

Peters, Jan Cornelius

**Conference Paper**

## Quantifying the effect of labor market size on learning externalities

56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria

**Provided in Cooperation with:**

European Regional Science Association (ERSA)

Suggested Citation: Peters, Jan Cornelius (2016) : Quantifying the effect of labor market size on learning externalities, 56th Congress of the European Regional Science Association: "Cities & Regions: Smart, Sustainable, Inclusive?", 23-26 August 2016, Vienna, Austria, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<http://hdl.handle.net/10419/174667>

**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

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

July 18, 2016

Paper prepared for '6th European Meeting of the Urban Economics Association'  
at the 56th ERSa Congress, August 23-26, 2016, Vienna, Austria

## Abstract

This paper provides strong empirical evidence that individual productivity significantly depends on the size of the local labor market where a worker previously acquired work experience. The analysis uses German micro data from the Institute for Employment Research (IAB) on new employment relationships within the period 2005 to 2011 and individual employment biographies from 1975 onwards. The results suggest that dynamic agglomeration benefits in general and learning externalities in particular play an important role, meaning workers accumulate significantly more valuable human capital by working in large than in small local 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 1 percent. At 10 years of experience the corresponding gain amounts to about 4.5 percent. A key factor seems to be the 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 Annetrin 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 elsewhere. Figure 1 illustrates disparities in regional wages for Germany. The difference between the largest local labor markets 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 U.S. and Combes et al. (2008) for France. This raises the question, why firms in agglomerations<sup>1</sup> pay high wages and do not relocate to regions where labor is cheaper.<sup>2</sup>



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 why workers might be more productive in urban than in non-urban labor markets. Most of the underlying theoretical arguments were already noted at the end of 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 sorting of more able workers into cities, arguing that higher wages in cities reflect higher abilities of 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 once dynamic

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

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

agglomeration benefits (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, and 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 an immediate wage increase by about one to two percent: see, 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, i.a., 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 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. 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 market was where experience was acquired. Pioneered by the work of De la Roca and Puga (2016), recent findings support this hypothesis. 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. Hence, a worker who previously acquired work experience in one of the largest cities receives today c.p. a higher wage than a worker who previously worked somewhere else. These results raise the question of how the value of experience exactly depends on the size of the local labor market where it was acquired. As far as I know, this has not been analyzed yet.

To provide such insights, I, in contrast to previous papers, explicitly estimate the elasticity of wage with respect to the size of the local labor market where work experience was acquired, employing a micro-econometric framework described by Combes and Gobillon (2015). The elasticity indicates by which rate the marginal value of experience increases if the labor market where experience was acquired becomes bigger.

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 where 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. The wage indicates how firms value experience when they hire a worker depending on the size of the

labor market where it was acquired. 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 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. As it presumably is especially skilled workers one learns from, 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. To the best of my knowledge the importance of the different characteristics with regard to the value of work experience has not been analyzed so far.

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 on learning externalities in agglomerations**

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) stresses 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 U.S., 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 emphasised 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 distinguishing 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. As per 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 mechanisms and analyzes 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 with city experience a faster wage growth compared to rural workers with no city experience. 'City experience' is based on a very rough definition, as it is a binary variable indicating "having worked in a city at some point" (D'Costa/Overman, 2014: 174). 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 things, 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 two types of firms: small and large, as well as two types of regions: rural and urban. 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.

"Currently one of the most complete studies on the topic" (Combes/Gobillon, 2015: 263) is a study by De la

Roca and Puga. As with previous authors, De la Roca and Puga (2016) control for unobserved individual characteristics that impact on the individual wage level and distinguish static and dynamic effects.<sup>3</sup> Additionally, they provide empirical evidence that strongly suggest that dynamic benefits depend on learning effects. To identify dynamic agglomeration benefits, they do not analyze individual wage growth after migrating from rural to urban areas as the studies discussed so far, but instead analyze the value of individual work experience, depending on the size of the city where it was acquired. They distinguish experience that was acquired in Madrid or Barcelona, i.e., the largest Spanish cities, experience acquired in Valencia, Sevilla, Zaragoza, the next largest cities, and experience acquired in the rest of the country. Their results show that experience acquired in Madrid or Barcelona has the highest value. In addition, experience acquired in Valencia, Sevilla or Zaragoza, is also more valuable than experience acquired in the rest of the county. 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 support 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 is highly portable. It does not significantly differ if 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) indicate that the urban wage growth premium is larger for young than for older workers. They employ different data sets and find stronger effects based on the National Longitudinal Survey of Youth (1979) that contains only information on white men aged 16 to 35. Results obtained by D'Costa and Overman (2014) support this suggestion.

De la Roca and Puga (2016) find that in particular 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) observe 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 U.S. 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

---

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

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. As noted by Heuermann et al. (2010) a related strand of literature exists that discusses the role of local learning externalities as well. 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 Duranton (2007) and Moretti (2004b).

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 which increases also the productivity of low-skilled workers through working with a larger stock of physical capital. 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).

As another reason on why localized human capital impacts on individual productivity this strand of literature refers to 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 as 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 U.S. where 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) and for Germany Heuermann (2011) and Braakmann (2009).

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. These studies often 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.



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 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 where 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 where 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 regional labor market regions in the period 1975-2011, this paper is, to the best of my knowledge, the first that 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 where experience was acquired and highly 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 analyse wages of new employment relationships making use of a micro-econometric framework described by Combes and Gobillon (2015). Suppose a worker  $i$  is hired by a firm at date  $t$  and the logarithm of the wage which the worker receives,  $w_{i,t}$ , is given by equation (1); compare Combes/Gobillon (2015: eq. 5.20).

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

$u_i$  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  $\tau$  and 0 else,  $v_{s,\tau,r}$  a time varying fixed effect for being at date  $\tau$  in region  $s$  and at date  $t$  in region  $r$ ,  $\mathbf{x}_{i,t}$  a vector of time-varying individual characteristics with parameter vector  $\beta$  and  $\varepsilon_{i,t}$  an error term. In the empirical analysis subscript  $t$  refers to dates between January 1, 2005 and December 31, 2011 and  $\tau = 1$  to January 1, 1975.

The individual fixed effect captures all unobserved time-invariant characteristics of a worker that determine her 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  $1(O_{i,\tau} = 1) \times v_{s,\tau,r}$ . Given that individual  $i$  was working at date  $\tau$  which is indicated by  $O_{i,\tau}$ , the parameter  $v_{s,\tau,r}$  may be interpreted as a weight of the work experience that  $i$  acquired at  $\tau$ . This marginal value of experience may depend on the labor market and date of acquirement, indicated by  $s$  and  $\tau$  respectively, and the labor market of usage,  $r$ .  $\sum_{\tau=1}^{t-1} 1(O_{i,\tau} = 1) \times v_{s,\tau,r}$  captures the wage premium that individual  $i$  receives at date  $t$  due to her entire work experience acquired until date  $t - 1$ .<sup>4</sup>

Imagine, I ignore that the value of experience presumably depends on the time and the location of its acquirement. 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.,  $E_{i,t} = \sum_{\tau=1}^{t-1} 1(O_{i,\tau} = 1)$ . 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) not.

In order to identify dynamic agglomeration benefits based on (1), Combes and Gobillon (2015) suggest a two-step procedure. The first step requires to estimate all parameters  $v_{s,\tau,r}$ . In a second step,  $\hat{v}_{s,\tau,r}$  is 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  $v_{s,\tau,r}$  which have to be estimated in the first step becomes very large (Combes/Gobillon, 2015).

De la Roca and Puga (2016) estimate a special case of equation (1) where they impose restrictions on  $v_{s,k,r}$  such that the number of parameters reduces. Using Spanish data, they distinguish experience only according to three classes of cities where 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 date when it was acquired. In this case, equation (1) simplifies to (3); compare De la Roca/Puga (2016: eq. 1).

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

$E_{i,s,t}$  denotes the amount of experience that was acquired until  $t - 1$  in, respectively, Madrid or Barcelona, Valencia, Sevilla or Zaragoza, 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.

<sup>4</sup> As 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, see, e.g., De la Roca/Puga (2016), Matano/Naticchioni (2016).

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 few groups of German local 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 replace  $v_{s,k,r}$  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 where experience was acquired.

As with De la Roca and Puga (2016), Andersson et al. (2014), Combes et al. (2008) and others, I assume that there is a log-linear relationship between labor market size  $Z_{s,\tau}$  and net-agglomeration gains and replace  $v_{s,\tau,r}$  by equation (4) which may be interpreted as a learning function capturing externalities of acquiring work experience in large local labor markets.<sup>5</sup>

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

In accordance with Duranton and Puga (2004), marginal learning benefits with regard to labor market size are assumed to be positive but decreasing.  $\delta_r \ln\left(\frac{Z_{s,\tau}}{\underline{Z}}\right)$  denotes the value of the additional human capital that a worker accumulates at date  $\tau$  by working in a labor market with size  $Z_{s,\tau}$ , given the labor market size exceeds threshold  $\underline{Z}$ .<sup>6</sup> The value of the accumulated human capital may vary according to the labor market where it is used. This is captured by  $\delta_r$ . For example, matching benefits may arise if the accumulated human capital is used in a large local labor market since it may be more likely there than in a small local labor market that a firm demands the specific knowledge that a worker previously accumulated. However, to obtain the baseline model of my empirical analysis, I assume that the value of accumulated human capital is the same in all regions, i.e.,  $\delta_r = \delta$  and  $v_{s,\tau,r} = v_{s,\tau}$ . Moreover, I rewrite the learning function such that I obtain:

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

with  $\gamma = -\delta \ln \underline{Z}$ . Inserting (5) into (1) leads to:

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

with  $z_{s,\tau} = \ln Z_{s,\tau}$ .<sup>7</sup> As in equation (2),  $E_{i,t}$  refers to individual experience, measured by the number of working days before date  $t$ .  $\sum_{\tau=1}^{t-1} 1(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau}$  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 where experience was acquired.

<sup>5</sup> 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 addresses that Duranton and Puga (2004) argue that the learning function depends on the local amount of *skilled* labor.

<sup>6</sup> I gratefully thank Johannes Bröcker for the suggestion to introduce  $\underline{Z}$  in the learning function.

<sup>7</sup> In equation (6), I assume that all local labor markets are larger than threshold  $\underline{Z}$ . This has to be verified after the regression analysis. If some local labor markets are smaller than  $\underline{Z}$ , the labor market size has to be set to  $\underline{Z}$  and an iterative procedure has to be applied to obtain the solution for equations (4) and (5), respectively.

The corresponding parameter  $\delta$  denotes the elasticity of wage received at date  $t$  with respect to the size of the labor market where individual  $i$  was working at date  $\tau$ . Accordingly, the wage rate at date  $t$  would be  $\delta E_{i,t}$  percent larger if all labor markets individual  $i$  was working in until date  $t$  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  $v_{s,\tau,r}$  is unity, given that individual  $i$  was employed at date  $\tau$  (see equation 1). Therefore, the estimated *elasticity of wage* with respect to the size of the labor market where experience was acquired can also be interpreted as the *elasticity of the value of experience* with respect to labor market size (compare equation 4). Based on  $\hat{\gamma}$  and  $\hat{\delta}$ ,  $\hat{Z}$  can be computed.

So far, I assumed that the marginal return to work experience does not depend on the amount of previously acquired experience. However, it is likely that the additional gain at date  $\tau$  is lower the more human capital already was accumulated. For example, results by De la Roca and Puga (2016) show that the marginal value of experience decreases in the number of days worked previously and additionally if experience was acquired in a large labor market. Taking this into account, I augment equation (4) such that  $v_{s,\tau,r}$  is given by equation (7) where  $\theta$  denotes the rate by which the marginal return to experience decreases in previously accumulated human capital with  $\theta \in (0; 1)$  and  $Q_{i,\tau,r(i,t)}$  denotes the value of human capital that individual  $i$  accumulated before date  $\tau$ .<sup>8</sup>

$$v_{r(i,\tau),\tau,r(i,t),i} = \begin{cases} \delta_{r(i,t)} \ln \left( \frac{Z_{r(i,\tau),\tau}}{\underline{Z}} \right) - \underbrace{\theta \sum_{k=1}^{\tau-1} 1(O_{i,k} = 1) \times v_{r(i,k),k,r(i,t),i}}_{Q_{i,\tau,r(i,t)}} & \text{if } Z_{r(i,\tau),\tau} > \underline{Z} \\ 0 & \text{else} \end{cases} \quad (7)$$

$Q_{i,\tau,r(i,t)}$  may also be written as (see, e.g., Sydsæter et al., 2008):

$$\begin{aligned} Q_{i,\tau,r(i,t)} &= \sum_{k=1}^{\tau-1} 1(O_{i,k} = 1) \times (1 - \theta)^{E_{\tau-1} - E_k} \delta_{r(i,t)} \max \left\{ \ln \left( \frac{Z_{r(i,k),k}}{\underline{Z}} \right); 0 \right\} \\ &= \gamma_{r(i,t)} \sum_{k=1}^{\tau-1} 1(O_{i,k} = 1) \times (1 - \theta)^{E_{\tau-1} - E_k} + \\ &\quad + \delta_{r(i,t)} \sum_{k=1}^{\tau-1} 1(O_{i,k} = 1) \times (1 - \theta)^{E_{\tau-1} - E_k} \max \left\{ \ln Z_{r(i,k),k}; \ln \underline{Z} \right\}. \end{aligned} \quad (8)$$

For  $k = \tau - 1$ , the weight  $(1 - \theta)^{E_{\tau-1} - E_k}$  equals one. Hence, the size of the labor market where 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,  $(1 - \theta)^{E_{\tau-1}}$ , addressing that the human capital acquired at this date, diminishes the amount of human capital that is additionally acquired at any day in the future.

In order to estimate  $\gamma_r$  and  $\delta_r$  based on the augmented learning function given by equation (7), I replace the term  $\sum_{\tau=1}^{t-1} 1(O_{i,\tau} = 1) \times v_{r(i,\tau),\tau,r(i,t)}$  in equation (1) by  $Q_{i,t,r(i,t)}$  as given by equation (8). Both denote the value of experience acquired before date  $t$ . I obtain:

$$w_{i,t} = u_i + \mu_{r(i,t),t} + \gamma_{r(i,t)} \sum_{k=1}^{\tau-1} 1(O_{i,k} = 1) \times (1 - \theta)^{E_{\tau-1} - E_k} +$$

<sup>8</sup> The marginal value of experience acquired in a region now varies across individuals. Therefore, I add the subscript  $i$ .

$$+ \delta_{r(i,t)} \sum_{k=1}^{\tau-1} 1(O_{i,k} = 1) \times (1 - \theta)^{E_{\tau-1} - E_k} \ln Z_{r(i,k),k} + \mathbf{x}'_{i,t} \boldsymbol{\beta} + \varepsilon_{i,t}. \quad (9)$$

This specification takes into account, that workers presumably accumulate less human capital per day the more human capital previously was acquired, which depends on the number of previous days at work and the size of the labor markets where previous employers were located. Since equation (9) is highly non-linear in  $\theta$ , I run the regression analysis several with different values for  $\theta$ . Once a value is imposed, the linear fixed effects estimator can be used to estimate the other parameters of the learning function. I check afterwards which specification minimizes the residual sum of squares. For  $\theta = 0$  equation (9) simplifies to the baseline model given by equation (6).<sup>9</sup>

After estimating a restricted specification with  $\gamma_{r(i,t)} = \gamma$  and  $\delta_{r(i,t)} = \delta$ , I let the parameters vary across different groups of local labor markets which are defined according to their size. Thereby, I test whether the value of experience depends not only on the size of the labor market where the experience was acquired, but also on the size of the labor market where it is used.

### 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 significantly 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 impact on wage.

When estimating equation (6), the within transformation that is applied to eliminate the individual fixed effect washes out all information on experience that was acquired before the date at which the first analyzed new employment relationship of worker  $i$  starts. Hence, in the baseline specification  $\hat{\delta}$  and  $\hat{\gamma}$  are only based on information that refers to the period 2005 to 2011. In contrast, estimating equation (9) which is based on the augmented learning function, requires information on all previous employment spells of the workers.

Also the risk of reverse causality should be of minor concern. The pivotal explanatory variable is the size of the labor markets where 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 where experience was acquired and where it is used are likely significantly correlated. However, the included region-year fixed effects control for all characteristics of the region where 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

<sup>9</sup> I gratefully thank Johannes Bröcker for the suggestion how to consider that the marginal value of experience decreases in the amount of previously accumulated human capital.

standard errors at the local industry level. However, this is not possible as the model includes individual fixed effects and workers are mobile between regions and industries. The standard errors that I report are clustered at the individual level. They are robust with regard to heteroskedasticity and serial correlation in the error terms (Wooldridge, 2013).

## 4 Data

### 4.1 Individual data set

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.<sup>10</sup> My data set comprises a 5 percent random sample of all employees with at least one notification to social security between 2005 and 2011.<sup>11</sup>

Since I need to observe the full employment biographies of the workers, I exclude all individuals for which it is not possible, among others individuals born before 1960 as the IEB contains only information on employment from 1975 onwards. A detailed description is provided in Appendix A. For the remaining sample of workers my data set captures all spells of employment subject to social security contributions. I use them to identify new employment relationships within the period between 2005 and 2011 focusing on new full-time employment subject to social security contributions outside the public sector and outside the temporary employment industry. I enrich the data set with information on the individual pre-employment status, meaning job-search status, benefits receipt, and participation in measures of active labor market policy which is also provided by the IEB.

For the new employment relationships I observe the corresponding gross daily wage as well as further information on the new job (e.g., 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 A.

Using all previous employment spells, I compute the individual labor market experience at the date at which the new employment relationship starts.<sup>12</sup> Moreover, I generate indicators that provide information on the recent labor market biography, i.e., the employment status and length of non-employment before the new employment relationship, unemployment benefits and participation in measures of active labor market policy. Detailed information on all variables that I use is provided by table A.1. Summary statistics can be found in table B.1.

---

<sup>10</sup>For a more detailed description of the IEB see vom Berge et al. (2013).

<sup>11</sup>Hamann et al. (2016) use the described data set to analyze the impact of labor market density on matching quality.

<sup>12</sup>Since the IEB does not contain information on the self-employed and civil servants, the calculated labor market experience refers to employment subject to social security contributions only.

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.<sup>13</sup>

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 control variable addressing that frequent job changes are discussed to improve the matching between firms and workers.<sup>14</sup> The establishment identifier allows me in addition 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 Size of local labor markets where experience was acquired

Di Addario and Patacchini (2008) show for Italy that the effect of local density on wages rapidly declines in distance. Only for population mass within a distance up to 12 kilometers, they find a significant impact on wages. It points out 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 regional labor market regions as spatial units for my analysis. On average they have a radius of about 27 km (table 1). The regions are supposed to represent integrated local labor markets containing NUTS3-regions connected through intense commuting. Therefore, it is reasonable to assume that workers exchange knowledge presumably within these regions and that their size in terms of employment is an appropriate measure for local learning opportunities. The definition of German local labor markets is taken from Kosfeld and Werner (2012), who define 141 regional labor market regions employing the factor analysis with oblique Oblimin rotation to German commuter structure between NUTS3-regions. One critical issue is that commuting patterns change over time, wherefore the chosen definition of regional labor market regions might be an inappropriate approximation of local labor markets decades ago. As a robustness check, I therefore repeat my estimations focussing on those workers who only acquired work experience 1995 or later, i.e., at least five years after reunification of East and West Germany.<sup>15</sup>

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

<sup>13</sup> Gould (2007) uses the same data set as the other authors, but does not provide summary statistics in his paper.

<sup>14</sup> Different units of one firm that are located in different municipalities are considered as independent establishments. Unfortunately, it is not possible to identify whether different establishments belong to the same firm.

<sup>15</sup> Please recall that the information on experience acquired before 2005 is only used when omitting individual fixed effects or when estimating equation (9) that is based on the augmented learning function given by equation (7) Furthermore, in the latter case the weight of the size of the labor market where experience was acquired is lower the more additional experience was acquired thereafter; see also equation (8).

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 which is defined as the local share of workers with a university degree. It varies between 2 percent and 20 percent (table 1).<sup>16</sup> The correlation between total regional employment and the local share of high-skilled workers is 0.38.

Table 1: Size of German regional labor market regions

	Local labor market size in terms of . . .		
	thousand employees <sup>†</sup>	share of high-skilled workers	radius in km <sup>‡</sup>
Minimum	13.8	2.3	10.3
Median	110.8	6.0	26.0
Mean	179.1	7.0	27.0
Maximum	1206.5	20.0	51.8
	(Berlin)	(Jena)	(Hamburg)

N=141.

<sup>†</sup> Number of employees subject to social security contributions in thousand. <sup>‡</sup> Under the assumption the regions are circular.

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.

### 4.3 Labor mobility

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

Descriptive statistics show that workers acquired work experience in either small or large local labor markets. Mobility between labor markets that significantly differ in size is rather low. In 62 percent the largest region in terms of employment where a worker acquired experience before the considered new employment relationship was less than twice as large as the smallest region where the same worker worked previously. In addition, in about 79 percent the standard deviation of the size of the labor market where a worker acquired work experience is less than half of the individual average size of the labor markets where experience was acquired.<sup>17</sup>

Table 2 compares the size of the labor markets where experience was acquired and where it is used. 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 where 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).

<sup>16</sup> The statistics in table 1 refer to region specific averages for the considered period, i.e., for West Germany 1975-2011, for East Germany 1993-2011. The share of high-skilled workers increased over time. The East German labor market of Jena still has the highest share of workers with a university degree if also for West German labor markets only information from 1993 onwards is considered.

<sup>17</sup> The standard deviation is measured on the individual level and refers to the variation in the size of the labor markets where a worker acquired work experience.



Table 2: Size of labor market where experience was acquired and where it is used

Share in %		Size of labor market where 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 where	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 where 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.

The sample mean of total work experience is 9.4 years (table B.1). On average 59 percent of this previous work experience (5.7 years) refer to employment in the same region, 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 where the new employer is located.

Taken together, workers tend to continue to work in the region where they acquired experience, or in a region of a comparable size. This implies, learning benefits in large local labor markets likely result in wage disparities between small and large local labor markets.

In order to test the hypothesis that the value of work experience depends not only on the size of the labor market where experience was acquired, but also on the size of the labor market where 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 where individual experience was acquired. Despite the rather low degree of mobility, there should be sufficient variation to test this hypothesis.

## 5 Results

### 5.1 Control variables

Table 3 summarizes the 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 this problem, I include individual fixed effects (FE, specification (2)) as proposed 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 ( $= e^{0.233} - 1$ )

Table 3: Control variables

	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) <sup>-1</sup> )	-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	525265		336286	
OLS: R <sup>2</sup> , FE: within R <sup>2</sup>	0.616		0.183	
Worker fixed effects	No		Yes	

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. All models include industry, occupation, and region-year fixed effects as well as total work experience and the pivotal variable  $\sum_{\tau=1}^{t-1} 1(O_{i,\tau} = 1) \times \ln z_{r(i,\tau),\tau}$ , see equation (6) and column (1) and (2) of table 5.

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 A), I include the length of the considered employment spell, 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 benefits 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, i.e., the number of different establishments a worker was working in until the considered employment spell 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 employment spell, the lower is the corresponding wage. Non-employment of more than one year results in a wage loss of about 4 percent. Following Mincer and Ofek (1982), a reasonable explanation is that human capital depreciates during periods of non-employment. However, it is important to note 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, my findings indicate that a worker receives a 2.5 percent lower wage if she was previously registered by the Federal Employment Agency as job seeker than if 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 receive a 0.8 percent higher 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 is in Germany paid only within the first year after the end of an employment spell of at least one year (with exceptions). Thereafter, 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 almost 2 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 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 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 in the local workforce.<sup>18</sup>

The agglomeration economies literature points out that also the local industry structure 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.<sup>19</sup> 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 wages.

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 a composite of an externality effect and an 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 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 (see e.g. Blanchflower/Oswald, 1990). I do obtain negative elasticities, even though the coefficient of the regional unemployment rate among low skilled labor is not statistically different from zero.

## 5.2 Main results

Table 4 summarizes results that I obtain adopting the empirical strategy suggested by De la Roca and Puga (2016) to analyse dynamic agglomeration benefits that is described by equation (3). In addition to the variables reported in table 4 and the control variables reported in table 3, 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, i.e., Berlin,

---

<sup>18</sup> 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. They are available from the author upon request.

<sup>19</sup> 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.

Hamburg and Munich.<sup>20</sup> 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 4: The value of labor market specific work experience

	(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 $\dagger$			-0.005*** (0.000)
Experience (Experience) <sup>2</sup>	0.016*** (0.000)	0.033*** (0.001)	0.065*** (0.001)
Experience Berlin\Hamburg\Munich, now elsewhere $\dagger$	0.060*** (0.005)	0.007 (0.006)	0.020 (0.016)
Experience Berlin\Hamburg\Munich $\times$ experience, now elsewhere $\dagger$			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 $\dagger$			-0.003*** (0.001)
New employment relationships	525265	336286	336286
OLS: R <sup>2</sup> , FE: within R <sup>2</sup>	0.613	0.180	0.191
Worker fixed effects	No	Yes	Yes

$\dagger$  Coefficient and standard error multiplied by 10. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. Work experience measured on a daily bases and expressed in years. All models including control variables as well as year, industry, occupation, and region fixed effects (see table 3).

Similar to De la Roca and Puga (2016) as well as 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. 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 include 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 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. For the value of experience acquired outside the large labor markets increases only marginally 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) Individuals with few work experience gain more from dynamic agglomeration benefits than more experienced workers. Moreover, the marginal value of experience

<sup>20</sup>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.

decreases in the amount of experience especially if the previous experience was acquired in the largest local labor markets. A reasonable explanation is that those workers c.p. already accumulated more human capital than individuals who worked elsewhere. (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 3 percent and 2 percent higher value than the first year acquired elsewhere.<sup>21</sup>

The results reported in table 4 and by other scholars raise the question of how the value of experience exactly depends on the size of the local labor market where it was acquired. To provide corresponding insights, I estimate the parameters of equations (6) and (9). The estimates of the pivotal parameters are summarized in table 5. 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 5 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 the 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 where experience was acquired at one day in the past, denoted by  $\delta$ , amounts to  $0.116 \times 10^{-4}$ . 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 4 indicate that this assumption does not hold. Therefore, column (3) of table 5 contains results for equation (9) which is based on the augmented learning function given by equation (7).<sup>22</sup> The significant differences in the parameter estimates between column (2) and (3) confirm that it is important

<sup>21</sup> As the largest local labor markets within the different countries and also the respective reference, i.e., the average local labor market within a country differ in size, the estimated earnings premia are only to a limited extent comparable across countries.

<sup>22</sup> Here I still assume that the value of experience does not vary depending on the labor market where the experience is used, i.e.,  $\delta_r = \delta$  and  $\gamma_r = \gamma$ .

Table 5: 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)
$\theta^\dagger$	0.000	0.000	2.358	3.138
New employment relationships	525265	336286	336286	214319
OLS: $R^2$ , FE: within $R^2$	0.616	0.183	0.195	0.260
Worker fixed effects	No	Yes	Yes	Yes

$\dagger$  Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. 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 (9) with  $\gamma_t = \gamma$  and  $\delta_t = \delta$ . It is based on the augmented learning function (equation (7)). The results in column (3) are obtained using the full sample. The specification reported column (4) is restricted to workers who acquired work experience only 1995 or later. All specifications including control variables as well as industry, occupation, and region-year fixed effects (see table 3).

to address that marginal returns to experience are decreasing, otherwise one obtains biased estimates for the parameters of interest. In the augmented learning function, the decrease in the marginal return to experience is captured by parameter  $\theta$ . For the wage at date  $t$  it follows that it is highly non-linear in this parameter (see equation (9)). To overcome this problem, I chose different values for  $\theta$  and estimated equation (9) 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  $\theta$  equal to  $2.358 \times 10^{-4}$ . 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.<sup>23</sup> Suppose a worker was employed at each day within the last 365 days, then a  $\theta$  of  $2.358 \times 10^{-4}$  means that the human capital which was acquired at date  $t - 365$  is weighted by 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.

For  $\delta$  I now obtain an elasticity of  $0.184 \times 10^{-4}$  and  $\hat{z}$  is almost zero. Both results are confirmed by the estimates reported in column (4). The latter are based on a reduced sample where I restrict the analysis to individuals who acquired experience only in 1995 or later. This robustness check takes into account that the chosen definition of regional 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, meaning the size of the labor markets where experience was acquired. However, it does not seem to be a severe problem here as the results in column (3) and (4) are very similar. It should be noted that they are based on different values for  $\theta$ . Using the reduced sample, I obtain the smallest residual sum of squares if I set  $\theta$  equal to  $3.138 \times 10^{-4}$  (see table B.3), meaning that human capital acquired 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 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 5, it illustrates for different levels of experience the elasticity of wage with respect to the size of the labor market where previous experience was acquired. Consider for

<sup>23</sup> The results that I obtain for other values of  $\theta$  are reported in table B.2. I also estimated a specification with a  $\theta$  close to zero (specification 22 in table B.2). That result corresponds to the result reported in column (2) of table 5.

example a worker with two years of work experience at date  $t$ . The corresponding elasticity is given by about 0.01. Hence, doubling the size of all labor markets where the two years of experience were acquired results in a 1 percent higher wage at date  $t$ . In contrast, at ten years of work experience the elasticity amounts to about 0.045, meaning that doubling the size of all labor markets where ten years of experience were acquired results in a productivity increase at date  $t$  by about 4.5 percent. At a very high level of experience the elasticity is even larger.<sup>24</sup> That the elasticity significantly increases in the amount of experience points out that dynamic agglomeration benefits 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)).

For levels of experience up to ten years, the estimated elasticities based on the reduced sample and the full sample, respectively, 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, it should be noted that 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 and their overall interpretation is clear: Dynamic agglomeration benefits have a highly significant impact of individual productivity. Since workers are rather immobile between labor markets of different size (see table 2), 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 following results 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 Importance of high-skilled labor and 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, it presumably is especially skilled workers one can learn from. Hence, an important question is to which extent dynamic agglomeration benefits may 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)) which was so far only measured in terms to total regional employment, in the following to be an aggregate of total regional employment, denoted by  $\text{emp}_{s,\tau}$ , and the local share of high-skilled workers,  $\frac{\text{skill}_{s,\tau}}{\text{emp}_{s,\tau}}$ :

$$\delta \ln(Z_{s,\tau}) = \rho_{\text{emp}} \ln(\text{emp}_{s,\tau}) + \rho_{\text{skill}} \ln\left(\frac{\text{skill}_{s,\tau}}{\text{emp}_{s,\tau}}\right). \quad (10)$$

In contrast to  $\delta$  which denotes the total dynamic agglomeration benefit,  $\rho_{\text{emp}}$  refers only to the direct benefit of working in large labor markets which stems from the larger number of workers, holding the share of high-skilled workers constant. The indirect benefit which stems from the larger share of high-skilled labor in urban labor markets is captured by  $\rho_{\text{skill}}$ . Both parameters are elasticities which denote the impact of the respective characteristic on the marginal value of experience and future wages.

<sup>24</sup>In general, the elasticity of the value of  $E$  days of experience with respect to the size of the labor markets where the experience was acquired is, according to the results reported in column (3) of table 5, given by  $0.186 \times 10^{-4} \times \sum_{k=1}^E (1 - 2.358 \times 10^{-4})^{k-1}$  (see also equation (8)).



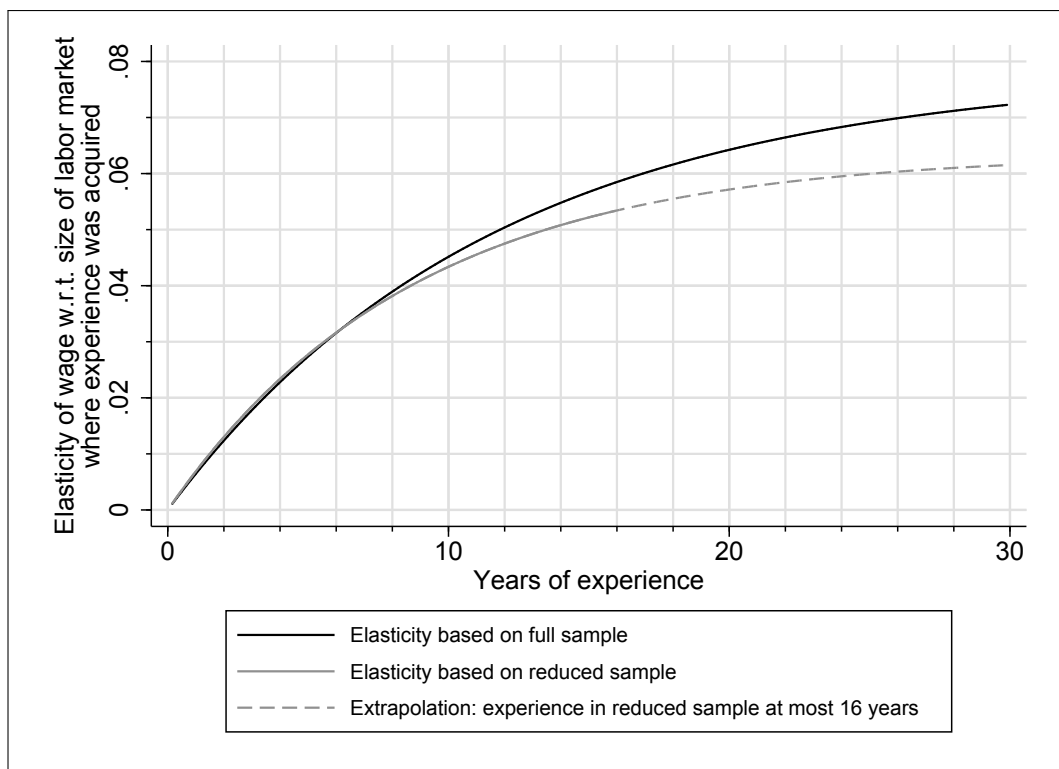


Figure 2: Magnitude of dynamic agglomeration benefits

Note: The graph illustrates the regression results reported in column (3) and (4) of table 5. The elasticity denotes the percentage increase in productivity at date  $t$ , given that all labor markets where previous work experience was acquired would have been one percent larger. The reduced sample contains only workers who acquired experience in 1995 or later.

As discussed in the literature review, some authors stress that it is important 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 3). The results reported in table 6 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 only be based on external effects.

Column (1) of table 6 summarizes results for the augmented learning function with  $Z_{s,\tau}$  being an aggregate measure as described by equation (10) that are obtained using the full sample of new employment relationships. The results reported in column (2) are based on the reduced sample which does not contain new employment relationships of workers who acquired experience before 1995. The reported results are obtained setting  $\theta$  equal to 2.724 and 3.138, respectively. Results for other values of  $\theta$  are summarized in table B.4 and table B.5. Both estimation results reported in table 6 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 highly significant impact is identified. For a worker with one year of work experience the estimated elasticities mean that doubling the share of high-skilled labor in all labor markets where the experience was acquired results in a 1.2 (column (2)) to 1.8 (column (1)) percent higher productivity ( $= \hat{\rho}_{\text{skill}} \sum_{k=1}^{365} (1 - \theta)^{k-1}$ ). It points out that local specialization on high-skilled labor is an important determinant of the value of work experience resulting in an increase of individual productivity. This also supports the view that the identified dynamic agglomeration benefits crucially depend on local learning externalities as it presumably is especially high-skilled labor one learns from.

Table 6: Estimates of the parameters of the learning function distinguishing the impact of total regional employment and the share of high-skilled labor

	(1)	(2)
$\hat{\rho}_{\text{emp}}^{\dagger}$	0.024 (0.028)	0.100*** (0.038)
$\hat{\rho}_{\text{skill}}^{\dagger}$	0.529*** (0.064)	0.336*** (0.096)
$\hat{\gamma}^{\dagger}$	3.443*** (0.476)	1.943*** (0.646)
$\theta^{\dagger}$	2.724	3.138
New employment relationships	336286	214319
within $R^2$	0.195	0.260

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. 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 (10). 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 3).

With respect 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 where 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 half of the dynamic agglomer-

ation benefit. The elasticity with respect to the number of workers in the local labor markets declines from  $0.2 \times 10^{-4}$  to  $0.1 \times 10^{-4}$  (compare column (4) of table 5 and column (2) of table 6), but it is still statistically significant and economically meaningful.

One explanation for the different results might refer to the different amount of experience the workers in both samples have. The workers in the reduced sample have on average 5 years of experience (at most 16 years) and the workers in the full sample on average more than 9 years (at most 36 years). Hence, the results may point out 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 where experience was acquired is weighted lower the more experience was acquired thereafter. Consider for example a worker who was employed each day within the previous 15 years. Then, a  $\theta$  of about 2.724 (see column (1) in table 6) means that the size of the labor market where experience was acquired fifteen years ago is weighted by only 0.22 ( $= (1 - 0.0002724)^{5479}$ ). In contrast, the size of the labor market where experience was acquired one year ago has a weight of 0.91. It points out that the size of the labor market where a worker acquired experience years ago has relative to the size of the labor market where recent experience was acquired a little impact on individual productivity at date  $t$ .<sup>25</sup> However, further research is needed to analyse 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.

Please note,  $\hat{\gamma}$  reported in table 5 and table 6 significantly differs. In the specifications reported in table 6  $\hat{\gamma}$  captures an additional effect. It now 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 emp and skill, respectively. It is now given by:  $\hat{\gamma} = -\hat{\rho}_{\text{emp}} \ln(\text{emp}) - \hat{\rho}_{\text{skill}} \ln(\text{skill})$ . Based on the regression result reported in table 6 it is not possible to compute emp and 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 elasticities reported in column (1), the maximum value that I obtain for the German labor market regions is 2.436. Based on the estimates reported in column (2) the maximum amounts to 0.759. Both values are smaller than the respective estimate for  $\gamma$ . This confirms my previous findings that learning takes place in all German labor markets.

A further test of the learning hypothesis is to analyse whether the value of experience is predominately determined by the size of the labor market where it was acquired or by the labor market where it is used. Results by De la Roca and Puga (2016) and the results reported in table 4 suggest that the value of experience acquired by working in large labor markets is highly portable to other regions. Hence, the findings support the hypothesis, that individuals learn more by working in large labor markets than working elsewhere. The parameter estimates reported in both columns of table 7 also support this interpretation of dynamic agglomeration benefits. They refer to equation 9. Now, I let the parameters  $\gamma$  and  $\delta$  differ according to the size of the labor market where the experience is used.<sup>26</sup> I distinguish three groups of labor markets. As large local labor markets I consider those

<sup>25</sup> Recall that the weight of the size of the labor market where experience was acquired decreases in the amount of experience that is acquired thereafter. This addresses that the marginal return to experience at any date is lower the more experience and human capital previously was acquired (see equations (7) and (8)).

<sup>26</sup> The specifications estimated here do not exactly correspond to equation 9. I do not impose restrictions on the parameters, to ensure that

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 local labor market regions. All other labor markets represent the reference group.<sup>27</sup>

Table 7: Estimates of the parameters of the learning function depending on the size of the labor market where 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)
$\theta^\dagger$	2.358	3.138
New employment relationships within $R^2$	336286 0.195	214319 0.260

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. Both columns summarize estimates of the parameters of the augmented learning function (equation (7)). The parameters are allowed to vary depending on the size of the labor market where 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.  $\Delta$  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 3).

Column (1) shows the results that are obtained based on the full sample, column (2) the results which are obtained using the reduced sample.<sup>28</sup>  $\hat{\delta}$ , the estimate of the elasticity of the marginal value of experience with respect to the size of the labor market where experience is acquired, differs only little between small, medium and large labor markets. For small and medium labor markets the difference is not statistically different from zero. For large labor markets, I observe an elasticity that is about one fifth smaller than the elasticity obtained for small and medium labor markets, meaning that wages in large labor markets are less elastic with respect to the size of the labor market where experience was acquired.<sup>29</sup> This finding is in accordance with my results reported in table 4: Consider two workers, each with one year of experience. One worker acquired the experience in the three largest German labor markets, the other somewhere else in Germany. If both workers

$\hat{\delta}_r$  and  $\hat{\gamma}_r$  obtained for small, medium and large labor markets result in the same estimate for  $\hat{Z}$ , the threshold beyond which working leads to the acquisition of human capital. Hence, in the estimated specifications  $\hat{Z}$  is allowed to vary depending on the size of the labor market where experience is used. However, the regression results show that the estimates of the threshold are all close to zero, statistically they are not significantly different.

<sup>27</sup> I also estimated specification where I require large labor markets to have a local workforce of, respectively, at least 500,000 employees and at least 1,000,000 employees. The obtained results are very similar to those reported in table 7.

<sup>28</sup> Results that I obtain setting  $\theta$  to other values than reported in table 7 can be found in tables B.6 and B.6, respectively.

<sup>29</sup> Recall that differences in the *wage level* between large, medium and small labor markets are captured by the region fixed effects, not by  $\delta$ .

use their experience in the largest German labor markets, the difference in their wages amounts to 0.4 percent ( $= (0.013 - 0.005) - (0.007 - 0.003)$ ). If both workers use their experience somewhere else in Germany, the wage difference is larger. It then amounts to 0.8 percent ( $= 0.013 - 0.005$ ). De la Roca and Puga (2016) obtain qualitatively the same finding for Spain.

This phenomenon might also be related to learning benefits in large labor markets. Suppose workers of type L who acquired experience in a large local labor market and workers of type S who acquired the same amount of experience in a small local labor market are imperfect substitutes since the accumulated knowledge differs. Then, workers of type L may receive a higher relative wage compared to type S workers if they use their large amount of human capital in a small labor market because type L workers are relatively seldom there (compare table 2). This results in a larger wage dispersion in small labor markets with respect to the size of the labor market where experience was acquired than if both types of worker use their accumulated human capital in a large labor market.

However, this effect should not be stressed too much. Quantitatively, the differences in the elasticities between large, medium and small labor markets are small. Overall the results confirm that the value of experience is predominantly determined by the size of the labor market where it was acquired, not by the size of the labor market where it is used supporting the hypotheses that the identified dynamic agglomeration benefits are strongly related to learning benefits.

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 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 benefits in urban labor markets and take them into account when they decided where to work. This would imply a positive selection. 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 would be 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 benefits. I analyze to which extent wages associated with new employment relationships depend on the size of the local labor markets where previous work experience was acquired. Based on an administrative data set capturing detailed information on individual employment biographies, I obtain similar findings for Germany as with De la Roca and Puga (2016) for Spain and Matano and Naticchioni (2016) for Italy. The value of work experience acquired in the largest local labor markets has a significantly higher value than experience acquired in the rest of the country, reflecting in higher future wages. However, in order to obtain more general results, this study goes a step further as previous papers and estimates the elasticity of wage with respect to the size of the labor market where experience was acquired.

The results clearly show that the benefit of having worked in large labor markets increases in the level of experience. For example, after two years of working the elasticity of wage with respect to the size of the labor market where 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 implies that doubling employment in all local labor markets where 30 years of experience were acquired results in an wage increase of about 6 to 7 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 large urbanized regions and non-urban labor markets.

Furthermore, 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 where the new employer is located, as well as for characteristics of the firm and the local industry. In addition, I control for various other channels of agglomeration economies, i.a., dynamic matching and provide evidence that the value of experience is predominately determined by the size of the labor market where experience was acquired. This finding also supports the hypotheses that working in large labor markets fosters the individual accumulation of knowledge which may be used in the future and also in other regions. The results even suggest that workers who acquired experience in large labor markets receive a slightly larger wage premium for their experience, relative to workers who acquired experience somewhere else, if they move to small or medium size labor markets. A similar result is reported by De la Roca and Puga (2016) using Spanish data. It may point to an imperfect substitutability of workers with experience acquired in large and small labor markets because of heterogeneity in the accumulated knowledge.

Further support for the learning hypothesis comes from the additional result that the share of high-skilled labor within the labor market where experience was acquired has a highly significant impact on the value of the acquired experience and future productivity. In order to provide new insights as to which characteristic of urban labor markets lead to the increasing gains of acquiring experience in large labor markets, I distinguish the effects of labor market size and the local share of high-skilled labor. In accordance with the literature on human capital externalities, the results suggest that it is especially high-skilled labor workers learn from. At least a significant part of the dynamic agglomeration benefit is related to the higher share of high-skilled labor in large urban labor markets. Once controlling for the share of high-skilled labor, the impact of labor market size in terms of total regional employment on the value of experience and future 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 one half, in other specifications its impact becomes insignificant. One explanation for the different results may be that the relative importance of labor market size and the local share of high-skilled labor varies depending on the career phase of a worker. However, 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 always increasing in labor market size, 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 Definition of new employment relationships and variables

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 for 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 East Germany before reunification, i.e. all workers for whom I do not observe an employment spell before 1990 and who were born before 1977. Additionally, I do not consider individuals who worked before 1993 in a regional 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 employment spells 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 labour market policy. In the latter case I cannot ensure that this employment is not publicly subsidised. 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 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 an other establishment, I require that the previous employment spell ends within 7 days. Furthermore, I exclude a new employment relationship, if it is the first employment spell of the person in her 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 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 interval regression, a generalisation 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 where 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 A.1 provides information on the definition of all variables used in this analysis.



Table A.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 where experience was acquired	Size of regional labor market regions according to Kosfeld and Werner (2012) where 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
Current labor market size	Logarithm of the number of workers subject to social security contributions that work in the particular local labor market where the new employer is located.*	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 categorical variable that combines information on the highest school leaving certificate, completed vocational training and university degree. For some employment spells, this information is missing. If so, I use the information from previous employment spells following Fitzenberger et al. (2005).	IEB
Gender		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"> <li>- unemployment benefits (Arbeitslosengeld I)</li> <li>- unemployment assistance (Arbeitslosengeld II / Arbeitslosenhilfe).</li> <li>- unemployed and registered as a job seeker</li> <li>- not unemployed, but registered as a job seeker</li> <li>- participating in active labor market policy programs.</li> </ul>	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)

Table A.1 continued

Variable	Definition	Source
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
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 employment spells 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.

## B Tables

Table B.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 where experience was acquired <sup>§</sup>								
ln(Number of employees)	12.510	0.848	9.512	14.187	12.493	0.841	9.512	14.187
ln(Share of high-skilled employment)	-2.413	0.419	-4.557	-1.125	-2.419	0.412	-4.331	-1.125
Tenure in month	5.990	3.585	0.033	12.000	5.784	3.540	0.033	12.000
ln(Number of previous employers) <sup>†</sup>	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) <sup>-1</sup> )	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) <sup>‡</sup>	1.855	0.420	0.294	2.838	1.855	0.419	0.294	2.838
ln(Local unemployment rate among skilled labor) <sup>‡</sup>	2.254	0.429	0.981	3.484	2.259	0.433	0.981	3.484
ln(Local unemployment rate among low-skilled labor) <sup>‡</sup>	3.403	0.375	2.245	4.293	3.402	0.375	2.245	4.293
Observations	525265				336286			

<sup>§</sup> 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 employment spells. <sup>†</sup> For less than 1 percent of the observations the number of previous employers exceeds 18, for less than 10 percent 9 previous employers. <sup>‡</sup> In the empirical analysis this variable is set to zero if the workers belongs to an other 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 B.2: Estimates of the parameters of the learning function based on different values for  $\theta$  and full sample, labor market size measured in terms of total regional employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\theta^\dagger$	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	336286	336286	336286	336286	336286	336286	336286	336286	336286	336286	336286
within R <sup>2</sup>	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)
$\theta^\dagger$	2.594	<b>2.358</b>	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)	<b>0.184***</b> <b>(0.020)</b>	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)	<b>0.0847</b> <b>(0.254)</b>	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	336286	<b>336286</b>	336286	336286	336286	336286	336286	336286	336286	336286	336286
within R <sup>2</sup>	0.195	<b>0.195</b>	0.195	0.195	0.194	0.194	0.193	0.192	0.191	0.185	0.183
RSS	8327.088	<b>8326.933</b>	8327.348	8328.358	8332.256	8338.680	8347.391	8358.395	8370.971	8428.451	8452.746

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. The table summarizes estimation results for equation (9) depending on the value  $\theta$  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*. *RSS* denotes the residual sum of squares. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3).

Table B.3: Estimates of the parameters of the learning function based on different values for  $\theta$  and reduced sample, labor market size measured in terms of total regional employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\theta^\dagger$	50.000	10.000	5.000	4.595	4.177	3.797	3.452	<b>3.138</b>	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)	<b>0.199***</b> <b>(0.025)</b>	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)	<b>-0.026</b> <b>(0.325)</b>	-0.045 (0.323)	-0.066 (0.320)	-0.088 (0.318)
N	214319	214319	214319	214319	214319	214319	214319	<b>214319</b>	214319	214319	214319
within R <sup>2</sup>	0.249	0.256	0.259	0.260	0.260	0.260	0.260	<b>0.260</b>	0.260	0.260	0.260
RSS	4792.050	4750.585	4728.430	4726.499	4724.810	4723.565	4722.783	<b>4722.464</b>	4722.471	4722.586	4722.793
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
$\theta^\dagger$	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	214319	214319	214319	214319	214319	214319	214319	214319	214319	214319	214319
within R <sup>2</sup>	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

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. The table summarizes estimation results for equation (9) depending on the value  $\theta$  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*. RSS denotes the residual sum of squares. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3).

Table B.4: Estimates of the parameters of the learning function based on different values for  $\theta$  and full sample, labor market size measured 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)
$\theta^\dagger$	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	<b>2.724</b>
$\hat{\rho}_{\text{emp}}^\dagger$	0.019 (0.093)	-0.049 (0.037)	-0.048 (0.031)	-0.044 (0.030)	-0.037 (0.030)	-0.027 (0.029)	-0.015 (0.029)	-0.001 (0.028)	0.006 (0.028)	0.015 (0.028)	<b>0.024</b> <b>(0.028)</b>
$\hat{\rho}_{\text{skill}}^\dagger$	1.651*** (0.229)	0.921*** (0.089)	0.798*** (0.073)	0.776*** (0.071)	0.746*** (0.070)	0.709*** (0.068)	0.666*** (0.067)	0.616*** (0.066)	0.589*** (0.065)	0.559*** (0.065)	<b>0.529***</b> <b>(0.064)</b>
$\hat{\gamma}^\dagger$	5.854*** (1.626)	6.428*** (0.640)	5.390*** (0.532)	5.215*** (0.522)	4.977*** (0.513)	4.704*** (0.504)	4.391*** (0.495)	4.039*** (0.487)	3.852*** (0.483)	3.647*** (0.480)	<b>3.443***</b> <b>(0.476)</b>
N	336286	336286	336286	336286	336286	336286	336286	336286	336286	336286	<b>336286</b>
within R <sup>2</sup>	0.179	0.189	0.194	0.194	0.195	0.195	0.195	0.195	0.195	0.195	<b>0.195</b>
RSS	8491.094	8385.554	8336.437	8332.986	8329.780	8327.287	8325.478	8324.326	8323.992	8323.799	<b>8323.760</b>
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
$\theta^\dagger$	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\rho}_{\text{emp}}^\dagger$	0.034 (0.028)	0.054** (0.027)	0.076*** (0.027)	0.099*** (0.027)	0.147*** (0.026)	0.194*** (0.026)	0.237*** (0.026)	0.273*** (0.026)	0.298*** (0.026)	0.196*** (0.026)	0.039 (0.025)
$\hat{\rho}_{\text{skill}}^\dagger$	0.495*** (0.064)	0.425*** (0.063)	0.351*** (0.062)	0.271*** (0.062)	0.105* (0.061)	-0.062 (0.061)	-0.217*** (0.061)	-0.353*** (0.062)	-0.450*** (0.063)	-0.209*** (0.066)	0.249*** (0.064)
$\hat{\gamma}^\dagger$	3.215*** (0.473)	2.754*** (0.466)	2.264*** (0.461)	1.751*** (0.456)	0.689 (0.448)	-0.366 (0.442)	-1.338*** (0.439)	-2.189*** (0.438)	-2.818*** (0.439)	-1.647*** (0.446)	0.972** (0.433)
N	336286	336286	336286	336286	336286	336286	336286	336286	336286	336286	336286
within R <sup>2</sup>	0.195	0.195	0.195	0.195	0.194	0.194	0.193	0.192	0.191	0.185	0.183
RSS	8323.840	8324.472	8325.641	8327.318	8332.097	8338.625	8346.720	8356.677	8368.269	8427.928	8451.958

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. The table summarizes estimation results for equation (9) depending on the value  $\theta$  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 (10).  $\hat{\rho}_{\text{emp}}$  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. *RSS* denotes the residual sum of squares. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3).

Table B.5: Estimates of the parameters of the learning function based on different values for  $\theta$  and reduced sample, labor market size measured 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)
$\theta^\dagger$	50.000	10.000	5.000	4.595	4.177	3.797	3.452	<b>3.138</b>	2.996	2.853	2.724
$\hat{\rho}_{\text{emp}}^\dagger$	-0.044 (0.108)	-0.064 (0.045)	0.017 (0.040)	0.032 (0.039)	0.049 (0.039)	0.067* (0.038)	0.084** (0.038)	<b>0.100***</b> <b>(0.038)</b>	0.108*** (0.037)	0.115*** (0.037)	0.122*** (0.037)
$\hat{\rho}_{\text{skill}}^\dagger$	1.799*** (0.273)	1.010*** (0.115)	0.652*** (0.101)	0.596*** (0.100)	0.529*** (0.099)	0.463*** (0.098)	0.398*** (0.097)	<b>0.336***</b> <b>(0.096)</b>	0.308*** (0.096)	0.280*** (0.096)	0.254*** (0.095)
$\hat{\gamma}^\dagger$	7.450*** (1.888)	5.897*** (0.781)	3.896*** (0.684)	3.562*** (0.676)	3.158*** (0.668)	2.748*** (0.660)	2.339*** (0.653)	<b>1.943***</b> <b>(0.646)</b>	1.757*** (0.643)	1.570** (0.639)	1.398** (0.636)
N	214319	214319	214319	214319	214319	214319	214319	<b>214319</b>	214319	214319	214319
within R <sup>2</sup>	0.250	0.256	0.260	0.260	0.260	0.260	0.260	<b>0.260</b>	0.260	0.260	0.260
RSS	4790.156	4746.970	4726.454	4724.818	4723.455	4722.509	4721.989	<b>4721.887</b>	4721.983	4722.180	4722.455
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
$\theta^\dagger$	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\rho}_{\text{emp}}^\dagger$	0.129*** (0.0369)	0.140*** (0.0365)	0.149*** (0.0362)	0.156*** (0.0358)	0.164*** (0.0351)	0.166*** (0.0345)	0.164*** (0.0339)	0.159*** (0.0333)	0.152*** (0.0327)	0.119*** (0.0307)	0.103*** (0.0298)
$\hat{\rho}_{\text{skill}}^\dagger$	0.229** (0.095)	0.186** (0.094)	0.150 (0.093)	0.121 (0.093)	0.0836 (0.091)	0.0680 (0.090)	0.0672 (0.088)	0.0759 (0.087)	0.0901 (0.085)	0.167** (0.080)	0.204*** (0.078)
$\hat{\gamma}^\dagger$	1.223* (0.633)	0.918 (0.626)	0.653 (0.620)	0.429 (0.614)	0.0996 (0.603)	-0.0948 (0.592)	-0.189 (0.581)	-0.217 (0.571)	-0.200 (0.562)	0.0673 (0.527)	0.231 (0.512)
N	214319	214319	214319	214319	214319	214319	214319	214319	214319	214319	214319
within R <sup>2</sup>	0.260	0.260	0.260	0.260	0.259	0.259	0.258	0.258	0.257	0.256	0.255
RSS	4722.876	4723.831	4725.041	4726.449	4729.632	4732.989	4736.239	4739.365	4742.231	4751.798	4755.475

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. The table summarizes estimation results for equation (9) depending on the value  $\theta$  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 (10).  $\hat{\rho}_{\text{emp}}$  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. *RSS* denotes the residual sum of squares. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3).

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\theta^\dagger$	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	336286	336286	336286	336286	336286	336286	336286	336286	336286	336286	336286
within R <sup>2</sup>	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
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
$\theta^\dagger$	2.594	<b>2.358</b>	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
Now large labor market											
$\Delta\hat{\delta}^\dagger$	-0.040*** (0.013)	<b>-0.039***</b> ( <b>0.012</b> )	-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)	<b>0.520***</b> ( <b>0.158</b> )	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)	<b>0.197***</b> ( <b>0.021</b> )	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)	<b>-0.075</b> ( <b>0.266</b> )	-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)	<b>-0.010</b> ( <b>0.013</b> )	-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)	<b>0.118</b> ( <b>0.155</b> )	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	336286	<b>336286</b>	336286	336286	336286	336286	336286	336286	336286	336286	336286
within R <sup>2</sup>	0.195	<b>0.195</b>	0.195	0.195	0.194	0.194	0.193	0.192	0.191	0.185	0.183
RSS	8326.474	<b>8326.318</b>	8326.729	8327.733	8331.610	8338.007	8346.684	8357.649	8370.182	8427.486	8451.725

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. The table summarizes estimation results for equation (9) depending on the value  $\theta$  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*. *RSS* denotes the residual sum of squares. 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.  $\Delta$  refers to the difference in a parameter in comparison to the reference group, i.e. medium size labor markets. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3).



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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
$\theta^\dagger$	50.000	10.000	5.000	4.595	4.177	3.797	3.452	<b>3.138</b>	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)	<b>-0.034*</b> <b>(0.020)</b>	-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)	<b>0.527**</b> <b>(0.256)</b>	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)	<b>0.197***</b> <b>(0.027)</b>	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)	<b>-0.0020</b> <b>(0.347)</b>	-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)	<b>0.012</b> <b>(0.021)</b>	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)	<b>-0.176</b> <b>(0.254)</b>	-0.173 (0.250)	-0.169 (0.246)	-0.165 (0.243)
N	214319	214319	214319	214319	214319	214319	214319	<b>214319</b>	214319	214319	214319
within R <sup>2</sup>	0.250	0.256	0.259	0.260	0.260	0.260	0.260	<b>0.260</b>	0.260	0.260	0.260
RSS	4789.369	4749.190	4727.687	4725.808	4724.171	4722.972	4722.232	<b>4721.950</b>	4721.973	4722.105	4722.327
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
$\theta^\dagger$	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	214319	214319	214319	214319	214319	214319	214319	214319	214319	214319	214319
within R <sup>2</sup>	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

<sup>†</sup> Coefficients and standard errors multiplied by 10,000. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at individual level in parentheses. The table summarizes estimation results for equation (9) depending on the value  $\theta$  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*. RSS denotes the residual sum of squares. 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.  $\Delta$  refers to the difference in a parameter in comparison to the reference group, i.e. medium size labor markets. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3).

## 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.
- 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 worker’ 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.
- 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 Arbeitsmarktregionen - 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 staatswirtschaftlicher 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.