acquired results in a 0.7 percent higher wage at date t . In contrast, after ten years of work experience the elasticity amounts to about 0.045, meaning that doubling the size of all labor markets in which ten years of experience were acquired results in a productivity increase at date t by about three percent. At a very high level of experience the elasticity is even larger.²⁶ That the elasticity significantly increases in the amount of experience points out that dynamic agglomeration economies play an important role. The marginal increase in the elasticity with respect to experience is smaller the higher the level of acquired experience since marginal returns to experience are decreasing (see equation (7)).



Figure 2: Magnitude of dynamic agglomeration benefits

Note: The figure illustrates the regression results reported in column (3) and (4) of table 6. For different levels of experience the graph denotes the percentage increase in productivity, given that the local workforce would have been one percent larger in all labor markets in which previous work experience was acquired. The reduced sample contains only workers who acquired experience in 1995 or later.

For levels of experience up to ten years, the estimated elasticities based on the reduced sample and the full sample are virtually the same. For higher levels of experience they differ. At a level of 30 years of experience the gap between the elasticities amounts to about 0.01. However, the reduced sample only contains individuals with work experience of at most 16 years. Therefore, it is not clear-cut which specification leads to the more reliable estimate of the elasticity at a high level of experience. Nevertheless, the results are still very similar

²⁶ In general, the elasticity of the value of E days of experience with respect to the size of the labor markets in which the experience was acquired is, according to the results reported in column (3) of table 6, given by $0.186 \times 10^{-4} \times \sum_{k=1}^E (1 - 2.358 \times 10^{-4})^{k-1}$ (see also equation (10)).

and their overall interpretation is clear: Dynamic agglomeration benefits have a highly significant impact on individual productivity. Since workers are rather immobile between labor markets of different size (see table 3), the dynamic gains should explain a significant part of regional wage disparities between urban and non-urban labor markets. Moreover, the identified agglomeration effects should be strongly related to learning externalities as I control for the number of job changes in the past which is supposed to be a proxy for dynamic matching.

The results discussed in the following sub-section provide further support for the learning hypothesis and help to understand which underlying mechanisms lead to the increasing gains of working in large labor markets.

5.3 The importance of high-skilled labor and the portability of accumulated human capital

Typically, large urbanized labor markets are not only characterized by a large number of workers but also by an above average share of high-skilled labor. In accordance with the literature on human capital externalities, presumably the especially skilled workers are the workers one learns from the most. Hence, an important question is to which extent dynamic agglomeration benefits can be explained by the higher share of high-skilled labor. In order to provide corresponding insights, I consider $Z_{s,\tau}$, the labor market size in the learning function (see equation (4) and (7)), in the following to be an *aggregate* of total regional employment, $\text{emp}_{s,\tau}$, and the local share of high-skilled workers, $\frac{\text{skill}_{s,\tau}}{\text{emp}_{s,\tau}}$, as given by equation (13). So far, $Z_{s,\tau}$ was only measured in terms of total regional employment.

$$\delta \ln(Z_{s,\tau}) = \rho_{\text{emp}} \ln(\text{emp}_{s,\tau}) + \rho_{\text{skill}} \frac{\text{skill}_{s,\tau}}{\text{emp}_{s,\tau}}. \quad (13)$$

In contrast to δ which denotes the total dynamic agglomeration benefit, ρ_{emp} refers only to the isolated, *direct* benefit of acquiring experience in a labor market with a large workforce in terms of total regional employment. It does not capture the benefit of acquiring experience in a labor market with an above average share of high-skilled labor. This *indirect* benefit is captured by ρ_{skill} . The literature on human capital externalities typically assumes that there is a linear relationship between the logarithmic wage and the local share of high-skilled labor (see, e.g., Moretti, 2004a, Heuermann, 2011). I do so as well.²⁷

As discussed in the literature review, some authors stress the importance to distinguish between productivity effects that are based on human capital externalities and those based on the imperfect substitution of workers belonging to different skill groups. However, these concerns refer to the interpretation of contemporaneous effects of a local employment's skill structure. I control for these contemporaneous effects by including region-year fixed effects as well as the local industry specific share of high- and low-skilled labor (see table 4). The results reported in table 7 refer to benefits of having *previously* worked in local labor markets with a high share of high-skilled workers. In contrast to contemporaneous effects, the gains discussed below should therefore only be based on external effects.

Table 7 summarizes results for the augmented learning function (equation (7)) with $Z_{s,\tau}$ being an aggregate measure as described by equation (13). The results reported in column (1) are obtained using the full sample of new employment relationships. The results reported in column (2) are based on the reduced sample which does

²⁷ I also estimated specifications where I assumed that $\ln(Z_{s,\tau})$ and thereby $w_{i,t}$ is linear in the *logarithmic* share of high-skilled labor. Qualitatively, I obtain the same results as discussed below. Furthermore, those results with respect to $\hat{\rho}_{\text{emp}}$ are statistically not significantly different from the corresponding results discussed below. The residual sum of squares differs only marginally as well.

not contain new employment relationships of workers who acquired experience before 1995.²⁸ Both specifications strongly suggest that at least a significant part of dynamic agglomeration benefits stem from the higher share of high-skilled labor. In both specifications a statistically as well as economically significant impact is identified. For a worker with 9.4 years of experience, which corresponds to the sample mean, a one standard deviation increase in the local share of high-skilled labor, i.e., an increase of about four percentage points, in all labor markets in which experience was acquired results in a 2.9 (column (2)) to 4.2 (column (1)) percent productivity increase ($\exp(\hat{\rho}_{\text{skill}} \sum_{k=1}^{3423} (1 - \theta)^{k-1} \times 0.04) - 1$). It shows that local specialization in high-skilled labor is an important determinant of the value of work experience and thereby of future productivity. This also supports the view that the identified dynamic agglomeration benefit crucially depends on local learning externalities as it presumably is especially high-skilled labor one learns from the most.

Table 7: Estimates of the parameters of the learning function distinguishing the impact of total regional employment and the share of high-skilled labor

	(1)	(2)
$\hat{\rho}_{\text{emp}}^{\dagger}$	0.029 (0.028)	0.077** (0.037)
$\hat{\rho}_{\text{skill}}^{\dagger}$	4.524*** (0.578)	3.389*** (0.766)
$\hat{\gamma}^{\dagger}$	1.637*** (0.318)	1.087*** (0.404)
θ^{\dagger}	2.724	3.138
New employment relationships	336,286	214,319
within R^2	0.195	0.260

[†] Coefficients and standard errors multiplied by 10,000.

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. Both columns summarize estimates of the parameters of the augmented learning function (equation (7)) with labor market size $Z_{s,\tau}$ being an aggregate of total regional employment and the local share of high-skilled labor as given by equation (13). The results reported in column (1) are obtained using the full sample, the results reported in column (2) using a reduced sample which does not contain new employment relationships of workers who acquired experience before 1995. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

With regard to the direct effect of labor market size on the value of experience and future productivity the estimation results are not clear-cut. According to the estimated elasticity reported in column (1) the impact of labor market size is virtually zero once controlling for the share of high-skilled labor within the labor markets in which experience was acquired. The elasticity obtained based on the reduced sample suggests that the higher share of high-skilled workers within large labor markets explains ‘only’ about 60 percent of the dynamic agglomeration benefit (compare column (4) of table 6 and column (2) of table 7).

One explanation for the different results might refer to the different amount of experience the workers in both samples have. Workers in the reduced sample have on average five years of experience (at most 16 years) and the workers in the full sample on average more than nine years (at most 36 years). Hence, the results may suggest that workers directly benefit from working in a labor market with a large workforce only at the beginning of their career. Later on, after they acquired a certain amount of experience and thereby human capital, only the skill level of the local labor force is important for further knowledge accumulation. That the positive effect of labor market size does not show up at all in the regression result when the full sample is considered might be due to the fact that the size of the labor market in which experience was acquired is weighted lower the more experience was acquired thereafter. Consider for example a worker who was employed each day

²⁸ The reported results are obtained setting θ equal to 2.724×10^{-4} and 3.138×10^{-4} , respectively. Results for other values of θ are summarized in table C.4 and table C.5.

within the previous 15 years. In this case, a θ of about 2.724×10^{-4} (see column (1) in table 7) means that the size of the labor market in which experience was acquired fifteen years ago is weighted by only factor 0.22 ($(1 - 0.0002724)^{5479}$). In contrast, the size of the labor market in which experience was acquired one year ago has a weight of 0.91. It points out that the size of the labor market in which a worker acquired experience years ago has little impact on individual productivity at date t , relative to the size of the labor market in which recent experience was acquired.²⁹ However, further research is needed to analyze in more detail whether the importance of labor market size in terms of total regional employment and the local share of high-skilled labor differ depending on the career phase of a worker.³⁰

Finally, I let the parameters of the learning function vary depending on the labor market in which the accumulated human capital is used. Thereby, I take into account that the productivity increase which comes from the acquired work experience might not be the same in all regions. More precisely, I estimate equation (12) again, but now I let the parameters of the learning function vary depending on whether the new employer is located in a small, medium, or large labor market. As large local labor markets I consider those with a local workforce of more than 750,000 employees: Berlin, Hamburg, Munich, Frankfurt, Düsseldorf, Stuttgart, Cologne. Labor markets with less than 100,000 employees are considered as small labor markets, i.e., 43 percent of all German local labor market regions. The remaining labor markets represent the reference group.³¹ Table 8 on page 29 captures the corresponding estimates for the pivotal parameters. The results reported in column (1) are obtained based on the full sample, the results in column (2) based on the reduced sample.³² The details of the result are discussed at the end of this sub-section.

Overall, both specifications indicate that the value of experience varies little between the considered groups of labor markets. This supports the hypothesis that the identified dynamic agglomeration benefits are related to learning externalities. A worker who previously worked in a large labor market is in comparison to a worker who previously worked somewhere else, *ceteris paribus*, more productive, independently of where they make use of their work experience. Similar as with De la Roca and Puga (2016: figure 3), figure 3 illustrates this exemplarily. The reference is a worker who acquired 30 years of work experience in a labor market with a size of 100,000 employees. The solid black line refers to a worker who acquired 30 years of experience in a labor market with a size of 1,000,000 employees. In line with the idea of dynamic agglomeration economies the gap in individual productivity significantly increases over time. Marginal benefits are decreasing because the individual knowledge stock growth faster in large labor markets which diminishes marginal returns (see equation 7). After 30 years the difference in productivity amounts to about 18 percent.

²⁹ Recall that the weight of the size of the labor market in which experience was acquired decreases in the amount of experience that is acquired thereafter. It takes into account that the marginal return to experience at any date is lower the more experience and human capital previously was acquired (see equation (7)).

³⁰ If $Z_{s,t}$ is considered to be an aggregate of total regional employment and the local share of high-skilled labor, $\hat{\gamma}$ depends not only on the minimum number of workers but also on the minimum share of high-skilled labor beyond which learning externalities arise, denoted by $\underline{\text{emp}}$ and $\underline{\text{skill}}$, respectively: $\hat{\gamma} = -\hat{\rho}_{\text{emp}} \ln(\underline{\text{emp}}) - \hat{\rho}_{\text{skill}} \ln(\underline{\text{skill}})$. Based on the regression result reported in table 7 it is not possible to compute $\underline{\text{emp}}$ and $\underline{\text{skill}}$ explicitly. However, it can be tested whether a specific labor market is large enough such that individuals accumulate human capital by working there. The condition is fulfilled given that $-\hat{\rho}_{\text{emp}} \ln(\text{emp}_{r,t}) - \hat{\rho}_{\text{skill}} \ln(\text{skill}_{r,t})$ is smaller than $\hat{\gamma}$. Based on the parameters reported in column (1), the maximum value that I obtain for the German labor market regions is -0.299 . Based on the estimates reported in column (2) the maximum amounts to -0.739 . Both values are significantly smaller than the respective estimate for γ . This confirms my previous findings that learning takes place in all German labor markets.

³¹ I also estimated specification where I require large labor markets to have a local workforce of at least 500,000 employees and at least 1,000,000 employees, respectively. The obtained results are very similar to those reported in table 8 and available upon request.

³² Results that I obtain setting θ to other values than those reported in table 8 can be found in tables C.6 and C.7, respectively.

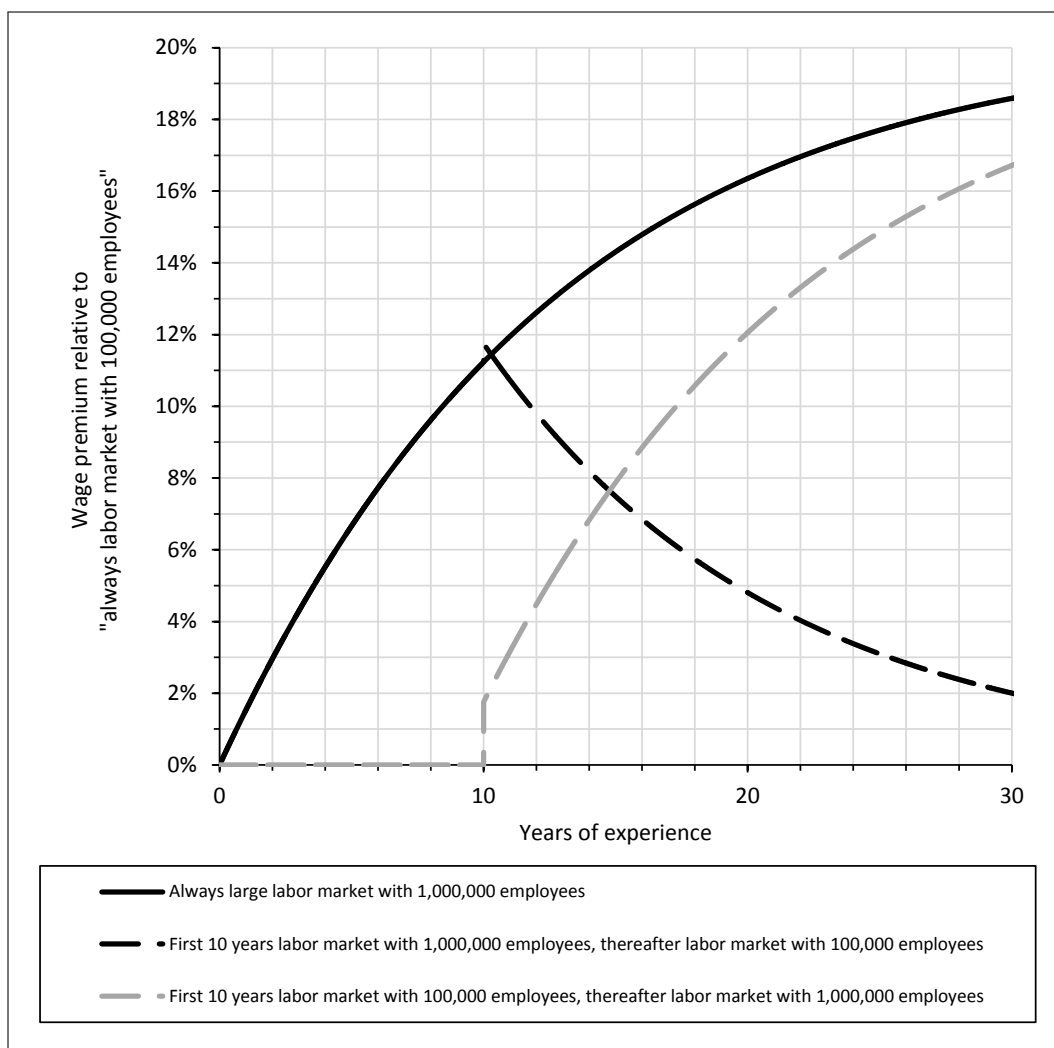


Figure 3: Labor productivity after migration

Note: Following De la Roca and Puga (2016: figure 3), this figure illustrates exemplarily the regression results reported in column (1) of table 8.

The dashed lines indicate what happens if a worker migrates after ten years from the larger to the smaller labor market or vice versa. The immediate changes in productivity as reflected in wage are, especially in comparison to the marked differences in the level of productivity, small. However, in the years after migration productivity adjusts. In the larger labor market the worker who initially worked in the smaller labor market catches-up. The intuition behind this development is straightforward. The worker who initially worked in the larger labor market, *ceteris paribus*, accumulated in the first ten years more knowledge than the worker who worked for ten years in the smaller labor market. Due to the differences in the individual knowledge stocks, the worker who migrates to the larger labor markets accumulates thereafter more additional human capital than the worker who always acquired experience in the larger labor market which results in the stronger productivity increase (compare equation 7). The same explanation applies to the adjustment process in the smaller labor market. Due to his or her larger knowledge stock the worker who comes from the larger labor market, *ceteris paribus*, learns less in the following years than the worker who always stayed in the smaller labor market. Therefore, the benefit of having previously worked in the larger labor market decreases the more time passes. That productivity in the years after migration adjusts, does not depend on whether or not I let the parameters of the learning function vary depending on where the previously acquired experience is used. The described dynamics are based on parameter θ in the augmented learning function (equation (7)). Since θ is larger than zero, the impact of the labor market where experience was acquired decreases in the amount of experience that is acquired thereafter.

Table 8: Estimates of the parameters of the learning function depending on the size of the labor market in which experience is used

	(1)	(2)
Now large labor market		
$\Delta \hat{\delta}^{\dagger}$	-0.039*** (0.012)	-0.034* (0.020)
$\Delta \hat{Z}$	-1.406 (1.814)	-0.970 (1.699)
Now medium size labor market (reference)		
$\hat{\delta}^{\dagger}$	0.197*** (0.021)	0.197*** (0.027)
\hat{Z}	1.466 (1.920)	1.010 (1.783)
Now small labor market		
$\Delta \hat{\delta}^{\dagger}$	-0.010 (0.013)	0.012 (0.021)
$\Delta \hat{Z}$	-0.668 (1.123)	1.341 (2.825)
θ^{\dagger}	2.358	3.138
New employment relationships	336,286	214,319
within R^2	0.195	0.260

[†] Coefficients and standard errors multiplied by 10,000.

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. Both columns summarize estimates of the parameters of the augmented learning function (equation (7)). \hat{Z} calculated based on delta method and $\hat{\delta}$ and $\hat{\gamma}$ (not reported). The parameters are allowed to vary depending on the size of the labor market in which the accumulated knowledge is used. Labor markets with a total number of at least 750,000 employees are considered as large labor markets, local labor markets with less than 100,000 employees as small labor markets. Δ refers to the difference in a parameter in comparison to the reference group, i.e., medium size labor markets. The results reported in column (1) are obtained using the full sample, the results reported in column (2) using a reduced sample which does not contain new employment relationships of workers who acquired experience before 1995. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

Turning to the parameter estimates reported in table 8 shows that the estimates which are obtained based on new employment relationships in small labor markets are of a comparable magnitude and not statistically different

from those obtained based on wages paid in medium size labor markets. With regard to large labor markets, I obtain an estimate for δ that is about one fifth smaller than the other estimates for δ . *Ceteris paribus*, it suggests that wages in large labor markets are slightly less elastic with regard to the size of the labor market in which previous experience was acquired than in the other labor markets.³³ This finding is in accordance with my results reported in table 5. In the three largest German labor markets the difference in the wage of a worker with one year of experience acquired in Berlin, Hamburg or Munich and the wage of a worker who acquired the experience somewhere else amounts to 0.4 percent ($= (0.013 - 0.005) - (0.007 - 0.003)$). Outside the three largest labor markets this difference amounts to 0.8 percent ($= 0.013 - 0.005$). De la Roca and Puga (2016) obtain a similar result for Spain. The phenomenon might also be related to learning externalities. Suppose workers who acquired experience in a large local labor market, let's call them workers of type L, and workers who acquired the same amount of experience in a medium or small local labor market, workers of type S, are imperfect substitutes since the accumulated knowledge differs. Then, workers of type L may receive a higher wage, relative to workers of type S, especially if they use their accumulated knowledge in a smaller labor market, in which workers of type L are relatively rare (compare table 3). In contrast, if both types of workers use their accumulated human capital in a large labor market, the wage difference presumably is smaller since there are many workers of type L and few of type S.

When discussing the difference in $\hat{\delta}$, it has to be taken into account that there also exists an opposite effect. For $\hat{\underline{Z}}$, defined as $\exp(-\hat{\gamma}/\hat{\delta})$, I obtain a smaller estimate as well when considering wages paid in large labor markets, although the difference in this parameter is not statistically significant. *Ceteris paribus*, this suggests that the value of each unit of experience is slightly higher if experience is used in a large labor market rather than somewhere else, see equations (4) and (7). One explanation may refer to a higher quality of matches between job seekers and vacancies in large than in other labor markets, e.g., due to a larger variety of available jobs. If a worker accumulated at each day when experience was acquired a small amount of human capital that he or she can make use of only if the new job is located in a large labor market because of the better match, it should result in a smaller estimate for $\hat{\underline{Z}}$.

However, it should be stressed here that the differences in $\hat{\delta}$ and $\hat{\underline{Z}}$ should not be overemphasized. Quantitatively, the differences in the parameter estimates are rather small depending on whether they are estimated based on wages paid in large, medium, or small labor markets. Furthermore, the effects go in opposite directions. As illustrated by figure 3, there is virtually no jump in the value of experience if a worker leaves a large labor market to work somewhere else. This implies that the opposite effects outweigh each other.³⁴ This is in accordance with the results reported in table 5 and by De la Roca and Puga (2016) which suggest that a worker from the largest labor markets of a country does not suffer from a decline in productivity if he or she uses the accumulated human capital somewhere else. For workers who migrate from a smaller to a larger labor market there might be a small immediate productivity increase as illustrated by figure 3 and reported in table 5 as well as by De la Roca and Puga (2016), maybe due to a better match between job seeker and vacancy. Nevertheless, in comparison to a worker who already previously worked in the larger labor market there is a significant gap in individual productivity which presumably is a result of learning externalities in the larger labor market.

³³ Recall that differences in static regional wage differentials between large, medium, and small labor markets are captured by the region fixed effects, not by $\delta_r(i, t)$.

³⁴ I also considered a case where a worker acquired experience in a labor market with a size of 750,000 employees and 1,400,000 employees, respectively. In the first case, there is again virtually no jump in the value of experience if the worker migrates to a labor market with a size of 100,000 employees. In the second case there is a small immediate productivity increase of about one percent.

Despite the various new insights this paper provides, there are some caveats to bear in mind. Although, I try to control for the second potential channel of dynamic agglomeration benefits, dynamic matching, by including the number of previous employers, there is no guarantee that this variable fully captures the effect. Furthermore, I cannot fully rule out that my analysis still suffers from selection effects, even though I control for various observable and unobservable characteristics of the workers. Imagine, workers have expectations on their individual learning opportunities in urban labor markets and took them into account when they decided where to work. This would imply a positive selection because especially those workers would have decided to acquire experience in a large labor market who expect to learn much. At least part of this selection should be captured by the individual fixed effects and by the time-variant individual characteristics like educational degree. However, if the expected individual learning opportunities depend on unobserved time-variant individual characteristics, the estimated benefits of acquiring experience in large local labor markets are likely biased upwards.

6 Conclusions

This paper provides empirical evidence as to why wages are higher in urban than in non-urban local labor markets. The focus lies on learning externalities which are discussed to be one mechanism behind dynamic agglomeration economies. I analyze to which extent wages associated with new employment relationships depend on the size of local labor markets in which previous work experience was acquired. Based on an administrative data set capturing detailed information on individual employment biographies, I first show for Germany, like De la Roca and Puga (2016) for Spain and Matano and Naticchioni (2016) for Italy, that the experience acquired in the largest local labor markets has a significantly higher impact on future productivity than experience acquired in the rest of the country, as reflected in higher wages. Thereafter, this study goes a step further as previous papers in order to obtain more general results. It estimates the elasticity of wage with respect to the size of the labor market in which experience was acquired. In accordance with the idea that dynamic agglomeration benefits accumulate over time, the elasticity increases with the level of experience. For example, after two years of working, the elasticity of wage with respect to the size of the labor markets in which experience was acquired amounts to about 0.01, after ten years to more than 0.04 and after 30 years to about 0.06 to 0.07. The latter, for example, implies that doubling employment in all local labor markets in which 30 years of experience were acquired results in a productivity increase of about four to five percent. Since workers are rather immobile between labor markets that differ in size, the identified dynamic agglomeration benefit should significantly contribute to wage disparities between urban and non-urban labor markets.

Moreover, the identified effect should be strongly related to learning externalities. I control for observable as well as unobservable characteristics of the worker and the region in which the new employer is located, as well as for characteristics of the firm and the local industry. I also take into account other channels of agglomeration economies, *inter alia*, dynamic matching. Furthermore, my results are also inline with previous findings that the value of experience is predominately determined by the size of the labor market in which experience was acquired. Initially, the productivity of workers as reflected in wages changes little when they migrate from a large labor market to a small one or vice versa. However, in the following years productivity adjusts. For workers who migrate to a smaller labor market, the benefit of having previously worked in a larger labor market declines the more time passes. Those workers who initially worked in the larger labor market and therefore have a larger stock of knowledge learn less, due to decreasing returns, than those workers who always worked

in the smaller labor. Therefore, the latter catch-up over time. On the contrary, workers who migrate to a larger labor market experience a strong productivity increase in the following years and catch-up to those who always worked in the larger labor market.

Further support for the learning hypothesis comes from an additional result: At least a significant part of the identified dynamic agglomeration benefit can be explained by the above average share of high-skilled labor within large local markets. I distinguish the impact of labor market size on the value of acquired experience and future productivity from the impact of the local share of high-skilled labor. Thereby, I provide new insights as to which characteristic of an urban labor market leads to the increasing gains of acquiring experience there. In accordance with the literature on human capital externalities, the results suggest that it is especially high-skilled labor workers learn from. Once controlling for the local share of high-skilled labor, the impact of labor market size in terms of total regional employment on the value of experience and future productivity as reflected in wages is significantly smaller as before. It is not clear-cut whether labor market size has a direct effect at all. In some specifications the impact of labor market size declines by about 60 percent, in other specifications the positive impact vanishes completely. Therefore, further research is needed to clarify whether learning benefits in large labor markets exclusively depend on the higher share of high-skilled labor or whether there is an additional direct effect that stems from the large number of workers in the local labor market. A further interesting question for future research is whether learning externalities are increasing in labor market size without bound, or whether the benefit decreases beyond some threshold since urban congestion may hinder the transmission of skills as discussed by, e.g., Duranton and Puga (2004).

Appendix

A The solution of equation (9)

The difference equation given by equation (9) is a special case of the general first-order difference equation $x_{t+1} = b_t + a_t x_t$. Therefore, the solution of equation (9) can be obtained based on the solution of the general difference equation.

General first-order difference equation,
see, e.g., Sydsæter et al. (2008: 394f.):

$$x_{t+1} = b_t + a_t x_t \quad t = 0, 1, 2, \dots$$

The first few values are given by:

$$\begin{aligned} x_1 &= b_0 + a_0 x_0 \\ x_2 &= b_1 + a_1 b_0 + a_1 a_0 x_0 \\ x_3 &= b_2 + a_2 b_1 + a_2 a_1 b_0 + a_2 a_1 a_0 x_0 \\ x_4 &= b_3 + a_3 b_2 + a_3 a_2 b_1 + a_3 a_2 a_1 b_0 + a_3 a_2 a_1 a_0 x_0 \end{aligned}$$

The solution is given by,
see Sydsæter et al. (2008: chap. 11.1, eq. 11):

$$x_t = \left(\prod_{s=0}^{t-1} a_s \right) x_0 + \sum_{k=0}^{t-1} \left(\prod_{s=k+1}^{t-1} a_s \right) b_k,$$

with the product $\prod_{s=0}^{t-1} a_s$ of zero terms being 1.

Difference equation given by (9),
omitting some indices for simplicity:

$$Q_{\tau+1} = \begin{cases} \delta \ln \left(\frac{Z_{r(\tau)}}{\underline{Z}} \right) + (1 - \theta) Q_{\tau} & \text{if } Z_{r(\tau)} > \underline{Z} \text{ and } O_{\tau} = 1 \\ Q_{\tau} & \text{else} \end{cases}$$

The first few values are given by:
(per definition, before day 1 no experience was acquired)

$$\begin{aligned} Q_1 &= 0 & (\text{see equation (8)}) \\ Q_2 &= \tilde{b}_1 \\ Q_3 &= \tilde{b}_2 + \tilde{a}_2 \tilde{b}_1 \\ Q_4 &= \tilde{b}_3 + \tilde{a}_3 \tilde{b}_2 + \tilde{a}_3 \tilde{a}_2 \tilde{b}_1 \end{aligned}$$

$$\begin{aligned} \text{with } Q_0 &= 0, \tilde{b}_0 = 0, \tilde{a}_{\tau} = 1 - I(O_{i,\tau} = 1) I(Z_{r(\tau)} > \underline{Z}) \theta, \\ \tilde{b}_{\tau} &= I(O_{i,\tau} = 1) \times \delta \max \left\{ \ln \left(\frac{Z_{r(\tau)}}{\underline{Z}} \right); 0 \right\} \end{aligned}$$

For $\tau = 2, 3, 4, \dots$, the solution is given by:

$$\begin{aligned} Q_{\tau} &= \sum_{k=1}^{\tau-1} (1 - \theta)^{\tilde{E}_{\tau} - \tilde{E}_{k+1}} \times \tilde{b}_k \\ &= \sum_{k=1}^{\tau-1} (1 - \theta)^{\tilde{E}_{\tau} - \tilde{E}_{k+1}} \times I(O_{i,k} = 1) \times \delta \max \left\{ \ln \left(\frac{Z_{r(k)}}{\underline{Z}} \right); 0 \right\}, \end{aligned}$$

where \tilde{E}_{τ} and \tilde{E}_{k+1} denote the number of days a worker was employed before date τ and $k+1$, respectively, in labor markets *larger than* \underline{Z} , wherefore $\tilde{E}_{\tau} \leq E_{\tau}$.

Proof by mathematical induction:

$$\begin{aligned} Q_2 &= (1 - \theta)^{\tilde{E}_2 - \tilde{E}_2} \times \tilde{b}_1 = \tilde{b}_1 \\ Q_{\tau+1} &= \tilde{b}_{\tau} + \tilde{a}_{\tau} Q_{\tau} \\ &= \tilde{b}_{\tau} + \tilde{a}_{\tau} \sum_{k=1}^{\tau-1} (1 - \theta)^{\tilde{E}_{\tau} - \tilde{E}_{k+1}} \times \tilde{b}_k \\ &= \tilde{b}_{\tau} + \sum_{k=1}^{\tau-1} (1 - \theta)^{\tilde{E}_{\tau+1} - \tilde{E}_{k+1}} \times \tilde{b}_k \\ &= \sum_{k=1}^{\tau} (1 - \theta)^{\tilde{E}_{\tau+1} - \tilde{E}_{k+1}} \times \tilde{b}_k \end{aligned}$$

□

B Further information on data

The units of observation in my analysis are new employment relationships within the period between 2005 and 2011. I restrict the analysis to new employment of individuals to whom information on the entire employment biography is available. As the IEB contains information on employment in West Germany only from 1975 onwards, I exclude all workers who were born before 1960. Reliable and complete information on employment in East Germany is only available from 1993 onwards. Therefore, I also exclude all workers who presumably worked in East Germany before reunification, i.e., all workers for whom I do not observe an spell of employment before 1990 and who were born before 1977. Additionally, I do not consider individuals who worked before 1993 in a labor market region which today contains parts of former East and West Germany. Furthermore, I restrict the analysis to workers with a German nationality. Since information on the place of birth is not available, it is the only possibility to exclude immigrants. This is necessary as for this group of individuals information on the entire previous work experience is not available.

In my analysis, I only consider new spells of employment with a length of at least seven days that refer to full-time employment subject to social security contributions outside the public sector and the temporary employment industry. I do not consider apprenticeships, nor are new employment relationships that start simultaneously with another employment relationship or with a measure of active labor market policy. In the latter case I cannot ensure that this employment is not publicly subsidized. Furthermore, new employment relationships with wages below two times the limit for marginal employment as well as recalls, i.e., cases in which a worker starts to work in an establishment in which he or she worked at least once during the previous 28 days, are not considered. If a worker is already employed at the starting date of the new employment relationship by another establishment, I consider the new employment relationship only if the previous spell of employment ends within 7 days. Furthermore, I exclude a new employment relationship if it is the first spell of employment in a person's life.

The dependent variable in my analysis is the wage of a new employment relationship. The first employment spell in the IEB of a new employment relationship ends, at the latest, by December 31 of the year in which the new employment relationship starts. By dividing total reported earnings by the length of the spell, daily wages are obtained which I use as the dependent variable. Information on actual working days or contract hours is not available. Firms report earnings only up to the upper limit for social security contributions. Therefore, the wage information in the IEB is right censored. I follow Reichelt (2015) and apply an interval regression, a generalization of tobit regression, to predict the wages above the threshold (about 6% of the observations). See Reichelt (2015) for a detailed description on how interval regression is applied to impute right-censored wages. For the imputation I use information on sex, age, nationality, educational level, industry and the region in which the establishment is located. The results of my regression analysis do not change when I use the reported wages as dependent variable instead of the imputed wages. Table B.1 provides information on the definition of all variables used in this analysis.

Table B.1: Variables - definitions and sources

Variable	Definition	Source
Gross daily wage	Daily wages are calculated by dividing the reported total earning from employment spell by the length of the spell. Wages above the upper limit for social security contributions are imputed (see Appendix A).	IEB
Size of local labor market in which experience was acquired	Size of regional labor market regions, defined according to Kosfeld and Werner (2012), in which work experience was acquired until the considered new employment relationship. Measured in terms of employment subject to social security contributions. The share of high-skilled labor refers to workers with an university degree.	IEB
Work experience	Length of previous employment spells measured on a daily basis. Marginal employment is not considered, nor are employment spells that refer to measures of active labor market policies. I also compute the work experience that was acquired in the largest German labor market regions, i.e., Berlin, Hamburg and Munich.	IEB
Tenure	The length of an employment spell in month that refers to a new employment relationship. The spell ends at the latest by December 31 of the year in which the new employment relationship starts.	IEB
ln(Number of employers)	The number of unique establishment identifiers until the considered new employment relationship, by person.	IEB
Educational level of the worker	A categorial variable that combines information on the highest school leaving certificate, completed vocational training and university degree. For some spells of employment this information is missing. If so, I use the information from previous employment spells following Fitzenberger et al. (2005).	IEB
Gender	Dummy variable distinguishing male and female workers.	IEB
Length of non-employment	The number of days between the beginning of the new employment relationship and the end of the previous employment spell.	IEB
Pre-employment status	Dummy variables referring to the 28 days before the considered transition to employment <ul style="list-style-type: none"> - unemployment benefits (Arbeitslosengeld I) - unemployment assistance (Arbeitslosengeld II / Arbeitslosenhilfe). - unemployed and registered as a job seeker - not unemployed, but registered as a job seeker - participating in active labor market policy programmes. 	IEB
Firm characteristics	Number of employees, share of workers with a university degree, share of workers with no completed vocational training/no university degree, share of workers younger than 30 years old, share of workers 50 years old or older. The information refers to the last reference date (June 30) before the considered transition.	Establishment History Panel (BHP)
Industry share	Logarithm of the employment share of the industry (2-digit level: 88 industries) in total regional employment.*	Employment statistics of the Federal Employment Agency (FEA)
Industrial diversity	Logarithm of the inverse Herfindahl index based on the employment shares of the different industries in total regional employment. The own industry is excluded when the inverse Herfindahl index is calculated.*	FEA

Table B.1 continued

Variable	Definition	Source
Human capital within the local industry	Share of workers with a university degree in total employment and share of workers without completed vocational training/university degree in the same industry and regional labor market.*	FEA
Skill-specific unemployment rate of the regional labor market	The share of persons registered as unemployed in the number of persons who are registered as unemployed or employed in the region. I distinguish three groups: persons with a university degree, persons with a completed vocational training, and persons without completed vocational training/university degree.*	(Un-)employment statistics of the Federal Employment Agency
Industry fixed effects	Fixed effects for 88 distinct industries (2-digit level according to the classification from 2008). In 2008, there was a change in the industry classification. If an establishment is observed before and after 2008, I assign the spells of employment from 2005–2007 to the industry that the firm reports in 2008 (or later). If an establishment identifier shows up only for 2005–2007, I use a correlation matrix between the old and new industry classification as described by Eberle et al. (2011).	IEB
Occupation fixed effects	Fixed effects for 335 distinct occupations.	IEB
Region-year fixed effects	Time varying fixed effects for the location of the establishment in which a person starts to work. The location refers to one of 141 functional labor markets which are defined according to commuting intensity between NUTS-3-regions (see Kosfeld/Werner, 2012).	IEB

* The information refers to June 30th of the previous year.

C Further results

Table C.1: Summary statistics

	Entire sample				Only workers with at least two new employment relationships			
	mean	sd	min	max	mean	sd	min	max
ln(imputed gross daily wage)	4.102	0.465	3.267	7.192	4.102	0.455	3.267	7.192
Total work experience in days	3423.791	2741.222	1.000	13403.001	3475.382	2644.871	1.000	13219.999
Experience acquired in Berlin / Hamburg / Munich in days	371.170	1216.804	0.000	13023.000	380.287	1199.576	0.000	13023.000
Size of local labor market in which experience was acquired [§]								
ln(Number of employees)	12.510	0.848	9.512	14.187	12.493	0.841	9.512	14.187
Share of high-skilled employment	0.100	0.040	0.011	0.383	0.099	0.039	0.013	0.325
Tenure in month	5.990	3.585	0.033	12.000	5.784	3.540	0.033	12.000
ln(Number of previous employers) [†]	1.191	0.784	0.000	4.331	1.325	0.758	0.000	4.331
Length of non-employment								
0-28 days (job-to-job transition)	0.552	0.497	0.000	1.000	0.561	0.496	0.000	1.000
29-92 days	0.142	0.350	0.000	1.000	0.161	0.368	0.000	1.000
93 days - 1 year	0.156	0.363	0.000	1.000	0.169	0.375	0.000	1.000
> 1 year	0.150	0.357	0.000	1.000	0.108	0.311	0.000	1.000
Pre-employment status								
Not registered as job seeker	0.584	0.493	0.000	1.000	0.552	0.497	0.000	1.000
Unemployed and registered as a job seeker	0.310	0.463	0.000	1.000	0.339	0.473	0.000	1.000
Not unemployed, but registered as a job seeker	0.106	0.307	0.000	1.000	0.109	0.312	0.000	1.000
Participation in measures of active labor market policy	0.060	0.238	0.000	1.000	0.061	0.240	0.000	1.000
Public assistance benefits								
No benefits	0.691	0.462	0.000	1.000	0.664	0.472	0.000	1.000
Unemployment benefit (ALG I)	0.235	0.424	0.000	1.000	0.268	0.443	0.000	1.000
Unemployment assistance (ALG II, ALHI)	0.074	0.261	0.000	1.000	0.068	0.252	0.000	1.000
Education:								
Secondary/intermediate school leaving certificate								
... without completed vocational training	0.083	0.276	0.000	1.000	0.071	0.257	0.000	1.000
... with completed vocational training	0.671	0.470	0.000	1.000	0.697	0.459	0.000	1.000
Upper secondary school leaving certificate								
... without completed vocational training	0.020	0.140	0.000	1.000	0.015	0.123	0.000	1.000
... with completed vocational training	0.101	0.302	0.000	1.000	0.099	0.299	0.000	1.000
Completion of a university of applied sciences	0.046	0.209	0.000	1.000	0.045	0.208	0.000	1.000
College/ university degree	0.079	0.270	0.000	1.000	0.072	0.258	0.000	1.000
Female worker	0.362	0.481	0.000	1.000	0.329	0.470	0.000	1.000
ln(Number of workers within the establishment)	3.951	1.936	0.000	*	3.830	1.883	0.000	*
Share high-skilled workers in establishment	0.112	0.200	0.000	1.000	0.103	0.193	0.000	1.000
Share low-skilled workers in establishment	0.152	0.204	0.000	1.000	0.155	0.209	0.000	1.000
Share of middle aged workers in establishment	0.523	0.176	-0.000	1.000	0.522	0.179	-0.000	1.000
Share of older workers in establishment	0.205	0.151	0.000	1.000	0.202	0.151	0.000	1.000
ln(Employment share of the industry within the region)	-3.528	1.053	-12.732	-0.855	-3.528	1.045	-12.732	-0.855
ln((Herfindahl index based on local industry shares) ⁻¹)	3.036	0.266	1.444	3.551	3.027	0.266	1.444	3.551
Share high-skilled workers in local industry	0.105	0.108	0.000	0.855	0.099	0.104	0.000	0.855
Share low-skilled workers in local industry	0.190	0.088	0.000	1.000	0.193	0.089	0.000	1.000
ln(Local unemployment rate among high-skilled labor) [‡]	1.855	0.420	0.294	2.838	1.855	0.419	0.294	2.838
ln(Local unemployment rate among skilled labor) [‡]	2.254	0.429	0.981	3.484	2.259	0.433	0.981	3.484
ln(Local unemployment rate among low-skilled labor) [‡]	3.403	0.375	2.245	4.293	3.402	0.375	2.245	4.293
Observations	525,265				336,286			

[§] Weighted average size of the labor markets where an individual acquired experience before the considered new employment relationship starts, computed on individual level and weighted by the respective length of the previous spell of employment. [†] For less than 1 percent of the observations the number of previous employers exceeds 18, for less than 10 percent 9 previous employers. [‡] In the empirical analysis this variable is set to zero if the considered worker belongs to another skill level. Therefore, the summary statistics refer only to transitions of, respectively, high-, medium- and low-skilled workers. * Due to data protection guidelines not reported. For less than 1 percent of the observations firm size exceeds about 7500 employees.

Table C.2: Estimates of the parameters of the learning function based on different values for θ and full sample,
labor market size in terms of total regional employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
$\hat{\delta}^\dagger$	0.506*** (0.064)	0.226*** (0.026)	0.195*** (0.022)	0.193*** (0.022)	0.191*** (0.021)	0.190*** (0.021)	0.189*** (0.021)	0.187*** (0.021)	0.187*** (0.021)	0.186*** (0.020)	0.186*** (0.020)
$\hat{\gamma}^\dagger$	-3.929*** (0.896)	0.807* (0.338)	0.409 (0.282)	0.358 (0.278)	0.301 (0.273)	0.250 (0.269)	0.204 (0.266)	0.166 (0.262)	0.149 (0.261)	0.133 (0.259)	0.120 (0.258)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.179	0.189	0.193	0.194	0.194	0.194	0.194	0.195	0.195	0.195	0.195
RSS	8493.665	8391.207	8342.917	8339.378	8335.950	8333.103	8330.811	8329.053	8328.385	8327.821	8327.413
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\delta}^\dagger$	0.185*** (0.020)	0.184*** (0.020)	0.183*** (0.020)	0.182*** (0.020)	0.179*** (0.019)	0.175*** (0.019)	0.170*** (0.019)	0.165*** (0.019)	0.159*** (0.018)	0.131*** (0.017)	0.116*** (0.016)
$\hat{\gamma}^\dagger$	0.106 (0.257)	0.0847 (0.254)	0.0684 (0.252)	0.0556 (0.250)	0.0363 (0.246)	0.0177 (0.242)	-0.00728 (0.239)	-0.0443 (0.235)	-0.0942 (0.232)	-0.391 (0.215)	-0.532** (0.206)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.195	0.195	0.195	0.195	0.194	0.194	0.193	0.192	0.191	0.185	0.183
RSS	8327.088	8326.933	8327.348	8328.358	8332.256	8338.680	8347.391	8358.395	8370.971	8428.451	8452.746

[†] Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (12) depending on the value θ is set to. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

Table C.3: Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
$\hat{\delta}^\dagger$	0.473*** (0.073)	0.229*** (0.031)	0.209*** (0.027)	0.207*** (0.027)	0.205*** (0.026)	0.203*** (0.026)	0.201*** (0.026)	0.199*** (0.025)	0.199*** (0.025)	0.198*** (0.025)	0.197*** (0.025)
$\hat{\gamma}^\dagger$	-3.022*** (1.006)	-0.053 (0.398)	0.056 (0.348)	0.060 (0.343)	0.050 (0.339)	0.033 (0.334)	0.008 (0.329)	-0.026 (0.325)	-0.045 (0.323)	-0.066 (0.320)	-0.088 (0.318)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.249	0.256	0.259	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260
RSS	4792.050	4750.585	4728.430	4726.499	4724.810	4723.565	4722.783	4722.464	4722.471	4722.586	4722.793
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\delta}^\dagger$	0.197*** (0.025)	0.195*** (0.024)	0.194*** (0.024)	0.192*** (0.024)	0.189*** (0.023)	0.187*** (0.023)	0.184*** (0.022)	0.181*** (0.022)	0.179*** (0.021)	0.169*** (0.020)	0.164*** (0.019)
$\hat{\gamma}^\dagger$	-0.116 (0.316)	-0.167 (0.312)	-0.222 (0.308)	-0.278 (0.304)	-0.389 (0.296)	-0.492* (0.289)	-0.582** (0.282)	-0.662** (0.276)	-0.729*** (0.271)	-0.919*** (0.252)	-0.976*** (0.245)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.260	0.260	0.260	0.260	0.259	0.259	0.258	0.258	0.257	0.256	0.255
RSS	4723.153	4724.016	4725.163	4726.530	4729.672	4733.017	4736.267	4739.401	4742.284	4752.006	4755.805

[†] Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (12) depending on the value θ is set to. The estimates are obtained using a reduced sample which do not contain workers who acquired experience before 1995. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

Table C.4: Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment and the local share of high-skilled labor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
$\hat{\rho}_{\text{emp}}^\dagger$	-0.102 (0.087)	-0.077** (0.036)	-0.047 (0.030)	-0.041 (0.030)	-0.031 (0.030)	-0.020 (0.029)	-0.007 (0.029)	0.008 (0.028)	0.015 (0.028)	0.024 (0.028)	0.029 (0.028)
$\hat{\rho}_{\text{skill}}^\dagger$	17.834*** (1.792)	9.061*** (0.780)	7.144*** (0.642)	6.866*** (0.629)	6.512*** (0.616)	6.118*** (0.604)	5.682*** (0.593)	5.203*** (0.583)	4.954*** (0.578)	4.683*** (0.573)	4.524*** (0.578)
$\hat{\gamma}^\dagger$	1.846* (1.057)	3.660*** (0.412)	2.739*** (0.346)	2.607*** (0.341)	2.445*** (0.335)	2.275*** (0.330)	2.095*** (0.326)	1.905*** (0.322)	1.809*** (0.320)	1.707*** (0.318)	1.637*** (0.317)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.179	0.189	0.194	0.194	0.195	0.195	0.195	0.195	0.195	0.195	0.195
RSS	8488.580	8383.570	8335.843	8332.583	8329.574	8327.248	8325.570	8324.503	8324.195	8324.016	8323.928
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\rho}_{\text{emp}}^\dagger$	0.042 (0.028)	0.060** (0.027)	0.080*** (0.027)	0.101*** (0.027)	0.142*** (0.027)	0.182*** (0.026)	0.215*** (0.026)	0.240*** (0.026)	0.253*** (0.026)	0.1527552*** (0.025)	0.038 (0.024)
$\hat{\rho}_{\text{skill}}^\dagger$	4.120*** (0.565)	3.531*** (0.558)	2.917*** (0.551)	2.286*** (0.545)	1.018* (0.537)	-0.186 (0.531)	-1.216** (0.529)	-2.031*** (0.530)	-2.527*** (0.531)	-0.572 (0.532)	2.087*** (0.517)
$\hat{\gamma}^\dagger$	1.496*** (0.315)	1.282*** (0.312)	1.061*** (0.309)	0.836*** (0.306)	0.386 (0.302)	-0.046 (0.298)	-0.427 (0.295)	-0.745** (0.291)	-0.966*** (0.289)	-0.588** (0.274)	0.184 (0.263)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.195	0.195	0.195	0.195	0.194	0.194	0.193	0.192	0.191	0.185	0.183
RSS	8324.057	8324.648	8325.753	8327.360	8332.053	8338.673	8347.095	8357.573	8369.709	8428.386	8451.833

[†] Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (12) depending on the value θ is set to. Here, labor market size $Z_{s,\tau}$ is considered to be an aggregate of total regional employment and the local share of high-skilled labor, see equation (13). $\hat{\rho}_{\text{emp}}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment* and $\hat{\rho}_{\text{skill}}$ the corresponding elasticity with respect to the local *share of high-skilled labor* where experience was acquired. Robust standard errors are given in parentheses which are clustered by worker.

***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

Table C.5: Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment and the local share of high-skilled labor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
$\hat{\rho}_{\text{emp}}^\dagger$	-0.184*	-0.099**	-0.005	0.011	0.028	0.045	0.061*	0.077*	0.084**	0.091**	0.093**
	(0.102)	(0.044)	(0.039)	(0.038)	(0.038)	(0.038)	(0.037)	(0.037)	(0.037)	(0.036)	(0.036)
$\hat{\rho}_{\text{skill}}^\dagger$	19.221***	9.635***	6.048***	5.521***	4.956***	4.404***	3.878***	3.389***	3.166***	2.945***	2.893***
	(2.132)	(0.947)	(0.817)	(0.806)	(0.795)	(0.785)	(0.775)	(0.766)	(0.762)	(0.758)	(0.765)
$\hat{\gamma}^\dagger$	3.069**	2.962***	2.011***	1.850***	1.664***	1.472***	1.278***	1.087***	0.996**	0.903**	0.860**
	(1.198)	(0.490)	(0.430)	(0.425)	(0.419)	(0.414)	(0.409)	(0.404)	(0.401)	(0.399)	(0.398)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.250	0.257	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260
RSS	4788.304	4745.478	4725.699	4724.162	4722.874	4721.996	4721.537	4721.490	4721.612	4721.834	4722.082
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\rho}_{\text{emp}}^\dagger$	0.103***	0.113***	0.122***	0.129***	0.137***	0.140***	0.138***	0.135***	0.130***	0.104***	0.091***
	(0.036)	(0.036)	(0.035)	(0.035)	(0.034)	(0.034)	(0.033)	(0.032)	(0.032)	(0.030)	(0.029)
$\hat{\rho}_{\text{skill}}^\dagger$	2.566***	2.230***	1.951***	1.726**	1.423**	1.270*	1.233*	1.258*	1.325**	1.756***	1.981***
	(0.750)	(0.743)	(0.736)	(0.729)	(0.716)	(0.704)	(0.692)	(0.682)	(0.672)	(0.635)	(0.619)
$\hat{\gamma}^\dagger$	0.731*	0.570	0.425	0.295	0.085	-0.068	-0.170	-0.240	-0.285	-0.327	-0.309
	(0.394)	(0.389)	(0.385)	(0.380)	(0.371)	(0.363)	(0.356)	(0.349)	(0.343)	(0.321)	(0.312)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.260	0.260	0.260	0.260	0.259	0.259	0.258	0.258	0.257	0.256	0.255
RSS	4722.570	4723.567	4724.813	4726.250	4729.475	4732.853	4736.108	4739.230	4742.089	4751.620	4755.288

[†] Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (12) depending on the value θ is set to. Here, labor market size $Z_{s,\tau}$ is considered to be an aggregate of total regional employment and the local share of high-skilled labor, see equation (13). $\hat{\rho}_{\text{emp}}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment* and $\hat{\rho}_{\text{skill}}$ the corresponding elasticity with respect to the local *share of high-skilled labor* where experience was acquired. Robust standard errors are given in parentheses which are clustered by worker.

***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

Table C.6: Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment, parameters allowed to vary depending on size of labor market where experience is used

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
Now large labor market											
$\Delta\hat{\delta}^\dagger$	0.129 (0.114)	-0.032 (0.029)	-0.043** (0.018)	-0.043** (0.017)	-0.043*** (0.016)	-0.043*** (0.015)	-0.042*** (0.015)	-0.041*** (0.014)	-0.041*** (0.014)	-0.041*** (0.013)	-0.040*** (0.013)
$\Delta\hat{\gamma}^\dagger$	1.794 (1.587)	0.735* (0.383)	0.652*** (0.238)	0.638*** (0.226)	0.622*** (0.213)	0.605*** (0.202)	0.587*** (0.191)	0.570*** (0.182)	0.562*** (0.177)	0.553*** (0.173)	0.545*** (0.169)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.491*** (0.021)	0.229*** (0.021)	0.204*** (0.021)	0.203*** (0.021)	0.202*** (0.020)	0.201*** (0.020)	0.200*** (0.019)	0.199*** (0.019)	0.199*** (0.019)	0.199*** (0.017)	0.198*** (0.017)
$\hat{\gamma}^\dagger$	-4.596*** (1.119)	0.707* (0.379)	0.276 (0.304)	0.220 (0.298)	0.158 (0.292)	0.102 (0.286)	0.054 (0.281)	0.012 (0.277)	-0.006 (0.275)	-0.023 (0.273)	-0.038 (0.271)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	-0.363*** (0.121)	-0.042 (0.031)	-0.016 (0.019)	-0.014 (0.018)	-0.013 (0.017)	-0.012 (0.016)	-0.012 (0.015)	-0.011 (0.015)	-0.011 (0.014)	-0.010 (0.014)	-0.010 (0.014)
$\Delta\hat{\gamma}^\dagger$	2.294 (1.518)	0.272 (0.375)	0.140 (0.233)	0.133 (0.222)	0.128 (0.209)	0.124 (0.198)	0.122 (0.188)	0.119 (0.178)	0.119 (0.174)	0.118 (0.170)	0.117 (0.166)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.179	0.189	0.193	0.194	0.194	0.194	0.194	0.195	0.195	0.195	0.195
RSS	8489.861	8389.722	8342.154	8338.658	8335.265	8332.446	8330.174	8328.429	8327.766	8327.205	8326.798
θ^\dagger	(12) 2.594	(13) 2.358	(14) 2.144	(15) 1.949	(16) 1.611	(17) 1.331	(18) 1.100	(19) 0.900	(20) 0.729	(21) 0.200	(22) 0.001
Now large labor market											
$\Delta\hat{\delta}^\dagger$	-0.040*** (0.013)	-0.039*** (0.012)	-0.038*** (0.012)	-0.037*** (0.011)	-0.035*** (0.010)	-0.033*** (0.010)	-0.032*** (0.009)	-0.030*** (0.009)	-0.029*** (0.008)	-0.024*** (0.007)	-0.023*** (0.007)
$\Delta\hat{\gamma}^\dagger$	0.536*** (0.165)	0.520*** (0.158)	0.504*** (0.151)	0.488*** (0.145)	0.460*** (0.134)	0.434*** (0.126)	0.411*** (0.119)	0.390*** (0.112)	0.371*** (0.107)	0.308*** (0.092)	0.283*** (0.086)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.198*** (0.021)	0.197*** (0.021)	0.196*** (0.021)	0.195*** (0.021)	0.192*** (0.020)	0.188*** (0.020)	0.184*** (0.019)	0.179*** (0.019)	0.173*** (0.019)	0.144*** (0.017)	0.128*** (0.017)
$\hat{\gamma}^\dagger$	-0.053 (0.269)	-0.075 (0.266)	-0.093 (0.263)	-0.107 (0.260)	-0.128 (0.255)	-0.147 (0.250)	-0.173 (0.246)	-0.210 (0.242)	-0.259 (0.238)	-0.547** (0.220)	-0.679*** (0.211)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	-0.010 (0.013)	-0.010 (0.013)	-0.010 (0.012)	-0.010 (0.012)	-0.010 (0.011)	-0.010 (0.010)	-0.010 (0.010)	-0.011 (0.009)	-0.011 (0.009)	-0.012* (0.007)	-0.012* (0.007)
$\Delta\hat{\gamma}^\dagger$	0.118 (0.162)	0.118 (0.155)	0.118 (0.148)	0.119 (0.142)	0.122 (0.132)	0.128 (0.123)	0.134 (0.116)	0.142 (0.110)	0.149 (0.105)	0.170* (0.090)	0.170** (0.084)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.195	0.195	0.195	0.195	0.194	0.194	0.193	0.192	0.191	0.185	0.183
RSS	8326.474	8326.318	8326.729	8327.733	8331.610	8338.007	8346.684	8357.649	8370.182	8427.486	8451.725

[†] Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (12) depending on the value θ is set to. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. Labor markets with a total number of at least 750,000 employees are considered as large labor markets, local labor markets with less than 100,000 employees as small labor markets. Δ refers to the difference in a parameter in comparison to the reference group, i.e. medium size labor markets. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

Table C.7: Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment, parameters allowed to vary depending on size of labor market where experience is used

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
Now large labor market											
$\Delta\hat{\delta}^\dagger$	0.044 (0.130)	-0.021 (0.035)	-0.032 (0.024)	-0.033 (0.023)	-0.033 (0.022)	-0.034 (0.021)	-0.034* (0.020)	-0.034* (0.020)	-0.034* (0.019)	-0.034* (0.019)	-0.034* (0.019)
$\Delta\hat{\gamma}^\dagger$	2.860 (1.804)	0.692 (0.459)	0.575* (0.310)	0.566* (0.298)	0.556* (0.286)	0.546** (0.275)	0.536** (0.265)	0.527** (0.256)	0.522** (0.252)	0.517** (0.248)	0.512** (0.245)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.450*** (0.096)	0.211*** (0.035)	0.200*** (0.030)	0.199*** (0.029)	0.198*** (0.028)	0.198*** (0.028)	0.197*** (0.028)	0.197*** (0.027)	0.196*** (0.027)	0.196*** (0.027)	0.196*** (0.027)
$\hat{\gamma}^\dagger$	-3.638*** (1.273)	0.076 (0.449)	0.132 (0.377)	0.126 (0.371)	0.106 (0.364)	0.078 (0.358)	0.042 (0.353)	-0.0020 (0.347)	-0.026 (0.345)	-0.053 (0.342)	-0.081 (0.339)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	-0.254* (0.140)	-0.010 (0.037)	0.010 (0.025)	0.010 (0.024)	0.011 (0.024)	0.012 (0.023)	0.012 (0.022)	0.012 (0.021)	0.012 (0.021)	0.012 (0.020)	0.012 (0.020)
$\Delta\hat{\gamma}^\dagger$	0.903 (1.754)	-0.114 (0.454)	-0.192 (0.307)	-0.192 (0.295)	-0.191 (0.283)	-0.188 (0.272)	-0.183 (0.263)	-0.176 (0.254)	-0.173 (0.250)	-0.169 (0.246)	-0.165 (0.243)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.250	0.256	0.259	0.260	0.260	0.260	0.260	0.260	0.260	0.260	0.260
RSS	4789.369	4749.190	4727.687	4725.808	4724.171	4722.972	4722.232	4721.950	4721.973	4722.105	4722.327
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
Now large labor market											
$\Delta\hat{\delta}^\dagger$	-0.034* (0.019)	-0.034* (0.018)	-0.034* (0.018)	-0.033* (0.017)	-0.033** (0.017)	-0.032** (0.016)	-0.032** (0.016)	-0.031** (0.015)	-0.031** (0.015)	-0.029** (0.014)	-0.029** (0.013)
$\Delta\hat{\gamma}^\dagger$	0.508** (0.241)	0.499** (0.234)	0.490** (0.228)	0.481** (0.223)	0.465** (0.214)	0.451** (0.206)	0.439** (0.200)	0.428** (0.195)	0.418** (0.190)	0.389** (0.177)	0.378** (0.172)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.195*** (0.027)	0.195*** (0.026)	0.194*** (0.026)	0.194*** (0.025)	0.192*** (0.025)	0.190*** (0.024)	0.189*** (0.024)	0.187*** (0.023)	0.185*** (0.023)	0.177*** (0.021)	0.172*** (0.021)
$\hat{\gamma}^\dagger$	-0.113 (0.337)	-0.174 (0.332)	-0.238 (0.327)	-0.303 (0.322)	-0.429 (0.314)	-0.545* (0.306)	-0.645** (0.299)	-0.733** (0.293)	-0.807*** (0.287)	-1.015*** (0.268)	-1.079*** (0.260)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	0.011 (0.020)	0.011 (0.019)	0.011 (0.019)	0.010 (0.018)	0.009 (0.018)	0.009 (0.017)	0.008 (0.017)	0.007 (0.016)	0.007 (0.016)	0.006 (0.015)	0.006 (0.014)
$\Delta\hat{\gamma}^\dagger$	-0.161 (0.239)	-0.152 (0.233)	-0.143 (0.227)	-0.135 (0.222)	-0.120 (0.213)	-0.106 (0.205)	-0.0955 (0.199)	-0.0865 (0.194)	-0.0792 (0.190)	-0.0605 (0.176)	-0.0553 (0.171)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.260	0.260	0.260	0.260	0.259	0.259	0.258	0.258	0.257	0.256	0.255
RSS	4722.701	4723.590	4724.761	4726.148	4729.323	4732.691	4735.959	4739.106	4741.997	4751.734	4755.535

[†] Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (12) depending on the value θ is set to. The estimates are obtained using a reduced sample which do not contain workers who acquired experience before 1995. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. Labor markets with a total number of at least 750,000 employees are considered as large labor markets, local labor markets with less than 100,000 employees as small labor markets. Δ refers to the difference in a parameter in comparison to the reference group, i.e. medium size labor markets. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 4).

References

- Andersson, M., Klaesson, J., Larsson, J. P. (2014). The sources of the urban wage premium by worker skills: Spatial sorting or agglomeration economies? *Papers in Regional Science*, 93(4), 727–747.
- vom Berge, P., Burghardt, A., Trenkle, S. (2013). Sample of integrated labour market biographies: Regional file 1975-2010 (SIAB-R 7510). FDZ-Datenreport 09/2013 (en), Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB), Nuremberg.
- Blanchflower, D. G., Oswald, A. J. (1990). The wage curve. *The Scandinavian Journal of Economics*, 92(2), 215–235.
- Braakmann, N. (2009). Are there social returns to both firm-level and regional human capital? Evidence from German social security data. Working Paper Series in Economics 143, University of Lüneburg, Institute of Economics.
- Ciccone, A., Peri, G. (2006). Identifying human-capital externalities: Theory with applications. *Review of Economic Studies*, 73(2), 381–412.
- Combes, P.-P., Duranton, G., Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2), 723–742.
- Combes, P.-P., Duranton, G., Gobillon, L. (2011). The identification of agglomeration economies. *Journal of Economic Geography*, 11(2), 253–266.
- Combes, P.-P., Gobillon, L. (2015). The empirics of agglomeration economies. In G. Duranton, J. V. Henderson, W. C. Strange (Eds.) *Handbook of Regional and Urban Economics*, vol. 5, (pp. 247–348). Elsevier.
- Combes, P.-P., Magnac, T., Robin, J.-M. (2004). The dynamics of local employment in France. *Journal of Urban Economics*, 56(2), 217–243.
- D’Costa, S., Overman, H. G. (2014). The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics*, 48(C), 168–179.
- Di Addario, S., Patacchini, E. (2008). Wages and the city. Evidence from Italy. *Labour Economics*, 15(5), 1040–1061.
- Driscoll, J. C., Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4), 549–560.
- Duranton, G. (2007). Human capital externalities in cities: Identification and policy issues. In R. J. Arnott, D. P. McMillen (Eds.) *A Companion to Urban Economics*, chap. 2, (pp. 24–39). Blackwell Publishing Ltd.
- Duranton, G., Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91(5), 1454–1477.
- Duranton, G., Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson, J. F. Thisse (Eds.) *Handbook of Regional and Urban Economics*, vol. 4, chap. 48, (pp. 2063–2117). Elsevier.
- Eberle, J., Jacobebbinghaus, P., Ludsteck, J., Witter, J. (2011). Generation of time-consistent industry codes in the face of classification changes - Simple heuristic based on the Establishment History Panel (BHP). FDZ-Methodenreport 05/2011, Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB), Nuremberg.
- Fitzenberger, B., Osikominu, A., Völter, R. (2005). Imputation rules to improve the education variable in the IAB employment subsample. FDZ-Methodenreport 03/2005, Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB), Nuremberg.

- Garloff, A., Roth, D. (2016). Regional age structure and young workers' wages. IAB-Discussion Paper 06/2016, Institute for Employment Research (IAB) - The Research Institute of the Federal Employment Agency, Nuremberg.
- Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics*, 46(2), 254–277.
- Glaeser, E. L., Maré, D. C. (2001). Cities and skills. *Journal of Labor Economics*, 19(2), 316–42.
- Gould, E. D. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *Review of Economic Studies*, 74(2), 477–506.
- Hamann, S., Niebuhr, A., Peters, J. C. (2016). Benefits of dense labour markets - Evidence from transitions to employment in Germany. IAB-Discussion Paper 13/2016, Institute for Employment Research (IAB) of the Federal Employment Agency (BA), Nuremberg.
- Heuermann, D. (2011). Human capital externalities in Western Germany. *Spatial Economic Analysis*, 6(2), 139–165.
- Heuermann, D., Halfdanarson, B., Suedekum, J. (2010). Human capital externalities and the urban wage premium: Two literatures and their interrelations. *Urban Studies*, 47(4), 749–767.
- Hoechle, D. (2007). Robust standard errors for panel regression with cross-section dependence. *The Stata Journal*, 7(3), 281–312.
- Jacobs, J. (1969). *The Economy of Cities*. New York: Random House.
- Jovanovic, B., Rob, R. (1989). The growth and diffusion of knowledge. *The Review of Economic Studies*, 56(4), 569–582.
- Kosfeld, R., Werner, A. (2012). Deutsche Arbeitsmarktreionen - Neuabgrenzung nach den Kreisgebietsreformen 2007-2011. *Raumforschung und Raumordnung*, 70, 49–64.
- Lehmer, F., Möller, J. (2010). Interrelations between the urban wage premium and firm-size wage differentials: a microdata cohort analysis for Germany. *The Annals of Regional Science*, 45(1), 31–53.
- List, F. (1838). *Das deutsche National-Transport-System in volks- und staatswirthschaftlicher Beziehung*. Altona/Leipzig: Johann Friedrich Hammerich.
- Marshall, A. (1890). *Principles of Economics*. Macmillan, London.
- Matano, A., Naticchioni, P. (2016). What drives the urban wage premium? Evidence along the wage distribution. *Journal of Regional Science*, 56(2), 191–209.
- Mincer, J., Ofek, H. (1982). Interrupted work careers: Depreciation and restoration of human capital. *The Journal of Human Resources*, 17(1), 3–24.
- Moretti, E. (2004a). Estimating the social return to higher education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics*, 121(1-2), 175–212.
- Moretti, E. (2004b). Human capital externalities in cities. In J. V. Henderson, J. F. Thisse (Eds.) *Handbook of Regional and Urban Economics*, vol. 4, chap. 51, (pp. 2243–2291). Elsevier.
- Moretti, E. (2004c). Workers' education, spillovers, and productivity: Evidence from plant-level production functions. *American Economic Review*, 94(3), 656–690.
- Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro unit. *The Review of Economics and Statistics*, 72(2), 334–38.
- Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science*, 50(1), 203–219.

- Reichelt, M. (2015). Using longitudinal wage information in linked data sets - The example of ALWA-ADIAB. FDZ-Methodenreport 01/2015, Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB), Nuremberg.
- De la Roca, J., Puga, D. (2016). Learning by working in big cities. *The Review of Economic Studies*, forthcoming.
- Roscher, W. (1878). *Ansichten der Volkswirtschaft aus dem geschichtlichen Standpunkte*, vol. 2. Leipzig/Heidelberg: C. F. Wintersche Verlagshandlung, 3 ed.
- Sydsæter, K., Hammond, P., Seierstad, A., Strøm, A. (2008). *Further mathematics for economic analysis*. Harlow: Financial Times Prentice Hall, 2 ed.
- Wheeler, C. H. (2006). Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics*, 60(2), 162–184.
- Wheeler, C. H. (2008). Local market scale and the pattern of job changes among young men. *Regional Science and Urban Economics*, 38(2), 101–118.
- Wooldridge, J. M. (2013). *Introductory Econometrics: A Modern Approach*. Mason, OH: South-Western, 5 ed.
- Yankow, J. J. (2006). Why do cities pay more? An empirical examination of some competing theories of the urban wage premium. *Journal of Urban Economics*, 60(2), 139–161.