

The Impact of Age Complementarities and
Learning Externalities on Labor Productivity:
Evidence from German Micro Data

Inaugural-Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Wirtschafts- und Sozialwissenschaften
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Diplom-Volkswirt
Jan Cornelius Peters
aus Geesthacht

Braunschweig, 2017

Erstbegutachtung:
Zweitbegutachtung:

Frau Professor Dr. Annetrin Niebuhr
Herr Professor Dr. Johannes Bröcker

Tag der mündlichen Prüfung:

27. Oktober 2016

Für Pia, Jon Hinrich und Svea Katrine

Vorwort

Die vorliegende Arbeit ist während meiner Tätigkeit am Institut für Arbeitsmarkt- und Berufsforschung (IAB) der Bundesagentur für Arbeit (BA) in der Einheit IAB Nord in Kiel entstanden. Im Kern besteht sie aus zwei in sich abgeschlossenen und in Alleinautorenschaft verfassten Abhandlungen, welche Kapitel 2 und Kapitel 3 dieser kumulativen Dissertation darstellen.

Kapitel 2 meiner Dissertation ist im Rahmen der von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekte „Diversität und Erfolg von Organisationen“ (NI 596/4-1) und „Diversität und individuelle Karrieren“ (NI 596/5-1) entstanden. Beide Projekte wurden von Frau Professor Dr. Annetrin Niebuhr (IAB Nord, CAU Kiel) und Frau Professor Dr. Monika Jungbauer-Gans (DZHW) geleitet und gemeinsam mit Frau Antje Buche und Herrn Andreas Eberl bearbeitet. Kapitel 3 der vorliegenden Arbeit fußt auf einer Zusammenarbeit mit Frau Anette Haas, Frau Silke Hamann, Frau Professor Dr. Annetrin Niebuhr und Herrn Malte Reichelt im Rahmen des IAB-Projektes „Arbeitskräftemobilität und regionale Disparitäten“. Ich habe durch die Zusammenarbeit viel gelernt und danke allen Beteiligten für ihre Anmerkungen zu meiner Arbeit.

Mein besonderer Dank gilt dabei meiner Doktormutter Frau Professor Dr. Annetrin Niebuhr. Ihr habe ich die Möglichkeit zu verdanken, in den letzten fünf Jahren an verschiedenen von ihr geleiteten Forschungs- und Beratungsprojekten mitzuarbeiten und Erfahrungen im wissenschaftlichen Arbeiten zu sammeln. Gleichzeitig hat sie mir große Freiheiten für die Erstellung meiner Dissertation eingeräumt und durch Anregungen sowie stets konstruktive Kritik entscheidend zum Gelingen dieser Arbeit beigetragen. Aus selbigem Grund möchte ich insbesondere auch meinem Zweitbetreuer, Herrn Professor Dr. Johannes Bröcker, danken. Die erhaltenen Hinweise und Vorschläge lieferten entscheidende Hilfestellungen, die Arbeit in die richtigen Bahnen zu lenken.

Ein großer Dank gilt außerdem Herrn Stefan Fuchs, der mir als Leiter des Regionalen Forschungsnetzes des IAB ebenfalls große Freiräume für die Fertigstellung dieser Dissertation gewährt hat. Für die freundliche und produktive Arbeitsatmosphäre im IAB Nord sowie für ihre Unterstützung jeglicher Art danke ich Frau Dr. Tanja Buch, Herrn Dr. Volker Kotte sowie Frau Andrea Stöckmann.

Darüber hinaus gilt mein Dank Herrn Tim Logan sowie Kolleginnen und Kollegen aus dem IAB Nürnberg und den anderen Einheiten des Forschungsnetzes, insbesondere dem Team um Herrn Ali Athmani für die Bereitstellung der verwendeten Forschungsdaten sowie Herrn Professor Dr. Uwe Blien und Frau Dr. Anja Rossen. Ferner danke ich meinen (ehemaligen) Kommilitonen aus dem PhD-Programm „Quantitative Economics“, allen voran Frau Veronika Harder, Frau Katerina Homolkova sowie Herrn Henning Meier, für die unterhaltsamen wie erfolgreichen Vorbereitungen auf die Prüfungen im Rahmen des Promotionsstudiums und ihre Anmerkungen zu dieser Arbeit.

Schließlich möchte ich meiner Ehefrau Pia von ganzem Herzen Dank aussprechen. Ihr starker Rückhalt und das aufgebrachte Verständnis in den letzten Wochen und Monaten haben einen wesentlichen Beitrag zum Abschluss dieser Arbeit beigetragen. Ein ganz besonderer Dank gilt außerdem meinen Eltern Christine und Hauke, deren Unterstützung ich mir stets gewiss sein kann.

Kiel, August 2016
Jan Cornelius Peters

Contents

Vorwort	iv
List of Tables	viii
List of Figures	x
List of Abbreviations	xi
1 Introduction	1
2 Do age complementarities affect labor productivity?	5
2.1 Introduction	6
2.2 Related literature	7
2.3 Theoretical framework	9
2.4 Empirical specification and econometric issues	11
2.5 Data	13
2.6 Results	15
2.7 Conclusion	20
Appendix	22
2.A Deriving the system of share equations	22
2.B Deriving the Hicks partial elasticity of complementarity and the factor price elasticity	23
2.C Deriving the bordered Hessian matrix	26
2.D Imputation of censored wages	27
2.E Further results	29

3	Quantifying the effect of labor market size on learning externalities	31
3.1	Introduction	32
3.2	Related literature	34
3.3	Methodology	39
3.3.1	Empirical model	39
3.3.2	Econometric issues	44
3.4	Data	45
3.4.1	Individual data set	45
3.4.2	The size of local labor markets	46
3.4.3	The mobility of labor	47
3.5	Results	49
3.5.1	Control variables	49
3.5.2	Main result - The benefit of acquiring experience in large labor markets	52
3.5.3	The importance of high-skilled labor and the portability of accumulated human capital	57
3.6	Conclusions	65
	Appendix	67
3.A	The solution of equation (3.9)	67
3.B	Further information on data	68
3.C	Further results	71
4	Summary of main results	78
	Bibliography	80
	Curriculum Vitae	85
	Erklärung zum selbstständigen Verfassen der Arbeit	89

List of Tables

Table 2.1	Summary statistics	15
Table 2.2	Estimated parameters and calculated elasticities, average effects assuming $\alpha_{iq} = \alpha_i$, $\beta_{ijq} = \beta_{ij}$	16
Table 2.3	Results of a model calculation	18
Table 2.4	Elasticities by skill level	19
Table 2.D.1	Results of an interval regression to partly impute the censored wages of employees liable to social security	28
Table 2.E.1	Results based on a reduced sample: only establishments with less than 50 employees, average effect assuming $\alpha_{iq} = \alpha_i$, $\beta_{ijq} = \beta_{ij}$	29
Table 2.E.2	Elasticities by skill level based on reduced sample: only establishments with less than 50 employees	30
Table 3.1	Size of German labor market regions	47
Table 3.2	The size of the largest labor market in which a worker acquired experience relative to size of the smallest labor market	48
Table 3.3	The size of labor market in which experience was acquired and in which it is used	49
Table 3.4	Results for control variables	50
Table 3.5	The value of labor market specific work experience	53
Table 3.6	Estimates of the parameters of the learning function	55
Table 3.7	Estimates of the parameters of the learning function distinguishing the impact of total regional employment and the share of high-skilled labor	59
Table 3.8	Estimates of the parameters of the learning function depending on the size of the labor market in which experience is used	64
Table 3.B.1	Variables - definitions and sources	69

Table 3.C.1	Summary statistics	71
Table 3.C.2	Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment	72
Table 3.C.3	Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment	73
Table 3.C.4	Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment and the local share of high-skilled labor	74
Table 3.C.5	Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment and the local share of high-skilled labor	75
Table 3.C.6	Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment, parameters allowed to vary depending on size of labor market where experience is used	76
Table 3.C.7	Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment, parameters allowed to vary depending on size of labor market where experience is used	77

List of Figures

Figure 2.1	Nested production framework	9
Figure 2.D.1	Wage distribution before (left figure) and after (right figure) imputation	27
Figure 3.1	Local labor market size and regional wages.	32
Figure 3.2	Magnitude of dynamic agglomeration benefits	57
Figure 3.3	Labor productivity after migration	62

List of Abbreviations

ALG I	Unemployment benefit (<i>Arbeitslosengeld I</i>)
ALG II	Unemployment assistance (<i>Arbeitslosengeld II</i>)
ALHI	Unemployment assistance (<i>Arbeitslosenhilfe</i>)
BeH	IAB employment history
BHP	Establishment History Panel
BMAS	Federal Ministry of Labor and Social Affairs
CES	Constant elasticity of substitution
GDP	Gross Domestic Product
FE	Fixed effects
FEA	Federal Employment Agency
IAB	Institute for Employment Research of the Federal Employment Agency
IABS	IAB Employment Samples
IEB	Integrated Employment Biographies
KldB 2010	The German Classification of Occupations 2010
ML	Maximum likelihood
N	Number of observations
NUTS-3	Nomenclature of territorial units for statistics, level 3 (<i>Kreise / kreisfreie Städte</i>)
OLS	Ordinary least squares
RSS	Residual sum of squares
translog	Transcendental logarithmic

Chapter 1

Introduction

This year the productivity of labor, defined as real gross domestic product (GDP) per hours worked, is expected to be on average 250 percent higher in OECD countries than it was in 1960, and 20 percent higher relative than the level in 2000. For Germany, the expected increase since 2000 amounts to about 9 percent (OECD, 2016). Labor productivity is a key factor of economic well-being, and it affects the distribution of income within societies. It is therefore apparent that understanding the driving forces of workers' productivity is highly relevant to policy makers in a social market economy as well as to all other actors on the labor market such as social partners and administrations.

It is an undisputed fact that individual characteristics like cognitive abilities, educational attainment and experience determine the productivity of individual workers. However, their productivity also depends on the context they are working in, for example the technology or the amount of capital that is available or labor market institutions. Further important characteristics of the environment are discussed by different strands of literature.

One such environmental characteristic is the composition of the labor force to which a worker belongs. It refers to the *imperfect substitutability* of labor. This literature emphasizes that workers are heterogeneous with respect to their skills and abilities and that skills and abilities of different workers may complement each other. It implies that not only the level of skills and abilities determine their value, but also the individual set compared to the skills and abilities of other workers. Following standard neoclassical theory, an individual worker is more productive the more unique her relevant individual skills and abilities are. Recently, this idea has gained increasing attention as it has important implications with regard to two trends which affect labor supply in several countries today, and in the future: migration and demographic change. In a debate on the impact of migration on wages, some scholars emphasize the necessity of taking into account that immigrants and native workers may complement each other (e.g., Ottaviano/Peri, 2012; Haas et al., 2013; Alesina et al., 2016). Other scholars stress the imperfect substitutability of workers with respect to the skill level (e.g., Moretti, 2004a; Ciccone/Peri, 2006). Still others, focus on the imperfect substitutability of workers that differ in age (e.g., Card/Lemieux, 2001; Morin, 2015; Garloff/Roth, 2016). My first paper, chapter 2 of this dissertation, adds to this literature.

There are good reasons to believe that differently aged workers complement each other in the production process. Referring to the concept of 'fluid' and 'crystallized intelligence', the psychologists John L. Horn and Raymond B. Cattell showed in a joint paper published in 1967 that comparative advantages exist between differently aged people. While young people are especially stronger in generating and recombining new knowledge, older people tend to have more experience and abilities to understand difficult situations as well as to use and apply existing skills. Younger people also typically have more physical strength and agility.

In Germany, like in many other European countries, demographic change will lead to a shift in the age structure of the workforce in the next decades. The imperfect substitutability of differently aged workers may have important implications for the relative productivity of age groups in the future. The aim of my first paper is to quantify the degree of complementarity of workers who differ in age. I estimate cross factor price elasticities which indicate how productivity of one age group changes in response to a change in the quantity of workers belonging to another age group. This not only enables the assessment of the impact of demographic change

on (relative) productivity of differently aged workers, but also the impact of any policy affecting relative labor supply, e.g., changes in retirement age.

In contrast to most previous studies on age complementarities which refer to the aggregate level, I analyze the imperfect substitutability of differently aged workers directly where production takes place, i.e., within firms. By estimating the parameters characterizing establishments' demand for workers of different age using micro data, I avoid 'ecological fallacy' resulting from interpreting aggregate results as if they pertain to behavior of firms. My analysis is based on a unique linked employer-employee data set capturing a 20 percent random sample of German establishments operating in 2012. I generated this data set based on the IAB employment history (BeH) which contains information on individual spells of employment in Germany. Taking into account the skill level of the workers, I analyze age complementarities within groups of low, medium, and high skilled workers by estimating translog cost functions. Thereby, I impose less restrictions on the estimated elasticities as compared with using CES-functions, which are often estimated in this context.

A second characteristic of a worker's environment that impacts on productivity is the local labor market. In Germany, there are not only marked differences in labor productivity between East and West Germany, but also on a more disaggregated level as reflected in significant regional wage disparities. The latter are a worldwide phenomenon.

The productivity effect of the local labor market was noted already by List (1838), Roscher (1878), and Marshall (1890) in the 19th century. In the recent literature, some scholars focus on benefits of working in a local labor market with a specialization in high-skilled labor (e.g., Moretti, 2004a; Ciccone/Peri, 2006; Heuermann, 2011). Local high-skilled labor is discussed as a factor that increases productivity of all workers within the labor market due to *human capital externalities* and the imperfect substitutability of differently skilled workers. Other scholars emphasize the productivity effect of the *size* of the local labor market, e.g., in terms of total regional employment (e.g., Ciccone/Hall, 1996; Glaeser/Maré, 2001; Combes et al., 2008). This strand of literature distinguishes static and dynamic agglomeration economies. Static agglomeration economies are supposed to result in an immediate productivity increase if a worker migrates from a small non-urban to a large urban labor market. The underlying mechanisms refer to, inter alia, sharing gains from variety and individual specialization, better matches between job seekers and vacancies, and improved production processes. Dynamic agglomeration economies let the productivity of workers increase over time, resulting in an urban wage growth premium. Two underlying mechanisms are discussed. One suggests that workers' opportunities to change jobs are larger in large labor markets than in small labor markets. The matching quality between workers and firms therefore further increases over time especially in urban labor markets. The second mechanism refers to *learning externalities*. It is argued that local learning opportunities based on meeting other workers and exchanging knowledge are larger in agglomerated labor markets than elsewhere because of the larger number of workers with different knowledge and ideas one can learn from. My second paper, chapter 3 of this dissertation, focuses on this channel which was already emphasized by List (1838).

If labor market size fosters the individual accumulation of human capital, this should reflect in the value of acquired work experience and future productivity, meaning that the productivity of a worker is higher the larger

the labor markets were in which he or she previously acquired experience. Recent studies confirm this hypothesis (e.g., Andersson et al., 2014; D’Costa/Overman, 2014; De la Roca/Puga, 2016; Matano/Naticchioni, 2016). They show that individuals who previously worked in the largest cities of a country receive higher wages today than other workers, all other things being equal. However, these studies do not provide *general* results on this issue. The main contribution of my second paper therefore is to estimate the elasticity of wage with respect to the size of the labor markets in which an individual worker previously acquired work experience. I identify transitions to full-time employment within the period 2005 to 2011 and analyze the wages associated with these new employment relationships. I take into account each previous spell of employment in order to identify where the respective worker so far acquired experience. The information is provided by the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB).

Taken together, the contribution of this dissertation to the ongoing discussions with regard to the determinants of labor productivity is twofold. Chapter 2 focuses on age complementarities, meaning that the individual productivity of a worker depends on his or her age relative to the age of the co-workers. Chapter 3 examines to which extent labor productivity of employees depends on the size of the labor market in which experience was acquired, referring to dynamic agglomeration economies with a focus on learning externalities. Both papers contain empirical analyses for which I used the statistical software Stata 13 and 14 (StataCorp LLC). The papers are self-contained and single authored. However, I have benefited from very helpful remarks and suggestions from especially my supervisors Professor Dr. Annekatrin Niebuhr and Professor Dr. Johannes Bröcker as well as from seminar participants at the IEB/UEA Summer School in Urban Economics 2016, the Annual Meeting of the German Economic Association (VfS) 2015, the Institute for Employment Research (IAB), the Institute for Regional Research, ifo Dresden, and the Kiel Institute for the World Economy (IfW). The main findings of both papers are summarized in chapter 4.

Chapter 2

Do age complementarities affect labor productivity? Evidence from German firm level data

This chapter has been published as Economics Working Paper No. 2016-10, Christian-Albrechts-Universität zu Kiel, Department of Economics, URL: <http://hdl.handle.net/10419/168033>.

An earlier version has been published in the series 'Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Empirical Labor Economics', URL: <https://www.econstor.eu/handle/10419/112941>.

2.1 Introduction

In the next decades, Germany, as well as many other European countries, will face a shift in the age structure of the workforce due to demographic change. According to a labor market projection commissioned by the Federal Ministry of Labor and Social Affairs, the labor force aged 15 to 24 years will decline from 2010 to 2030 by almost one million persons, and the labor force aged 25 to 54 years by almost five million. On the contrary, the labor supply of workers older than 55 years will increase by almost three million individuals (BMAS, 2013). Considering a certain degree of complementarity between the workers of each group, it is to be expected that such a shift in age structure leads to downward pressure on the productivity, and thereby on the wage rate, of older workers because their share in total employment increases. In contrast, the productivity and the wage rate of younger and middle aged workers presumably will increase since they become relatively rare. Quantitatively, these effects depend on the degree of complementarity between workers that are of a different age. The larger the degree of complementarity between them, the larger is the effect of demographic change on the (relative) productivity of age groups.

Complementary relationships between differently aged workers are supposed to exist as the sets of characteristics relevant in the production process differ across younger and older workers (Garloff/Roth, 2016). Younger workers are stronger in generating and recombining new knowledge, as found by Horn and Cattell (1967) who refer to this as ‘fluid intelligence’. Furthermore, they typically have more physical strength and agility. In contrast, older workers tend to have more experience and abilities to understand difficult situations as well as to use and apply existing skills, i.e., ‘crystallized intelligence’ (Horn/Cattell, 1967; Börsch-Supan/Weiss, 2007; Börsch-Supan, 2013). Empirical results by, e.g., Göbel and Zwick (2013) support that differences in abilities cause complementarities between younger and older workers. Based on German linked employer-employee data, they find that the productivity of younger and older workers in establishments with mixed-age working teams is significantly higher than in other establishments. They relate the productivity gains to the knowledge exchange between different age groups and to benefits of sharing tasks according to comparative advantages (Göbel/Zwick, 2013).

In this paper, I quantify the degree of complementarity of workers that belong to different age groups by estimating cross factor price elasticities. These indicate how the wage rate of age group i changes due to a change in the quantity of workers that belong to age group j . This is of special interest, as it provides information on the impact of demographic change, as well as of any policy that impacts on relative labor supply, such as changes in retirement age, on relative wages. Based on the estimated elasticities, I thereby apply model calculations and show how the expected demographic change might impact on relative wages of younger, middle aged, and older workers in Germany.

In contrast to most previous papers, my analysis is based on a linked employer-employee data set. This allows the investigation of age complementarities on establishment level, i.e., directly at the workplace. It is likely that this yields better estimates of the parameters characterizing establishments’ labor demand than aggregate data (Hamermesh, 1996). Most other studies analyze complementarities between different groups of workers using

data referring to the regional or national level. These studies might suffer from ecological fallacy as aggregate results are interpreted as if they pertain to behavior of firms (see discussion by, e.g., Naz et al., 2015).

Taking into account the interaction of skill and age, I analyze age complementarities on the establishment level within groups of low, medium, and high skilled workers by estimating translog cost functions. Thereby, I impose less restrictions on the elasticities as compared with using CES-functions, which are often used in this context. One basic assumption of a CES-function is that the elasticity between any two groups of workers that differ in age is the same, independently of how old the workers within the two groups are. In contrast, I provide empirical evidence that the degree of complementarity varies between age groups.

The remainder of this paper is structured as follows. In section 2.2 I review the related literature. In section 2.3 I describe the theoretical framework of my empirical analysis, in section 2.4 the empirical specification and in section 2.5 the data set. In section 2.6 I discuss the results of my empirical analysis, and in section 2.7 I set out my conclusions.

2.2 Related literature

Several studies assess the imperfect substitutability of differently aged workers by analyzing cohort effects (e.g., Freeman, 1979; Welch, 1979; Berger, 1985; Sapozhnikov/Triest, 2007; Morin, 2015). Under the hypothesis that workers of different age are imperfect substitutes, an increase in the relative size of an age group has a negative effect on the wage rate of the members of that group. For Germany, the impact of cohort size on cohort specific wages was recently analyzed by Garloff and Roth (2016). They estimate the effect of changes in the share of the youth in total regional employment on the wages of youth workers within the regional labor market. As expected, their findings indicate that wages of younger workers decrease if their share in total regional employment increases.

Similarly, Card and Lemieux (2001) relate the raise in the cohort specific college wage premium to cohort specific shifts in the relative supply of highly educated workers. They observe for the United States, the United Kingdom, and Canada that the college wage premium for younger workers significantly increased relative to the college wage premium for older workers. The empirical analysis indicates that this can be explained by the observation that the age-group specific relative supply of college-educated labor grew faster in the group of older workers than in the group of younger workers.

In their analysis Card and Lemieux (2001) employ a nested CES-framework. This strategy has been adopted by several scholars. Recently, there are several studies focusing on the impact of immigration on wages of natives (e.g., Borjas, 2003; Ottaviano/Peri, 2012; Haas et al., 2013). They apply a nested CES-framework and aim at estimating the degree of complementarity between natives and immigrants within groups of workers that are similar with respect to education and work experience.

Fitzenberger and Kohn (2006) use the CES-framework described by Card and Lemieux (2001) to analyze complementarities between differently skilled workers for Germany. They estimate the parameters of a production

function which refers to the national level using information on a random sample of employees based on the IAB employment subsample (IABS). When analyzing complementarities between differently skilled workers, Fitzenberger and Kohn (2006) take into account the imperfect substitutability of age groups, as they model the amount of differently skilled labor as CES-subaggregates of the skill-specific employment quantities of individuals in different age groups. Their findings confirm that differently aged employees are imperfect substitutes.

However, one drawback of the nested CES-framework applied by Card and Lemieux (2001), Fitzenberger and Kohn (2006) and others is that it is based on “strong functional form assumptions” (Fitzenberger/Kohn, 2006: 24). By modelling skill-specific employment quantities as CES-subaggregates of different age groups, it is assumed that the elasticity of substitution between any pair of differently aged workers is the same, i.e., the degree of complementarity between younger and middle aged workers is the same as between younger and older workers as well as between middle aged and older workers (Hamermesh/Grant, 1979; Ottaviano/Peri, 2012). Suppose, for example the abilities and the experience of a worker change monotonically with her age, then this assumption likely is too restrictive.

As an alternative to a CES-function, a more flexible translog-function may be used to describe the imperfect substitutability of differently aged workers in the production process. It allows the degree of complementarity between age groups to vary. So far, there are only very few studies estimating age complementarities based on translog-functions. Grant and Hamermesh (1981) distinguish between younger and different types of older labor. Their findings for the United States indicate that the degree of complementarity between the different groups of workers indeed differs. Similarly, Arntz and Gregory (2014) find that differently aged workers are complements in the knowledge production process using regional data for Germany. However, this is only true for younger and older workers. Younger and middle aged workers as well as middle aged and older workers seem to be perfect substitutes.

One drawback of the studies by Grant and Hamermesh (1981) and Arntz and Gregory (2014) is that they disregard the interaction of skill and age. As stressed by Fitzenberger and Kohn (2006) it is important to analyze age complementarities within groups of workers that have a comparable skill level.

Taking this into account, I distinguish workers according to their skill level and approximate skill specific labor input by a flexible translog-function rather than by a more restrictive CES-function allowing the degree of complementarity to vary between age groups. In order to quantify the parameters that describe firms’ demand for workers of different age, I estimate an establishment level cost function based on a linked employer-employee data set. Estimating a cost function is based on the weaker assumption that firms minimize costs rather than assuming profit maximization (Hamermesh/Grant, 1979).¹

¹ Hamermesh and Grant (1979) also note that one drawback of estimating a cost function is that it requires to invert the matrix of estimated coefficients to obtain the elasticities of complementarity. A problem arises if one parameter has a very high standard error as this reflects in the standard error of the computed elasticities. However, my analysis does not suffer from this.

2.3 Theoretical framework

To identify complementarities between different age groups, I apply a nested production framework which is in line with neoclassical production theory. This well established approach (see, e.g., Card/Lemieux, 2001; Borjas, 2003; Fitzenberger/Kohn, 2006; Ottaviano/Peri, 2012; Haas et al., 2013) is based on the assumption that the demand for different types of labor is separable from the demand for other inputs (Hamermesh, 1996). As illustrated by figure 2.1, the top level describes the output of establishment f , Y_f , as a function of technology A_f , intermediate products X_f , capital κ_f , and labor L_f .

The labor aggregate L_f is distinguished according to the qualification of the workers (index: q). On the next level the skill specific labor input L_{qf} itself is partitioned into groups that differ according to the age of the workers (index: $i \in \{y, m, o\}$): *younger workers* (y), *middle aged workers* (m), *older workers* (o). It is important to note that L_{qf} is not the pure sum of, respectively, low, medium, and high skilled workers employed by the establishment, but compound labor input, that also depends on complementarities between differently aged workers with skill level q . This paper's aim is to quantify the degree of complementarity between differently aged workers. Therefore, the empirical analysis focuses on the analysis of the third level of the nested production framework, i.e., on the age composition of skill-specific labor inputs.

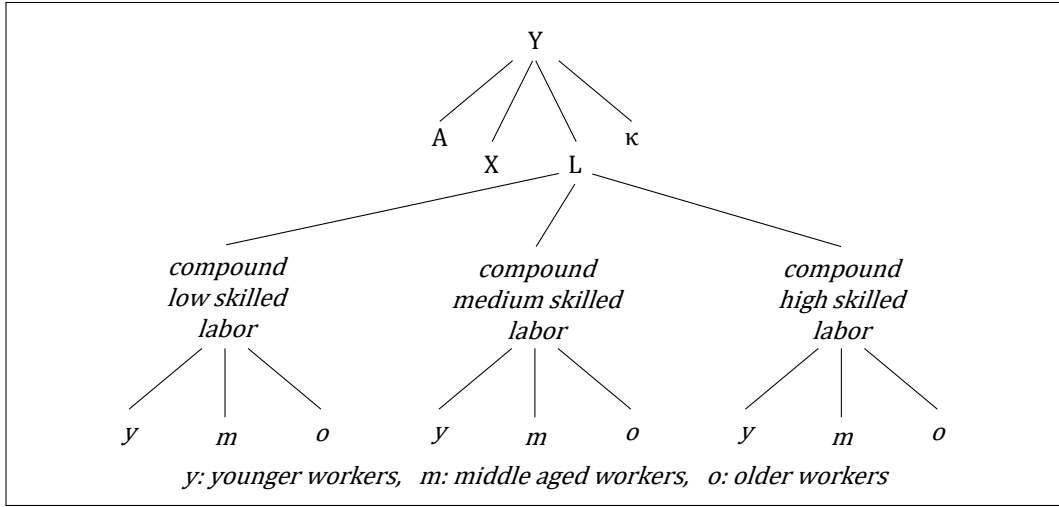


Figure 2.1: Nested production framework

Assuming establishment f choose L_{iqf} , the amount of labor with age i and skill level q , such that the costs to accumulate L_{qf} units of compound labor with skill level q at wage rate vector \mathbf{w}_{qf} are minimized, then, the total costs of L_{qf} are given by $C_q(\mathbf{w}_{qf}, L_{qf})$. As suggested by Christensen et al. (1973) and others, $\ln(C_q(\mathbf{w}_{qf}, L_{qf}))$ can be approximated by a function that is quadratic in the logarithms, see also Greene (2012):

$$\ln(C_q(\mathbf{w}_{qf}, L_{qf})) = \alpha_q + \sum_i \alpha_{iq} v_{iqf} + \frac{1}{2} \sum_i \sum_j \beta_{ijq} v_{iqf} v_{jqf} + l_{qf}$$

with $v_{iqf} := \ln(w_{iqf})$ and $l_{qf} := \ln(L_{qf})$ (2.1)

I assume that $C_q(\mathbf{w}_{qf}, L_{qf})$ is linear homogeneous in L_{qf} . This assumption implies that the cost minimizing, relative composition of L_{qf} with respect to the different age groups is independent of the level of compound labor input with skill level q . Its cost minimizing composition is given by (2.2) which is a system of share equations that depend on the same parameters as (2.1).² S_{iqf} is the value share of age group i within skill group q and establishment f and the parameters β_{ijq} can be interpreted as share elasticities which indicate the skill level specific response of i 's value share to an increase of j 's wage rate (Jorgenson, 1986).

$$S_{iqf}(\mathbf{w}_{qf}) \equiv \frac{L_{iqf}^* w_{iqf}}{\sum_j L_{jqf}^* w_{jqf}} = \alpha_{iq} + \sum_j \beta_{ijq} v_{jqf}, \quad (2.2)$$

Estimates of (2.2) enable computation of the Hicks partial elasticities of complementarity, c_{ijq} , and factor price elasticities, θ_{ijq} (Hicks, 1970; Sato/Koizumi, 1973). In the context of my analysis, the elasticities provide information on age complementarities by skill level. The derivation of the elasticities is given in appendix 2.B.³

In terms of a translog cost function c_{ijq} is given by:

$$c_{ijq} = \frac{|\mathbf{K}_q^{ij}|}{|\tilde{\mathbf{K}}_q|}, \quad (2.3)$$

where $|\tilde{\mathbf{K}}_q|$ is the determinant of matrix $\tilde{\mathbf{K}}_q$ and $|\mathbf{K}_q^{ij}|$ is the cofactor of K_{ijq} . Matrix $\tilde{\mathbf{K}}_q$ is a transformation of the bordered Hessian matrix⁴ of $C_q(\mathbf{w}_{qf}, L_{qf})$:

$$\tilde{\mathbf{K}}_q = \left[\begin{array}{c|ccc} 0 & k_{1q} & \dots & k_{nq} \\ \hline k_{1q} & & & \\ \vdots & & k_{ijq} & \\ k_{nq} & & & \end{array} \right]$$

with

$$\begin{aligned} k_{iq} &= S_{iq}, \\ k_{ijq} &= \beta_{ijq} + S_{iq} S_{jq} \quad , \text{ if } i \neq j \\ k_{iiq} &= \beta_{iiq} + S_{iq} S_{iq} - S_{iq}. \end{aligned} \quad (2.4)$$

Based on c_{ijq} the factor price elasticity is given by:

$$\theta_{ijq} = \frac{\partial \ln(w_{iq})}{\partial \ln(L_{jq})} = S_{jq} c_{ijq}. \quad (2.5)$$

² (2.2) has been derived by taking the partial derivative of (2.1) with respect to v_{iqf} and by applying Shephard's Lemma. See appendix 2.A for a detailed description.

³ As the elasticities depend on the cost shares, the elasticities vary across observations. I omit the index f in the description of the elasticities for simplicity. In the empirical analysis the elasticities are computed at mean predicted shares.

⁴ For details see appendix 2.B. The bordered Hessian matrix and its derivation are given in appendix 2.C.

If $c_{ijq} > 0 (< 0)$ the age groups i and j within skill group q are called quantity complements (substitutes) since an increase in L_{jq} leads to an increase (decline) in w_{iq} . θ_{ijq} measures the percentage change in w_{iq} in response to a one percent increase in L_{jq} given wage rates are flexible (Sato/Koizumi, 1973).

2.4 Empirical specification and econometric issues

Since estimates of translog cost functions like (2.2) do not, per se, fulfill the properties of a microeconomic cost function, parameter constraints have to be imposed. To maintain consistency with Young's theorem of integrable functions, symmetry constraints have to be fulfilled, i.e., $\beta_{ijq} = \beta_{jiq}$, $i \neq j$. In addition, $C_q(\mathbf{w}_{qf}, L_{qf})$ has to be monotone as well as linear homogeneous and concave in w_{iq} (Jorgenson, 1986).

Monotonicity implies that the (predicted) cost shares have to be non-negative. However, since $C_q(\mathbf{w}_{qf}, L_{qf})$ is quadratic in the logarithms of the wage rates, it is not possible to impose parameter constraints that ensure monotonicity (Jorgenson, 1986). Thus, whether this property is fulfilled has to be checked after the estimation.

The other two properties, linear homogeneity and concavity, however, can be imposed. Linear homogeneity requires that the parameters α_{iq} add up to unity and the parameters β_{ijq} add up to zero:

$$\sum_i \alpha_{iq} = 1, \quad \sum_i \beta_{ijq} = 0. \quad (2.6)$$

Concavity of $C_q(\mathbf{w}_{qf}, L_{qf})$ implies that its Hessian matrix \mathbf{H}_q is negative semi-definite. Diewert and Wales (1987) show that this is fulfilled provided that $C(\mathbf{w}_{qf}, L_{qf}) > 0$, if and only if the matrix \mathbf{K}_q (as defined in (2.4)) is negative semi-definite.

However, since \mathbf{H}_q and \mathbf{K}_q depend on the data, i.e., on the cost shares S_{iqf} , a different Hessian matrix and thus a different matrix \mathbf{K}_q is obtained for each observation. As shown by Diewert and Wales (1987), imposing concavity globally destroys the flexibility of the translog function and may result in biased estimates as restrictions on own and cross price elasticities are imposed. Therefore, Ryan and Wales (2000) suggest to impose concavity locally at a normalization point. If the sample mean is chosen as normalization point, i.e., all wage rates are centered around their respective mean, the logarithmic wage rates are zero at this point and S_{iqf} equals α_{iq} , see (2.2), such that \mathbf{H}_q and \mathbf{K}_q in this point only depend on parameters.⁵ The ij th element of \mathbf{K}_q is then given by:

$$k_{ijq} = \beta_{ijq} - \alpha_{iq}\delta_{ij} + \alpha_{iq}\alpha_{jq} \quad (2.7)$$

with $\delta_{ij} = 1$ if $i = j$ and 0 otherwise, compare (2.4). Hence, in this point parameter restrictions can be imposed.

⁵ Please note, centering the wage rates does not affect the parameter estimates, except the estimates of α_{iq} .

Concavity of the cost function in the neighborhood of the normalization point can be attained by setting $-\mathbf{K}_q = \mathbf{D}\mathbf{D}'_q$ (Ryan/Wales, 2000), where \mathbf{D} is a lower triangular matrix.⁶ Solving (2.7) for β_{ijq} and replacing $-k_{ijq}$ by the ij th element of $\mathbf{D}\mathbf{D}'_q$ ⁷ yields non linear constraints that impose concavity of $C_q(\mathbf{w}_{qf}, L_{qf})$ in the neighborhood of the normalization point (Ryan/Wales, 2000: 255):

$$\beta_{ijq} = -(\mathbf{D}\mathbf{D}')_{ijq} + \alpha_{iq}\delta_{ijq} - \alpha_{iq}\alpha_{jq}. \quad (2.8)$$

If the parameter constraint given by (2.8) as well as the symmetry constraint, i.e., $\beta_{ijq} = \beta_{jiq}$, are inserted into the system of share equations given by (2.2) and random disturbances added, (2.2) changes to (2.9).⁸ Since the parameters β_{ijq} are replaced by (2.8), the parameters d_{ijq} are estimated instead of β_{ijq} . It is worth noting that the number of parameters d_{ijq} is equal to the number of parameters β_{ijq} . The non-linear constraint given by (2.8) ensures that the parameters β_{ijq} are in a certain relationship to each other such that the Hessian matrix of $C_q(\mathbf{w}_{qf}, L_{qf})$ is negative semi-definite, at least in the reference point.

$$\begin{aligned} S_{yqf} &= \alpha_{yq} + (-d_{yyq}d_{yyq} + \alpha_{yq} - \alpha_{yq}\alpha_{yq})v_{yqf} + (-d_{ymq}d_{yyq} - \alpha_{yq}\alpha_{mq})v_{mqf} + (-d_{yyq}d_{yoq} - \alpha_{yq}\alpha_{oq})v_{oqf} + \varepsilon_{yqf} \\ S_{mqf} &= \alpha_{mq} + (-d_{ymq}d_{yyq} - \alpha_{yq}\alpha_{mq})v_{yqf} + (-d_{ymq}d_{ymq} - d_{mmq}d_{mmq} + \alpha_{mq} - \alpha_{mq}\alpha_{mq})v_{mqf} \\ &\quad + (-d_{yoq}d_{ymq} - d_{moq}d_{mmq} - \alpha_{mq}\alpha_{oq})v_{oqf} + \varepsilon_{mqf} \\ S_{oqf} &= \alpha_{oq} + (-d_{yyq}d_{yoq} - \alpha_{yq}\alpha_{oq})v_{yqf} + (-d_{yoq}d_{ymq} - d_{moq}d_{mmq} - \alpha_{mq}\alpha_{oq})v_{mqf} \\ &\quad + (-d_{yoq}d_{yoq} - d_{moq}d_{moq} - d_{ooq}d_{ooq} + \alpha_{oq} - \alpha_{oq}\alpha_{oq})v_{oqf} + \varepsilon_{oqf}. \end{aligned} \quad (2.9)$$

The indices y , m , and o refer to younger, middle aged, and older workers respectively, that belong to skill group q within establishment f . ε_{iqf} are unobservable random disturbances for the cost shares.

Since the cost shares S_{yqf} , S_{mqf} , and S_{oqf} sum up to unity for each observation qf , the sum of the error terms $\sum_i \varepsilon_{iqf}$ is equal to zero for all observations (Jorgenson, 1986). Hence, the error covariance matrix is singular. Greene (2012) therefore suggests to estimate just $n - 1$ share equations. Furthermore, he notes that computing maximum likelihood estimates of the parameters ensures invariance with respect to the choice of which share equation is dropped.⁹

As stressed for example by Grant and Hamermesh (1981), one only derives unbiased estimates of a cost function, or a system of share equations that is based on factor prices, if the latter are exogenous to the firm. Hamermesh (1996) therefore argues that the cost function would ideally be embedded in a system of equations that also describes the labor supply reaction or workers' choices. As an alternative, strictly exogenous variables might be used to generate instruments for the factor prices in the cost function. However, to find such vari-

⁶ The idea of employing the Cholesky factorization to impose concavity goes back to Lau (1978).

⁷ Here: $(\mathbf{D}\mathbf{D}')_{\mathbf{q}} = \begin{bmatrix} d_{yyq}^2 & d_{yyq}d_{myq} & d_{yyq}d_{oyq} \\ d_{yyq}d_{myq} & d_{myq}^2 + d_{mmq}^2 & d_{myq}d_{oyq} + d_{mmq}d_{omq} \\ d_{yyq}d_{oyq} & d_{myq}d_{oyq} + d_{mmq}d_{omq} & d_{oyq}^2 + d_{omq}^2 + d_{ooq}^2 \end{bmatrix}$.

⁸ The adding-up constraints $\sum_i \alpha_{iq} = 1$ and $\sum_i \beta_{ijq} = 0$ are directly imposed when estimating the system of equations.

⁹ I estimate (2.9) using Stata's iterative feasible generalized least squares estimator which is equivalent to ML estimation with multivariate normal disturbances and also invariant to the choice of the estimated share equations (see Stata's manual *nlsur - Estimation of nonlinear systems of equations*, page 9, <http://www.stata.com/manuals13/rnl sur.pdf>).

ables is already a severe problem for studies that consider homogeneous labor. In the case of heterogeneous labor it is “insolvable” (Hamermesh, 1996: 70). Therefore, one has to decide whether to treat quantities as exogenous and to estimate a production function, or to treat wage rates as exogenous and to estimate a cost function. Hamermesh (1996) argues that in the case of microeconomic data it is reasonable to assume the latter since small units like establishments face perfect elastic labor supply in the long run. As a robustness check, I exclude all establishments with 50 employees or more to address that large establishments may have local monopsony power to set wages.

2.5 Data

To estimate the parameters of (2.9) that are needed to calculate c_{ij} and θ_{ij} , I use detailed information on a 20 percent random sample of all establishments located in Germany that employed at least one worker subject to social security contributions on June 30, 2012. The public sector, temporary employment agencies, and crop and animal production, hunting and related service activities are excluded. The considered number of establishments amounts to 508,514. For these establishments my raw data set contains all spells of employment that were reported by the establishments in 2012, i.e., 9,124,309 individual employment spells. The information on the spells of employment are taken from the IAB employment history (BeH).

The BeH is based on the integrated notification procedure for health, pension, and unemployment insurance. Employers are required to submit notifications to the responsible social security agencies concerning all of their employees covered by social security as well as marginally employed at least once a year. The Federal Employment Agency collects and edits these notifications, which the IAB subsequently integrates into the BeH. Thus, the BeH contains very reliable information on all spells of employment as long as the workers are not exempt from social security contributions, like civil servants and self-employed persons (about 12 percent of total employment in Germany).

Among other things, each observation in the BeH provides information on individual and job characteristics as well as on the length of the specific spell of employment and the paid gross wage sum during this time period. Due to the social security threshold, the wage information is right-censored and in 2012 only reported up to €67,200 per year in West Germany and €57,600 per year in East Germany. To avoid biased estimates, I partly impute the wages of employees subject to social security contributions with estimated wages.¹⁰

Aggregating the information provided by the 9 million individual spells of employment from the BeH, a unique establishment level data set is generated.¹¹ The units of observation are skill groups within the considered establishments. As I am interested in composition effects of the skill groups at the establishment level with regard to the age of the workers, I exclude very small skill groups, which I define as groups with less than five

¹⁰ I follow Reichelt (2015) and apply an interval regression to estimate the (daily) wages above the thresholds, see appendix 2.D.

¹¹ Six percent of the employment spells are not considered. I exclude apprentices from the analysis. Furthermore, employment spells of interns, and student trainees are only considered as long as the employees have at least a bachelors degree and have an occupation that requires a vocational training of at least two years or higher education.

members, i.e., $\sum_i L_{iqf} < 5$. The final number of observations that are used in the empirical analysis amounts to 290,334 skill groups within 208,868 establishments.¹²

To distinguish the employees of an establishment according to their skill level, i.e., low skilled, medium skilled, and high skilled, the requirement level of the individual job is used.¹³ I define workers with an occupation that does not require a vocational training of at least two years as low skilled. Medium skilled workers are those with an occupation that requires a vocational training of at least two years, and workers with an occupation that requires at least a bachelor's degree, a qualification as master craftsman or technician, or an equivalent technical school or college graduation are considered as high skilled.

I use the requirement level of the job instead of the highest educational degree of a worker which is also provided in the BeH in order to identify low skilled, medium skilled, and high skilled labor for two reasons. First, the requirement level enables assignment of workers to the different skill groups according to the complexity of their actual task and not according to their formal level of qualification. Haas et al. (2013) show that workers with the same formal educational degree, but with different job requirements, are imperfect substitutes. Second, 27 percent of the nine million spells of employment do not provide information on the educational degree of the worker, whereas the information on the requirement level is missing only in two percent of the employment spells.¹⁴

For each skill group within an establishment, I compute its individual composition with regard to the age of the workers. To be more specific, I compute the establishment-specific cost share of age group i within skill group q and establishment f , S_{iqf} . Unfortunately, the BeH does not provide information on non-wage labor costs. Therefore, the computed cost shares refer to wage costs only. Summary statistics are provided by table 2.1. In addition to the cost shares, the wages of the different age-skill groups, w_{iqf} , that enter (2.9) on the right hand side are calculated. w_{iqf} indicates how much establishment f paid in 2012 on average per day to an employee with age i and skill level q . For the calculation of w_{iqf} only spells employment are used that refer to full-time employment as more detailed information on working hours is not available. Moreover, when calculating w_{iqf} only spells of employment with a length of at least eight days and a gross-daily wage above the limit of marginally employed are considered. In 40 percent of the cases it is not possible to calculate w_{iqf} since establishment f does not employ workers with age i and skill level q whose spells of employment fulfill the criteria above. In these cases $w_{iqr(f)}$ is used as an approximation of w_{iqf} , where $w_{iqr(f)}$ is the average gross daily wage of group iq in region $r(f)$, i.e., the local labor market as defined by Kosfeld and Werner (2012) in which establishment f is located. $w_{iqr(f)}$ is calculated based on information from the Employment Statistics of the Federal Employment Agency. It can be interpreted as the daily wage that establishment f would have to pay to employ one unit of type iq labor in the local labor market where it is located.

¹² Without excluding skill groups with less than five members, the number of observations would be 885,377.

¹³ The requirement level of a job is the second dimension of the German Classification of Occupations 2010 (KldB 2010). A detailed description can be found in Paulus and Matthes (2013).

¹⁴ If the requirement level is missing but the highest educational degree is not, then the later is used to assign the workers to one of the skill groups. This reduces the share of employment spells that cannot be assigned to one of the three groups to 0.07%. In 65.8% of the cases in which information on the requirement level as well as on the highest educational degree is available the requirement level accords with the reported highest educational degree. In 16.6% the reported highest educational degree is higher and in 17.6% lower than the requirement level.

The age group to which an employee belongs is defined by his or her day of birth that is also given in the BeH. In accordance with Mahlberg et al. (2013), employees younger than 30 years are considered as *younger workers*, workers aged 30 to 49 years as *middle aged workers* and workers that are 50 years old or older as *older workers*.¹⁵

Table 2.1: Summary statistics

	Mean	SD	Min	Max
Groups of low skilled labor (N = 64405)				
Cost shares				
Younger labor	0.215	0.244	0.000	1.000
Middle aged labor	0.459	0.259	0.000	1.000
Older labor	0.326	0.266	0.000	1.000
ln(Average gross daily wages)				
Younger labor	3.976	0.227	2.558	5.621
Middle aged labor	4.155	0.295	2.565	6.268
Older labor	4.229	0.273	2.570	6.127
Groups of medium skilled labor (N = 173313)				
Cost shares				
Younger labor	0.205	0.213	0.000	1.000
Middle aged labor	0.501	0.232	0.000	1.000
Older labor	0.294	0.238	0.000	1.000
ln(Average gross daily wages)				
Younger labor	4.203	0.273	2.560	6.710
Middle aged labor	4.336	0.365	2.558	6.214
Older labor	4.424	0.351	2.559	6.460
Groups of high skilled labor (N = 52617)				
Cost shares				
Younger labor	0.121	0.178	0.000	1.000
Middle aged labor	0.561	0.237	0.000	1.000
Older labor	0.318	0.252	0.000	1.000
ln(Average gross daily wages)				
Younger labor	4.506	0.286	2.570	6.368
Middle aged labor	4.823	0.417	2.568	7.441
Older labor	4.960	0.408	2.567	6.959

2.6 Results

I start with a regression where I pool the observations that belong to different skill groups. Thus, by now I assume that the parameters and elasticities do not differ across skill levels. The first and the second column of table 2.2 contain the estimated parameters of (2.9), i.e., the system of share equations including the parameter constraints ensuring that the cost function is concave. As the adding-up constraints, see (2.6), are imposed when estimating the model, not all parameters are explicitly reported.¹⁶ Inserting the reported parameters into (2.8)

¹⁵ 0.4 percent of the nine million spells of employment are not assigned to one of the age groups since the worker is implausible old (above 75 years).

¹⁶ Please note, the adding-up constraints given by (2.6) require that $\sum_i \beta_{ij}$ is zero. The corresponding constraints with regard to d_{ij} are obtained by combining the adding-up constraints with the constraints that are imposed to ensure that the cost function is concave which is given by (2.8).

and (2.6), i.e., the concavity and the adding-up constraints, the parameters β_{ij} are obtained. They are reported in the third column of table 2.2. Based on these parameters, the Hicks partial elasticities of complementarity, c_{ij} , and factor price elasticities, θ_{ij} and θ_{ji} , are calculated using (2.3) and (2.5), respectively. As θ_{ij} depends on the cost share of age group j and θ_{ji} on the cost share of age group i , θ_{ij} is not symmetric, i.e., $\theta_{ij} = \theta_{ji}$ only holds if $S_{iq} = S_{jq}$.

Table 2.2: Estimated parameters and calculated elasticities, average effects assuming $\alpha_{iq} = \alpha_i$, $\beta_{ijq} = \beta_{ij}$

i	j	α_i	d_{ij}	β_{ij}	c_{ij}	θ_{ij}	θ_{ji}
y	y	0.194*** (0.000)	0.455*** (0.002)	-0.050*** (0.001)	-3.273*** (0.022)	-0.628*** (0.004)	-0.628*** (0.004)
m	m	0.503*** (0.000)	0.528*** (0.001)	-0.066*** (0.001)	-0.801*** (0.004)	-0.403*** (0.002)	-0.403*** (0.002)
o	o			-0.137*** (0.001)	-1.394*** (0.006)	-0.426*** (0.002)	-0.426*** (0.002)
y	m		-0.192*** (0.002)	-0.010*** (0.001)	0.999*** (0.008)	0.502*** (0.004)	0.192*** (0.002)
y	o			0.061*** (0.001)	0.413*** (0.008)	0.126*** (0.002)	0.079*** (0.002)
m	o			0.076*** (0.001)	0.690*** (0.004)	0.211*** (0.001)	0.347*** (0.002)
N		290335					

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the sample mean. Monotonicity and concavity of the cost function are satisfied at all data points. The concavity constraint is locally imposed at the sample mean. The elasticities are calculated at mean predicted shares.

Since I center the wage rate of each age-skill group around their mean, the results for α_i indicate that, at the reference point where the wage rates correspond to their respective mean, the cost share of younger workers amounts to about 20 percent, the cost share of middle aged workers to about 50 percent, and, hence, the cost share of older workers to about 30 percent. Furthermore, the results in column three indicate that the cost share of each age group declines, given an increase in the wage rate of the respective age group. This shows that firms react on increasing relative wages by substituting one age group by another. However, there is also a negative relationship between the cost share of younger workers and the wage rate of middle aged workers (and vice versa due to the symmetry constraint). This negative cross effect points to the imperfect substitutability of younger and middle aged workers. However, positive cross price elasticities, per se, do not necessarily imply that those groups of workers are perfect substitutes.¹⁷

The results in column four to six indicate how the productivity of age groups change in response to changes in the amount of age specific labor input. As expected, the marginal productivity of an age group declines if its size increases. The effect is largest for younger workers which is in line with findings by Welch (1979). He observes that the negative effect of cohort size on wages of the same cohort declines over the career.

¹⁷ It follows from the adding-up constraint given by (2.6) that the parameters β_{ym} as well as β_{mo} need to be positive if β_{ym} and β_{ii} are negative.

The cross effects indicate that all three age groups are complementary input factors, i.e., all age groups benefit from working together with differently aged workers. But the degree of complementarity differs. This points out that flexible functional forms, like a translog function, are needed to quantify the degree of complementarity between different age groups. A CES framework, for example, is too restrictive as the elasticities of substitution between any two pairs of age groups are assumed to be the same.

My results show that the highest degree of complementarity is observed for younger and middle aged workers, the lowest for younger and older workers. With regard to younger workers this indicates that they especially benefit of working together with middle aged workers. One reasonable explanation is that younger workers benefit from the working experience of middle aged workers. At the same time it might be beneficial that the knowledge base and the abilities of younger and middle aged workers are not too different which might enhance cooperation. An alternative interpretation of the cross effects is that it is more difficult to substitute middle aged workers by either younger or older workers, than younger by older workers (and vice versa) as middle aged workers have a combination of cognitive abilities of younger workers and experience of older workers. This combination make them especially valuable, what in turn decreases substitutability.

I use the calculated own and cross factor price elasticities, θ_{ij} , to simulate how the demographic change might affect the productivity of the three age groups due to the expected shift in relative labor supply (compare Grant/-Hamermesh, 1981). It is assumed that the shift in relative labor supply reflects in a corresponding shift in the age structure of (skill-specific) employment. Furthermore, the model calculation is based on the assumption that the productivity of compound skill-specific labor in terms of the value of its marginal product, $P \frac{\partial Y}{\partial L} \frac{\partial L}{\partial L_q}$ where P denotes the price of one unit of output Y , remains constant for all three skill groups.¹⁸ In the model calculation, I examine a scenario where the number of younger workers decreases by 21 percent, and the number of middle aged workers by 16.5 percent. The number of older workers is assumed to increase by 3.9 percent.¹⁹

The results of the model calculation are summarized in table 2.3. They suggest that the productivity of younger workers will increase most (by five percent) due to the assumed changes in the age structure of the workforce. The own factor price elasticity of younger workers (θ_{yy}) would predict an productivity increase of the wage rate by more than 13 percent in response to the assumed decline of younger labor by 21 percent. Further upward pressure on the wage of younger workers comes from the increasing number of older workers as it lets the relative amount of young workers decline. In contrast, the decreasing number of middle aged workers puts downward pressure on the wages of younger workers due to the complementary relationship. In total, the productivity of younger workers is therefore expected to rise ‘only’ by about five percent.

Similarly, the decline in the number of middle aged workers and the increase in the number of older workers let the productivity of middle aged workers increase. The decreasing number of younger workers, however, puts downward pressure on the wage of middle aged workers. The total effect is expected to be a productivity

¹⁸ This assumption implies that the marginal costs of compound skill-specific labor are assumed to remain unchanged since at the cost minimum the marginal value product equals marginal costs (compare appendix 2.B: for the model calculation $d\phi$ is assumed to be zero).

¹⁹ This scenario is based on the relative change of the population aged 15-29 years, 30-49 years, and 50-70 years that is expected by the German Federal Statistical Office until 2030 (DESTATIS, 2009). Changes in labor participation rates are not considered.

increase of about three percent. For older workers the model calculation predicts a strong decrease in productivity, i.e., a decline by nine percent. Predominantly this is caused by the decreasing number of younger and middle aged workers, wherefore the relative amount of older workers increases. The ceteris paribus effect of the increasing number of older workers suggests a reduction in productivity by less than two percent.

Table 2.3: Results of a model calculation

Expected percentage change in productivity of ...		
younger workers	middle aged workers	older workers
5.4	3.4	-9.0
(0.06)	(0.02)	(0.04)

Effects based on factor price elasticities (θ_{ij} , θ_{ji} in table 2.2) and the scenario: number of younger workers -21%, middle aged workers -16.5%, older workers: +3.9%. Standard errors in parentheses, calculation based on Delta-method.

By relaxing the assumption that the elasticities do not vary across skill levels, heterogeneous effects across skill levels become visible, see table 2.4. Own price elasticities are highest in absolute terms within groups of high-skilled labor, especially for younger high-skilled labor. It indicates that a relatively small decline in the number of younger high-skilled workers leads to a relatively large productivity increase of the remaining younger workers.

The cross effects are heterogeneous across skill groups as well. The degree of complementarity between younger and middle aged workers is highest within groups of high-skilled labor. The same result is obtained for the degree of complementarity between middle aged and older workers. This supports the findings of Welch (1979) who observes that the degree of complementarity between cohorts increases by skill level.

Surprisingly and in contrast to the expectation that especially within groups of high skilled labor cognitive abilities of younger workers and experience of older workers complement each other, I find that the degree of complementarity between younger and older workers decreases in the skill level. The presence of younger workers seems to have a higher importance for the productivity of older workers (and vice versa) in groups of low skilled workers than in other skill groups. One reasonable explanation for the high degree of complementarity between younger and older workers with jobs that require only little specific training is that some of their tasks are often connected with physical burden. Hence, older low skilled workers might benefit from working together with younger workers as tasks could be shared according to the individual physical capacity. In addition, especially within groups of low skilled workers with tasks that do not require specific training, general labor market experience might still enhance productivity. Therefore, younger low skilled workers might benefit from older low skilled workers who presumably have at least some general working experience.

Table 2.4: Elasticities by skill level

i	j	low skilled labor			medium skilled labor			high skilled labor		
		c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}
y	y	-2.936*** (0.053)	-0.630*** (0.011)	-0.630*** (0.011)	-2.257*** (0.014)	-0.463*** (0.003)	-0.463*** (0.003)	-7.870*** (0.183)	-0.954*** (0.022)	-0.954*** (0.022)
m	m	-0.818*** (0.010)	-0.376*** (0.005)	-0.376*** (0.005)	-0.707*** (0.004)	-0.354*** (0.002)	-0.354*** (0.002)	-1.042*** (0.019)	-0.585*** (0.010)	-0.585*** (0.010)
o	o	-1.242*** (0.014)	-0.405*** (0.005)	-0.405*** (0.005)	-1.324*** (0.006)	-0.390*** (0.002)	-0.390*** (0.002)	-1.847*** (0.021)	-0.587*** (0.007)	-0.587*** (0.007)
y	m	0.890*** (0.019)	0.409*** (0.009)	0.191*** (0.004)	0.768*** (0.006)	0.385*** (0.003)	0.157*** (0.001)	1.892*** (0.053)	1.061*** (0.030)	0.229*** (0.006)
y	o	0.678*** (0.022)	0.221*** (0.007)	0.146*** (0.005)	0.265*** (0.006)	0.078*** (0.002)	0.054*** (0.001)	-0.337*** (0.043)	-0.107*** (0.014)	-0.041*** (0.005)
m	o	0.566*** (0.009)	0.185*** (0.003)	0.260*** (0.004)	0.669*** (0.004)	0.197*** (0.001)	0.335*** (0.002)	1.119*** (0.018)	0.355*** (0.006)	0.628*** (0.010)
N		64405			173313			52617		

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the respective sample mean. Monotonicity and concavity of the cost function are satisfied at more than 99 percent of the respective data points. The concavity constraint is locally imposed at the respective sample mean. The elasticities are calculated at mean predicted shares.

The high degree of substitutability between younger and older high-skilled labor supports the view of Fitzenberger and Kohn (2006). They argue that university graduates have in general (independent of their age) high competence to solve problems which reduces complementarities. Nevertheless, it remains surprising that, within groups of high skilled labor, younger and older workers are rather substitutes than complements, but younger and middle aged workers as well as middle aged and older workers complement each other. But the finding is robust.

As stressed earlier, my estimates can only be interpreted as causal effects if wage rates are exogenous to the firms. As a robustness check, I repeat the estimations with a reduced sample that only covers rather small establishments, defined as establishments with less than 50 employees. The additional results are summarized in table 2.E.1 and table 2.E.2 in appendix 2.E. Qualitatively the results are similar to those discussed in this section. Quantitatively, the parameter estimates and calculated elasticities that are based on the reduced sample are ten percent to 20 percent smaller than those that are based on the full sample. As a further robustness check, I restrict my sample to skill groups with, on the establishment level, at least ten members. Again I observe qualitatively similar results, even though the number of observations reduces significantly.²⁰

2.7 Conclusion

In this study, I analyze complementarities between differently aged workers. This is of special interest since demographic change leads to a shift in the workforce age structure in Germany as well as in many other European countries which likely affects (relative) productivity of workers. Taking into account the interaction of skill and age, I analyze age complementarities within groups of workers with a comparable skill level. Contrary to previous papers, my analysis is based on a linked employer-employee data set. This allows the investigation of age complementarities on establishment level, i.e., directly at the workplace. Studies operating at the aggregate level might suffer from ecological fallacy because their results are interpreted as if they pertain to behavior of firms. Moreover, in contrast to most other studies, I provide empirical evidence on the way the degree of complementarity varies between age groups. Often it is assumed that the substitutability of different age groups is constant, meaning that it does not depend on the age groups that are considered. My findings suggest that this assumption does not hold.

The results of this study indicate that workers who belong to different age groups are complements and that the degree of complementarity differs across the considered age groups and by skill level. Even though, for each skill level the largest degree of complementarity is identified between younger and middle aged workers which especially applies to the group of high-skilled labor. In contrast, younger and older high-skilled workers seem to be substitutes rather than complements. This surprises, as Arntz and Gregory (2014) find that cognitive abilities of younger workers and experience of older workers complement each other in generating innovations, which typically involves high-skilled labor.

²⁰ The additional results are available from the author upon request.

Based on my results, I simulate how the expected demographic change impacts on the productivity of the different age groups. Due to the complementary relationship between older workers and the two other groups, the expected decline in younger and middle aged workers puts pressure on the wages of older workers. Under the assumption that wages are completely flexible and that the overall productivity of compound labor remains constant, one could for this age group expect a decline in wages of about nine percent until 2030. On the contrary, my results show that the wage of younger and middle aged workers is expected to rise by about five percent and three percent, respectively, since their relative shares decline.

Appendix

2.A Deriving the system of share equations

The partial derivative of $\ln(C_q(\mathbf{w}_{qf}, L_{qf}))$, as given by (2.1), with respect to $\ln(w_{iqf})$ corresponds to

$$\frac{\partial \ln(C_q(\mathbf{w}_{qf}, L_{qf}))}{\partial \ln(w_{iqf})} = \alpha_{iq} + \sum_j \beta_{ijq} \ln(w_{jqf}). \quad (2.10)$$

$\frac{\partial \ln(C_q(\mathbf{w}_{qf}, L_{qf}))}{\partial \ln(w_{iqf})}$ is the elasticity of $C_q(\mathbf{w}_{qf}, L_{qf})$ with regard to w_{iqf} . Thus:

$$\frac{\partial \ln(C_q(\mathbf{w}_{qf}, L_{qf}))}{\partial \ln(w_{iqf})} = \frac{\frac{dC_q(\mathbf{w}_{qf}, L_{qf})}{C_q(\mathbf{w}_{qf}, L_{qf})}}{\frac{dw_{iqf}}{w_{iqf}}} = \frac{dC_q(\mathbf{w}_{qf}, L_{qf})}{dw_{iqf}} \frac{w_{iqf}}{C_q(\mathbf{w}_{qf}, L_{qf})}. \quad (2.11)$$

According to Shephard's Lemma, see appendix 2.C: $\frac{dC_q(\mathbf{w}_{qf}, L_{qf})}{dw_{iqf}} = L_{iqf}^*$, where L_{iqf}^* is the cost minimizing number of type iq workers given the price vector \mathbf{w}_{qf} and L_{qf} . Hence, the cost minimizing share S_{iqf} of type iq workers in type q workers amounts to:

$$S_{iqf}(\mathbf{w}_{qf}) = \frac{L_{iqf}^* w_{iq}}{\sum_j L_{jqf}^* w_{jqf}} = \alpha_{iq} + \sum_j \beta_{ijq} \ln(w_{jqf}). \quad (2.12)$$

2.B Deriving the Hicks partial elasticity of complementarity and the factor price elasticity

The derivation of the Hicks partial elasticity of complementarity and of factor price elasticity are provided by, e.g., Sato and Koizumi (1973). As the elasticities are of central importance for my empirical analysis, I sketch it here in more detail. For simplicity, I omit the indices q and f .

Sato and Koizumi (1973) start with Shephard's Lemma. Accordingly, the partial derivative of a cost function with respect to factor price of input i corresponds to the conditional factor demand of i :

$$\underbrace{\frac{\partial C(\mathbf{w}, L)}{\partial w_i}}_{C_i} = L_i(\mathbf{w}, L). \quad (2.13)$$

Assuming the cost function is linear homogeneous in L , the corresponding total differential is given by, compare appendix 2.C:

$$C_i \frac{dL}{L} + \sum_j C_{ij} dw_j = dL_i. \quad (2.14)$$

This equation can be transferred into a system of total differentials and expressed in matrix form:

$$\underbrace{\begin{bmatrix} 0 & C_1 & \dots & C_n \\ C_1 & C_{11} & \dots & C_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ C_n & C_{n1} & \dots & C_{nn} \end{bmatrix}}_{\tilde{\mathbf{H}}} \underbrace{\begin{bmatrix} \frac{dL}{L} \\ dw_1 \\ \vdots \\ dw_n \end{bmatrix}}_{d\mathbf{w}} = \underbrace{\begin{bmatrix} Ld\phi \\ dL_1 \\ \vdots \\ dL_n \end{bmatrix}}_{d\mathbf{L}}$$

where $\tilde{\mathbf{H}}$ denotes the bordered Hessian matrix of the cost function, see appendix 2.C The first row refers to the change of minimum total costs due to changes in factor prices. ϕ denotes the Lagrange multiplier of cost minimization which in equilibrium equals marginal costs (Sato/Koizumi, 1973).

It follows that $d\mathbf{w} = \tilde{\mathbf{H}}^{-1} d\mathbf{L}$. Hence, the partial derivative of w_i with respect to L_j is given by:

$$\frac{\partial w_i}{\partial L_j} = (\tilde{\mathbf{H}}^{-1})_{ij} \equiv \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}, \quad (2.15)$$

where $|\mathbf{H}^{ij}|$ denotes the cofactor of C_{ij} and $|\tilde{\mathbf{H}}|$ the determinant of $\tilde{\mathbf{H}}$.

Therefore, the factor price elasticity is given by

$$\frac{\partial \ln(w_i)}{\partial \ln(L_j)} = \frac{\partial w_i}{\partial L_j} \frac{L_j}{w_i} = \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|} \frac{L_j}{w_i}. \quad (2.16)$$

By multiplying the right hand side by $\frac{w_j C}{w_j C}$, the factor price elasticity can be expressed as a function of j 's cost share, S_j , and the Hicks partial elasticity of complementarity, c_{ij} :

$$\frac{\partial \ln(w_i)}{\partial \ln(L_j)} = \frac{L_j w_j}{C} \underbrace{\frac{C}{w_i w_j}}_{S_j} \underbrace{\frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}}_{c_{ij}}. \quad (2.17)$$

As I consider a translog specification of a cost function, the Hicks partial elasticity of complementarity simplifies to:

$$c_{ij} = \frac{|\mathbf{K}^{ij}|}{|\tilde{\mathbf{K}}|}, \quad (2.18)$$

where $\tilde{\mathbf{K}}$ is a transformation of $\tilde{\mathbf{H}}$:

$$\tilde{\mathbf{K}} = \begin{bmatrix} 0 & K_1 & \dots & K_n \\ K_1 & K_{11} & \dots & K_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ K_n & K_{n1} & \dots & K_{nn} \end{bmatrix}$$

with, compare appendix 2.C,

$$\begin{aligned} K_i &= C_i \frac{w_i}{C(\mathbf{w}, L)} \\ &= S_i \end{aligned} \quad (2.19)$$

$$\begin{aligned} K_{ii} &= C_{ii} \frac{w_i^2}{C(\mathbf{w}, L)} \\ &= S_i^2 - S_i + \beta_{ii} \end{aligned} \quad (2.20)$$

$$\begin{aligned} K_{ij} &= C_{ij} \frac{w_i w_j}{C(\mathbf{w}, L)} \\ &= S_i S_j + \beta_{ij}. \end{aligned} \quad (2.21)$$

That in the case of a translog cost function it holds

$$\frac{|\mathbf{K}^{ij}|}{|\tilde{\mathbf{K}}|} = \frac{C}{w_i w_j} \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}, \quad (2.22)$$

can be seen if $\tilde{\mathbf{K}}$ is expressed in matrix notation:

$$\tilde{\mathbf{K}} = \frac{\widehat{\mathbf{W}} \widehat{\mathbf{H}} \widehat{\mathbf{W}}}{C(\mathbf{w}, L)},$$

where $\widehat{\mathbf{W}}$ is a diagonal matrix:

$$\widehat{\mathbf{W}} \equiv \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & w_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n \end{bmatrix}.$$

Hence:

$$\tilde{\mathbf{K}}^{-1} = C(\mathbf{w}, L) \widehat{\mathbf{W}}^{-1} \tilde{\mathbf{H}}^{-1} \widehat{\mathbf{W}}^{-1} \quad (2.23)$$

and

$$\frac{|\mathbf{K}^{ij}|}{|\tilde{\mathbf{K}}|} \equiv (\mathbf{K}^{-1})_{ij} = C(\mathbf{w}, L) \frac{1}{w_i} (\mathbf{H}^{-1})_{ij} \frac{1}{w_j} \equiv \frac{C}{w_i w_j} \frac{|\mathbf{H}^{ij}|}{|\tilde{\mathbf{H}}|}. \quad (2.24)$$

2.C Deriving the bordered Hessian matrix $\tilde{\mathbf{H}}$ of the cost function $C(\mathbf{w}, L)$

In the following, I derive the (bordered) Hessian matrix of the translog cost function given by equation (2.1). For simplicity, I omit the indices q and f . The result can also be found in, e.g., Hamermesh and Grant (538ff. 1979).

$$\tilde{\mathbf{H}} = \left[\begin{array}{c|ccc} 0 & C_1 & \dots & C_n \\ \hline C_1 & & & \\ \vdots & & \mathbf{H} & \\ C_n & & & \end{array} \right]$$

$$\mathbf{H} = \begin{bmatrix} C_{11} & \dots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \dots & C_{nn} \end{bmatrix}$$

$$\begin{aligned} C(\mathbf{w}, L) &= \exp \left[a_0 + \sum_i \alpha_i \ln(w_i) + \frac{1}{2} \sum_i \sum_j \beta_{ij} \ln(w_i) \ln(w_j) + l \right] \\ C_i = \frac{\partial C(\mathbf{w}, L)}{\partial w_i} &= C(\mathbf{w}, L) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right), \text{ with } \beta_{ij} = \beta_{ji} \\ &= \frac{C(\mathbf{w}, L)}{w_i} \left[\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right] \\ &= \frac{C(\mathbf{w}, L)}{w_i} S_i (= L_i^*, \text{ Shephard's Lemma}) \end{aligned} \quad (2.25)$$

$$\begin{aligned} C_{ii} = \frac{\partial^2 C(\mathbf{w}, L)}{\partial w_i \partial w_i} &= C(\mathbf{w}, L) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right) \\ &\quad + C(\mathbf{w}, L) \left(-\alpha_i \frac{1}{w_i} \frac{1}{w_i} + \left(-\sum_j \beta_{ij} \frac{1}{w_i} \frac{1}{w_i} \ln(w_j) + \beta_{ii} \frac{1}{w_i} \frac{1}{w_i} \right) \right) \\ &= C(\mathbf{w}, L) \left[\frac{1}{w_i} \left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) \frac{1}{w_i} \left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) + \frac{1}{w_i^2} \left(-\alpha_i - \sum_j \beta_{ij} \ln(w_j) + \beta_{ii} \right) \right] \\ &= \frac{C(\mathbf{w}, L)}{w_i^2} [S_i^2 - S_i + \beta_{ii}] \end{aligned} \quad (2.26)$$

$$\begin{aligned} C_{ij} = \frac{\partial^2 C(\mathbf{w}, L)}{\partial w_i \partial w_j} &= C(\mathbf{w}, L) \left(\alpha_i \frac{1}{w_i} + \sum_j \beta_{ij} \frac{1}{w_i} \ln(w_j) \right) \left(\alpha_j \frac{1}{w_j} + \sum_i \beta_{ij} \frac{1}{w_j} \ln(w_i) \right) + C(\mathbf{w}, L) \beta_{ij} \frac{1}{w_i w_j}, \quad i \neq j \\ &= C(\mathbf{w}, L) \frac{1}{w_i} \left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) \frac{1}{w_j} \left(\alpha_j + \sum_i \beta_{ij} \ln(w_i) \right) + C(\mathbf{w}, L) \beta_{ij} \frac{1}{w_i w_j} \\ &= \frac{C(\mathbf{w}, L)}{w_i w_j} \left[\left(\alpha_i + \sum_j \beta_{ij} \ln(w_j) \right) \left(\alpha_j + \sum_i \beta_{ij} \ln(w_i) \right) + \beta_{ij} \right] \\ &= \frac{C(\mathbf{w}, L)}{w_i w_j} [S_i S_j + \beta_{ij}] \end{aligned} \quad (2.27)$$

2.D Imputation of censored wages

Since the wages in the BeH are only reported up to the social security threshold (2012: €67,200 per year in West Germany, €57,600 in East Germany), I have to estimate those wages which are above the censoring limit. I use additional information from the BeH, i.e., the age, sex, and nationality of the workers as well as requirement level of the job, the employment status, the establishment size, the industry (2-digit level), the occupation (3-digit level), and the type of region the employees are working in.

Following Reichelt (2015), I apply interval regression to estimate the wages above the threshold and add an error term to the estimates, i.e., a random value from a normal distribution. As the wage information is right censored, the interval regression is equivalent to tobit regression. Since the estimated wage has to be larger than the censoring limit, the random variable is drawn from a truncated distribution as suggested by Gartner (2005). Wages of certain groups like marginal employed, interns, and student trainees should not be affected by the social security threshold. They are therefore excluded from the interval regression, whose results are summarized in table 2.D.1. Figure 2.D.1 shows the wage distribution before and after imputation.

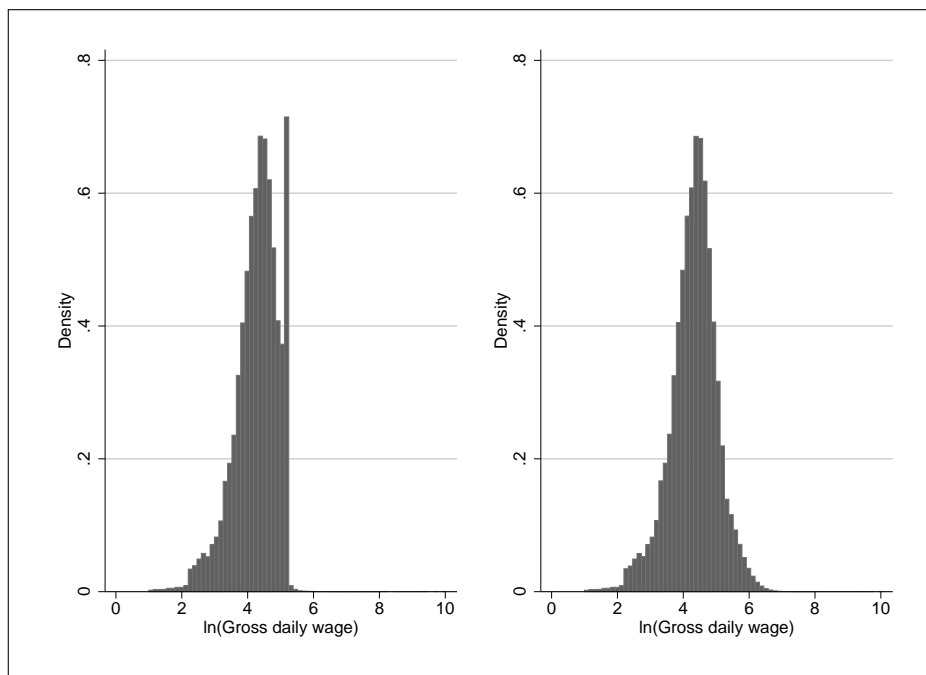


Figure 2.D.1: Wage distribution before (left figure) and after (right figure) imputation

Table 2.D.1: Results of an interval regression to partly impute the censored wages of employees liable to social security

Dependent variable: ln(Censored gross daily wage)		
Age	0.040***	(311.32)
Age ²	-0.000***	(-254.03)
Female	-0.189***	(-345.38)
German	0.074***	(96.50)
Requirement level according to KldB 2010 (see Paulus/Matthes, 2013), reference: Unskilled or semi-skilled activities		
Specialist activities	0.204***	(283.20)
Complex specialist activities	0.408***	(409.92)
Highly complex activities	0.609***	(553.25)
Parttime	-0.527***	(-967.73)
Employment status, reference: Employees liable to social security without special characteristics		
Employees in partial retirement	0.118***	(73.02)
Causal worker	0.466***	(94.22)
Employed pensioner (subject to social security contributions)	-0.349***	(-104.63)
Seaman	0.313***	(28.86)
Seaman in partial retirement	0.706***	(12.59)
Establishment size: ln(Number of employees)	0.076***	(609.69)
Type of region, reference: Agglomerations in West Germany		
Urbanized region in West Germany	-0.038***	(-75.40)
Rural region in West Germany	-0.049***	(-60.32)
Agglomeration in East Germany	-0.127***	(-166.33)
Urbanized region in East Germany	-0.211***	(-240.32)
Rural region in East Germany	-0.227***	(-202.68)
Constant	2.423***	(74.32)
Total number of observations	6,173,784	
Censored observations	525,152	
Uncensored observations	5,648,632	
Log-likelihood full model	-4.523e+06	
Log-likelihood intercept only	-6.870e+06	

t-statistics in parentheses, significance levels: * p<0.05, ** p<0.01, *** p<0.001. Beside the explanatory variables reported in this table industry fixed effects (2-digit level: 82 industries) and occupation fixed effects (3-digit level: 144 occupations) are considered. Marginal employed, interns, and student trainees are not considered as their wage should not be affected by the social security threshold.

2.E Further results

Table 2.E.1: Results based on a reduced sample: only establishments with less than 50 employees, average effect assuming $\alpha_{iq} = \alpha_i$, $\beta_{ijq} = \beta_{ij}$

i	j	α_i	d_{ij}	β_{ij}	c_{ij}	θ_{ij}	θ_{ji}
y	y	0.205*** (0.000)	0.505*** (0.002)	-0.092*** (0.002)	-2.599*** (0.042)	-0.524*** (0.016)	-0.524*** (0.003)
m	m	0.502*** (0.001)	0.543*** (0.001)	-0.093*** (0.002)	-0.730*** (0.007)	-0.367*** (0.002)	-0.367*** (0.002)
o	o			-0.169*** (0.002)	-1.339*** (0.006)	-0.396*** (0.002)	-0.396*** (0.002)
y	m		-0.220*** (0.002)	0.008*** (0.001)	0.854*** (0.007)	0.429*** (0.008)	0.172*** (0.001)
y	o			0.084*** (0.001)	0.322*** (0.007)	0.095*** (0.002)	0.065*** (0.001)
m	o			0.085*** (0.001)	0.658*** (0.004)	0.194*** (0.001)	0.331*** (0.002)
N		226290					

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the sample mean. Monotonicity and concavity of the cost function are satisfied at all data points. The concavity constraint is locally imposed at the sample mean. The elasticities are calculated at mean predicted shares.

Table 2.E.2: Elasticities by skill level based on reduced sample: only establishments with less than 50 employees

i	j	low skilled labor			medium skilled labor			high skilled labor		
		c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}	c_{ij}	θ_{ij}	θ_{ji}
y	y	-2.320*** (0.044)	-0.524*** (0.010)	-0.524*** (0.010)	-2.081*** (0.013)	-0.434*** (0.003)	-0.434*** (0.003)	-5.203*** (0.113)	-0.730*** (0.016)	-0.730*** (0.016)
m	m	-0.706*** (0.009)	-0.329*** (0.004)	-0.329*** (0.004)	-0.679*** (0.004)	-0.340*** (0.002)	-0.340*** (0.002)	-0.869*** (0.016)	-0.488*** (0.009)	-0.488*** (0.009)
o	o	-1.151*** (0.014)	-0.356*** (0.004)	-0.356*** (0.004)	-1.298*** (0.006)	-0.377*** (0.002)	-0.377*** (0.002)	-1.711*** (0.021)	-0.509*** (0.006)	-0.509*** (0.006)
y	m	0.767*** (0.017)	0.357*** (0.008)	0.173*** (0.004)	0.723*** (0.006)	0.362*** (0.003)	0.151*** (0.001)	1.427*** (0.038)	0.802*** (0.021)	0.200*** (0.005)
y	o	0.540*** (0.019)	0.167*** (0.006)	0.122*** (0.004)	0.248*** (0.006)	0.071*** (0.002)	0.052*** (0.001)	-0.243*** (0.034)	-0.072*** (0.010)	-0.034*** (0.005)
m	o	0.502*** (0.009)	0.155*** (0.003)	0.234*** (0.004)	0.651*** (0.004)	0.189*** (0.001)	0.326*** (0.002)	0.967*** (0.016)	0.288*** (0.005)	0.543*** (0.009)
N		45302			147292			33696		

Standard errors in parentheses, calculation based on Delta-method, significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The data is normalized at the respective sample mean. Monotonicity and concavity of the cost function are satisfied at more than 99 percent of the respective data points. The concavity constraint is locally imposed at the respective sample mean. The elasticities are calculated at mean predicted shares.

Chapter 3

Quantifying the effect of labor market size on learning externalities

This chapter has been published as Economics Working Paper No. 2017-06, Christian-Albrechts-Universität zu Kiel, Department of Economics, URL: <http://hdl.handle.net/10419/167623>.

Earlier versions have been published as Economics Working Paper No. 2016-11, Christian-Albrechts-Universität zu Kiel, Department of Economics, URL: <http://hdl.handle.net/10419/144741> and ERSA conference paper ersa16p615, European Regional Science Association, URL: http://www-sre.wu.ac.at/ersa/ersaconfs/ersa16/Paper615_JanCorneliusPeters.pdf, respectively.

3.1 Introduction

There is extensive empirical evidence that wages in large local labor markets are significantly higher than wages in smaller ones. Figure 3.1 illustrates disparities in regional wages for Germany. The difference in the regional wage level of the largest labor market region and the smallest amounts to more than 40 percent. Similar significant disparities are observed for other countries; see, e.g., Glaeser and Maré (2001) for the United States and Combes et al. (2008) for France. This raises the question, why firms in agglomerations¹ pay high wages and do not relocate to regions in which labor is cheaper.²



Figure 3.1: Local labor market size and regional wages.

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

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

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

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

The first explanation refers to the sorting of more able workers into cities, arguing that higher wages in cities reflect higher abilities of its workers. Regarding the importance of sorting, there still is an ongoing debate. While Combes et al. (2008) conclude ‘sorting matters’, De la Roca and Puga (2016) find that it is less important once dynamic agglomeration economies (see below) are considered. D’Costa and Overman (2014) in turn argue that sorting plays a role in explaining why wages in agglomerations grow faster.

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

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

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

The latter implies that the impact of a labor market's size on wage is expected to be lower the more time has passed since experience was acquired in that specific labor market.

Using administrative data from the Institute for Employment Research (IAB) on individual employment biographies, I am able to follow a worker's employment history back to 1975 and observe the size of all regional labor markets in Germany in which he or she acquired work experience. This information is used to estimate the extent to which wages depend on the size of the labor market where previous work experience was acquired. The wage information refers to individual new employment relationships in Germany starting between 2005 and 2011. These wages contain important information as to how firms value work experience when they hire a worker depending on the size of the labor market in which it was acquired. By considering the wages associated with new employment relationships, I reduce the risk of an omitted variable bias since other factors which are unobserved by the econometrician gain importance for productivity as tenure increases, like professional development training offered by the firm (Hamann et al., 2016).

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

This paper also takes into account that large local labor markets are typically not only characterized by a large workforce but also by an above average share of high-skilled labor. Presuming that those workers one learns from the most are the skilled ones, the question arises 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.

The remainder of the paper is structured as follows. In section 3.2 I review the corresponding literature with a focus on learning externalities. In section 3.3 I describe the methodology of my empirical analysis and in section 3.4 the data set. In section 3.5 I discuss the results of my empirical analysis and finally, in section 3.6 I set out my conclusions.

3.2 Related literature

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

In his paper 'Learning in Cities' Glaeser (1999: 255) state that "the primary informational role of cities may not be in creating cutting edge technologies, but rather in creating learning opportunities for everyday people".

He argues that agglomeration fosters individual learning opportunities since the latter increase with the rate of new contacts between individuals and this is highest in a dense urban environment.

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

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

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

In contrast, Wheeler (2006) considers both: between-job and within-job wage growth. He finds evidence in favor of the coordination hypothesis, as labor market density has a positive impact on between-job wage growth. However, this result is not robust when individual fixed effects are included in the analysis. With regard to within-job wage growth, Wheeler (2006) finds no significant impact of the local labor market, using three different measures of local market scale. D'Costa and Overman (2014), who control for unobserved individual characteristics that impact on wage growth, obtain a similar result for Great Britain, when they compare wage growth of workers who move from rural to urban areas with workers who stay in rural areas. However, they observe for rural workers who "worked in a city at some point" (D'Costa/Overman, 2014: 174) a faster wage

growth compared to rural workers with no city experience. Once city experience is taken into account, they find that workers who have either current or past experience in an urban environment benefit from faster wage growth than workers who have never had any urban experience. Furthermore, their results indicate that the wage growth of rural workers with past urban experience are based on within-job wage growth. D’Costa and Overman (2014) interpret this finding as evidence in favor of the learning hypothesis. Individuals with past urban experience are able to make use of their accumulated knowledge even after relocating to rural areas.

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

“Currently one of the most complete studies on the topic” (Combes/Gobillon, 2015: 263) is the one by De la Roca and Puga (2016).³ As with previous authors, they control for unobserved individual characteristics that impact on the individual wage level and distinguish static and dynamic effects. De la Roca and Puga (2016) provide empirical evidence that strongly suggest that dynamic benefits depend on learning effects. In order to identify dynamic agglomeration economies, they do not analyze individual wage growth after migrating from rural to urban areas like the studies discussed so far, but instead analyze the value of individual work experience, depending on the size of the city in which it was acquired. They distinguish experience that was acquired in three groups of Spanish cities. Their results show that experience acquired in the largest Spanish cities has the highest value. In addition, experience acquired in the next largest cities is also more valuable than experience acquired in the rest of the 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 evidence for the learning hypothesis as a driver of the urban wage growth premium. They find that the additional value of experience acquired in the biggest Spanish cities does not significantly change when a worker leaves the largest cities to work somewhere else.

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

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

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

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

The literature discussed so far focuses on externalities arising from the spatial concentration of economic activity and employment. However, a related strand of literature exists that discusses the role of local learning externalities as well (Heuermann et al., 2010). This strand analyzes the external effects of localized human capital. It focuses on the composition of the local workforce with respect to skills, rather than on labor market size. Comprehensive overviews are given by Moretti (2004b) and Duranton (2007).

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

Empirical evidence of productivity effects of local human capital is provided by, e.g., Moretti (2004c). He finds that the value added of manufacturing plants grows faster in those cities in the United States in which the share

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

of college graduates increases at a higher rate than in other cities. However, most frequently human capital externalities are analyzed by comparing wages of workers across cities with differing levels of human capital (Heuermann et al., 2010), among others Moretti (2004a) for the United States and Heuermann (2011) as well as Braakmann (2009) for Germany.

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

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

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

Taken together, the existing literature on dynamic agglomeration economies predominantly analyzes whether workers migrating from non-urban to urban areas experience thereafter a faster wage growth than workers who remain in non-urban areas (Glaeser/Maré, 2001; Yankow, 2006; Lehmer/Möller, 2010; D'Costa/Overman, 2014; Matano/Naticchioni, 2016). Even though the results are heterogeneous, the overall impression is that wage growth in urban areas is in fact stronger than in non-urban areas. In addition, there is also evidence that workers who migrate from urban to non-urban areas thereafter still benefit from their work experience acquired in the urban labor market. This is typically interpreted as support for the learning hypothesis. However, the magnitude of dynamic gains, depending on the size of the local labor market in which experience was acquired, has not been analyzed by these studies.

Inspired by the work of De la Roca and Puga (2016), a few studies recently estimated the value of work experience by cities. They distinguish two or three classes of large cities and compare the value of work experience acquired there to the value of work experience acquired in the rest of the country. In line with the learning hypothesis they find a positive relationship between the value of city-class experience and the size of the cities within the class. This indicates that a worker accumulates more human capital per unit of experience the larger the local labor market is in which experience is acquired. However, general conclusions with regard to the impact of city size on learning benefits are difficult as these studies distinguish only very few cities.

Furthermore, the underlying mechanisms remain unclear, i.e., which characteristics of urban labor markets are important for the emergence of learning externalities. Is it a scale effect, meaning the large number of workers who provide different ideas and knowledge or, as suggested by the literature on human capital externalities, the typically high share of high-skilled labor in urban labor markets?

Distinguishing experience acquired in 141 German labor market regions in the period 1975-2011, this paper explicitly estimates the impact of labor market size on dynamic agglomeration benefits taking into account the depreciation of human capital and the composition of a region's local labor force with respect to the skill level. The identified effects are supposed to be strongly related to learning externalities. I focus on wages associated with transitions to new employment relationships that indicate how firms value previous work experience when they hire a worker. Furthermore, I control for various other channels of agglomeration economies and show that the value of experience predominately is determined by the size of the labor market in which experience was acquired and portable to other regions.

3.3 Methodology

3.3.1 Empirical model

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

In my analysis, I make use of a micro-econometric framework described by Combes and Gobillon (2015). Suppose a worker i is hired by an establishment at date t and the logarithm of the wage which the worker receives, $w_{i,t}$, is given by equation (3.1); compare Combes and Gobillon (2015: eq. 5.20).

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times v_{r(i,\tau),\tau,r(i,t)} + \mathbf{x}'_{i,t} \boldsymbol{\beta} + \varepsilon_{i,t} \quad (3.1)$$

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

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

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

Imagine, I ignore that the value of experience depends on the time and the location of its acquirement and on where it is used, then, $v_{s,\tau,r} = v$ and (3.1) reduces to a frequently applied Mincer type wage equation:

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + vE_{i,t} + \mathbf{x}'_{i,t}\beta + \varepsilon_{i,t}, \quad (3.2)$$

where $E_{i,t}$ denotes the total amount of work experience which is given by the number of days of employment until date $t - 1$, i.e., $E_{i,t} = \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1)$. In both equations, (3.1) and (3.2), v refers to the marginal value of work experience. In (3.1) it is allowed to vary across labor markets, in (3.2) it is not.

In order to identify dynamic agglomeration economies based on (3.1), Combes and Gobillon (2015) suggest a two-step procedure. The first step requires to estimate $v_{s,\tau,r}$ for all combinations of s , t and r , in addition to all other parameters of equation (3.1). In a second step, the estimated parameters $\hat{v}_{s,\tau,r}$ are regressed on local characteristics such as labor market size to quantify its impact on the marginal value of experience. A severe problem of this approach is that the number of parameters which have to be estimated in the first step amounts to the square of the number of labor market regions times the number of considered dates (Combes/Gobillon, 2015). Moreover, it is not possible to take into account that the marginal return to experience acquired in labor market s at date τ may vary across individuals depending on the amount of previously acquired experience (see below).

De la Roca and Puga (2016) estimate a special case of equation (3.1) where they impose restrictions on $v_{s,\tau,r}$ such that the number of parameters reduces. Using Spanish data, they distinguish experience only according to three classes of cities in which it was acquired: in Madrid/Barcelona, Valencia/Sevilla/Zaragoza, or elsewhere in Spain. Furthermore, they assume that the value of experience acquired in one of the classes does not depend on the day when it was acquired. In this case, equation (3.1) simplifies to (3.3); compare De la Roca and Puga (2016: eq. 1).

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + \sum_{s=1}^3 v_{s,r(i,t)} E_{i,s,t} + \mathbf{x}'_{i,t}\beta + \varepsilon_{i,t} \quad (3.3)$$

$E_{i,s,t}$ denotes the amount of experience that was acquired until $t - 1$ in, respectively, Madrid or Barcelona, Valencia, Sevilla, or Zaragoza, or elsewhere in Spain. This approach enables to compare the marginal value of experience acquired in different groups of cities to assess the magnitude of dynamic agglomeration benefits.

⁶ Because information on self-employment is not available, I only consider spells of employment subject to social security contribution when measuring work experience. In this specification, I assume that human capital which was accumulated while working does not depreciate over time; compare, e.g., De la Roca and Puga (2016), Matano and Naticchioni (2016). This assumption is relaxed below.

However, to obtain general results with regard to the impact of labor market size on dynamic gains is difficult. The two-step procedure suggested by Combes and Gobillon (2015) requires in the first step not only few parameters $v_{s,\tau,r}$ to be estimated, but many.

I adopt the strategy by De la Roca and Puga (2016) as a point of departure and estimate a model such as (3.3) where I distinguish only between two groups of German labor markets. Thereafter, I apply an alternative estimation strategy that is also discussed by Combes and Gobillon (2015) in order to obtain more general results. I express $v_{s,\tau,r}$ in (3.1) by a function of labor market size, equivalent to the second step of the two-step procedure and estimate to which extent wages associated with new employment relationships depend on the size of the labor market in which experience was acquired. At this point, I assume that the value of experience is only determined by the size of the labor market where experience was acquired and does not vary depending on where it is used, i.e., $v_{s,\tau,r} = v_{s,\tau}$. This assumption is relaxed at the end of my analysis. Furthermore, I follow Combes et al. (2008), Andersson et al. (2014), De la Roca and Puga (2016), and others, and assume that there is a log-linear relationship between net-agglomeration gains and labor market size denoted by $Z_{s,\tau}$. More precisely, I assume that $v_{s,\tau}$ is given by equation (3.4). It may be interpreted as a learning function capturing externalities of acquiring work experience in large local labor markets.⁷

$$v_{s,\tau} = \begin{cases} \delta \ln \left(\frac{Z_{s,\tau}}{\underline{Z}} \right) & \text{if } Z_{s,\tau} > \underline{Z} \\ 0 & \text{else} \end{cases} \quad (3.4)$$

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

$$v_{s,\tau} = \begin{cases} \gamma + \delta \ln Z_{s,\tau} & \text{if } Z_{s,\tau} > \underline{Z} \\ 0 & \text{else,} \end{cases} \quad (3.5)$$

with $\gamma = -\delta \ln \underline{Z}$. Given that $Z_{s,\tau}$ is always larger than \underline{Z} , inserting (3.5) into (3.1) leads to:

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + \gamma E_{i,t} + \delta \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau} + \mathbf{x}'_{i,t} \beta + \varepsilon_{i,t}, \quad (3.6)$$

with $z_{s,\tau} = \ln Z_{s,\tau}$ and $E_{i,t}$ being individual work experience measured by the number of days a worker was employed before date t (compare equation (3.2)). After estimating equation (3.6) it has to be verified whether

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

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

all local labor markets are indeed larger than \hat{Z} which is given by $\exp(-\hat{\gamma}/\hat{\delta})$. If some local labor markets are smaller than \hat{Z} , the size of the respective labor market has to be set to \hat{Z} (leaving $E_{i,t}$ unchanged) and an iterative procedure has to be applied in order to obtain the solution for equation (3.4).

$\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau}$ in equation (3.6) may be interpreted as the total number of local learning opportunities of individual i until date $t - 1$ depending on the size of the labor markets in which experience was acquired. The corresponding parameter δ denotes the elasticity of wage with respect to the size of the labor market in which individual i was working at one day in the past. Accordingly, the productivity at date t as reflected in wage would be $\delta E_{i,t}$ percent higher if all labor markets individual i was working in before would have been on average one percent larger. For interpretation it is important to note that the partial derivative of $w_{i,t}$ with respect to $v_{s,\tau}$ is unity, given that individual i was employed at date τ (see equation 3.1). Therefore, the estimated *elasticity of wage* with respect to the size of the labor market in which experience was acquired can also be interpreted as the *elasticity of the marginal value of experience* with respect to labor market size (compare equation 3.4).

So far, I implicitly assumed that human capital does not depreciate over time. However, there are good reasons to believe that its value decreases the more time passes. This might be caused by, i.a., changes in the skills demanded for a particular job due to technological change, shifts in the demand for particular occupations due to changes in the industry structure, or since certain knowledge and skills get lost because they are insufficiently used (De Grip/Van Loo, 2002). Taking this into account, I augment the learning function such that the value of work experience acquired by working at date τ , $I(O_{i,\tau} = 1) \times v_{s,\tau,r}$ in equation (3.1), is given by equation (3.7).⁹ θ denotes the rate by which previously acquired human capital depreciates per day and $Q_{i,\tau}$ is the value of the human capital, as given by equation (3.8), that individual i accumulated before date τ evaluated at the beginning of τ .¹⁰ Hence, $-\theta Q_{i,\tau}$ denotes the daily loss in the value of previously accumulated human capital. It implies that the marginal change in the value of the individual stock of human capital at date τ is positive if and only if individual i acquires a sufficient amount of new human capital at date τ by working in a sufficiently large labor market.

$$I(O_{i,\tau} = 1) \times v_{s,\tau,i} = \begin{cases} \delta \ln\left(\frac{Z_{s,\tau}}{\hat{Z}}\right) - \theta Q_{i,\tau} & \text{if } Z_{s,\tau} > \hat{Z} \text{ and } O_{i,\tau} = 1 \\ -\theta Q_{i,\tau} & \text{else} \end{cases} \quad (3.7)$$

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

$$Q_{i,\tau} = \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times v_{r(i,k),k,i} \quad (3.8)$$

for $\tau > 1$.

⁹ In equation (3.7), I still assume that the value of previously acquired work experience is the same in all labor markets.

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

Since $I(O_{i,\tau} = 1) \times v_{s,\tau,i}$ is defined as $Q_{i,\tau+1} - Q_{i,\tau}$, equation (3.7) corresponds to the first order difference equation given by (3.9).¹¹ Its solution is given by equation (3.10). For details see appendix 3.A.

$$Q_{i,\tau+1} = \begin{cases} \delta \ln\left(\frac{Z_{s,\tau}}{\underline{Z}}\right) + (1 - \theta)Q_{i,\tau} & \text{if } Z_{s,\tau} > \underline{Z} \text{ and } O_{i,\tau} = 1 \\ (1 - \theta)Q_{i,\tau} & \text{else} \end{cases} \quad (3.9)$$

$$Q_{i,\tau} = \delta \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times \max\left\{\ln\left(\frac{Z_{r(i,k),k}}{\underline{Z}}\right); 0\right\} \times (1 - \theta)^{\tau-k-1} \quad (3.10)$$

Hence, the value of human capital acquired before date τ , $Q_{i,\tau}$, is proportional to the weighted sum of the size of all labor markets in which previous work experience was acquired. For $k = \tau - 1$, the weight $(1 - \theta)^{\tau-k-1}$ equals one, meaning that the size of the labor market in which a worker acquired work experience at the previous day is not discounted. In contrast, the size of the labor market in which the first day of experience was acquired has the lowest weight, $(1 - \theta)^{\tau-2}$, addressing that the value of the human capital acquired at this day declined each day after acquisition by rate θ .

In order to obtain estimates for θ , δ , and \underline{Z} based on the augmented learning function given by equation (3.7), I again replace $\ln \underline{Z}$ by $-\gamma/\delta$ such that equation (3.10) can be written as:

$$Q_{i,\tau} = \gamma \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times (1 - \theta)^{\tau-k-1} + \delta \sum_{k=1}^{\tau-1} I(O_{i,k} = 1) \times \max\{\ln Z_{r(i,k),k}; \ln \underline{Z}\} \times (1 - \theta)^{\tau-k-1}. \quad (3.11)$$

Replacing the term $\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times v_{r(i,\tau),\tau,r(i,t)}$ in equation (3.1) by this expression yields equation (3.12).¹² Recall that both, $\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times v_{r(i,\tau),\tau,r(i,t)}$ as well as $Q_{i,t}$, denote the value of experience acquired before date t .

$$w_{i,t} = u_i + \mu_{r(i,t),y(t)} + \gamma \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times (1 - \theta)^{t-\tau-1} + \delta \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau} \times (1 - \theta)^{t-\tau-1} + \mathbf{x}'_{i,t} \beta + \varepsilon_{i,t} \quad (3.12)$$

According to equation (3.12) the productivity of worker i at date t , as reflected in wage $w_{i,t}$, depends, like in equation (3.6), on the size of all labor markets in which i worked previously. However, (3.12) takes into account that having recently benefited from acquiring experience in a large labor market may increase today's productivity stronger than agglomeration economies experienced years ago. The crucial parameter is θ . The larger θ is, the relatively larger is the productivity effect of the size of the labor markets in which recent experience

¹¹ I gratefully thank Johannes Bröcker for pointing this out and to give me the decisive hint how to obtain the solution of (3.9) referring to Sydsæter et al.

¹² Again, I assume that all labor markets are larger than threshold \underline{Z} . This assumption has to be verified after an estimate for \underline{Z} is obtained.

was acquired. Since equation (3.12) is highly non-linear in θ , maximum likelihood estimation is applied. For $\theta = 0$, equation (3.12) simplifies to the baseline model given by equation (3.6).

After estimating equations (3.6) and (3.12) which are based on the assumption that the value of experience is the same in all labor markets, I let the parameters of the learning function vary across different groups of local labor markets which are defined according to their size. This provides evidence as to whether the value of experience depends only on the size of the labor market in which experience was acquired or also on the size of the labor market in which it is used.

3.3.2 Econometric issues

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

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

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

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

3.4 Data

3.4.1 Individual data set

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

The information on labor market biographies is taken from the Integrated Employment Biographies (IEB) of the Institute for Employment Research (IAB). Among other information, the IEB contains very reliable micro data on employment which comes from the integrated notification procedure for health, pension and unemployment insurance.¹⁴ My data set comprises a 5 percent random sample of all employees with at least one notification to social security between 2005 and 2011.¹⁵

I exclude all individuals for which it is not possible to observe the full employment biography. The setup of my analysis requires information on all locations where previous work experience was acquired.¹⁶ Among others, I exclude all individuals born before 1960 because the IEB contains only information on employment from 1975 onwards. A detailed description is provided in appendix 3.B. For the remaining sample of workers my data set captures all spells of employment subject to social security contributions. I use them to identify transitions to employment within the period between 2005 and 2011 focusing on new full-time employment outside the public sector and outside the temporary employment industry. For the new employment relationships I observe the corresponding gross daily wage as well as further information on the new job (e.g., kind of occupation) and the worker (e.g., age, educational attainment and sex). The wage information in the IEB is right censored, as firms report earnings only up to the upper limit for social security contributions. Therefore, I partly impute the wages, see appendix 3.B.

Using information on all previous spells of employment, I compute the individual labor market experience at the date at which the new employment relationship starts. Moreover, I generate important control variables that provide information on the recent labor market biography with respect to the pre-employment status, length of non-employment before the transition to employment, unemployment benefits, and participation in measures of active labor market policy. The information is also taken from the IEB. Detailed information on all variables that I use is provided by table 3.B.1 in appendix 3.B. Summary statistics can be found in table 3.C.1 in appendix 3.C.

Descriptive statistics indicate that a large share of the considered new employment relationships refers to rather young workers with few years of labor market experience. One likely explanation is that workers change jobs

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

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

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

more frequently when they are young to find the job they like most as noted by Wheeler (2008). The mean work experience in my data set amounts to about 9.4 years. The analyses by Glaeser and Maré (2001), Wheeler (2006), Yankow (2006) as well as Gould (2007) are based on a data set with a similar sample mean.¹⁷

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

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

3.4.2 The size of local labor markets

Di Addario and Patacchini (2008) show for Italy that the effect of local density on wages rapidly declines in distance. Only for population mass within a distance up to 12 kilometers, they find a significant impact on wages. It indicates that agglomeration benefits depend on the immediate environment.

Taking this into account and addressing that learning externalities are thought to crucially depend on interaction between individuals, I choose labor market regions as spatial units for my analysis. They are supposed to represent integrated local labor markets and have on average a radius of about 27 km (see table 3.1). Their definition is taken from Kosfeld and Werner (2012), who define 141 regional labor market regions employing a factor analysis to German commuter structure between NUTS-3 regions. Because the regions contain NUTS-3 regions connected through intense commuting, I assume that workers exchange knowledge exclusively within these regions and that their size in terms of employment is an appropriate measure for local learning opportunities.

One critical issue is that commuting patterns change of over time, wherefore the chosen definition of labor market regions might be an inappropriate approximation of local labor markets decades ago. However, when estimating equation (3.6) which is based on the simplified learning function, the within transformation that is applied to eliminate the individual fixed effect wipes out all information on experience that was acquired before the date at which the first analyzed new employment relationship of worker i started. Hence, in the baseline specification $\hat{\delta}$ and $\hat{\gamma}$ are only based on information that refers to the period 2005 to 2011. Estimating equation (3.12) which is based on the augmented learning function, requires information on all previous employment spells of a worker. In this case, the size of the labor market in which experience was acquired is

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

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

weighted lower the more additional experience was acquired thereafter; see also equation (3.10). Nevertheless, as a robustness check, I repeat my estimations focusing on those workers who only acquired work experience in 1995 or later, i.e., at least five years after reunification of East and West Germany. Overall, the results based on the full sample are confirmed as discussed in the next section.

Table 3.1: Size of German labor market regions

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

N=141.

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

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

Local learning externalities should crucially depend on the amount of localized knowledge and ideas. I approximate this amount by the number of employees working in a labor market. The data is taken from the Employment History of the IAB and refers to annual regional employment at June 30th. The size of the labor market regions in terms of employment varies between 14 thousand and 1,2 million employees (region specific averages, see table 3.1).

After identifying the effect of total regional employment on the value of experience as reflected in higher future wages, I analyze to which extent the dynamic agglomeration benefit may be explained by an above average share of high-skilled labor within large urbanized labor markets. I define this as the local share of workers with a university degree which varies between two percent and 20 percent (see table 3.1). The correlation between total regional employment and the local share of high-skilled workers is 0.38.

3.4.3 The mobility of labor

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

Descriptive statistics show that a worker acquired work experience in either small or large local labor markets. Mobility between labor markets that significantly differ in size is rather low. Table 3.2 provides information on the ratio of the size of the largest labor market and the size of the smallest labor market in which a worker

acquired experience before one of the considered new employment relationships. In more than 60 percent of the observations the largest labor market was less than twice as large as the smallest labor market, and in less than ten percent more than ten times as large.¹⁹

Table 3.2: The size of the largest labor market in which a worker acquired experience relative to size of the smallest labor market

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

N=525,265.

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

Table 3.3 compares the size of the labor markets where experience was acquired and where it is used distinguishing five categories of labor market regions according to their size. In 61 percent of all considered new employment relationships the new employer is located in a region that belongs to the same category as the average region in which experience was acquired (main diagonal), in 25 percent the current region is larger than previous ones (upper triangle) and in 14 percent smaller (lower triangle).

Workers are not only rather immobile between local labor markets that belong to different categories as defined in table 3.3, but also between labor markets that are of a comparable size. The sample mean of total work experience is 9.4 years (see table 3.C.1). On average, 59 percent of this previous work experience (5.7 years) refer to employment in the same labor market in which the new employer is located. The median amounts to 79 percent. 36 percent of the workers acquired all their work experience in that particular region. On the opposite, 25 percent of the workers were never previously employed in the region in which the new employer is located.

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

In order to test the hypothesis that the value of work experience depends not only on the size of the labor market in which experience was acquired, but also on the size of the labor market in which it is used, it requires that at

¹⁹ Recall that the largest labor market is more than ten times as large as the median labor market (see table 3.1).

least some workers start to work in a region that is of a different size to the average region in which individual experience was acquired. Despite the rather low degree of mobility, there presumably is sufficient variation to test this hypothesis. Corresponding results are reported at the end of the next section. Before this, the next section discusses the results for the control variables and the impact of the size of the labor market in which experience was acquired on experience's value and future productivity as reflected in wages.

Table 3.3: The size of labor market in which experience was acquired and in which it is used

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

N=525,265.

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

3.5 Results

3.5.1 Control variables

Table 3.4 summarizes the regression results for the control variables that I use in my analysis. Column 1 refers to ordinary least squares estimation (OLS). Since the available information on educational degrees are only imperfect measures of skills, OLS results are likely biased. To address that workers might sort on unobserved abilities into large labor markets, I include individual fixed effects (FE, specification (2)) as introduced by Glaeser and Maré (2001). This leads to a reduction of my sample, since 188,979 observations refer to workers, for whom I only observe one new employment relationship between 2005 and 2011. The fixed effects estimation is therefore based on 336,286 observations for 126,862 employees.

The comparison of OLS and FE results shows that the OLS estimates are in most cases biased upwards. However, the sign of the estimated coefficients is in both models almost always the same and as expected. The larger the highest educational degree of a worker, the larger is the wage at the beginning of the considered new employment relationship. For example, workers with a university degree receive a 26 percent ($e^{0.233} - 1$) higher wage than workers with a secondary/intermediate school leaving certificate and a completed vocational training.

Since the wage rate that I use as a dependent variable refers to the average wage rate that is paid until December 31 of the year in which the employment relationship starts (see appendix 3.B), I include the length of the

considered spell of employment, measured in months. It confirms that tenure affects the wage rate positively, the monthly increase amounts to 0.8 percent.

Table 3.4: Results for control variables

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

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include industry, occupation, and region-year fixed effects as well as total work experience and the pivotal variable $\sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times z_{r(i,\tau),\tau}$, see equation (3.6) and column (1) and (2) of table 3.6.

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

With respect to the pre-employment status of a worker, the results show that the longer the period of non-employment before the considered new spell of employment, the lower is the corresponding wage. Non-employment of more than one year results in a wage loss of about four percent. Following Mincer and Ofek (1982), a reasonable explanation is that non-employment accelerates the depreciation of human capital. Since I assume in the learning function given by equation (3.7) that the work experience depreciates at a constant rate irrespective of a worker's employment status, it is worth noting that Mincer and Ofek (1982) also provide evidence that non-employment only has a temporary negative effect on individual's human capital, meaning that eroded human capital is restored after an individual returns to work.

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

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

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

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

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

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

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

Table 3.5 summarizes results that I obtain adopting the empirical strategy suggested by De la Roca and Puga (2016) to analyze dynamic agglomeration benefits that is described by equation (3.3). In addition to the vari-

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

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

ables reported in table 3.5 and the control variables reported in table 3.4, all models contain industry, occupation, and region-year fixed effects. As with De la Roca and Puga (2016) for Spain and Matano and Naticchioni (2016) for Italy, I focus on the benefit of acquiring experience in Germany's largest local labor markets: Berlin, Hamburg and Munich.²² I also include the amount of overall experience in the analysis so that the estimated effect of experience acquired in Berlin, Hamburg or Munich refers to the difference in the value of experience acquired there and experience acquired elsewhere in Germany.

Table 3.5: The value of labor market specific work experience

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

[†] Coefficient and standard error multiplied by 10.

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. Work experience measured on a daily bases and expressed in years. All models include control variables as well as year, industry, occupation, and region fixed effects (see table 3.4).

Similar to De la Roca and Puga (2016) as well as to Matano and Naticchioni (2016), I find that one year of experience acquired in the largest German local labor markets has a significantly higher value than experience acquired in the rest of the country. The inclusion of individual fixed effects in specification (2) let the value of experience increase, indicating a negative correlation of unobserved abilities and work experience.²³ The magnitude of the agglomeration benefit is discussed below. In order to test whether the value of experience depends on where it is used, I consider similar to De la Roca and Puga (2016) interaction effects between the experience variables and an indicator for moving to the respective other group of labor markets. The corresponding results of the fixed effects model in column (2) suggest that the value of experience acquired in

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

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

the largest labor markets does not change when transferring it to smaller labor markets. The value of experience acquired outside the large labor markets increases slightly if it is used in Berlin, Hamburg or Munich.

Following De la Roca and Puga (2016), specification (3) additionally contains the square of experience to let the marginal value of experience decay with more experience. Furthermore, interaction effects of experience acquired in the largest labor markets and overall experience are included. They allow for heterogeneous effects for less and more experienced workers (De la Roca/Puga, 2016). Qualitatively, I obtain the same results as De la Roca and Puga (2016): (i) Experience acquired in the largest local labor markets has a significantly higher value than experience acquired elsewhere. (ii) The marginal gain of working in one of the largest labor markets is higher for individuals with less work experience than for more experienced workers. This is particularly true for workers who previously worked elsewhere in the country. A reasonable explanation is that, *ceteris paribus*, the latter workers so far accumulated less human capital than individuals who worked in the largest labor markets. (iii) The value of experience acquired in the largest labor markets is highly portable to smaller labor markets which strongly supports the hypothesis that learning externalities play an important role. (iv) Experience acquired in the rest of the country has a higher value if it is used in the largest local labor markets than in the rest of the country. However, the location where experience was acquired has a stronger impact on its value than the location where it is used. The first year of experience acquired in Berlin, Hamburg or Munich increases wages by about 1.3 percent ($e^{0.01338-0.00047} - 1$) relative to having worked elsewhere and independently of the new job location. In comparison, the value of the first year of experience acquired outside the largest local labor markets increases by less than 0.7 percent if the worker moves to Berlin, Hamburg or Munich. Qualitatively, the results are the same as obtained by De la Roca and Puga (2016) for Spain. Quantitatively, the identified agglomeration benefit for the largest German agglomerations is somewhat smaller than the dynamic agglomeration gains obtained by De la Roca and Puga (2016) for Madrid and Barcelona and by Matano and Naticchioni (2016) for Rome and Milan. Their results indicate that the value of the first year work experience acquired in the largest cities of the considered country has, respectively, a three percent and two percent higher value than the first year acquired elsewhere.²⁴

The results reported in table 3.5 and by other scholars raise the question of how the value of experience exactly depends on the size of the local labor market in which it was acquired. To provide corresponding insights, I estimate the parameters of equations (3.6) and (3.12). The estimates of the pivotal parameters are summarized in table 3.6. Again, I observe that the value of experience is biased downwards if I do not control for unobserved characteristics by means of individual fixed effects.

Column (1) and (2) of table 3.6 contain results for equation (3.6) referring to the simplified learning function given by equation (3.4). It is assumed that the marginal value of experience does not depend on the amount of previous work experience and already accumulated human capital. The findings support the hypothesis that labor market size fosters the individual accumulation of knowledge. Labor market size positively impacts on the value of the acquired experience which reflects in higher future wages. According to the specification with

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

individual fixed effects, the elasticity of wage with respect to the size of the labor market in which experience was acquired at one day in the past, denoted by δ , amounts to 0.116×10^{-4} . This effect may also be interpreted as the elasticity of the marginal value of work experience with respect to labor market size (see equation (3.4)).

Table 3.6: Estimates of the parameters of the learning function

	(1)	(2)	(3)	(4)
$\hat{\delta}^\dagger$	0.059*** (0.002)	0.116*** (0.016)	0.222*** (0.020)	0.236*** (0.026)
\hat{z}	140.101*** (33.936)	97.489 (111.521)	25.506 (22.027)	90.753 (80.944)
θ^\dagger			3.452	4.177
New employment relationships	525,265	336,286	336,286	214,319
OLS: R ² , FE: within R ²	0.616	0.183	0.197	0.261
Worker fixed effects	No	Yes	Yes	Yes

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

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. \hat{z} calculated based on delta method and $\hat{\delta}$ and $\hat{\gamma}$ (not reported). Column (1) and (2) refer to the baseline model given by equation (3.6) which is based on the simplified learning function (equation (3.4)). Column (3) and (4) report results for equation (3.12). It is based on the augmented learning function (equation (3.7)). The results in column (3) are obtained using the full sample. The specification reported in column (4) is restricted to workers who acquired work experience only 1995 or later. All specifications include control variables as well as industry, occupation, and region-year fixed effects (see table 3.4).

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 (3.4)). 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 human capital does not depreciate over time. However, this assumption is too restrictive as indicated by the results reported in column (3) of table 3.6. They refer to equation (3.12) which is based on the augmented learning function given by equation (3.7). The significant differences in the parameter estimates between column (2) and (3) as well as the notable increase in the within R² confirm that it is important to address that a worker's human capital acquired at some point in the past has a lower value the more time passed. This effect is captured in the augmented learning function by θ . Equation (3.12) which I estimate in order to determine the parameters of the augmented learning function is highly non-linear in this parameter. Therefore, maximum likelihood estimation is applied in order to obtain an estimate for θ . Once a value is imposed for θ , a linear fixed-effects estimator can be applied to estimate the other parameters. Column (3) contains the results that I obtain when θ is set equal to 3.452×10^{-4} which maximizes the fit of the regression model.²⁵ A θ of 3.452×10^{-4} means that the human capital which was acquired at date $t - 365$ is weighted by factor $0.882 (= (1 - 0.0003452)^{364})$. The weight is smaller than

²⁵ The results that are obtained for other values of θ are reported in table 3.C.2. Starting with $\theta = 1 \times 10^{-4}$ (not reported in table 3.C.2), I gradually increased / decreased θ by 10 percent. I also estimated a specification with θ close to zero. The result corresponds to the result reported in column (2) of table 3.6 (compare with specification 22 in table 3.C.2).

unity since the knowledge that was acquired at date $t - 365$ lost in value since the day of acquirement. In contrast, the weight of the knowledge that was acquired at day $t - 1$ is unity.

With respect to δ , I now obtain an elasticity of 0.222×10^{-4} , and \hat{Z} is still significantly smaller than the smallest German local labor market. Both estimates are confirmed by the results reported in column (4). The latter are based on a reduced sample where I restrict the analysis to individuals who acquired experience not before 1995. This robustness check takes into account that the chosen definition of labor market regions might be an inappropriate approximation for the shape of local labor markets decades ago which would result in a measurement error in the pivotal explanatory variable, i.e., the size of the labor markets in which experience was acquired. However, it does not seem to be a severe problem here as the results reported in column (3) and (4) are very similar, although a different value of θ minimizes the residual sum of squares. Using the reduced sample, I obtain the smallest residual sum of squares if I set θ equal to 4.177×10^{-4} (see table 3.C.3), meaning that human capital acquired, e.g., at date $t - 365$ is weighted by 0.859. Furthermore, the reduction of the sample let the fit of the model increase. Now it explains more than one quarter of wage's within variation.

The nature of dynamic agglomeration economies in general, and learning externalities in particular, is that benefits of working in large labor markets accumulate over time. Based on the regression results reported in columns (3) and (4) of table 3.6, figure 3.2 illustrates for different levels of experience the elasticity of wage with respect to the (cumulated) size of all the labor markets in which previous experience was acquired. The elasticity is increasing in the level of experience since the benefit of a one percent increase in the size of *all* labor markets in which experience was acquired is larger the higher the level of experience is.²⁶ Consider for example a worker with two years of work experience at date t . The corresponding elasticity is given by about 0.01, indicating that doubling the size of all labor markets in which the two years of experience were acquired results in a 0.7 percent higher wage at date t . For a worker with 10 years of work experience the benefit of having acquired all his or her experience in a one percent larger labor markets is larger. At this level of experience the elasticity amounts to about 0.045. Hence, doubling the size of all labor markets in which ten years of experience were acquired results in a productivity increase at date t by about three percent. At a very high level of experience the elasticity is even larger. The depreciation of accumulated human capital implies that the dynamic agglomeration gain converges towards an upper bound. If a worker has a sufficiently large amount of work experience, the size of the labor market in which the first years of experience were acquired has only a negligible impact on today's productivity and wage since the human capital acquired at the beginning of the individual working life is (almost) fully depreciated.

For levels of experience up to ten years, the estimated elasticities based on the reduced sample and the full sample are virtually the same. For higher levels of experience they slightly differ. At a level of 30 years of experience the gap between the point estimates amounts to about 0.006. However, the reduced sample only contains individuals with work experience of at most 16 years. Therefore, it is not clear-cut which specification leads to the more reliable estimate of the elasticity at a high level of experience. Nevertheless, the results

²⁶In general, the elasticity of the value of E days of experience with respect to the size of the labor markets in which the experience was acquired is, according to the results reported in column (3) of table 3.6, given by $0.222 \times 10^{-4} \times \sum_{k=1}^E (1 - 3.452 \times 10^{-4})^{k-1}$ (see also equation (3.10)).

are still very similar and their overall interpretation is clear: Dynamic agglomeration benefits have a highly significant impact on individual productivity. Since workers are rather immobile between labor markets of different size (see table 3.3), the dynamic gains should explain a significant part of regional wage disparities between urban and non-urban labor markets. Moreover, the identified agglomeration effects should be strongly related to learning externalities as I control for the number of job changes in the past which is supposed to be a proxy for dynamic matching.

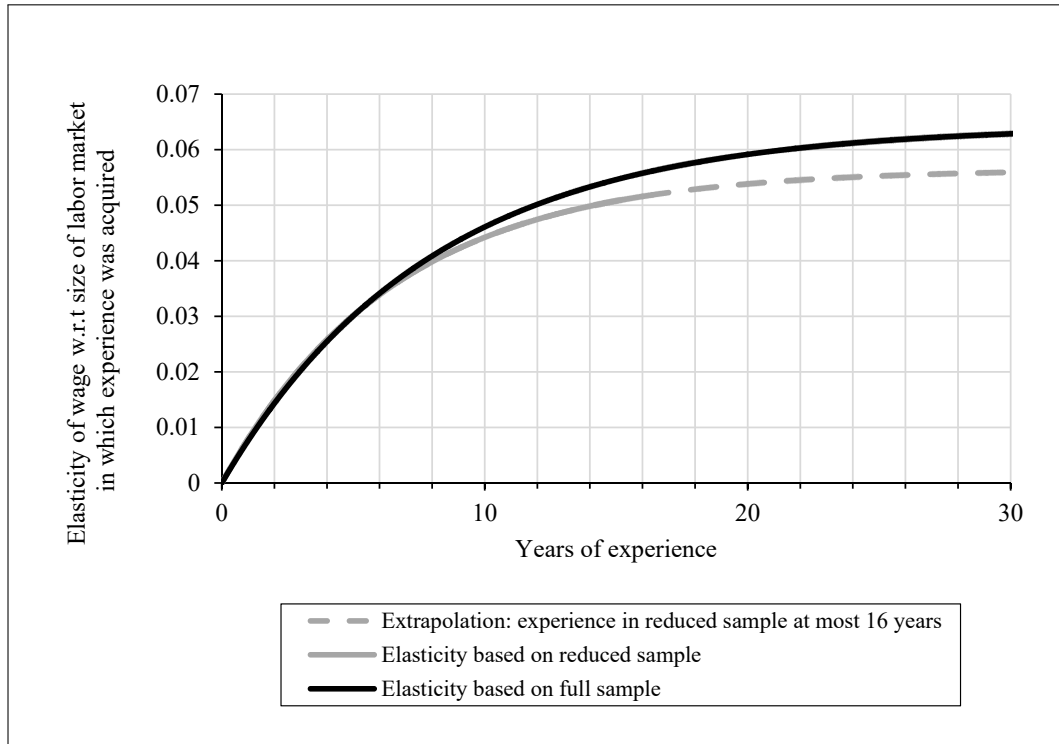


Figure 3.2: Magnitude of dynamic agglomeration benefits

Note: The figure illustrates the regression results reported in column (3) and (4) of table 3.6. For different levels of experience the graph denotes the percentage increase in productivity, given that the local workforce would have been one percent larger in all labor markets in which previous work experience was acquired. The reduced sample contains only workers who acquired experience in 1995 or later.

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

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

Typically, large urbanized labor markets are not only characterized by a large number of workers but also by an above average share of high-skilled labor. In accordance with the literature on human capital externalities, presumably the especially skilled workers are the workers one learns from the most. Hence, an important question is to which extent dynamic agglomeration benefits can be explained by the higher share or larger

number of high-skilled workers. In order to provide corresponding insights, I consider $Z_{s,\tau}$, the labor market size in the learning function (see equation (3.4) and (3.7)), in the following to be an *aggregate* of total regional employment, $\text{emp}_{s,\tau}$, and the local share of high-skilled workers, $\frac{\text{hskill}_{s,\tau}}{\text{emp}_{s,\tau}}$, as given by equation (3.13). So far, $Z_{s,\tau}$ was only measured in terms of total regional employment. The regression model that corresponds to equation (3.13) is given by equation (3.14) where $\tilde{\gamma}$ refers not only to the (unknown) minimum number of workers but also to an (unknown) minimum share of high-skilled labor beyond which learning externalities arise, denoted by $\underline{\text{emp}}$ and $\frac{\text{hskill}}{\underline{\text{emp}}}$, respectively: $\tilde{\gamma} = -\rho_{\text{emp}} \ln(\underline{\text{emp}}) - \rho_{\text{hskill}} \ln\left(\frac{\text{hskill}}{\underline{\text{emp}}}\right)$.²⁷

$$\delta \ln(Z_{s,\tau}) = \rho_{\text{emp}} \ln(\text{emp}_{s,\tau}) + \rho_{\text{hskill}} \ln\left(\frac{\text{hskill}_{s,\tau}}{\text{emp}_{s,\tau}}\right) \quad (3.13)$$

In contrast to δ which denotes the total dynamic agglomeration gain, ρ_{emp} refers only to the isolated benefit of acquiring experience in a labor market with a large workforce in terms of total regional employment. It does not capture the benefit of acquiring experience in a labor market with an above average share of high-skilled labor. This latter benefit is captured by ρ_{hskill} .

$$\begin{aligned} w_{i,t} = & u_i + \mu_{r(i,t),y(t)} + \tilde{\gamma} \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times (1 - \theta)^{t-\tau-1} + \\ & + \rho_{\text{emp}} \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times \ln(\text{emp}_{r(i,\tau),\tau}) \times (1 - \theta)^{t-\tau-1} + \\ & + \rho_{\text{hskill}} \sum_{\tau=1}^{t-1} I(O_{i,\tau} = 1) \times \ln\left(\frac{\text{hskill}_{r(i,\tau),\tau}}{\text{emp}_{r(i,\tau),\tau}}\right) \times (1 - \theta)^{t-\tau-1} + \mathbf{x}'_{i,t} \beta + \varepsilon_{i,t} \end{aligned} \quad (3.14)$$

The definition of labor market size given by equation (3.13) enables not only testing whether the size of a local labor market in terms of total regional employment positively impacts on the individual accumulation of human capital holding the composition of the local labor force with regard to skills constant and whether local high-skilled labor is more important for the individual accumulation of human capital than lower skilled labor. These hypotheses are confirmed if $\hat{\rho}_{\text{emp}}$ and $\hat{\rho}_{\text{hskill}}$, respectively, are significantly larger than zero. But in addition, it also enables testing whether it is solely the absolute amount of high-skilled labor in a local labor market that determines the amount of human capital which is acquired by working in the region. Equation (3.13) and, analogously, equation (3.14) can be rewritten such that total regional employment and the absolute number of high-skilled labor rather than the share of the latter are simultaneously considered as measures of labor market size:

$$\delta \ln(Z_{s,\tau}) = (\rho_{\text{emp}} - \rho_{\text{hskill}}) \ln(\text{emp}_{s,\tau}) + \rho_{\text{hskill}} \ln(\text{hskill}_{s,\tau}). \quad (3.15)$$

From equation (3.15) it follows the following: If $\hat{\rho}_{\text{emp}}$ is not statistically different from $\hat{\rho}_{\text{hskill}}$ when estimating

²⁷I also estimated specifications where I assumed that $\ln(Z_{s,\tau})$ and thereby $w_{i,t}$ is linear in the share of high-skilled labor which is typically assumed in the literature on human capital externalities (see, e.g., Moretti, 2004a; Heuermann, 2011). Qualitatively, I obtain the same results as discussed below. Furthermore, those results with respect to $\hat{\rho}_{\text{emp}}$ are statistically not significantly different from the corresponding results discussed below. The residual sum of squares differs only marginally as well.

equation (3.14), the estimation results indicate that, holding the absolute number of high-skilled workers in a local labor market constant, a change in total regional employment, which has by construction to be caused by a change in the local employment of non-high-skilled workers, does not affect the amount of acquired human capital as reflected in future wages. This would imply that solely the *absolute amount* of local high-skilled labor is relevant for the analyzed dynamic productivity gains.²⁸

The results that are obtained when estimating equation (3.14) are summarized in Table 3.7. They refer to the augmented learning function (equation (3.7)) with $Z_{s,\tau}$ being an aggregate measure as described by equation (3.13).²⁹ As discussed in the literature review, some authors stress the importance to distinguish between productivity effects that are based on human capital externalities and those based on the imperfect substitution of workers belonging to different skill groups. However, these concerns refer to the interpretation of contemporaneous effects of a local employment's skill structure. I control for these contemporaneous effects by including region-year fixed effects as well as the local industry specific share of high- and low-skilled labor (see table 3.4). The results reported in table 3.7 refer to benefits of having *previously* worked in local labor markets with a high share of high-skilled workers. In contrast to contemporaneous effects, the gains discussed below should therefore only be based on external effects.

Table 3.7: Estimates of the parameters of the learning function distinguishing the impact of total regional employment and the share of high-skilled labor

	(1)	(2)
$\hat{\rho}_{\text{emp}}^{\dagger}$	0.228*** (0.029)	0.164*** (0.038)
$\hat{\rho}_{\text{hskill}}^{\dagger}$	-0.020 (0.068)	0.244** (0.096)
$\hat{\gamma}^{\dagger}$	-0.841* (0.485)	0.377 (0.648)
θ^{\dagger}	3.452	4.177
New employment relationships	336,286	214,319
within R^2	0.197	0.261

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

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. Both columns summarize estimates of the parameters of the augmented learning function (equation (3.7)) 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 (3.13). The results reported in column (1) are obtained using the full sample, the results reported in column (2) using a reduced sample which does not contain new employment relationships of workers who acquired experience before 1995. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

The results reported in column (1) are obtained using the full sample of new employment relationships. The results reported in column (2) are based on the reduced sample which does not contain new employment rela-

²⁸ Whether solely the absolute amount of high-skilled labor matters, could alternatively be tested by measuring local labor market size in terms of the absolute number of high-skilled workers and the absolute number of the other (non-high-skilled) workers: $\delta \ln(Z_{s,\tau}) = \omega_{\text{hskill}} \ln(\text{emp}_{s,\tau} - \text{hskill}_{s,\tau}) + \omega_{\text{hskill}} \ln(\text{hskill}_{s,\tau})$. In order to test, whether the local amount of high-skilled labor is of more importance for the analyzed productivity gains than the local amount of non-high-skilled labor, it then has to be tested whether ω_{hskill} is statistically larger than ω_{hskill} after estimating these elasticities.

²⁹ The reported results are obtained setting θ equal to 3.452×10^{-4} and 4.177×10^{-4} , respectively. Results for other values of θ are summarized in table 3.C.4 and table 3.C.5.

tionships of workers who acquired experience before 1995. The regression results are ambiguous. The results in column (1) suggest, that it is the amount of total regional employment in the labor markets in which a worker acquired work experience that determines his or her future productivity. The composition of the local labor force with respect to the skill level seems not to have an important effect. The null hypothesis that ρ_{hskill} is zero cannot be rejected at conventional levels. The sign of the estimated parameter is even negative. This surprises as one presumably learns more from high- than from low-skilled labor. The latter view, contrasting with the results in column (1), is, however, supported by the results reported in column (2). Using the restricted sample, a statistically and economically highly significant effect is identified not only for the labor market size in terms of total regional employment, but also for the local share of high-skilled labor. Hence, these results indicate that, holding total employment in the labor markets in which work experience was acquired constant, wages today are higher the larger the share of high-skilled workers in the local labor markets were. Vice versa, holding the composition of a local labor force with respect to skills constant, an increasing size of the local labor force has a positive effect on future productivity as well.

Comparing the estimated elasticities $\hat{\rho}_{\text{skill}}$ and $\hat{\rho}_{\text{emp}}$ shows that they are of comparable magnitude. Moreover, the hypotheses that the elasticities are equal to each other cannot be rejected at conventional levels (p-value=0.526). Therefore, the regression results reported in column (2) furthermore suggest that it is solely the absolute amount of high-skilled labor in the local labor markets in which experience was acquired previously that determines the productivity of a worker. For an increase in the total number of workers which is caused by an increasing number of non-high-skilled workers, meaning that the number of high-skilled workers remains constant, a statistically significant effect on future wages cannot be identified since $\hat{\rho}_{\text{emp}}$ is not statistically different from $\hat{\rho}_{\text{skill}}$ (see equation (3.15)).

Based on the results reported in table 3.7, it is not possible to compute the thresholds $\frac{\text{emp}}{\text{emp}}$ and $\frac{\text{hskill}}{\text{emp}}$ explicitly. However, for each considered region-year-combination the term $-\hat{\rho}_{\text{emp}} \ln(\text{emp}_{r,t}) - \hat{\rho}_{\text{hskill}} \ln\left(\frac{\text{hskill}_{r,t}}{\text{emp}_{r,t}}\right)$ is smaller than $\hat{\gamma}$. It again indicates that working in any of the local labor markets in Germany has a positive impact on the future productivity of an individual worker, even though it remains ambiguous whether the benefit increases in the size of the entire local labor force or only in the local (absolute) amount of high-skilled labor.

Finally, I let the parameters of the learning function vary depending on the labor market in which the accumulated human capital is used. As in the baseline specifications, I use the local labor market size in terms of total regional employment in order to quantify dynamic agglomeration benefits. I do not distinguish whether the gain from acquiring experience in a large labor market is related to the large number of workers or the above average share of high-skilled labor. In contrast, I take into account that the productivity increase which comes from the acquired work experience might not be the same in all regions. More precisely, I estimate equation (3.12) again, but now I let the parameters of the learning function vary depending on whether the new employer is located in a small, medium, or large labor market. As large local labor markets I consider those with a local workforce of more than 750,000 employees: Berlin, Hamburg, Munich, Frankfurt, Düsseldorf, Stuttgart, Cologne. Labor markets with less than 100,000 employees are considered as small labor markets, i.e., 43 percent of all German

local labor market regions. The remaining labor markets represent the reference group.³⁰ Table 3.8 on page 64 captures the corresponding estimates for the pivotal parameters. The results reported in column (1) are obtained based on the full sample, the results in column (2) based on the reduced sample.³¹ The details of the result are discussed at the end of this sub-section.

Overall, both specifications indicate that the value of experience varies little between the considered groups of labor markets. This supports the hypothesis that the identified dynamic agglomeration benefits are related to learning externalities. A worker who previously worked in a large labor market is in comparison to a worker who previously worked somewhere else, *ceteris paribus*, more productive, independently of where they make use of their work experience. Similar as with De la Roca and Puga (2016: figure 3), figure 3.3 illustrates this exemplarily. The reference is a worker who acquired 30 years of work experience in a labor market with a size of 100,000 employees. The solid black line refers to a worker who acquired 30 years of experience in a labor market with a size of 1,000,000 employees. In line with the idea of dynamic agglomeration economies the gap in individual productivity significantly increases over time. However, marginal benefits are decreasing. This is caused by the depreciation of previously accumulated human capital. The worker in the larger labor market accumulated in the past more human capital which is why he or she experiences an in absolute terms larger gross loss in the value of previously accumulated human capital than the worker in the smaller labor market (see equation 3.7). After 30 years the difference in productivity amounts to about 17 percent.

The dashed lines indicate what happens if a worker migrates after ten years from the larger to the smaller labor market or vice versa. The immediate changes in productivity as reflected in wage are, especially in comparison to the marked differences in the level of productivity, small. However, in the years after migration productivity adjusts. In the larger labor market the worker who initially worked in the smaller labor market catches-up. The intuition behind this development is straightforward. Due to the depreciation of accumulated human capital, it becomes less important in which local labor market experience was acquired at the beginning of the individual career the more time passes (compare equation 3.7). 20 years after migrating to the respective other regional labor market, the productivity differential relative to the worker who always stayed in the destination amounts only to about one percent.

³⁰ I also estimated specification where I require large labor markets to have a local workforce of at least 500,000 employees and at least 1,000,000 employees, respectively. The obtained results are very similar to those reported in table 3.8 and available upon request.

³¹ Results that I obtain setting θ to other values than those reported in table 3.8 can be found in tables 3.C.6 and 3.C.7, respectively.

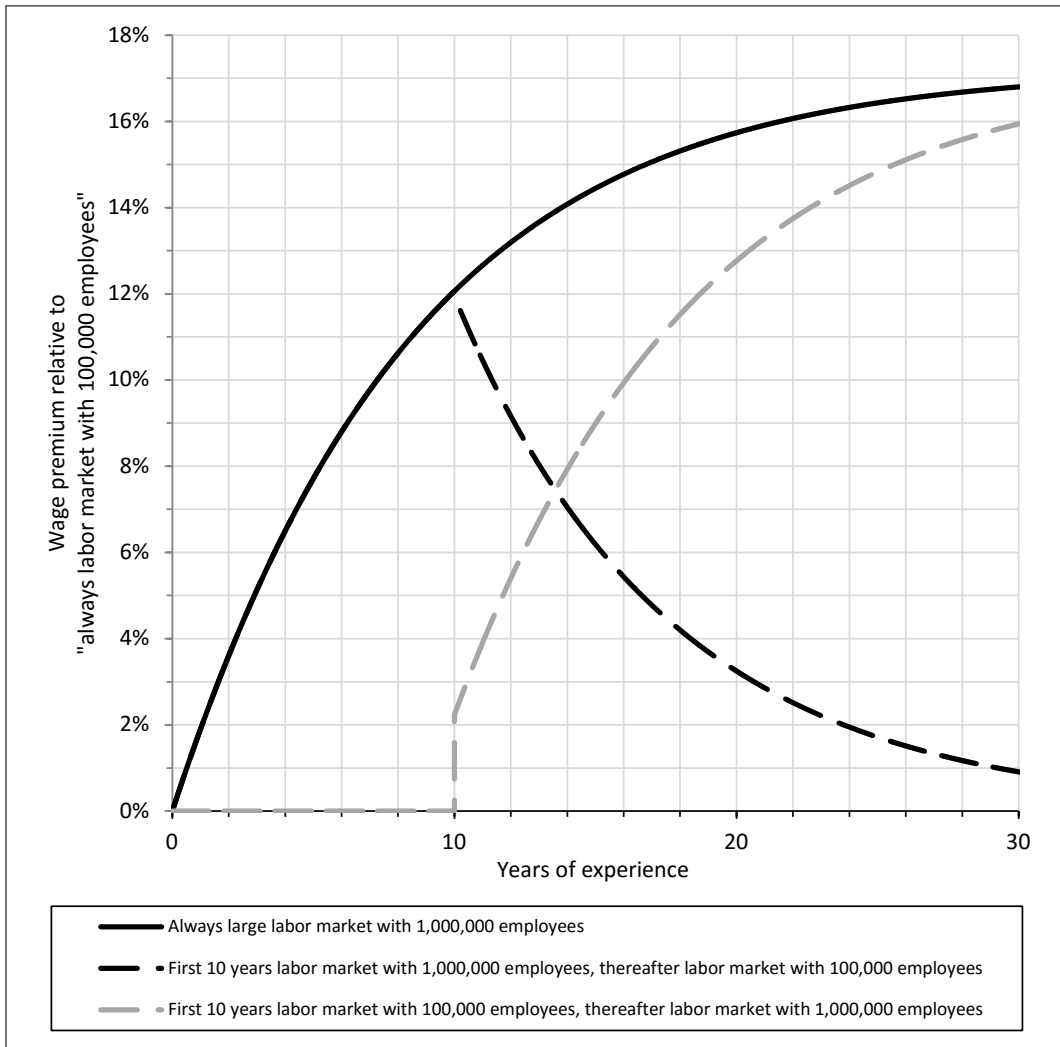


Figure 3.3: Labor productivity after migration

Note: Following De la Roca and Puga (2016: figure 3), this figure illustrates exemplarily the regression results reported in column (1) of table 3.8.

Turning to the parameter estimates reported in table 3.8 shows that the estimates which are obtained based on new employment relationships in small labor markets are of a comparable magnitude and not statistically different from those obtained based on wages paid in medium size labor markets. With regard to large labor markets, I obtain an estimate for $\hat{\delta}$ that is about one fifth smaller than the other estimates for δ . *Ceteris paribus*, it suggests that wages in large labor markets are slightly less elastic with regard to the size of the labor market in which previous experience was acquired than in the other labor markets.³² This finding is in accordance with my results reported in table 3.5. In the three largest German labor markets the difference in the wage of a worker with one year of experience acquired in Berlin, Hamburg or Munich and the wage of a worker who acquired the experience somewhere else amounts to about 0.6 percent ($= (0.0134 - 0.0005) - (0.0071 - 0.0003)$). Outside the three largest labor markets this difference amounts to about 1.3 percent ($= 0.0134 - 0.0005$). De la Roca and Puga (2016) obtain a similar result for Spain. The phenomenon might also be related to learning externalities. Suppose workers who acquired experience in a large local labor market, let's call them workers of type L, and workers who acquired the same amount of experience in a medium or small local labor market, workers of type S, are imperfect substitutes since the accumulated knowledge differs. Then, workers of type L may receive a higher wage, relative to workers of type S, especially if they use their accumulated knowledge in a smaller labor market, in which workers of type L are relatively rare (compare table 3.3). In contrast, if both types of workers use their accumulated human capital in a large labor market, the wage difference presumably is smaller since there are many workers of type L and few of type S.

When discussing the difference in $\hat{\delta}$, it has to be taken into account that there also exists an opposite effect. For $\hat{\underline{z}}$, defined as $\exp(-\hat{\gamma}/\hat{\delta})$, I obtain a smaller estimate as well when considering wages paid in large labor markets, although the difference in this parameter is not statistically significant. *Ceteris paribus*, this suggests that the value of each unit of experience is slightly higher if experience is used in a large labor market rather than somewhere else, see equations (3.4) and (3.7). One explanation may refer to a higher quality of matches between job seekers and vacancies in large than in other labor markets, e.g., due to a larger variety of available jobs. If a worker accumulated at each day when experience was acquired a small amount of human capital that he or she can make use of only if the new job is located in a large labor market because of the better match, it should result in a smaller estimate for $\hat{\underline{z}}$.

However, it should be stressed here that the differences in $\hat{\delta}$ and $\hat{\underline{z}}$ should not be overemphasized. Quantitatively, the differences in the parameter estimates are rather small depending on whether they are estimated based on wages paid in large, medium, or small labor markets. Furthermore, the effects go in opposite directions. As illustrated by figure 3.3, there is virtually no jump in the value of experience if a worker leaves a large labor market to work somewhere else. This implies that the opposite effects outweigh each other.³³ This is in accordance with the results reported in table 3.5 and by De la Roca and Puga (2016) which suggest that a worker from the largest labor markets of a country does not suffer from a decline in productivity if he or

³² Recall that differences in static regional wage differentials between large, medium, and small labor markets are captured by the region fixed effects, not by $\delta_{r(i,t)}$.

³³ I also considered a case where a worker acquired experience in a labor market with a size of 750,000 employees and 1,400,000 employees, respectively. In both cases, there is again virtually no jump in the value of experience if the worker migrates to a labor market with a size of 100,000 employees.

she uses the accumulated human capital somewhere else. For workers who migrate from a smaller to a larger labor market there might be a small immediate productivity increase as illustrated by figure 3.3 and reported in table 3.5 as well as by De la Roca and Puga (2016), maybe due to a better match between job seeker and vacancy. Nevertheless, in comparison to a worker who already previously worked in the larger labor market there is a significant gap in individual productivity which presumably is a result of learning externalities in the larger labor market.

Table 3.8: Estimates of the parameters of the learning function depending on the size of the labor market in which experience is used

	(1)	(2)
Now large labor market		
$\Delta \hat{\delta}^\dagger$	-0.044*** (0.016)	-0.037 (0.024)
$\Delta \hat{Z}$	-41.357 (33.423)	-70.444 (72.390)
Now medium size labor market (reference)		
$\hat{\delta}^\dagger$	0.236*** (0.022)	0.228*** (0.028)
\hat{Z}	45.866 (37.650)	76.967 (79.588)
Now small labor market		
$\Delta \hat{\delta}^\dagger$	-0.025 (0.017)	-0.001 (0.026)
$\Delta \hat{Z}$	-27.928 (26.775)	18.587 (78.715)
θ^\dagger	3.452	4.177
New employment relationships within R ²	336,286 0.195	214,319 0.260

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

Note: Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. Both columns summarize estimates of the parameters of the augmented learning function (equation (3.7)). \hat{Z} calculated based on delta method and $\hat{\delta}$ and $\hat{\gamma}$ (not reported). The parameters are allowed to vary depending on the size of the labor market in which the accumulated knowledge is used. Labor markets with a total number of at least 750,000 employees are considered as large labor markets, local labor markets with less than 100,000 employees as small labor markets. Δ refers to the difference in a parameter in comparison to the reference group, i.e., medium size labor markets. The results reported in column (1) are obtained using the full sample, the results reported in column (2) using a reduced sample which does not contain new employment relationships of workers who acquired experience before 1995. All models including control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

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

degree. However, if the expected individual learning opportunities depend on unobserved time-variant individual characteristics, the estimated benefits of acquiring experience in large local labor markets are likely biased upwards.

3.6 Conclusions

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

The identified effect should be strongly related to learning externalities. I control for observable as well as unobservable characteristics of the worker and the region in which the new employer is located, as well as for characteristics of the firm and the local industry. I also take into account other channels of agglomeration economies, *inter alia*, dynamic matching. Furthermore, my results are also in line with previous findings that the value of experience is predominately determined by the size of the labor market in which experience was acquired rather than by the labor market in which it is used. Initially, the productivity of workers as reflected in wages changes little when they migrate from a large labor market to a small one or vice versa. However, in the following years productivity adjusts presumably due to the depreciation of the accumulated human capital. For workers who migrate to a smaller labor market, the benefit of having previously worked in a larger labor market declines the more time passes. On the contrary, workers who migrate to a larger labor market experience a strong relative productivity increase in the following years and catch-up to those who always worked in the larger labor market.

Finally, this paper aims at analyzing as to which characteristic of an urban labor market leads to the increasing gains of acquiring experience there. More precisely, I distinguish the impact of labor market size on the value

of acquired experience and future productivity from the impact of the local share of high-skilled labor. Thereby, I address that it presumably is especially high-skilled labor workers learn from. However, the results are ambiguous. Some specifications suggest that it is the local labor market size in terms of total regional employment which determines future productivity and wages. According to these results the local workforce composition with regard to skills does not affect the value of acquired work experience. In contrast, other specifications indicate that the value of work experience depends *only* on the number of high-skilled workers in the local labor market in which the experience was acquired. Due to these contradictory results, further research is needed to investigate the importance of a local labor force's skill composition for the growth of a worker's productivity over working life. A further interesting question for future research is whether dynamic agglomeration economies are increasing in labor market size without bound, or whether the benefit decreases beyond some threshold since, inter alia, urban congestion may hinder the transmission of skills as discussed by, for example, Duranton and Puga (2004).

Appendix

3.A The solution of equation (3.9)

The difference equation given by equation (3.9) is a special case of the general first-order difference equation $x_{t+1} = b_t + a_t x_t$. Therefore, the solution of equation (3.9) can be obtained based on the solution of the general difference equation.

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

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

The first few values are given by:

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

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

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

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

Difference equation given by (3.9),
omitting some indices for simplicity:

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

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

$$\begin{aligned} Q_1 &= 0 && \text{(see equation (3.8))} \\ Q_2 &= \tilde{b}_1 \\ Q_3 &= \tilde{b}_2 + \tilde{a}_2 \tilde{b}_1 \\ Q_4 &= \tilde{b}_3 + \tilde{a}_3 \tilde{b}_2 + \tilde{a}_3 \tilde{a}_2 \tilde{b}_1 \end{aligned}$$

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

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

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

Proof by mathematical induction:

$$\begin{aligned} Q_2 &= (1 - \theta)^{2-1-1} \times \tilde{b}_1 = \tilde{b}_1 \\ Q_{\tau+1} &= \tilde{b}_{\tau} + (1 - \theta) Q_{\tau} \\ &= \tilde{b}_{\tau} + \tilde{a}_{\tau} \sum_{k=1}^{\tau-1} (1 - \theta)^{\tau-k-1} \times \tilde{b}_k \\ &= \tilde{b}_{\tau} + \sum_{k=1}^{\tau-1} (1 - \theta)^{(\tau+1)-k-1} \times \tilde{b}_k \\ &= \sum_{k=1}^{\tau} (1 - \theta)^{(\tau+1)-k-1} \times \tilde{b}_k \end{aligned}$$

□

3.B Further information on data

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

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

The dependent variable in my analysis is the wage of a new employment relationship. The first employment spell in the IEB of a new employment relationship ends, at the latest, by December 31 of the year in which the new employment relationship starts. By dividing total reported earnings by the length of the spell, daily wages are obtained which I use as the dependent variable. Information on actual working days or contract hours is not available. Firms report earnings only up to the upper limit for social security contributions. Therefore, the wage information in the IEB is right censored. I follow Reichelt (2015) and apply an interval regression, a generalization of tobit regression, to predict the wages above the threshold (about 6% of the observations). See Reichelt (2015) for a detailed description on how interval regression is applied to impute right-censored wages. For the imputation I use information on sex, age, nationality, educational level, industry and the region in which the establishment is located. The results of my regression analysis do not change when I use the reported wages as dependent variable instead of the imputed wages. Table 3.B.1 provides information on the definition of all variables used in this analysis.

Table 3.B.1: Variables - definitions and sources

Variable	Definition	Source
Gross daily wage	Daily wages are calculated by dividing the reported total earning from employment spell by the length of the spell. Wages above the upper limit for social security contributions are imputed (see Appendix 3.A).	IEB
Size of local labor market in which experience was acquired	Size of regional labor market regions, defined according to Kosfeld and Werner (2012), in which work experience was acquired until the considered new employment relationship. Measured in terms of employment subject to social security contributions. The share of high-skilled labor refers to workers with an university degree.	IEB
Work experience	Length of previous employment spells measured on a daily basis. Marginal employment is not considered, nor are employment spells that refer to measures of active labor market policies. I also compute the work experience that was acquired in the largest German labor market regions, i.e., Berlin, Hamburg and Munich.	IEB
Tenure	The length of an employment spell in month that refers to a new employment relationship. The spell ends at the latest by December 31 of the year in which the new employment relationship starts.	IEB
ln(Number of employers)	The number of unique establishment identifiers until the considered new employment relationship, by person.	IEB
Educational level of the worker	A categorial variable that combines information on the highest school leaving certificate, completed vocational training and university degree. For some spells of employment this information is missing. If so, I use the information from previous employment spells following Fitzenberger et al. (2005).	IEB
Gender	Dummy variable distinguishing male and female workers.	IEB
Length of non-employment	The number of days between the beginning of the new employment relationship and the end of the previous employment spell.	IEB
Pre-employment status	Dummy variables referring to the 28 days before the considered transition to employment <ul style="list-style-type: none"> - unemployment benefits (Arbeitslosengeld I) - unemployment assistance (Arbeitslosengeld II / Arbeitslosenhilfe). - unemployed and registered as a job seeker - not unemployed, but registered as a job seeker - participating in active labor market policy programmes. 	IEB
Firm characteristics	Number of employees, share of workers with a university degree, share of workers with no completed vocational training/no university degree, share of workers younger than 30 years old, share of workers 50 years old or older. The information refers to the last reference date (June 30) before the considered transition.	Establishment History Panel (BHP)

Table 3.B.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 spells of employment from 2005–2007 to the industry that the firm reports in 2008 (or later). If an establishment identifier shows up only for 2005–2007, I use a correlation matrix between the old and new industry classification as described by Eberle et al. (2011).	IEB
Occupation fixed effects	Fixed effects for 335 distinct occupations.	IEB
Region-year fixed effects	Time varying fixed effects for the location of the establishment in which a person starts to work. The location refers to one of 141 functional labor markets which are defined according to commuting intensity between NUTS-3-regions (see Kosfeld/Werner, 2012).	IEB

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

3.C Further results

Table 3.C.1: Summary statistics

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

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

Table 3.C.2: Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.853
$\hat{\delta}^\dagger$	0.802*** (0.067)	0.300*** (0.026)	0.240*** (0.022)	0.235*** (0.021)	0.230*** (0.021)	0.226*** (0.021)	0.222*** (0.020)	0.218*** (0.020)	0.216*** (0.020)	0.215*** (0.020)	0.215*** (0.020)
$\hat{\gamma}^\dagger$	-3.567*** (0.856)	-1.080*** (0.325)	-0.766*** (0.272)	-0.751*** (0.268)	-0.737*** (0.263)	-0.727*** (0.260)	-0.718*** (0.256)	-0.710*** (0.253)	-0.706*** (0.252)	-0.702*** (0.250)	-0.702*** (0.250)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.184	0.192	0.196	0.196	0.196	0.197	0.197	0.197	0.197	0.196	0.196
RSS	8433.749	8350.796	8314.283	8312.030	8310.187	8309.065	8308.629	8308.834	8309.152	8309.630	8309.630
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\delta}^\dagger$	0.211*** (0.020)	0.208*** (0.019)	0.205*** (0.019)	0.202*** (0.019)	0.196*** (0.019)	0.190*** (0.019)	0.183*** (0.019)	0.176*** (0.018)	0.168*** (0.018)	0.134*** (0.017)	0.116*** (0.016)
$\hat{\gamma}^\dagger$	-0.693*** (0.248)	-0.683*** (0.246)	-0.671*** (0.244)	-0.656*** (0.242)	-0.621*** (0.239)	-0.580** (0.236)	-0.536** (0.233)	-0.493** (0.231)	-0.457** (0.228)	-0.444** (0.215)	-0.532*** (0.206)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.196	0.196	0.196	0.196	0.195	0.194	0.193	0.192	0.191	0.185	0.183
RSS	8310.969	8312.805	8315.103	8317.830	8324.474	8332.594	8341.885	8352.720	8364.726	8423.750	8452.718

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

Note: The table summarizes estimation results for equation (3.12) depending on the value θ is set to. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.853
$\hat{\delta}^\dagger$	0.765*** (0.076)	0.294*** (0.030)	0.244*** (0.026)	0.240*** (0.026)	0.236*** (0.026)	0.232*** (0.025)	0.229*** (0.025)	0.226*** (0.025)	0.224*** (0.025)	0.222*** (0.025)	0.222*** (0.025)
$\hat{\gamma}^\dagger$	-4.019*** (0.962)	-1.409*** (0.382)	-1.103*** (0.332)	-1.084*** (0.328)	-1.065*** (0.324)	-1.047*** (0.320)	-1.031*** (0.316)	-1.017*** (0.313)	-1.010*** (0.311)	-1.004*** (0.309)	-1.004*** (0.309)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.254	0.259	0.261	0.261	0.261	0.261	0.261	0.261	0.261	0.260	0.260
RSS	4762.913	4729.010	4719.557	4719.247	4719.119	4719.221	4719.544	4720.074	4720.405	4720.798	4720.798
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\delta}^\dagger$	0.219*** (0.024)	0.216*** (0.024)	0.213*** (0.024)	0.210*** (0.024)	0.204*** (0.023)	0.199*** (0.023)	0.194*** (0.022)	0.190*** (0.022)	0.185*** (0.021)	0.171*** (0.020)	0.164*** (0.019)
$\hat{\gamma}^\dagger$	-0.991*** (0.306)	-0.980*** (0.303)	-0.970*** (0.300)	-0.961*** (0.296)	-0.947*** (0.290)	-0.939*** (0.284)	-0.934*** (0.279)	-0.934*** (0.274)	-0.936*** (0.269)	-0.960*** (0.252)	-0.976*** (0.245)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.260	0.260	0.260	0.260	0.259	0.259	0.258	0.258	0.257	0.256	0.255
RSS	4721.700	4722.758	4723.953	4725.260	4728.115	4731.146	4734.150	4737.182	4740.100	4751.033	4755.799

\dagger Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (3.12) depending on the value θ is set to. The estimates are obtained using a reduced sample which do not contain workers who acquired experience before 1995. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

Table 3.C.4: Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment and the local share of high-skilled labor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
$\hat{\rho}_{\text{emp}}^\dagger$	0.377*** (0.099)	0.233*** (0.038)	0.237*** (0.031)	0.234*** (0.030)	0.232*** (0.030)	0.230*** (0.029)	0.228*** (0.029)	0.227*** (0.028)	0.227*** (0.028)	0.228*** (0.028)	0.228*** (0.028)
$\hat{\rho}_{\text{hskill}}^\dagger$	1.385*** (0.240)	0.220** (0.094)	0.011 (0.075)	0.003 (0.073)	-0.005 (0.071)	-0.012 (0.06911)	-0.020 (0.068)	-0.030 (0.066)	-0.036 (0.065)	-0.043 (0.065)	-0.043 (0.065)
$\hat{\gamma}^\dagger$	4.753*** (1.676)	0.250 (0.649)	-0.697 (0.527)	-0.733 (0.516)	-0.767 (0.505)	-0.801 (0.495)	-0.841* (0.485)	-0.895* (0.476)	-0.928** (0.473)	-0.968** (0.469)	-0.968** (0.469)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.185	0.193	0.196	0.196	0.196	0.197	0.197	0.197	0.197	0.196	0.196
RSS	8432.18	8350.51	8314.28	8312.03	8310.19	8309.06	8308.62	8308.82	8309.14	8309.61	8309.61
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\rho}_{\text{emp}}^\dagger$	0.230*** (0.027)	0.233*** (0.027)	0.238*** (0.027)	0.244*** (0.026)	0.259*** (0.026)	0.278*** (0.026)	0.296*** (0.026)	0.313*** (0.026)	0.325*** (0.026)	0.218*** (0.026)	0.0393 (0.025)
$\hat{\rho}_{\text{hskill}}^\dagger$	-0.060 (0.064)	-0.081 (0.063)	-0.106* (0.062)	-0.135** (0.061)	-0.205*** (0.061)	-0.285*** (0.060)	-0.367*** (0.0603)	-0.445*** (0.0609)	-0.506*** (0.0618)	-0.272*** (0.0659)	0.248*** (0.0644)
$\hat{\gamma}^\dagger$	-1.064** (0.462)	-1.184*** (0.456)	-1.330*** (0.451)	-1.499*** (0.446)	-1.897*** (0.440)	-2.348*** (0.435)	-2.802*** (0.434)	-3.229*** (0.434)	-3.555*** (0.435)	-2.086*** (0.444)	0.968** (0.433)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.196	0.196	0.196	0.196	0.195	0.194	0.194	0.193	0.191	0.186	0.183
RSS	8310.92	8312.72	8314.95	8317.57	8323.86	8331.41	8339.93	8349.91	8361.19	8422.85	8451.93

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

Note: The table summarizes estimation results for equation (3.12) depending on the value θ is set to. Here, labor market size $Z_{s,\tau}$ is considered to be an aggregate of total regional employment and the local share of high-skilled labor, see equation (3.13). $\hat{\rho}_{\text{emp}}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment* and $\hat{\rho}_{\text{hskill}}$ the corresponding elasticity with respect to the local *share of high-skilled labor* where experience was acquired. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

Table 3.C.5: Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment and the local share of high-skilled labor

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.724
$\hat{\rho}_{\text{emp}}^\dagger$	0.291*** (0.112)	0.156*** (0.0447)	0.159*** (0.0388)	0.161*** (0.0384)	0.164*** (0.0379)	0.167*** (0.0375)	0.170*** (0.0372)	0.173*** (0.0368)	0.174*** (0.0367)	0.175*** (0.0365)	0.175*** (0.0365)
$\hat{\rho}_{\text{hskill}}^\dagger$	1.576*** (0.276)	0.465*** (0.113)	0.286*** (0.0983)	0.267*** (0.0973)	0.244** (0.0964)	0.222** (0.0955)	0.200** (0.0947)	0.178* (0.0940)	0.168* (0.0937)	0.158* (0.0934)	0.158* (0.0934)
$\hat{\gamma}^\dagger$	5.298*** (1.890)	1.330* (0.762)	0.585 (0.663)	0.489 (0.655)	0.377 (0.648)	0.263 (0.642)	0.148 (0.636)	0.0356 (0.630)	-0.0178 (0.627)	-0.0723 (0.625)	-0.0723 (0.625)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.254	0.259	0.261	0.261	0.261	0.261	0.261	0.261	0.261	0.260	0.260
RSS	4761.59	4728.25	4719.16	4718.90	4718.82	4718.97	4719.33	4719.90	4720.25	4720.66	4720.66
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
$\hat{\rho}_{\text{emp}}^\dagger$	0.178*** (0.0362)	0.180*** (0.0359)	0.181*** (0.0356)	0.182*** (0.0353)	0.182*** (0.0347)	0.179*** (0.0342)	0.174*** (0.0336)	0.168*** (0.0331)	0.160*** (0.0326)	0.122*** (0.0307)	0.103*** (0.0298)
$\hat{\rho}_{\text{hskill}}^\dagger$	0.139 (0.0927)	0.121 (0.0921)	0.106 (0.0915)	0.0930 (0.0909)	0.0747 (0.0897)	0.0662 (0.0885)	0.0661 (0.0873)	0.0728 (0.0861)	0.0844 (0.0849)	0.160** (0.0799)	0.204*** (0.0775)
$\hat{\gamma}^\dagger$	-0.173 (0.620)	-0.265 (0.615)	-0.345 (0.610)	-0.413 (0.605)	-0.507 (0.595)	-0.548 (0.586)	-0.544 (0.577)	-0.504 (0.569)	-0.437 (0.560)	-0.0117 (0.527)	0.231 (0.512)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.260	0.260	0.260	0.260	0.259	0.259	0.258	0.258	0.257	0.256	0.255
RSS	4721.59	4722.68	4723.89	4725.21	4728.08	4731.12	4734.12	4737.15	4740.05	4750.84	4755.47

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

Note: The table summarizes estimation results for equation (3.12) depending on the value θ is set to. Here, labor market size $Z_{s,\tau}$ is considered to be an aggregate of total regional employment and the local share of high-skilled labor, see equation (3.13). $\hat{\rho}_{\text{emp}}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment* and $\hat{\rho}_{\text{hskill}}$ the corresponding elasticity with respect to the local *share of high-skilled labor* where experience was acquired. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

Table 3.C.6: Estimates of the parameters of the learning function based on different values for θ and full sample, labor market size in terms of total regional employment, parameters allowed to vary depending on size of labor market where experience is used

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.853
Now large labor market											
$\Delta\hat{\delta}^\dagger$	-0.002 (0.135)	-0.047 (0.034)	-0.047** (0.021)	-0.047** (0.0206)	-0.046** (0.019)	-0.045** (0.017)	-0.044** (0.016)	-0.043** (0.016)	-0.042** (0.015)	-0.042** (0.015)	-0.042** (0.015)
$\Delta\hat{\gamma}^\dagger$	1.398 (1.761)	0.903** (0.439)	0.708** (0.270)	0.685** (0.256)	0.660** (0.240)	0.636** (0.227)	0.613** (0.214)	0.590** (0.202)	0.580** (0.197)	0.569** (0.192)	0.569** (0.192)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.703** (0.094)	0.302** (0.031)	0.252** (0.024)	0.248** (0.023)	0.243** (0.023)	0.239** (0.022)	0.236** (0.022)	0.232** (0.022)	0.231** (0.021)	0.229** (0.021)	0.229** (0.021)
$\hat{\gamma}^\dagger$	-2.578** (1.176)	-1.168** (0.384)	-0.936** (0.300)	-0.925** (0.294)	-0.916** (0.287)	-0.908** (0.281)	-0.902** (0.276)	-0.895** (0.271)	-0.892** (0.269)	-0.888** (0.267)	-0.888** (0.267)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	-0.348** (0.146)	-0.086** (0.036)	-0.037* (0.022)	-0.034 (0.021)	-0.030 (0.020)	-0.027 (0.019)	-0.025 (0.018)	-0.023 (0.017)	-0.022 (0.016)	-0.021 (0.016)	-0.021 (0.016)
$\Delta\hat{\gamma}^\dagger$	3.328* (1.745)	0.894** (0.435)	0.417 (0.267)	0.382 (0.252)	0.348 (0.237)	0.318 (0.224)	0.293 (0.211)	0.271 (0.199)	0.261 (0.194)	0.252 (0.189)	0.252 (0.189)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.185	0.193	0.196	0.196	0.196	0.197	0.197	0.197	0.197	0.197	0.197
RSS	8430.50	8348.84	8313.46	8311.28	8309.51	8308.44	8308.04	8308.27	8308.60	8309.09	8309.09
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
Now large labor market											
$\Delta\hat{\delta}^\dagger$	-0.041** (0.014)	-0.040** (0.013)	-0.038** (0.013)	-0.037** (0.012)	-0.035** (0.011)	-0.034** (0.010)	-0.032** (0.010)	-0.031** (0.009)	-0.029** (0.009)	-0.025** (0.007)	-0.023** (0.007)
$\Delta\hat{\gamma}^\dagger$	0.549** (0.182)	0.530** (0.173)	0.512** (0.165)	0.495** (0.157)	0.463** (0.144)	0.436** (0.134)	0.412** (0.125)	0.390** (0.118)	0.371** (0.112)	0.308** (0.0928)	0.283** (0.0860)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.226** (0.021)	0.223** (0.021)	0.220** (0.020)	0.217** (0.020)	0.210** (0.020)	0.204** (0.019)	0.197** (0.019)	0.190** (0.019)	0.182** (0.019)	0.146** (0.017)	0.128** (0.017)
$\hat{\gamma}^\dagger$	-0.879** (0.263)	-0.869** (0.260)	-0.856** (0.257)	-0.840** (0.254)	-0.803** (0.249)	-0.759** (0.245)	-0.713** (0.242)	-0.667** (0.239)	-0.628** (0.235)	-0.601** (0.220)	-0.679** (0.211)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	-0.019 (0.015)	-0.018 (0.014)	-0.017 (0.013)	-0.016 (0.013)	-0.014 (0.012)	-0.014 (0.011)	-0.013 (0.010)	-0.013 (0.010)	-0.0125 (0.009)	-0.0127* (0.008)	-0.0125* (0.007)
$\Delta\hat{\gamma}^\dagger$	0.236 (0.179)	0.222 (0.170)	0.210 (0.162)	0.201 (0.155)	0.186 (0.142)	0.177 (0.132)	0.171 (0.123)	0.169 (0.116)	0.168 (0.110)	0.172* (0.091)	0.170** (0.084)
N	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286	336,286
within R ²	0.196	0.196	0.196	0.196	0.195	0.194	0.193	0.192	0.191	0.186	0.183
RSS	8310.43	8312.27	8314.56	8317.28	8323.90	8331.98	8341.23	8352.02	8363.98	8422.80	8451.70

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

Note: The table summarizes estimation results for equation (3.12) depending on the value θ is set to. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. Labor markets with a total number of at least 750,000 employees are considered as large labor markets, local labor markets with less than 100,000 employees as small labor markets. Δ refers to the difference in a parameter in comparison to the reference group, i.e. medium size labor markets. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

Table 3.C.7: Estimates of the parameters of the learning function based on different values for θ and reduced sample, labor market size in terms of total regional employment, parameters allowed to vary depending on size of labor market where experience is used

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
θ^\dagger	50.000	10.000	5.000	4.595	4.177	3.797	3.452	3.138	2.996	2.853	2.853
Now large labor market											
$\Delta\hat{\delta}^\dagger$	0.010 (0.151)	-0.040 (0.040)	-0.038 (0.026)	-0.038 (0.025)	-0.037 (0.024)	-0.037 (0.023)	-0.036 (0.022)	-0.036* (0.021)	-0.036* (0.021)	-0.036* (0.021)	-0.036* (0.021)
$\Delta\hat{\gamma}^\dagger$	1.429 (1.983)	0.912* (0.519)	0.680** (0.342)	0.656** (0.328)	0.632** (0.313)	0.610** (0.300)	0.589** (0.288)	0.571** (0.277)	0.562** (0.272)	0.554** (0.267)	0.554** (0.267)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.598*** (0.106)	0.267*** (0.036)	0.233*** (0.029)	0.230*** (0.029)	0.228*** (0.028)	0.225*** (0.028)	0.223*** (0.028)	0.221*** (0.027)	0.220*** (0.027)	0.219*** (0.027)	0.219*** (0.027)
$\hat{\gamma}^\dagger$	-2.216* (1.334)	-1.161** (0.453)	-1.001*** (0.369)	-0.994*** (0.363)	-0.989*** (0.357)	-0.985*** (0.351)	-0.982*** (0.345)	-0.980*** (0.340)	-0.979*** (0.338)	-0.978*** (0.335)	-0.978*** (0.335)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	-0.212 (0.167)	-0.039 (0.043)	-0.005 (0.028)	-0.003 (0.027)	-0.001 (0.026)	0.001 (0.025)	0.002 (0.024)	0.003 (0.023)	0.004 (0.022)	0.004 (0.022)	0.004 (0.022)
$\Delta\hat{\gamma}^\dagger$	1.570 (2.000)	0.269 (0.517)	-0.012 (0.340)	-0.029 (0.326)	-0.045 (0.311)	-0.058 (0.298)	-0.068 (0.286)	-0.075 (0.275)	-0.078 (0.270)	-0.081 (0.265)	-0.081 (0.265)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.254	0.259	0.261	0.261	0.261	0.261	0.261	0.261	0.261	0.261	0.261
RSS	4760.30	4727.15	4718.60	4718.37	4718.33	4718.51	4718.90	4719.49	4719.84	4720.26	4720.26
	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)
θ^\dagger	2.594	2.358	2.144	1.949	1.611	1.331	1.100	0.900	0.729	0.200	0.001
Now large labor market											
$\Delta\hat{\delta}^\dagger$	-0.0351* (0.020)	-0.0347* (0.019)	-0.0343* (0.019)	-0.0339* (0.018)	-0.0332* (0.017)	-0.0326** (0.017)	-0.0321** (0.016)	-0.0315** (0.016)	-0.0311** (0.015)	-0.0294** (0.014)	-0.0287** (0.013)
$\Delta\hat{\gamma}^\dagger$	0.538** (0.258)	0.524** (0.250)	0.511** (0.242)	0.500** (0.236)	0.479** (0.224)	0.462** (0.215)	0.448** (0.207)	0.436** (0.200)	0.425** (0.195)	0.391** (0.178)	0.378** (0.172)
Now medium size labor market (reference)											
$\hat{\delta}^\dagger$	0.217*** (0.026)	0.215*** (0.026)	0.212*** (0.026)	0.210*** (0.025)	0.206*** (0.025)	0.202*** (0.024)	0.198*** (0.024)	0.195*** (0.023)	0.191*** (0.023)	0.178*** (0.021)	0.172*** (0.021)
$\hat{\gamma}^\dagger$	-0.977*** (0.331)	-0.976*** (0.326)	-0.976*** (0.322)	-0.976*** (0.318)	-0.978*** (0.310)	-0.983*** (0.304)	-0.990*** (0.297)	-0.999*** (0.292)	-1.009*** (0.286)	-1.056*** (0.268)	-1.079*** (0.260)
Now small labor market											
$\Delta\hat{\delta}^\dagger$	0.005 (0.021)	0.006 (0.021)	0.006 (0.020)	0.006 (0.019)	0.007 (0.019)	0.007 (0.018)	0.007 (0.017)	0.006 (0.017)	0.006 (0.016)	0.006 (0.015)	0.006 (0.014)
$\Delta\hat{\gamma}^\dagger$	-0.084 (0.256)	-0.086 (0.248)	-0.087 (0.241)	-0.087 (0.234)	-0.085 (0.223)	-0.082 (0.214)	-0.079 (0.206)	-0.075 (0.200)	-0.071 (0.194)	-0.059 (0.178)	-0.055 (0.171)
N	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319	214,319
within R ²	0.260	0.260	0.260	0.260	0.259	0.259	0.258	0.258	0.257	0.256	0.255
RSS	4721.21	4722.30	4723.53	4724.86	4727.76	4730.82	4733.84	4736.89	4739.82	4750.76	4755.53

\dagger Coefficients and standard errors multiplied by 10,000.

Note: The table summarizes estimation results for equation (3.12) depending on the value θ is set to. The estimates are obtained using a reduced sample which do not contain workers who acquired experience before 1995. $\hat{\delta}$ denotes the elasticity of wage with respect to the size of the labor markets where experience was acquired in terms of *total regional employment*. Labor markets with a total number of at least 750,000 employees are considered as large labor markets, local labor markets with less than 100,000 employees as small labor markets. Δ refers to the difference in a parameter in comparison to the reference group, i.e. medium size labor markets. Robust standard errors are given in parentheses which are clustered by worker. ***, ** and * indicate significance at the 1, 5 and 10 percent level. All models include control variables as well as worker, industry, occupation, and region-year fixed effects (see table 3.4).

Chapter 4

Summary of main results

This dissertation sheds light on two determinants of individual labor productivity. Both refer to the environment a worker is working in. In the first paper, chapter 2, I focus on the impact of a workforce age structure on individual productivity taking into account that differently aged workers complement each other in the production process. In the second paper, chapter 3, I address that the size of the local labor market in which an individual works determines productivity. The focus here lies on dynamic agglomeration gains that accumulate over time, meaning that labor market size enhances the individual accumulation of knowledge and thereby a worker's productivity in the future.

The results of the first paper confirm that workers belonging to different age groups are complementary factors. The analysis is based on a cross sectional linked employer-employee data set from 2012. By estimating translog cost functions, I show that the degree of complementarity between differently aged workers differs depending on the age and the skill level of the workers. Complementarities especially arise between younger and middle aged workers. The highest degree of complementarity is observed between younger high skilled and middle aged high skilled labor. Model calculations indicate that the shift in the age structure that is expected for Germany will change the relative productivity of age groups. Accordingly, the productivity of younger and middle aged workers will increase by, respectively, about five percent and three percent, whereas the productivity of older workers will significantly decline by about nine percent caused by their increasing share in the workforce. The significant piece of information that changes in a workforce age structure affect (relative) productivity of differently aged workers should be considered in any discussion of the impact of demographic change, and in formulating policy that impacts on relative labor supply, such as changes in retirement age.

The results of the second paper strongly indicate that individual productivity significantly depends on the size of the local labor market in which a worker previously acquired work experience. The analysis uses micro data on transitions to employment within the period 2005 to 2011 and individual employment biographies from 1975 onwards. My findings suggest that dynamic agglomeration economies in general, and learning externalities in particular, play an important role in explaining individual labor productivity. Workers accumulate significantly more valuable human capital by working in larger local labor markets than by working in smaller ones. Doubling local employment in all labor markets in which experience was acquired, increases the productivity of an employee with two years of experience by more than 0.7 percent as reflected in wages. After 10 years of experience the corresponding gain amounts to about three percent. The identified dynamic agglomeration economies are supposed to reflect in the local wage level since workers are rather immobile between local labor markets that differ in size. Therefore, my results provide additional evidence as to why significant regional wage disparities between urban and non-urban labor markets exist in many countries.

Bibliography

- Alesina, A., Harnoss, J., Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2), 101–138.
- 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.
- Arntz, M., Gregory, T. (2014). What old stagers could teach us: Examining age complementarities in regional innovation systems. ZEW Discussion Papers 14-050, ZEW - Zentrum für Europäische Wirtschaftsforschung / Center for European Economic Research.
- 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.
- Berger, M. C. (1985). The effect of cohort size on earnings growth: A reexamination of the evidence. *Journal of Political Economy*, 93(3), 561–573.
- Blanchflower, D. G., Oswald, A. J. (1990). The wage curve. *The Scandinavian Journal of Economics*, 92(2), 215–235.
- BMAS (2013). Arbeitsmarktprognose 2030 Eine strategische Vorausschau auf die Entwicklung von Angebot und Nachfrage in Deutschland. Bundesministerium für Arbeit und Soziales Referat Information, Publikation, Redaktion (Ed.), Bonn.
- Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics*, 118(4), 1335–1374.
- Börsch-Supan, A. (2013). Myths, scientific evidence and economic policy in an aging world. *The Journal of the Economics of Ageing*, 1–2, 3–15.
- Börsch-Supan, A., Weiss, M. (2007). Productivity and age: Evidence from work teams at the assembly line. mea Discussion Paper 148-2007, mea - Munich Center for the Economics of Aging.
- 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.
- Card, D., Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2), 705–746.
- Christensen, L. R., Jorgenson, D. W., Lau, L. J. (1973). Transcendental logarithmic production frontiers. *The Review of Economics and Statistics*, 55(1), 28–45.

- Ciccone, A., Hall, R. E. (1996). Productivity and the density of economic activity. *American Economic Review*, 86(1), 54–70.
- 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.
- De Grip, A., Van Loo, J. (2002). The economics of skills obsolescence: A review. In A. De Grip, J. Van Loo, K. Mayhew (Eds.) *The Economics of Skills Obsolescence*, vol. 21 of *Research in Labor Economics*, chap. 1, (pp. 1–26). Emerald Group Publishing Limited.
- DESTATIS (2009). Bevölkerung Deutschlands bis 2060 - 12. koordinierte Bevölkerungsvorausberechnung. Statistisches Bundesamt (Ed.), Wiesbaden.
- Di Addario, S., Patacchini, E. (2008). Wages and the city. Evidence from Italy. *Labour Economics*, 15(5), 1040–1061.
- Diewert, W. E., Wales, T. J. (1987). Flexible functional forms and global curvature conditions. *Econometrica*, 55(1), 43–68.
- Driscoll, J. C., Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4), 549–560.
- Duranton, G. (2007). Human capital externalities in cities: Identification and policy issues. In R. J. Arnott, D. P. McMillen (Eds.) *A Companion to Urban Economics*, chap. 2, (pp. 24–39). Blackwell Publishing Ltd.
- Duranton, G., Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91(5), 1454–1477.
- Duranton, G., Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson, J. F. Thisse (Eds.) *Handbook of Regional and Urban Economics*, vol. 4, chap. 48, (pp. 2063–2117). Elsevier.
- Eberle, J., Jacobebbinghaus, P., Ludsteck, J., Witter, J. (2011). Generation of time-consistent industry codes in the face of classification changes - Simple heuristic based on the Establishment History Panel (BHP). FDZ-Methodenreport 05/2011, Research Data Centre of the Federal Employment Agency (BA) at the Institute for Employment Research (IAB), Nuremberg.
- Fitzenberger, B., Kohn, K. (2006). Skill wage premia, employment, and cohort effects: Are workers in Germany all of the same type? IZA Discussion Papers 2185, IZA - Institute for the Study of Labor.

- 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.
- Freeman, R. B. (1979). The effect of demographic factors on age-earnings profiles. *The Journal of Human Resources*, 14(3), 289–318.
- Garloff, A., Roth, D. (2016). Regional age structure and young workers' wages. IAB-Discussion Paper 06/2016, Institute for Employment Research (IAB) - The Research Institute of the Federal Employment Agency, Nuremberg.
- Gartner, H. (2005). The imputation of wages above the contribution limit with the German IAB employment sample. FDZ Methodenreport 02/2005, The Research Data Centre (FDZ) of the Federal Employment Agency in the Institute for Employment Research, 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.
- Göbel, C., Zwick, T. (2013). Are personnel measures effective in increasing productivity of old workers? *Labour Economics*, 22(C), 80–93.
- Gould, E. D. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *Review of Economic Studies*, 74(2), 477–506.
- Grant, J. H., Hamermesh, D. S. (1981). Labor market competition among youths, white women and others. *The Review of Economics and Statistics*, 63(3), pp. 354–360.
- Greene, W. H. (2012). *Econometric Analysis*. Upper Saddle River, NJ: Pearson, 7 ed. International Edition.
- Haas, A., Lucht, M., Schanne, N. (2013). Why to employ both migrants and natives? A study on task-specific substitutability. *Journal for Labour Market Research*, 46(3), 201–214.
- 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.
- Hamermesh, D. S. (1996). *Labor Demand*. Princeton, New Jersey: Princeton University Press.
- Hamermesh, D. S., Grant, J. (1979). Econometric studies of labor-labor substitution and their implications for policy. *Journal of Human Resources*, 14(4), 543–562.
- 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.
- Hicks, J. (1970). Elasticity of substitution again: Substitutes and complements. *Oxford Economic Papers*, 22(3), pp. 289–296.
- Hoechle, D. (2007). Robust standard errors for panel regression with cross-section dependence. *The Stata Journal*, 7(3), 281–312.

- Horn, J. L., Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica*, 26, 107–129.
- Jacobs, J. (1969). *The Economy of Cities*. New York: Random House.
- Jorgenson, D. W. (1986). Econometric methods for modeling producer behavior. In Z. Griliches, M. D. Intriligator (Eds.) *Handbook of Econometrics*, vol. 3, chap. 31, (pp. 1841–1915). Elsevier.
- 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.
- Lau, L. J. (1978). Testing and imposing monotonicity, convexity and quasiconvexity constraints. In M. Fuss, D. McFadden (Eds.) *Production Economics: A Dual Approach to Theory and Applications*, vol. 1 of *History of Economic Thought Chapters*, chap. A4. McMaster University Archive for the History of Economic Thought.
- Lehmer, F., Möller, J. (2010). Interrelations between the urban wage premium and firm-size wage differentials: a microdata cohort analysis for Germany. *The Annals of Regional Science*, 45(1), 31–53.
- List, F. (1838). *Das deutsche National-Transport-System in volks- und staatswirthschaftlicher Beziehung*. Altona/Leipzig: Johann Friedrich Hammerich.
- Mahlberg, B., Freund, I., Cuaresma, J. C., Prskawetz, A. (2013). Ageing, productivity and wages in Austria. *Labour Economics*, 22, 5 – 15.
- 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.
- Morin, L.-P. (2015). Cohort size and youth earnings: Evidence from a quasi-experiment. *Labour Economics*, 32, 99 – 111.
- 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.
- Naz, A., Niebuhr, A., Peters, J. C. (2015). What's behind the disparities in firm innovation rates across regions? Evidence on composition and context effects. *The Annals of Regional Science*, 55(1), 131–156.

- OECD (2016). Labour productivity forecast (indicator). Doi: 10.1787/cb12b189-en (Accessed on 21 July 2016).
- Ottaviano, G. I. P., Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association*, 10(1), 152–197.
- Paulus, W., Matthes, B. (2013). The German classification of occupations 2010 - Structure, coding and conversion table. FDZ-Methodenreport 08/2013, The Research Data Centre (FDZ) of the Federal Employment Agency in the Institute for Employment Research, Nuremberg.
- 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.
- Ryan, D. L., Wales, T. J. (2000). Imposing local concavity in the translog and generalized Leontief cost functions. *Economics Letters*, 67(3), 253–260.
- Sapozhnikov, M., Triest, R. K. (2007). Population aging, labor demand, and the structure of wages. Working Papers 07-8, Federal Reserve Bank of Boston.
- Sato, R., Koizumi, T. (1973). On the elasticities of substitution and complementarity. *Oxford Economic Papers*, 25(1), pp. 44–56.
- Sydsæter, K., Hammond, P., Seierstad, A., Strøm, A. (2008). *Further mathematics for economic analysis*. Harlow: Financial Times Prentice Hall, 2 ed.
- Welch, F. (1979). Effects of cohort size on earnings: The baby boom babies' financial bust. *Journal of Political Economy*, 87(5), pp. S65–S97.
- 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.

Curriculum Vitae

Personal Data

Jan Cornelius Peters, born February 14, 1985 in Geesthacht, Germany

Education

- | | |
|-----------------|--|
| since 10/2011 | Participant in the PhD-Program “Quantitative Economics”,
Christian-Albrechts-Universität zu Kiel |
| 06/2016 | Summer School in Urban Economics for PhD students,
Barcelona Institute of Economics (IEB) and
the Urban Economics Association (UEA) |
| 06/2011 | Diploma in Economics (<i>Diplom-Volkswirt</i>),
Christian-Albrechts-Universität zu Kiel,
Thesis: “Determinanten der interregionalen Arbeitskräftewanderung von
Hochqualifizierten” (in English: “Determinates of high-skilled interregional migration”)
Supervisor: Professor Dr. Annekatrin Niebuhr |
| 10/2005–06/2011 | Studies in Economics at Christian-Albrechts-Universität zu Kiel,
Focus on Spatial Economics and Public Economics / Social Policy |
| 01/2009–06/2009 | Studies in Economics as Erasmus Exchange Student
at the University of Jönköping, Sweden |
| 08/2008–12/2008 | Studies in Economics as Erasmus Exchange Student
at Åbo University, Turku, Finland |
| 2005 | Abitur, Otto-Hahn-Gymnasium, Geesthacht |

Positions

- since 07/2011 Researcher at the Institute for Employment Research (IAB),
IAB Northern Germany in Kiel
- 10/2010–06/2011 Student Assistant at the Institute for Employment Research (IAB)
- 06/2010–08/2010 Internship at Hamburgisches WeltWirtschaftsInstitut (HWWI)
- 07/2009–06/2010 Student Assistant at the Institute for Regional Research, CAU Kiel
- 08/2007–09/2007 Internship at “Behörde für Soziales, Gesundheit und Verbraucherschutz”,
City of Hamburg
- 05/2007–07/2008 Student Assistant at the Institute for Statistics & Econometrics, CAU Kiel

Publications in refereed journals

- What’s behind the disparities in firm innovation rates across regions? Evidence on selection and context effects. with A. Naz, A. Niebuhr. *The Annals of Regional Science*, 2015, Vol. 55, No.1, 131-156.
- Does cultural diversity of migrant employees affect innovation? with C. Ozgen, A. Niebuhr, P. Nijkamp, J. Poot. *International Migration Review*, 2014, Vol. 48, No. S1.
- Diversität und Erfolg von Organisationen. with A. Buche, M. Jungbauer-Gans, A. Niebuhr. *Zeitschrift für Soziologie*, 2013 Jg. 42, H. 6, 483-501.

Working Papers

- Quantifying the effect of labor market size on learning externalities. Economics Working Paper No. 2016-11, Christian-Albrechts-Universität zu Kiel, Department of Economics.
- Do age complementarities affect labor productivity? Evidence from German firm level data. Economics Working Paper No. 2016-10, Christian-Albrechts-Universität zu Kiel, Department of Economics.
- Benefits of dense labour markets - Evidence from transitions to employment in Germany. with S. Hamann, A. Niebuhr. Economics Working Paper No. 2016-07, Christian-Albrechts-Universität zu Kiel, Department of Economics.
- Benefits of dense labour markets - Evidence from transitions to employment in Germany. with S. Hamann, A. Niebuhr. IAB-Discussion Paper, 13/2016.

Other Publications

Beschäftigungs- und Arbeitslosigkeitsentwicklung in Hamburg - Gutachten. with Ch. Boll, A. Niebuhr, D. Nitt-Drießelmann, A. H. Otto, A. Rossen, 2016.

Zur Integration von Flüchtlingen in den Arbeitsmarkt in Schleswig-Holstein - Schriftliche Anhörung des Wirtschaftsausschusses des Schleswig-Holsteinischen Landtages. with T. Buch, St. Fuchs, A. Niebuhr, A. Stöckmann. IAB-Stellungnahme, 02/2016.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Region Schleswig-Holstein Süd-West - Befunde für die Teilregion Kreis Steinburg. with V. Harder, J. Wiegand, H. Herrmann, H. Lüth, B. Bauer, D. Ollech, A. Niebuhr. IAB-Regional, IAB Nord, 04/2014.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Region Schleswig-Holstein Süd-West - Befunde für die Teilregion Kreis Dithmarschen. with V. Harder, J. Wiegand, H. Herrmann, H. Lüth, B. Bauer, D. Ollech, A. Niebuhr. IAB-Regional, IAB Nord, 03/2014.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Arbeitsmarktregion Lübeck - Befunde für die Teilregion Kreis Ostholstein. with V. Harder, J. Wiegand, H. Herrmann, H. Lüth, B. Bauer, D. Ollech, A. Niebuhr. IAB-Regional, IAB Nord, 02/2014.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Arbeitsmarktregion Lübeck - Befunde für die Teilregion Stadt Lübeck. with V. Harder, J. Wiegand, H. Herrmann, H. Lüth, B. Bauer, D. Ollech, A. Niebuhr. IAB-Regional, IAB Nord, 01/2014.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Arbeitsmarktregion Kiel - Befunde für die Teilregion Rendsburg-Eckernförde. with H. Lüth, J. Puckelwald, H. Herrmann, D. Ollech, A. Niebuhr. IAB-Regional, IAB Nord, 05/2013.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Arbeitsmarktregion Kiel - Befunde für die Teilregion Neumünster. with H. Lüth, J. Puckelwald, H. Herrmann, D. Ollech, A. Niebuhr. IAB-Regional, IAB Nord, 04/2013.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Arbeitsmarktregion Kiel - Befunde für die Teilregion Kreis Plön. with H. Lüth, J. Puckelwald, H. Herrmann, D. Ollech, A. Niebuhr. IAB-Regional, IAB Nord, 03/2013.

Analyse der Übergänge aus Arbeitsuche in Beschäftigung in der Arbeitsmarktregion Kiel - Befunde für die Teilregion Stadt Kiel. with H. Lüth, J. Puckelwald, H. Herrmann, A. Gillam, A. Niebuhr. IAB-Regional, IAB Nord, 02/2013.

Analyse der Berücksichtigung eines Wanderungsindikators im Rahmen der Abgrenzung des GRW-Fördergebiets - Gutachten für die Gemeinschaftsaufgabe „Verbesserung der regionalen Wirtschaftsstruktur“ im Auftrag des Hessischen Ministeriums für Wirtschaft, Verkehr und Landesentwicklung. with T. Buch, S. Hamann, H. Meier, A. Niebuhr, J. Puckelwald. IAB-Forschungsbericht, 04/2011.

Conferences and Workshops

2016: 3rd Geography of Innovation Conference, Toulouse; IEB/UEA Summer School in Urban Economics for PhD students, Barcelona; IfR Doctoral Seminar, Kiel.

2015: Annual Conference of the “Verein für Socialpolitik”, Münster; Summer Conference of the “Gesellschaft für Regionalforschung (GfR)”, Kiel; ifo Workshop “Arbeitsmarkt und Sozialpolitik”, Dresden; IfR Doctoral Seminar, Kiel and Haderslev.

2014: IfR Doctoral Seminar, Kiel and Schleimünde; Federal Employment Agency, Heide and Lübeck.

2013: NORFACE Conference, London; Winter Seminar of the “Gesellschaft für Regionalforschung (GfR)”, Matri/Osttirol; IfR Doctoral Seminar, Kiel; Brown-Bag-Seminar, Kiel; Federal Employment Agency, Heide and Lübeck.

2012: Annual Conference of the European Regional Science Association (ERSA), Bratislava; Winter Seminar of the “Gesellschaft für Regionalforschung (GfR)”, Matri/Osttirol; IfR Doctoral Seminar, Tellow and Schleimünde; Conference of the “Deutsche Gesellschaft für Soziologie”, Bochum; Federal Employment Agency, Kiel; International Workshop of the Institute for Employment Research (IAB), Nuremberg.

Work in Progress

Agglomeration economies and firm innovation - Evidence for different innovation types, with A. Niebuhr, A. Schmidke.

From rags to riches: The effects of local labor market density on social mobility, with M. Reichelt, A. Haas.

Labour diversity and firm’s innovation - Evidence from Germany, with A. Niebuhr.

Awards

Prize of the “Gesellschaft für Regionalforschung (GfR)” (German speaking section of the European Regional Science Association (ERSA)) for the best presentation of a PhD student at the Winter Seminar 2012.

Faculty award for best diploma thesis in economics (“Erich Schneider-Preis”), 2012.

Take-Maracke-Förderpreis (1st prize) of the “Wirtschaftswissenschaftlicher Club am Institut für Weltwirtschaft” for a seminar paper on the topic “Analysis of the regional labor supply until 2050”, 2010.

Kiel, August 2016
Jan Cornelius Peters

Erklärung zum selbstständigen Verfassen der Arbeit

Ich erkläre hiermit, dass ich meine Doktorarbeit mit dem Titel *The Impact of Age Complementarities and Learning Externalities on Labor Productivity: Evidence from German Micro Data* selbstständig und ohne fremde Hilfe angefertigt habe. Alle von anderen Autoren wörtlich übernommenen Stellen, wie auch die sich an die Gedanken anderer Autoren eng anlehnenden Ausführungen der aufgeführten Beiträge wurden besonders gekennzeichnet und die Quellen nach den mir angegebenen Richtlinien zitiert.

Kiel, den 02.08.2016

(Jan Cornelius Peters)