THE UNIVERSITY OF CHICAGO

ASSESSING THE CITY SIZE EARNINGS PREMIUM IN GERMANY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

BY

GITA KHUN JUSH

CHICAGO, ILLINOIS

DECEMBER 2017

Assessing the City Size Earnings Premium in Germany

Gita Khun Jush

Abstract

I estimate the city size earnings elasticity for Germany, accounting for possible selection of more productive workers into bigger labor markets. The short-run elasticity for Germany suggests that a worker moving from rural areas to medium sized cities can expect his earnings to increase by 6.1 percent on impact. In the long run, workers' earnings in bigger cities increase by even more, due to *exposure effects*: work experience gained in bigger cities increases earnings by more than the kind earned in smaller cities. For instance, working 30 years in medium sized cities, as opposed to the rural areas of Germany, adds an extra 3.1 percent to earnings. To better understand the result, I decompose the differences into the sources of earnings growth: growth through job mobility and on the job. I find that workers in bigger labor markets experience fewer earnings losses when going through unemployment, but also fewer gains on-the-job. Further, the Job-Unemployment-Job rate is lower in bigger cities and Job-Job changes are similar in frequency and earnings change conditional on mobility. Taken together, I interpret the findings as evidence for thick labor market effects leading to lower unemployment risk for workers. Learning mechanisms, as sources for agglomeration economies, are likely less relevant due to the lower on-the-job earnings gains in bigger cities.

This study uses the weakly anonymous Sample of Integrated Labour Market Biographies (Years 1975 - 2010). Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access.

Contents

List of Figures

List of Tables

1 Introduction

In this paper, I analyze earnings differences between German cities, using nominal, average quarterly earnings as the measure and investigate agglomeration mechanisms working through the labor market. I use linked employer-employee data from workers in West Germany to confirm the City Size Earnings Premium (CSEP). I use full-time employed workers to account for possible differences in work hours, since detailed information on hours worked is not available in my dataset. Following the literature, I assess differences in nominal earnings, not adjusted for differences in the cost of living, to measure agglomeration economies.[1]

I exploit the panel dimension of the data in several ways. First, at each point in time the location of work is reported, such that I can construct variables for work experience that are location specific. Further, I can control for unobserved, time invariant worker characteristics that might bias the results, by including worker fixed effects in the estimation. The result is a moderate, positive elasticity of earnings with respect to city size of 0.0089, which is smaller than a similar result by Roca & Puga (2017) for Spain of 0.0223. To understand the number, I compare earnings in rural areas with around 140,000 inhabitants in the bottom quartile of Germany's city size distribution, to medium sized cities with 1.1 million inhabitants in the top quartile. The elasticity translates to 6.1 percent higher earnings, on impact of moving from the bottom to the top quartile of the city size distribution (*impact effect*).

The CSEP estimation also reveals that the estimate should be viewed as a short-run elasticity. Experience gained in bigger cities leads to larger earnings growth than experience gained in smaller cities. Hence, the more work experience is gathered in bigger cities, the more earnings respond in the long run. I call this the *exposure effect*. The exposure effect is estimated to be economically significant, and can amount to 50 percent of the short run gains. A worker, who gains 30 years of work experience in the top quartile, as opposed to the bottom quartile of the city size distribution, ends up with 3.1 percent higher earnings, solely due to the exposure effect. Taken together, the worker will have 9.2 percent higher earnings after 30 years in the top quartile.

To understand possible mechanisms leading to the finding, I decompose overall earnings gains into their possible sources. The data allow me to decompose overall earnings growth into changes when a worker changes jobs, i.e. at instances of job mobility, or on-the-job. I distinguish between changing jobs, when going through unemployment and changing jobs directly, without intermittent unemployment. The differences in earnings patterns by city size can be summarized in three stylized facts:

1. Earnings losses due to unemployment spells are smaller in bigger cities. This is partly due to a lower likelihood for workers in bigger cities to change jobs; in particular, changing jobs with intermittent unemployment is less likely in bigger cities.

2. There are no statistically significant differences between big and small German cities regarding patterns of job mobility without intermittent unemployment. Neither the rate of job-job

[1] See next section for more details on the use of nominal vs. real earnings measures.

mobility, nor the earnings change conditional on it differ by city size.

3. On-the-job earnings growth is smaller in bigger cities, undoing part of the unemployment related differences.

I discuss five mechanisms leading to agglomeration economies, in light of my empirical findings. First, I assume a fixed relationship between productivity and earnings and discuss how well higher match quality, reduced human capital losses due to fewer unemployment spells, labor pooling and learning can explain the described observations. These mechanisms are commonly discussed in the literature on Urban Economics. But the relationship between productivity and earnings might itself be influenced by city size, especially in a frictional labor market where employers have monopsony power. I end by addressing how monopsony power can influence the different explanations. A more detailed discussion can be found in section (5.5).

First, thick labor market effects could lead to better equilibrium match qualities in bigger cities. The empirical findings are consistent with theories of higher match quality in bigger cities. There are several reasons for higher match qualities in bigger cities. If search frictions are reduced it is easier for firms and workers to find each other. Better matches may also result from higher levels of specialization of workers and firms. Further, better matches can be a consequence of more intense searching by workers, due to higher returns to search in bigger markets. Higher idiosyncratic match quality leads to less labor market mobility since workers and firms find the best match quickly. The lower mobility rates are in line with this finding, as are the higher overall income levels. It is less clear why better match qualities, alone, would result in less on-the-job earnings growth, however.

Second, human capital loss during unemployment spells can lead to differences in the accumulation process across local labor markets. Work experience is included in wage regressions, to reflect that occupation specific human capital levels of workers increase as workers gain more experience. But human capital also depreciates, when workers are unemployed. If workers are less unemployed in bigger cities, the same amount of work experience amounts to higher occupation specific levels of human capital, since less of it has depreciated. Hence, earnings will be higher, because workers are more productive. My finding of lower mobility rates, especially due to unemployment, is consistent with this mechanism. If the depreciation rates in big and small cities are comparable, more unemployment would results in lower earnings. Controlling for the origin of experience accounts for the periods of unemployment the worker has experienced. Lower levels of earnings growth in bigger cities, may then be related to decreasing returns effects of human capital: at lower human capital level, human capital grows more steeply. These predictions are in line with the findings I present.

Third, thick labor markets allow for labor pooling, one reason for lower unemployment rates in bigger cities. This mechanism has difficulties rationalizing the empirical findings. According to this theory, employers in bigger cities have access to a flatter labor supply curve, allowing less costly adjustment of their labor force in case of firm specific demand shocks. If negative idiosyncratic shocks are similarly common in big and small markets, labor pooling mechanisms would lead to more worker mobility in bigger markets. Labor pooling is named as one of the reasons for lower unemployment rates in bigger cities. While my findings regarding unemployment are in line with this theory, the overall lower likelihood of changing jobs in bigger cities is inconsistent.

Fourth, it has been suggested that thick labor markets lead to increased learning by workers, through knowledge spillovers. This mechanism is inconsistent with the empirical work I present. Suppose that learning increases the marginal productivity of workers and that employers compensate workers for increased productivity in similar ways across local labor markets. Steeper human capital growth should be measurable as steeper earnings growth on the job. If this were the true reason why work experience is worth more in bigger cities, then controlling for the source of the work experience amounts to controlling for the speed of human capital accumulation across locations. The finding that on-the-job earnings growth is lower in bigger cities, is inconsistent with this suggested mechanism. Further, this explanation would have to be combined with other assumptions regarding the specificity of human capital, to be able to account for lower mobility rates.

Finally, in frictional labor markets, monopsony power of firms likely changes with city size, affecting how worker productivity is mapped into earnings. In light of differences of monopsony power, the impact and exposure effects I measure, can also be interpreted in an alternative way. First, higher monopsony power of firms in smaller cities can translate into lower bargaining power of workers, because workers have fewer outside options. Hence, even if productivity were the same across cities of differing sizes, workers in bigger cities can earn more, because they can bargain up the portion of match surplus they receive. Second, higher experience and specialization levels increase monopsony powers of firms, since higher specialization reduces the attractiveness of firms that would have been otherwise attractive to a worker. This mechanism by itself could lead to measured exposure effects, since increased experience in bigger cities would result in less bargaining power for the worker, and hence a smaller share of match surplus. This mechanism can explain the observed earnings patterns; it is more difficult to rationalize the observed mobility patterns.

To summarize, the empirical findings particularly support two groups of mechanisms commonly suggested in the literature; they are inconsistent with simple learning and pooling mechanisms. Mechanisms related to labor pooling and match quality have some power in explaining mobility patterns; the differences in on-the-job earnings growth in bigger cities, however, remain to be explained.

Before explaining the data in section (3), I discuss the related literature. I present the estimation of the CSEP in section (4). The decomposition of the CSEP into job mobility and on-the-job earnings change are discussed in section (5).

2 Agglomeration Economies: Mechanisms & Evidence

A large literature has evolved around measuring the productive advantages of bigger cities, by quantifying the City Size Earnings Premium (CSEP), i.e. the premium in nominal earnings a worker can expect upon moving to a bigger city (see for example Ciccone & Hall (1996), Combes *et al.* (2008) or Baum-Snow & Pavan (2011)). Different productivity measures have been used, to establish the correlation between city size and productivity. Combes *et al.* (2012) show, using balance sheet and capital data of firms in France, that firms in bigger cities have higher TFP.

Agglomeration economies are a possible reason for the productivity differences between big and small cities. Many economic interactions are subject to increasing returns to scale, which could lead to higher productivity levels when more individuals interact. Such effects of city size on productivity are referred to as agglomeration economies, without further specifying the mechanisms by which they arise.

Agglomeration economies leading to productivity differences between locations can persist in spatial equilibrium and can be measured, using nominal earnings. In the simplest spatial equilibrium model, local welfare levels are equalized across cities, such that there are no incentives to move between labor markets. A measure of worker welfare is the real wage: the wage paid locally, adjusted by the cost of living. We expect real earnings to be equalized across different local labor markets. Since the cost of living, due to the prices of non-traded goods (housing and services) is higher in bigger cities, workers have to be compensated with higher earnings in spatial equilibrium. Therefore, differences in nominal earnings are not necessarily eliminated by workers. This finding is supported by data; for example, in the US, Glaeser *et al.* (2008) show that dispersion in real earnings is indeed much smaller than the dispersion in nominal earnings.

Yet, the question, of why firms in bigger cities are willing to pay higher nominal earnings, remains. If firms are mobile and there are no productivity differences between cities, there are strong incentives for firms to move location, until differences in nominal earnings are eliminated. As discussed by Moretti (2004), productivity differences across locations are offset by higher nominal earnings of the same size. They can be used to assess agglomeration economies.

Using nominal earnings as the measure for agglomeration economies can lead to biased results, because nominal earnings also reflect productivity differences due to reasons other than agglomeration mechanisms. Besides agglomeration, nominal earnings reflect the selection of more productive firms through tougher competition and the sorting of higher skilled workers into bigger cities. Combes *et al.* (2012) distinguish between higher firm productivity in cities due to competitive selection and agglomeration forces; they find little evidence for selection in bigger markets. But there is ample evidence for sorting of more skilled workers into bigger markets (see for example Combes *et al.* (2008), Behrens *et al.* (2014)). While it is often possible to control for a large range of observed worker demographics, some characteristics, such as innate worker productivity, remain unobserved and can bias the results.

Glaeser & Mare (2001) and Combes *et al.* (2010a) are two studies assessing the CSEP using worker fixed effects. The fixed effects control for unobserved and time invariant worker characteristics. Both papers, and follow-up papers using the same approach, find a significant drop in the estimated CSEP when worker fixed effects are included. The drop is 35% for France (Combes *et al.* (2010a)), 47% for Spain (Roca & Puga (2017)) and 66% for Italy (Mion & Naticchioni (2009)). The drop is generally interpreted as evidence for sorting of more productive workers into bigger cities, that can not be controlled for, using observed worker demographics. I estimate a drop of 54% for Germany.

More recently, Roca & Puga (2017) challenge this interpretation. Unobserved worker productivity can emerge as workers interact with bigger labor markets. In this case, worker fixed effects confound worker differences in productivity that are innate, with those emerging when some workers work in bigger cities. The drop, when including worker fixed effects over-controls: it assigns some of the variation due to agglomeration effects to worker selection. This can be remedied by controlling directly for the differences in where the workers' labor market experience was gained. Such measures are rarely available in aggregate datasets that do not allow following workers over time, but can be calculated using linked employer-employee datasets, such as the German dataset I use for this study. When including location specific work experience, the drop in the estimate of the CSEP, when including worker fixed effects is greatly reduced from 54% to merely 9%. I conclude that the effect of unobserved sorting on the CSEP is mild, once the effects of working in bigger cities are accounted for.

In general, the micro-foundations of agglomeration forces are classified into three types: sharing, matching and learning mechanisms. Sharing is related to the availability of a large variety of intermediate inputs and proximity to final consumers; matching mechanisms are also sometimes referred to more loosely as thick labor market effects; and learning mechanisms are related to knowledge spillovers. Each of the mechanisms leads to higher city level productivity, the bigger the city. Depending on the mechanism, aggregate productivity might be correlated with population size or with population density. If the mechanism relies on the number of workers interacting, population will be the measure of agglomeration at the macro level. If the average distance between workers is crucial, population density matters more. Duranton & Puga (2004) and Moretti (2011) provide useful overviews of the three groups of micro-foundations.

At the core of sharing mechanisms are the models of New Economic Geography, emanating from Krugman's work on models of variety and increasing returns (for example Krugman (1991)). Bigger cities can support a larger variety of non-traded, specialized intermediate inputs. Final good producers have higher productivity if a larger variety of intermediate inputs is available; each intermediate good producer provides a different variety to the market, paying a fixed cost of production. Since intermediate firms have to recover the fixed cost, if they enter the market, the scale of the market determines how many firms can be supported in equilibrium. Hence, the more consumers there are, the more varieties can be produced, and the more productive the final good producer. Note that this mechanism can also lead to higher levels of specialization among firms, as the number of consumers increases. Such mechanisms lead to higher average productivity of firms in bigger cities, and are part of the average city effect I measure.

Thick labor market effects are one of the earliest micro-foundations of agglomeration, as discussed by (Marshall (1890)). Whenever search frictions are an important feature of interaction on the market, the thickness of the market might play a role. Market thickness affects two features of matches: match quality and the likelihood with which matches form. The two factors are not completely independent: higher likelihood of matches can lead to better equilibrium matches, especially when workers can search for new jobs while working. Higher match quality can translate into higher earnings and longer tenure; higher likelihood protects firms from loosing income due to unfilled vacancies and workers from unemployment.

There is evidence that thick markets can affect the match quality in markets where search is important. Ngai & Tenreyro (2014) show that housing market thickness fluctuates seasonally, affecting housing prices and the match quality of houses to buyers. A large literature tries to establish increasing returns in the matching function, i.e. that the likelihood of matches increases as the number of workers and firms increases. Petrongolo & Pissarides (2001) survey the literature, concluding that there is little evidence for increasing returns. Petrongolo & Pissarides (2006) follows up on the survey, showing that thick markets, also affecting equilibrium match quality, can make workers choosier in their acceptance decision when being offered a job. The endogenous match quality response of workers can compensate for the lower search friction. At the aggregate level, the matching function will look like a constant returns function, but earnings will reflect market thickness effects. Petrongolo & Pissarides (2006) find supporting evidence for the hypothesis in the UK: differences in re-employment earnings, after workers experience unemployment, are higher in London than in smaller UK cities. Bleakley & Lin (2012) find evidence that young workers are more likely to change occupation and industry in thicker US labor markets; on the other hand, older workers, are more likely to stay in the same occupation and industry, the bigger the labor market. The finding supports the notion that thicker markets allow for job shopping early on in the career, but lead to better matches and job stability as workers gather more work experience.

Finally, knowledge spillovers, even though difficult to measure, are often used as explanation for higher productivity levels in bigger cities. The basic idea is that human capital benefits individuals, other than the owner. Moretti (2004), for instance, uses plant level data to compare differences in the productivity gains of plants in cities with a large inflow of high-skilled workers, to the productivity gains in cities with a small increase in high-skilled workers. After controlling for plant level skills, he finds that plants gain more, the bigger the influx of high-skilled workers. Other work has used the likelihood of citations in big and small cities to get at spillovers (for instance Carlino *et al.* (2007)).

The focus in this paper is on thick labor market effects. I measure the concepts of match likelihood and match quality, using some unique features of the employer-employee dataset and find evidence suggesting some of the discussed thick labor market mechanisms are at play. First, a measure for equilibrium match quality in the local labor market is the duration of a match (see Hagedorn & Manovskii (2013) for an application to business cycles). Hence higher job mobility can be a sign of lower match qualities, especially if coinciding with lower earnings. I estimate job mobility and earnings changes at such mobility instances to assess if match quality is likely to differ between labor markets of different sizes. I find differences, that I interpret as thick market effects: mobility rates are lower in bigger cities, while earnings are higher.

Second, the data allow me to distinguish between mobility with and without unemployment. Therefore, I assess if job stability is a feature of thick labor markets. I find that workers in bigger cities experience unemployment less frequently, leading to smaller earnings losses due to unemployment. Additionally, the longer workers stay in bigger markets, the less likely they experience unemployment, again in line with thick market effects discussed in the literature.

Finally, I also estimate on-the-job earnings growth. On-the-job earnings growth has been linked to human capital accumulation on the job (see for instance Bagger *et al.* (2011)). If thick markets lead to more learning, we would expect on-the-job growth to reflect such thick market effects. In particular, better match quality might affect learning on the job. I do not find such effects; on the contrary, on-the-job earnings gains are smaller in bigger cities. This finding is better explained by the ability of employers in bigger cities to ensure workers against unemployment.

3 Data

The labor market data stems from a German linked employer-employee dataset. I lay out the measures for describing labor market biographies in detail and describe sample restrictions. The geographic unit of analysis is the commuting zone (CZ). I conclude the section by presenting summary statistics for the main variables of the analysis.

3.1 Labor Market Data

3.1.1 The Sample of Integrated Labour Market Biographies (SIAB)

The analysis is based on a German linked employer-employee dataset, the *Sample of Integrated Labour Market Biographies* (SIAB). The SIAB is a two percent, random sample drawn from the universe of all employment subject to social security in Germany.[2] The data cover full-time and part-time employment in West Germany since 1975, and in the re-unified Germany since 1990. Once a worker is selected into the sample, all her labor market records are included in the SIAB, allowing tracking of her entire labor market history, back and forward in time. Further, if a worker is employed, the data record the unique, temporally consistent establishment identifier of the employer. The data come at a daily precision, in spell form.

One limitation of the data is that the SIAB excludes self-employment, regular students and civil servants. Therefore, intermittent employment of a worker in these categories will show up as a missing observation. I refer to these observations as "out of sample observations."

Out of sample observations are not to be confused with unemployment. I define unemployment as an instance of a worker not linked to an establishment, where she receives unemployment benefits and is registered as searching for employment. This definition is consistent with the official definition of unemployment of the Federal Employment Agency.

Aside from the unique worker and establishment identifiers, each employment spell also reports the average daily earnings for the worker. For the duration of the spell, the employer reports the number of employment days and the total compensation for the spell. The information is translated into a value for average daily earnings in the SIAB. Hence, if the compensation scheme is altered, a new spell between the same worker and establishment will appear in the dataset. To be able to compare earnings over time, I use the CPI adjustment provided by the Federal Statistical Office.

[2]The data are provided by the Research Data Center of the German Federal Employment Agency.

The data report establishment characteristics, demographics of workers and information on the match. On the establishment side, I use information on industry, occupation of the job, location and tenure of employment. I use industry indicators at the 2-digit level, corresponding to the ISIC 2-digit classification. I use occupation indicators at the 1-digit level following the classification of occupations by Blossfeld (1983). Blossfeld (1983) constructs broad occupational groups that are easy to interpret; occupations are classified such that the range of activities conducted is similar, and the skill intensities are homogeneous. The location of the establishment is known at the *Kreis*-level.[3] I use the information to define the location of the worker as his work location. Finally, the spell data allow me to calculate the tenure of a worker with a particular establishment.

On the worker side, I observe education, age and gender. I work with three levels of education: high-school degree, vocational degree and college degree. Due to the way the data are presented, I can further calculate the precise work experience of each worker in the sample, back to 1975. I use information on the total work experience of the worker; since the location of each job is known, I also generate work experience variables, specific to the CZ population categories defined below, depending on where the work experience was gained.

3.1.2 From Spell Data to Quarterly Dataset

Using the spell data, I generate a quarterly dataset. Beside the worker and establishment information, I generate indicators for Job-Job (JJ), Job-Unemployment-Job (JUJ) , and on-the-Job (OTJ) earnings growth events. I define a JJ event as an instance of establishment ID change for the same worker, without intermittent unemployment. A JUJ event is the same as a JJ with intermittent unemployment. At OTJ events both establishment and worker IDs are unchanged, but the earnings of the match change.

To get the quarterly earnings measure, I add up all earnings that fell within the quarter and divide by the number of employment days. If any of the events - JJ, JUJ or OTJ - takes place, I generate a quarterly dummy variable recording that. Finally, conditional on an event I record the earnings change associated with the event.

3.1.3 Sample Restrictions

I restrict the sample in several ways.

- Full-time and main employment: The dataset does not record the hours worked. In order to track comparable employment across different CZs I restrict the sample to full-time employment. If a person holds several jobs at a time, I choose the main employment, as the highest paying job.

- Time period 1995-2005: First, the geographic unit at which the SIAB is available is not economically meaningful. As explained below, I rely on the calculations of Eckey *et al.* (2006) to aggregate the *Kreise* to Commuting Zones (CZs). Since their analysis uses data from 2004

[3] Kreis is a governmental unit, see below for more detailed explanation.

I choose a window including that year. Second, to avoid confounding effects related to the reunification of Eastern and Western Germany, unrelated to long run agglomeration forces, I start the analysis five years after reunification. Finally, to avoid confounding effects related to the run-up of the 2007/08 recession the analysis ends in 2005. Together, these restrictions lead to the window of analysis.

- Western Germany: Differences in the trend and level of both employment and earnings between Western Germany and the former GDR still exist; employment and earnings remain significantly lower in the former GDR. These differences are mostly unrelated to agglomeration forces. To analyze comparable workers, I focus on Western Germany.
- Male workers: The life cycle patterns of earnings and employment differ greatly by gender. Labor market biographies of women are significantly changed around the ages of child bearing. Further, part-time and marginal employment constitutes a much larger portion of female employment, which I exclude due to the lack of information on hours worked. To avoid confounding differences stemming from the limitation of the data with differences in mobility patterns of women, I focus on male employment only.
- Ages 20-58: Early retirement in Germany starts in the late 50s. To avoid confounding effects due to differences in labor force participation I cut the sample at age 58. Similarly, the work life starts around the age of 20.[4]

3.2 Commuting Zones (CZs)

The geographic unit of analysis is the Commuting Zone (CZ) introduced by Eckey *et al.* (2006). Their classification, considering that many German, administrative datasets, including the SIAB, are published at the *Kreis*-level, provides a crosswalk for aggregating data at the *Kreis*-level to the CZ-level. *Kreise* are historically and politically determined units, often cutting city centers from their surrounding suburbs or not reflecting that formerly distinct *Kreise* have grown into one labor market. The CZ classification is a partition covering the entire area of Germany. Using data on the commuting patterns across different *Kreise*, Eckey *et al.* (2006) cluster Kreise to CZs such that the resulting CZs are not connected by significant commuter flows. They impose two constraints: commuting times may not exceed a threshold of 45-60min (depending on the population of the *Kreis*) and no CZ has a population below 50,000. The classification leads to relatively large CZs in large economic centers like Berlin, Munich and the Ruhr valley. A map of the CZs is shown in Figure (3.1).

The classification has two main advantages. First, it is one of the two most frequently used classifications for spatial, economic analysis of German data. This makes my study comparable to other studies. Other classifications, not using commuter flows, are better suited for analyses where political boundaries are more important, for instance studies on public investments (see Eckey *et al.*

[4]The exception are workers with a college degree, who usually enter the labor market at the age of 25. For workers with a college degree the age restriction is 25-58.

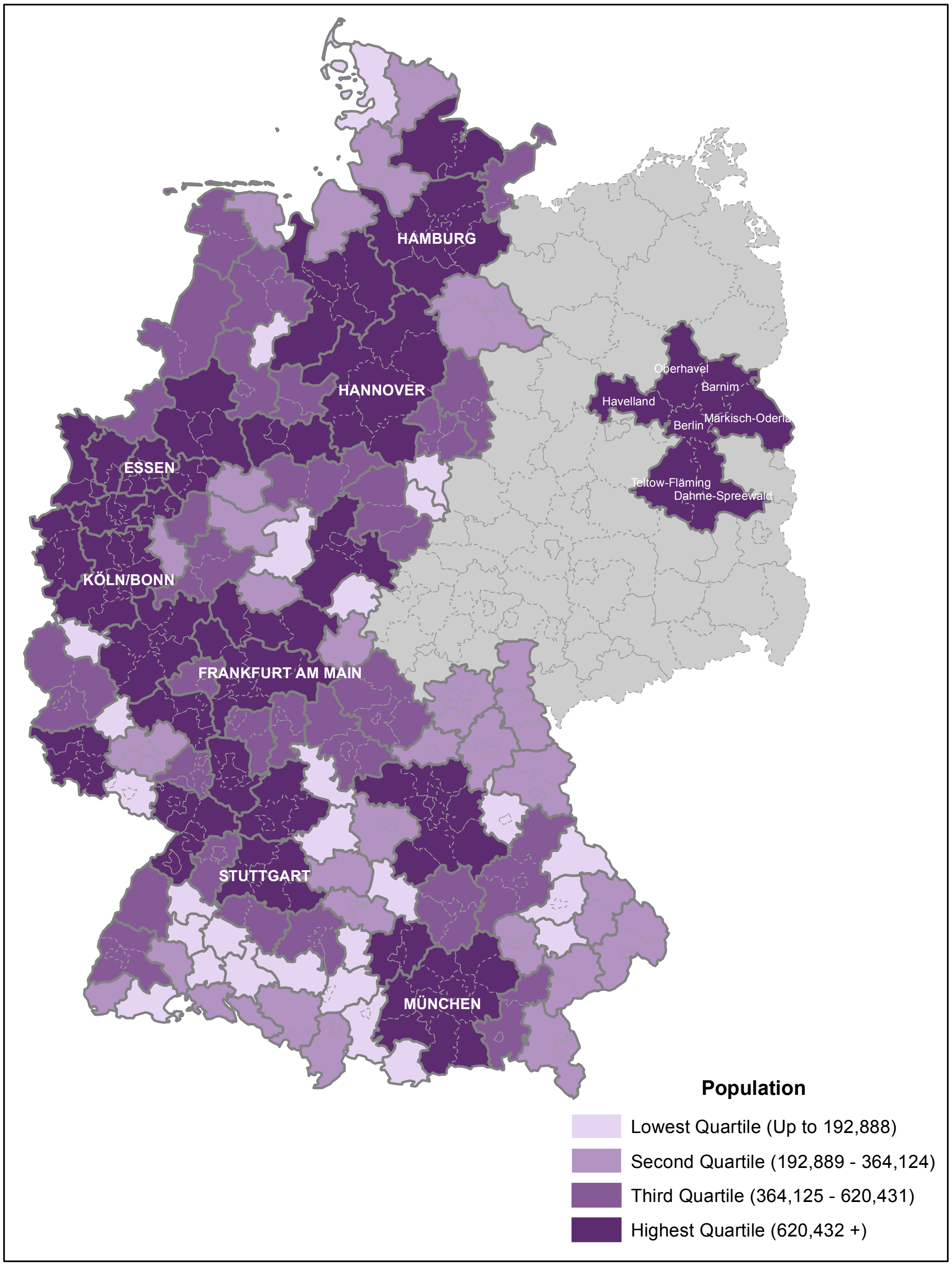


Figure 3.1: Geographic delineation of local labor markets in West Germany and 2005 population levels; thick, grey lines are CZ boundaries, dashed lines are Kreis boundaries; gray areas are excluded from the analysis.

(2007)). Second, the classification uses commuter data from 2004, which lies within the period of analysis of this study (1995-2005). Previous classifications rely on data from 1992. Since 1992 constitutes a period of transformation in Germany and is outside of my period of analysis, I use the more recent classification.[5]

I use two measures of CZ size from Germany's Federal Statistical Office, population and population density. Population and population density are positively correlated at the CZ level (correlation coefficient of 0.56 in 2005, and steady over time), yet the city size earnings premium and the mechanisms yielding it are affected differently by either measure. Population is a measure of the size of overall economic activity, population density reflects the average distance between economic actors. Conceptually, population is a better measure for agglomeration economies if the main mechanisms are related to increasing returns in population size, i.e. the number of economic agents interacting determines the magnitude of agglomeration economies; we refer to this type of agglomeration economies as "urbanization economies". Population density matters more if the proximity of economic agents is the main driver of agglomeration economies. Population density has been criticized in the past as a measure of agglomeration economies, since it varies significantly depending on the geographic unit; this is particularly true for Germany's CZs (see Figure (3.2) below). More recent papers have reverted to using population density, but at smaller geographic units, reflecting the distance between economic agents at the city's core. To reflect both possible mechanisms, I use the population of the CZ as well as the population density of the most densely populated *Kreis* of the CZ as measures of agglomeration.

The distribution of log population and log core population density are shown in Figure (3.3). There are 113 CZs in the sample. CZ population ranges from 63,801 in Daun (Rhineland-Palatinate) to 4,436,046 in Berlin. For some of the analysis below I divide the CZs in four groups depending on population (CZ quartiles). The red lines in Figure (3.3) indicate the cut-off for each group. The areas with the highest core densities are in the Ruhr valley (Essen, Dusseldorf, Hagen, etc.) and these areas are also among the most populous.

3.3 Summary Statistics

A large literature documents the selection of particular kinds of establishments and workers into bigger or more densely populated cities (for instance Baum-Snow & Pavan (2011) or Combes *et al.* (2008)). I address this issue in the analysis below. Here, to better understand how CZs of different population sizes differ in terms of observable worker characteristics, I present basic summary statistics of the key explanatory variables I use. The statistics also help better understand the data and the sample I selected. The sample I end up with has 9,300,000 person-quarter observations. It covers the years 1995 to 2005, with roughly 900,000 observations per year. The following tables

[5] The classification has one main limitation, relevant to this study. The population floor at 50,000 may be too restrictive. It may lead to artificially big regions, especially in rural parts of Germany. I am interested in measuring agglomeration economies, i.e. the effect market size has on labor market biographies of workers. Therefore, if truly rural areas are included as part of a bigger CZ, due to the size restriction alone, this might wash out some effects of agglomeration.

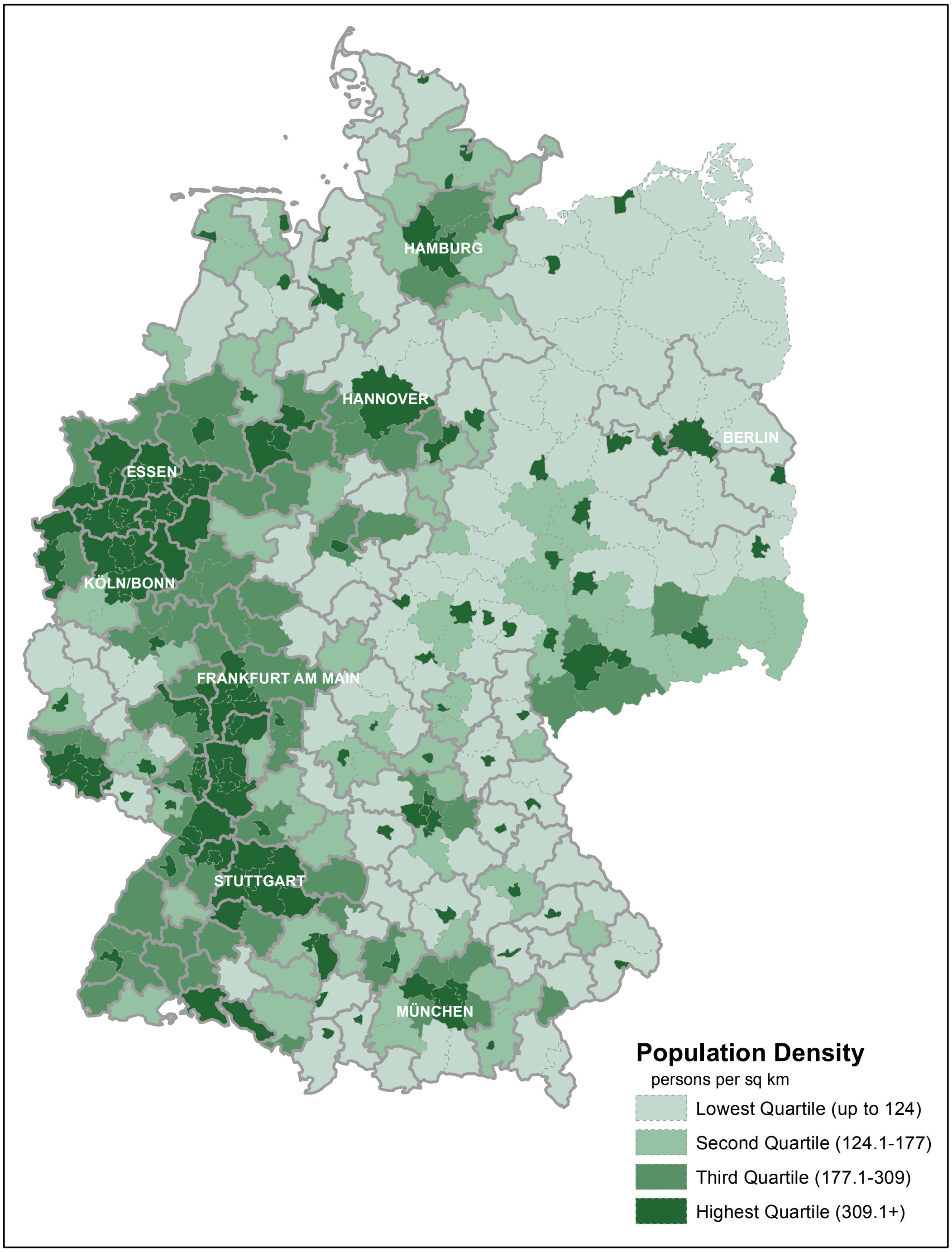
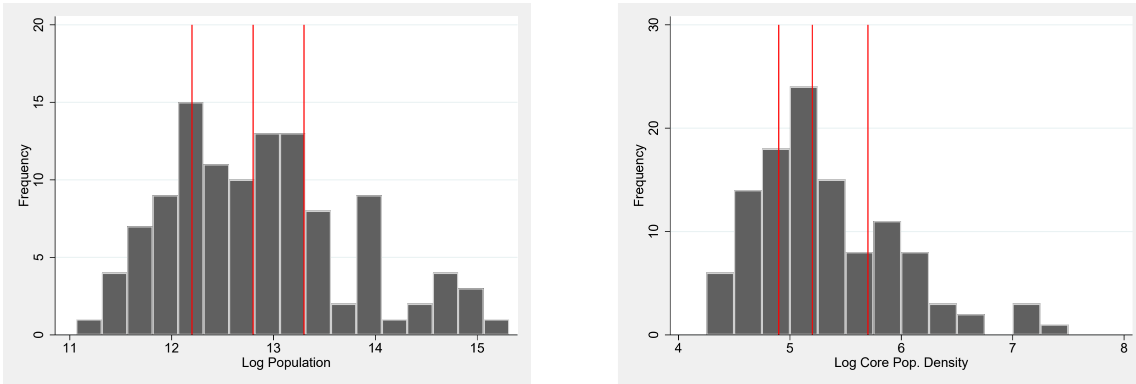


Figure 3.2: Population density at the Kreis-level with thick gray CZ boundaries; CZs consisting of several Kreise, tend to have a core Kreis with much higher population density.



(a) median = 12.7; inter quartile range= 1.2

(b) median = 5.1; inter quartile range= 0.8

Figure 3.3: Log population and core population density in 2005 by CZ

present summary data for the year 2005.

I classify workers into three education categories: high-school or less (low), vocational training (medium) and college degree or more (high). Germany has different types of high-schools which last 9, 10 and 13 years, respectively. I bundle the three high-school types. The biggest education group on the German labor market are workers with vocational training. Most vocational degrees can be acquired with either of the high-school degrees and they span a range of industries and occupations, with varying skill intensity. As shown in Table (3.1), almost 3/4 of the German working population have vocational degrees. The share of workers with vocational degrees is larger in smaller CZs. Instead, bigger cities have larger populations of workers with low and high education levels. This pattern, that medium skilled workers are relatively more present in smaller cities has also been reported in Eeckhout *et al.* (2010) and Eeckhout *et al.* (2014), for instance.

Table 3.1: Number of workers in 2005, in each education group, by population size quartile.

	Education Distribution		
	all CZs	CZs low pop.	CZs high pop.
high-school or less	11.7%	10.7%	12.2%
vocational training	74.2%	81.5%	71.6%
college degree	14.1%	7.7%	16.2%
N	228,096	12,570	147,811

Blossfeld (1983) classifies occupations into three broad categories - Production, Service and Administration occupations - which are easy to interpret and have been shown to have explanatory power for labor market dynamics related to gender, inequality, and social and economic mobility[6] (see Hout & DiPrete (2006)). More than half of workers in Germany work in production related occupations. Production occupations are much more common in smaller CZs, consistent with trends in many industrialized countries, where expensive, specialized labor can be economized on

[6]The 1987 classification is updated by the Research Data Center of the German Federal Employment Agency, using the German Mikrozensus.

by moving manufacturing firms out of the city. About a fifth of workers hold service occupations. Service occupations are more concentrated in bigger CZs. The remaining quarter of occupations are in administrative occupations, which, similar to service occupations, are more prevalent in bigger CZs. The details are listed in Table (3.2).

Table 3.2: Number of workers in 2005, in each occupation group, by population size quartile.

	Occupational Distribution		
	all CZs	CZs low pop.	CZs high pop.
Production			
Agricultural occupations	1.5%	1.6%	1.5%
Unskilled manual occupations	18.4%	24.9%	16.2%
Skilled manual occupations	23.3%	29.9%	21.1%
Technicians	6.8%	6.6%	6.8%
Engineers	5.1%	3.1%	5.8%
Services			
Unskilled services	14.3%	13.3%	14.6%
Skilled services	2.0%	1.4%	2.3%
Semiprofessions	2.5%	2.0%	2.6%
Proffesions	1.6%	1.0%	1.7%
Administration			
Unskilled commercial/administrational occ.	4.0%	3.0%	4.3%
Skilled commercial/administrational occ.	16.6%	10.9%	18.5%
Managers	3.3%	1.9%	3.9%
N	220,823	12,236	142,943

Table (3.3) summarizes employment in Germany by industry of the establishments. The most pronounced differences in industry composition between CZs with high and low population are in Mining/Manufacturing, Construction and Real Estate. While Real Estate is more present in high population CZs, Mining/Manufacturing and Construction are more prevalent in lower population CZs.

Finally, Table (3.4) shows average experience and tenure. The average experience in the sample is 12.7 years, and is slightly declining with increasing CZ population. Similarly, average firm tenure is 5.8 years and also declining in CZ population, already hinting at differences in job mobility patterns across CZs of different population.

4 City Size Earnings Premium

Working in bigger and more densely populated cities is associated with earning more. In this section, I parse out how much of the differences in average earnings between CZs are due to sorting of higher productivity workers into bigger CZs and how much is due to agglomeration forces increasing productivity. I first follow a common strategy in the literature of estimating average CZ earnings

Table 3.3: Establishment characteristics of workers in 2005, by industry and population size quartile.

	Industry Distribution		
	all CZs	CZs low pop.	CZs high pop.
Agriculture	1.2 %	1.3%	1.1%
Mining/Manufacturing	36.9%	50.3%	32.3%
Energy/ Water Supply	1.3%	0.9%	1.4%
Construction	8.9%	11.3%	8.2%
Trade/ Foodservice Industry	15.1%	13.1%	15.9%
Transportation	7.3%	5.0%	8.1%
Finance	3.4%	2.1%	4.0%
Real Estate	13.4%	5.8%	15.8%
Public Administration	9.4%	8.7%	9.5%
Administration	3.3%	1.6%	3.8%
N	213,451	11,926	137,883

Table 3.4: Average experience and tenure in 2005, by population size quartile.

	Experience	Tenure
quartile 1 (smallest)		
mean	13.0	6.3
sd	8.3	6.7
quartile 2		
mean	12.7	5.9
sd	8.2	6.4
quartile 3		
mean	12.9	6.1
sd	8.3	6.6
quartile 4 (biggest)		
mean	12.7	5.6
sd	8.2	6.3
Total		
mean	12.7	5.8
sd	8.2	6.4
N		2649,924

using worker fixed effects. Including worker fixed effects amounts to controlling for time invariant, unobserved worker productivity. Upon including worker fixed effects in the estimation, the city size earnings premium declines from 0.0214 to 0.0098 (see Table (4.2)). The interpretation is that innately more productive workers sort into bigger markets. [interpret the number]

But not all unobserved productivity differences are time invariant. In particular, the unobserved productivity of workers can be related to different labor market experiences. I find evidence that exposure to bigger labor markets increases earnings and changes worker productivity over time. In that case, controlling via fixed effects alone does not eliminate the bias. It falsely attributes effects of the market to workers, inflating the estimate, unless unobserved worker productivity is accounted for.

A way to test the relevance of exposure effects is to include variables, measuring the experience gained in big and small CZs, as opposed to including overall work experience, only. The linked employer-employee data allow me to do so. Therefore, I assess if exposure to bigger labor markets, as measured by big CZ work experience, has an effect on earnings, beyond the expected effect of total work experience. I show workers tend to earn more, the more experience they gain in bigger CZs. The finding changes the interpretation of the earnings premium; it becomes the premium that is gained by workers on impact of moving to a bigger CZ. Over time, however, the long run elasticity is likely higher, since working in bigger cities increases worker productivity.

The estimation relies on the labor market experiences of workers who move between CZs. To control for time invariant, unobserved worker effects I compare earnings from big and small CZ employment of the same worker. If movers are otherwise similar to stayers, this provides a quasi experimental set-up to assess how much of the earnings differences are truly due to CZ size. However, there is evidence that this is not true for all workers. Movers tend to have higher estimated innate productivity. But the differences are moderate for worker groups that have higher moving rates, suggesting that moving costs might be important to understand the patterns.

4.1 Descriptive Findings

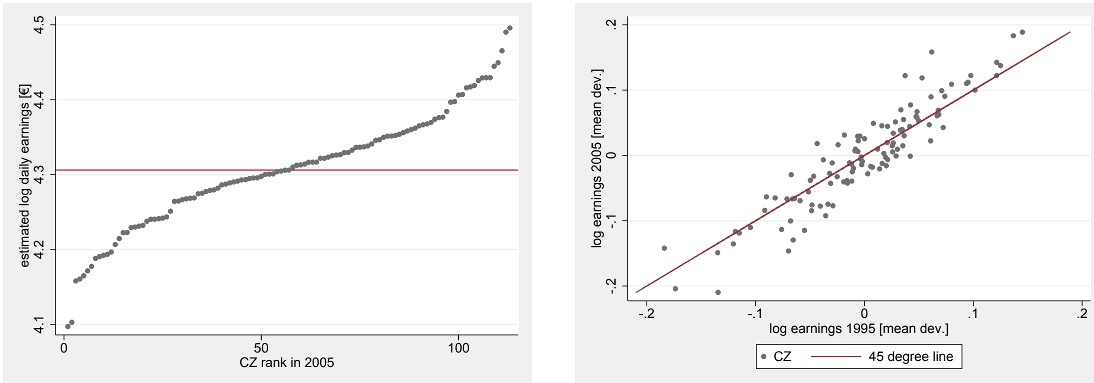
Unconditional, average earnings vary greatly and systematically across CZs in Germany. High earnings CZs remain so over time. CZs with higher earnings yield higher earnings for all education groups. The differences in earnings are correlated with CZ population size, with an elasticity of 0.0216 in 2005. I show evidence, suggesting that earnings differences across CZs are related to exposure to bigger labor markets.

4.1.1 Systematic variation in average earnings across CZs

The difference between the highest and lowest earnings CZ amounts to 0.40 log-points, roughly the magnitude of the average differences in earnings between workers with low and medium education levels. Figure (4.1a) shows the variation in average log earnings, by plotting the CZs' earnings against their earnings rank in 2005. The Figure shows large variation in unconditional, average log daily earnings across CZs in Germany.

The differences in earnings across CZs are also persistent. CZs with high average earnings in 1995, remained so in 2005. To see this, I plot the deviation of 2005 earnings from their mean against 1995 mean deviations of earnings (Figure (4.1b)). The CZ earnings line up along the 45 degree line. The correlation is 0.927 (Table (4.1a)). Agglomeration forces are able to explain slow moving differences between locations.

The differences in earnings across CZs are systematic by education: CZs with high average earnings for workers with vocational degrees tend to yields higher earnings for workers with high-school and college degrees as well. I plot the deviation from their mean of high-school and college earnings against earnings of workers with vocational degrees on the x-axis (Figure (4.2)). The CZs again line up along the 45 degree line, better so for college graduates than for workers with



(a) Variation in log earnings in 2005, each CZ is a dot. The horizontal line at 4.3€ marks earnings of the median CZ.

(b) Auto-correlation of average CZ earnings (deviations from the year's mean) between 1995 and 2005.

Figure 4.1: Variation and auto-correlation of CZ earnings in Germany.

vocational degrees. Hence it is at least plausible that differences in earnings might not be purely due to selection warranting a deeper investigation of the differences.

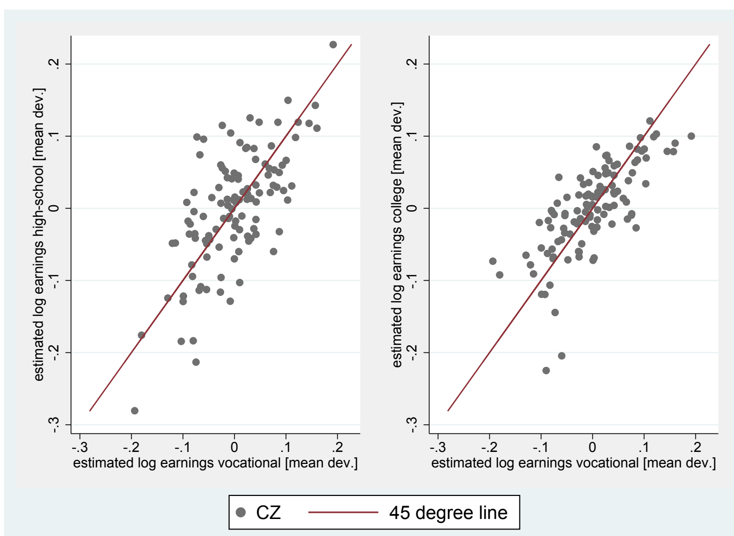


Figure 4.2: Correlation of high-school and college earnings with vocational earnings (deviations from CZ education mean), by CZ, in 2005

Table 4.1: Correlation tables corresponding to Figure (4.1b) and (4.2), respectively.

(a) Earnings auto-correlation at CZ level over time. N=113

	low	medium	high
low	1		
medium	0.674***	1	
high	0.398***	0.738***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Earnings correlation by education at CZ level over time. N=113

	1995	2000	2005
1995	1		
2000	0.946***	1	
2005	0.927***	0.966***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Finally, to establish the relation of unconditional, average CZ earnings and measures of agglomeration, I plot average CZ earnings against log population in Figure (4.3). The elasticity of average earnings with respect to population is 0.0216 in line with findings by Combes *et al.* (2008), Roca & Puga (2017) or Rosenthal & Strange (2004) if a bit smaller. The elasticity is stable for the years I analyze (0.0290 in 1995 and 0.0206 in 2000).

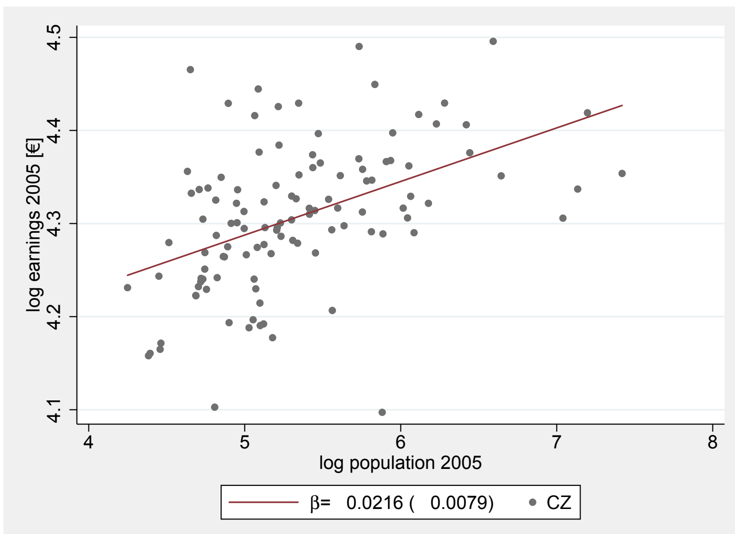


Figure 4.3: Log earnings against log population with regression line.

4.1.2 The dynamic aspect of earnings differences by population

Considering merely average earnings hides the dynamic evolution of earnings differences across CZs in Germany. Earnings of workers with little work experience do not vary much across locations; but as workers acquire more work experience, earnings start to diverge. This points to the importance of labor market interactions between workers and employers as a source for the earnings premium paid in bigger cities.

To see this, and before estimating the effect more rigorously, first, I focus on workers who never moved between CZs. Looking at workers who do not move limits the role of sorting of higher productivity workers into certain labor markets. Further, to somewhat control for skill levels, I only consider workers with a vocational degree.[7] On this sample, I run the following regression for a cross section of workers in 1995-2005

$$w_{it} \quad = \quad \sum_{j=1}^{30}\sum_{i=1}^{4}\delta_{ijt}\times q_{it}\times D'_{ijt}+X'_{ijt}\beta+\epsilon_{it}$$

where w_{it} are log earnings of person i at time t, q_{it} are dummies for the quartile of the CZ distribution the worker works in and D_{ijt} are dummies for the years of experience from 1 to 30. X_{it} is a vector containing tenure, tenure squared, occupation×industry and time dummies. I plot the experience dummies for the lowest and highest quartiles of the CZ distribution, which represent the average earnings in the biggest and smallest CZs as a function of work experience, conditional of firm tenure and the industry and occupation the worker is working in, and cleared of time trends.

The results are shown in Figure (4.4). While average earnings for workers with little work experience are higher in smaller cities (0.08 log-points), workers in bigger cities experience higher earnings growth throughout their careers. At 30 years of experience the gap between big and small cities is reversed, such that workers in bigger cities earn 0.09 log-points more than workers in smaller cities. The pattern holds similarly, when analyzing a cohort of workers over time, as opposed to a cross-section.[8]

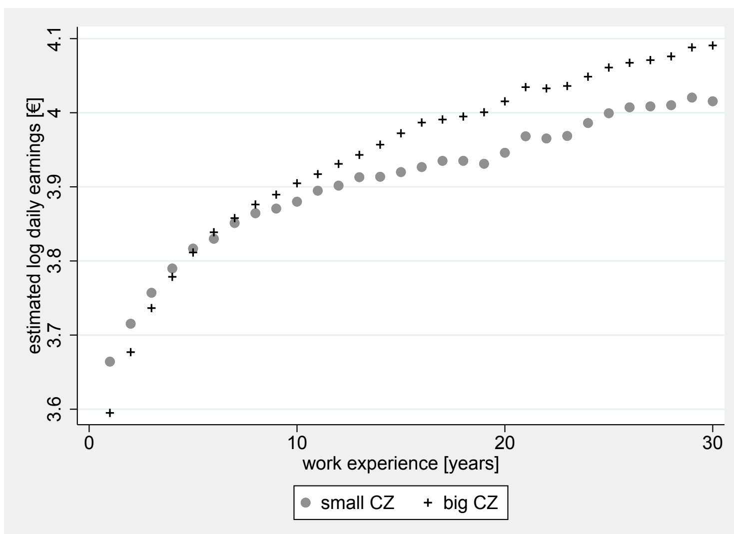


Figure 4.4: Differences in average earnings of men with vocational training, as work experience increases. Dots represent average earnings in the to quartile of the CZ population distribution, pluses represent the lowest quartile.

Even though the estimates here are likely to be biased due to selection biases, the results motivate a more careful analysis of whether exposure effects exist in the German labor market.

[7]In 1995-2005, 57% of workers with vocational degrees gained all their work experience in the same CZ.

[8]In an F-test, earnings differences at 0-10 years of experience are statistically smaller than differences at 20-30 years of experience.

4.2 Estimation Strategy

A key issue when estimating the city size earnings premium is selection of more productive workers into bigger CZs. Studies such as Combes *et al.* (2008) show that failing to control for worker characteristics biases the earnings premium upwards, since higher skilled workers tend to sort into bigger cities. This is consistent with the distribution of observed worker characteristics presented in section (3.3) above for the German market. Further, some dimension of worker skill, not perfectly captured by the available variables, might be unobserved. Combes *et al.* (2008) find evidence that unobserved skill can bias the earnings premium, as well. To interpret the earnings premium as due to market interactions, as opposed to worker selection, the estimation strategy needs to address selection on, both, observed and unobserved worker productivity.

The descriptive evidence above suggests that differences in worker productivity might arise due to differences in market interactions in big and small cities. Hence, it is plausible that upon moving to a bigger city, earnings would not respond immediately. But over time, as the worker gains more experience in the bigger CZ, earnings grow more steeply than they would have, had the worker remained in the smaller city. This effect is easy to confound with unobserved productivity levels, in the absence of appropriate variables to control for location specific work experience. It is crucial to distinguish between the two phenomena, however, since one should be interpreted as due to market interactions, while the other is true selection bias.

To address these issues, I follow an estimation strategy similar to Roca & Puga (2017). The estimation of the premium happens in two steps. First, I estimate average CZ log earnings, controlling for observed and unobserved confounding factors, such as innate worker productivity. In the second step, I regress the CZ log earnings estimates on log population and core population density, to get the elasticity of earnings with respect to CZ size.

To illustrate the effect of unobserved worker characteristics and the dynamic gains of working in a bigger CZ, I build up to my preferred estimates in Table (4.3) below. The first estimation strategy ignores the potential origin dependence of work experience. Then I include origin specific experience terms to estimate potential exposure effects.

4.2.1 Static estimation

The most basic first stage estimation is static and only controls for observed worker characteristics:

$$w_{ict} = \mu_c + X_{it}'\beta + \epsilon_{ict} \tag{1}$$

w_{ict} are log earnings of person i at time t, c indexes the CZs, and ϵ_{ict} is the error term. X_{it} is a vector of observed person characteristics containing experience and experience squared, tenure and tenure squared, and dummies for education, industry × occupation and the quarter of occupation t. μ_c are CZ fixed effects, i.e. the average log CZ earnings after controlling for observable worker and job characteristics. This is the interpretation of the dummies I will refer to throughout the paper.

In the second stage I run the following regression:

$$\mu_c \quad = \quad \alpha + \xi_1 \, p_c + \xi_2 \, d_c + \eta_c \tag{2}$$

where p_c and d_c are the log population and log core population density of CZ c and ξ_i are the population and density elasticities of earnings across German CZs.

Equation (1) only leads to unbiased estimates of μ_c if unobserved worker skills are uncorrelated with the error term ϵ_{ict} and if experience is location unspecific, i.e. experience has the same impact on earnings, independent of where it was gained. Suppose, instead, that the true data generating process is

$$w_{ict} \quad = \quad \tilde{\mu}_c + \gamma_i + \widetilde{X}_{it}\widetilde{\beta} + \widetilde{\epsilon}_{ict} \tag{3}$$

where γ_i is the unobserved portion of worker i's productivity, leading to higher earnings for person i; $\widetilde{X}_{it}$ are the same control variables as above, excluding time invariant variables, such as education.[9] In this case, the estimates for average CZ earnings from equation (1), lead to biased CZ effects μ_c. The bias of μ_c estimates is positive if workers with higher unobserved skill levels γ_i are more likely to work in the CZ, and negative otherwise. If higher skilled workers tend to sort in bigger markets, the bias would lead to overestimation of the ξ_i.

Using the panel dimension of the data, together with two assumptions about unobserved worker skill, allows controlling for the confounding effects. First, assume that all of the unobserved worker skill is reflected in earnings at each point in time. This assumption means that earnings are informative about worker productivity, observed and unobserved. Second, assuming the unobserved skill portion is time invariant, it can be eliminated from the estimation by using worker fixed effects. If the same worker is observed more than once over time, subtracting the worker's life time earnings from each observation eliminates the unobserved skill component from his earnings.

Several studies have used this strategy. Combes *et al.* (2008) use the method to estimate the city size earnings premium in France and find that the premium drops by half from 0.0625 to 0.0336 once worker fixed effects are included in the estimation. The drop is interpreted as strong evidence for sorting on unobserved worker characteristics.

I estimate, both, equation (2) and (3) to show the drop in the earnings premium. Notice that worker fixed effects γ_i and CZ fixed effects $\tilde{\mu}_c$ are not separately identified, unless the person moves between CZs; hence, in the presence of worker fixed effects, $\tilde{\mu}_c$ are identified by movers between CZs. It is plausible that movers between CZs differ from non-movers in other ways as well. If that is the case, the estimation using worker fixed effects will still lead to biased results, if non-movers have a lower tendency to experience exposure effects. While there is no way to test whether this is true or not, I can test if movers and non-movers have similar unobserved productivity levels by comparing their predicted fixed effects. I will analyze the worker fixed effects below.

[9] In the estimation of this equation, $\widetilde{X}$ excludes education since for the vast majority of workers, education does not change over the work life. Education coefficients would be identified only by workers who change education categories, which is only a small, not representative sample of workers.

4.2.2 Dynamic estimation

Next, suppose the true data generating process also includes location specific experience effects:

$$w_{ict} = \hat{\mu}_c + X_{it}\widehat{\beta} + \gamma_i + \sum_c \left[\delta_{1k(c)} \times e_{ik(c)t} + \delta_{2k(c)} \times e^2_{ik(c)t}\right] + \widehat{\epsilon}_{ict} \tag{4}$$

where $k(c)$ is the unique quartile of the CZ population distribution assigned to CZ c and $e_{ik(c)t}$ is location specific experience. Compared to the previous estimates, equation (4) accounts for two sources of bias. First, the μ_c from equation (1) may be biased due to unobserved, time invariant worker skill γ_i, similar to before. Here, the μ_c are also biased, due to the effects of location specific experience, a potential omitted variable bias.

Exposure effects can be thought of as unobserved worker productivity that emerges from labor market interactions of workers in big CZs: by working in bigger CZs workers increase their earnings potential. The increase in productivity can stem from an increase in human capital, better matching between workers and jobs, or any other mechanism of agglomeration forces. As such, this type of unobserved worker productivity is not time invariant and cannot be controlled for by including worker fixed effects. Including only worker fixed effects, when the data generating process is as in equation (4), leads to biased μ_c and worker fixed effects that confound the effects of market and the worker's unobserved productivity.

Additionally, the interpretation of the drop in the CZs' size elasticity, when including worker fixed effects in the static estimation, changes. If the drop is largely due to differences in productivity induced by market interactions, such differences should be reflected in the elasticity. In particular, one needs to distinguish between short term and long term elasticities. While the short run elasticity measures the effect of working in bigger CZs on impact, the long term elasticity might be higher due to the dynamic gains of working in bigger CZs.

To assess the relative importance of the two biases I estimate the dynamic equation (4) with and without including worker fixed effects.

4.3 City Size Earnings Premium

4.3.1 Static results

Table (4.2), column (1) presents the first stage results for the static estimation without worker fixed effects; column (3) shows results for the static estimation with worker fixed effects. Columns (2) and (4) show the corresponding second stage results. Starting with column (1), earnings in Germany increase with experience, with diminishing returns, a classic Ben Porath finding. University graduates earn 42.8 percent[10] and workers with vocational degrees 13.8 prcent more than workers with a high-school degree or less. Earnings also increase with tenure at the establishment with decreasing returns.

Table (4.2), column (2) estimates the elasticity of earnings with respect to city size, using the

[10] The difference in wage levels W is computed as follows: $W^{Univ} - W^{HS} = \exp(0.3561) - 1$

Table 4.2: Estimation of equation (1) in column (1) and equation (3) in column (3). The regressions in (1) and (3) also include occupation× industry dummies and time fixed effects. Columns (2) and (4) use CZ estimates of regressions in (1) and (3), respectively.

	(1) Log earnings	(2) CZ coeff. (1)	(3) Log earnings	(4) CZ coeff. (3)
Log population		0.0214		0.0098
		(0.0077)		(0.0052)
Log core density		-0.0038		-0.0005
		(0.0057)		(0.0042)
Experience	0.0303		0.0866	
	(0.0002)		(0.0008)	
Experience^2	-0.0006		-0.0007	
	(0.0000)		(0.0000)	
University	0.3561			
	(0.0024)			
Vocational	0.1290			
	(0.0016)			
Firm tenure	0.0275		0.0101	
	(0.0002)		(0.0001)	
Firm tenure^2	-0.0009		-0.0004	
	(0.0000)		(0.0000)	
Constant	3.5633	-0.3135	2.9000	-0.0745
	(2.9118)	(0.0800)	(0.0161)	(0.0506)
pers. fixed effects	no	N/A	yes	N/A
CZ fixed effects	yes	N/A	yes	N/A
N	9.3e+06	113	9.3e+06	113
R^2	0.4208	0.0995	0.0710	0.0621

Standard errors in parentheses

static CZ fixed effects. While the core population density of the CZ plays a minor role, the elasticity with respect to population size is 0.0214.[11] To understand the magnitude, I compare the estimated difference between earnings of a worker in CZs the bottom and top of the popultion distribution. CZs in the bottom have a population of around 140,000; these are relatively rural areas, often in the South of Germany, like Straubing. CZs in the top have a population of around 1,100,000; they consist of cities like Dortmund, Munster or Aachen, including their suburbs. Therefore, moving from a quarter below to a quarter above the median population is associated with an on-impact gain of 14.8 percent in earnings.

While the qualitative patterns in Table (4.2), column (3), after including worker fixed effects, are similar to those in column (1), the magnitudes of the experience and tenure effects change. The coefficient on tenure halves to 0.0101, and the coefficient of experience more than doubles to 0.0866, suggesting a positive correlation for worker fixed effects with respect to tenure, and a positive correlation with respect to experience. If the worker fixed effects are interpreted as innate, time invariant worker productivity, the estimates suggest more able workers tend to have less work

[11] This estimate is smaller than the estimates in the literature for Italy (0.0221 in Mion & Naticchioni (2009)), France (0.051 in Combes *et al.* (2010b)), Spain (0.046 in Roca & Puga (2017)) and the US (0.041 in Glaeser & Resseger (2010)).

experience and more tenure at their establishments.

Using the CZ earnings estimates from column (3) yields a greatly reduced city size earnings elasticity of 0.0098. Moving from Straubing to Dortmund, for instance, only leads to a 6.7 percent increase in earnings. The reduction of the elasticity is consistent with the findings in Combes *et al.* (2008) and Roca & Puga (2017). According to Combes *et al.* (2008) a significant portion of the differences in earnings across CZs are explained by differences in the innate worker productivity, and not by agglomeration effects resulting from labor market interactions. Therefore, the inclusion of worker fixed effects reduces the CZ premiums and leads to a smaller elasticity. Next I show that location specific experience is an omitted, changing this interpretation.

4.3.2 Dynamic results

The results of the dynamic estimation are presented in Table (4.2). First, consider columns (3) and (4) in Table (4.3). The first stage estimates in column (3) include worker fixed effects and allow for experience to be location specific. In particular, additional to controlling for the total number of years of work experience, I control separately for the years of experience gained in the top, second and third quartiles of the CZ population distribution. The estimated overall experience and tenure effects of the dynamic estimation in Table (4.3) column (3) are quantitatively similar to the estimates in Table (4.2) column (3) above.

The key difference is the extra effect of experience gained in bigger CZs. For instance, an extra year of experience gained in the top quartile of the CZ population distribution, holding overall experience constant, increases earnings by 0.0040 log-points. Extra experience in the second quartile still has a positive earnings effect, but the effect tapers off, such that in the second CZ quartile there is no effect statistically different from zero.

To understand the magnitude, I compare a worker who works 30 years in Straubing, to a worker who works 30 years in Dortmund. The comparison mimics the reduced form differences presented in Figure (4.4). Earnings between the two workers differ solely due to the experience effect by 3.0 percent.[12] On top of this difference there is the on-impact effect: the elasticity estimated in column (4) of 0.0089 translates into an earnings difference between Straubing and Dortmund on-impact of 6.1 percent. A third of the earnings difference between the workers earnings emerges due to market interactions over the workers career.

Notice that the difference in interpretation barely affects the on-impact estimate; the static estimate is only slightly higher (0.0098 vs. 0.0089). Even though the estimates do not change significantly, it is important to appreciate the change in interpretation of the coefficient. The earnings premium measured as suggested above captures the premium the worker experiences on impact, in the short run. Since working in bigger CZs increases earnings by more than working in smaller CZs, the earnings premium increases over time, the more a worker spends in bigger CZs.

To understand the relative importance biases from sorting on unobserved ability vs. location specific experience compare the results in Table (4.3) columns (3) and (4) to those in columns

[12]The difference is calculated as $\exp\left(0.0040 \times 30 - 0.0001 \times 30^2\right) - 1 = 0.0305$.

Table 4.3: Estimation of equation (4) without worker fixed effects in column (1) and and with worker fixed effects in column (3). The regressions in (1) and (3) also include occupation × industry dummies and time fixed effects. Columns (2) and (4) use CZ estimates of regressions in (1) and (3), respectively.

	(1) Log earnings	(2) CZ premium (1)	(3) Log earnings	(4) CZ premium (3)
Log population		0.0079		0.0089
		(0.0076)		(0.0052)
Log core density		-0.0049		-0.0006
		(0.0056)		(0.0042)
Experience	0.0256		0.0838	
	(0.0006)		(0.0010)	
Experience^2	-0.0005		-0.0006	
	(0.0000)		(0.0000)	
Experience top quartile CZs (pop)	0.0060		0.0040	
	(0.0006)		(0.0007)	
Experience top quartile CZs (pop) × experience	-0.0001		-0.0001	
	(0.0000)		(0.0000)	
Experience 3rd quartile CZs (pop)	0.0036		0.0016	
	(0.0007)		(0.0008)	
Experience 3rd quartile CZs (pop) × experience	-0.0001		-0.0001	
	(0.0000)		(0.0000)	
Experience 2nd quartile CZs (pop)	0.0015		-0.0006	
	(0.0007)		(0.0008)	
Experience 2nd quartile CZs (pop) × experience	-0.0000		-0.0000	
	(0.0000)		(0.0000)	
Firm tenure	0.0275		0.0101	
	(0.0002)		(0.0001)	
Firm tenure^2	-0.0009		-0.0004	
	(0.0000)		(0.0000)	
University	0.3560			
	(0.0024)			
Vocational	0.1291			
	(0.0016)			
Constant	3.5638	-0.1209	2.9003	-0.0622
	(1.0623)	(0.0792)	(0.0161)	(0.0505)
pers. fixed effects	no	N/A	yes	N/A
CZ fixed effects	yes	N/A	yes	N/A
N	9.3e+06	113	9.3e+06	113
R^2	0.4210	0.0087	0.0710	0.0512

Standard errors in parentheses

(1) and (2). Columns (1) and (2) show estimates including location specific experience terms, but exclude worker fixed effects. Similar to the corresponding comparison above, experience is negatively correlated with ability, tenure is positively related. Relative to column (3) the location specific experience coefficients are larger in column (1), indicating that higher ability workers tend

to have more big city work experience.

Column (2) shows that even without worker fixed effects, the CZ premium drops significantly once location specific experience terms are included. Including the worker fixed effects actually leads to an increase in the point estimate of the population effect, suggesting workers might sort negatively into bigger cities. The interpretation of the importance of CZ changed.

4.3.3 Analysis of worker fixed effects by mover status

The results on the premium when average CZ earnings are estimated using worker fixed effects are based on comparisons of the earnings of workers who move between CZs. If workers who move location are intrinsically and in unobserved ways different from workers who do not move, then it is not possible to apply results obtained for movers to non-movers. I argue, first, that selection on unobserved worker effects is not crucial to the estimation. Further, if including worker fixed effects were important to control for unobserved worker productivity, the differences between movers and stayers are milder, the higher the productivity levels.

To argue the first point, I generate a variable measuring the share of work experience gained in the top CZ quartile of population, as of 2005. The higher the share, the more time the worker has spent in the biggest CZ group. I correlate this variable with the estimated worker fixed effects to assess if workers with higher unobserved productivity are more likely to work in bigger CZs. The correlation is low at 0.0312. The low correlation is consistent with the finding that the second stage estimates change very little, in response to including worker fixed effect. From these two findings I conclude that unobserved worker skills affect estimates only mildly.

Nevertheless, there is some evidence suggesting that movers might differ somewhat from stayers. Table (4.4) shows the average of worker fixed effects by education group and by mover status. A stayer is a worker who never changes CZ over his entire work life. Using the categories of mover and stayer I look at the distribution of time invariant worker productivity to understand if the two groups differ. I use the fixed effects from the dynamic estimation, since the fixed effects from the static estimation confound time invariant and innate worker productivity.

Movers in general have higher innate productivity (i.e. a higher average fixed effect) than stayers. The difference is relatively small for college graduates and workers with vocational degrees (0.0147 and 0.0322, respectively), but is 0.1257 for workers with high-school or less.

The differences can be understood in light of a simple spatial equilibrium model. Suppose there are two labor markets, A and B. First, assume there are no mobility costs and workers can freely move between the locations. Any differences in real earnings across locations, not pertaining to innate productivity, will be eliminated in equilibrium. Movers are only observed if workers experience idiosyncratic taste shocks or if one market experiences a shock and workers move to restore equilibrium. In either case, any worker gains from moving, and there is no selection of movers. This is true even if there is heterogeneity between workers.

Next suppose there are mobility costs. If there is heterogeneity among workers, only workers who gain enough from moving will move. Hence, there will be selection in who moves. Assuming

Table 4.4: Worker fixed effects by mover type and education

	Worker Fixed Effects – log earnings	
	stayer	mover
≤ high-school		
mean	-0.0386	0.0871
N	42,109	25,990
vocational		
mean	-0.0439	-0.0117
N	147,017	111,021
≥ college		
mean	0.5140	0.4993
N	17,286	23,453

that workers with higher gains are also higher productivity workers, we expect to see higher fixed effects workers to move.

Getting back to Table (4.4), we see that the difference between movers and stayers among workers with high school education is large. At the same time, these workers have much lower mobility rates than workers with higher education levels (42.4% for high-school, 57.0% for vocational and 61.8% for college). This suggests that moving costs might be more prohibitive or higher for low educated workers, leading to stronger selection into mover status among them. For this education group, it seems that applying the mover results to the entire population might be more problematic. But among medium and high educated workers the results can be applied to at least the movers, which is a large portion of the population, and it is not obvious that non-movers are strongly selected, either.

5 Mechanisms Leading to the CSEP

After documenting the positive effect of exposure to bigger labor markets on earnings, in this section, I assess possible sources for the exposure effect. Why do earnings grow more when workers gain experience in bigger Commuting Zones (CZs)? This also allows me to assess the plausibility of the omitted variable explanation, and to dig deeper into the mechanism for the observed earnings differences. Therefore, I shift focus from earnings levels to growth. I investigate if other characteristics of the labor market experience of workers in bigger CZs, such as job mobility and the sources of earnings growth (on the job or between jobs) differ by CZ size. The details of earnings growth can help distinguish between some of the mechanisms proposed in the literature and discussed in section (2).

To do so, I divide the growth process of earnings into four mutually exclusive parts. First, earnings may change on the job (OTJ earnings change). Second, earnings change when workers change jobs, without going through an unemployment spell (Job-Job or JJ mobility). Third, workers can go through unemployment (Job-Unemployment-Job or JUJ mobility). Fourth, workers' labor market histories can include experiences outside of the sample, for instance, if they choose to be self-employed for a period of time. To obtain an earnings growth measure that reflects all possible sources earnings growth I include Job-out of sample-Job (JOJ) events in the analysis. To understand thick market effects related to unemployment and match quality I also analyze mobility differences between CZs.

I first present reduced form findings confirming that earnings growth is larger in bigger CZs. After controlling for worker selection, I find JUJ and OTJ events to be most important for understanding differences in CZ earnings. Workers in bigger CZs lose less earnings at JUJ events. On the other hand, and almost offsetting the benefits due to mobility in bigger cities, OTJ earnings growth is lower in bigger CZs. I discuss the findings in more detail at the end of the section.

5.1 Descriptive findings

I start by describing the patterns of average earnings change, without controlling for person characteristics, using the grouping of CZs into quartiles of the population distribution. Table (5.2a) shows the decomposition of earnings into JJ, JUJ, OTJ and JOJ events for the four CZ groups. The first column reports average quarterly earnings changes, independent of the source. Consistent with the finding above regarding the earnings premium, average quarterly earnings growth in more populous CZs is larger by 0.0004 log-points.

The next four columns show the earnings changes by source of change: JJ mobility, JUJ mobility, earnings change on the job and JOJ mobility. Notice, that the earnings changes by source sum up to the overall change in each CZ group. In general, earnings grow at instances of JJ mobility and on the job; they decline when a worker goes through unemployment. This finding is consistent with findings by Jolivet *et al.* (2006) for Europe and the US. Further, there are barely any differences between on-the-job earnings change in CZs of different population size. But, in bigger CZs, earnings decline less due to JUJ events and increase more due to JJ events.

Table (5.2b) reports the average likelihood of JJ and JUJ events, by CZ group; Table (5.2c) reports the increment at the event. Likelihood times increment yields the overall earnings change. OTJ earnings change happens about once a year; it is more than ten times as likely as JJ or JUJ events. JJ events are more frequent in bigger CZs. Even though the earnings increment at each event is smaller, the higher frequency leads to a larger share of earnings growth due to JJ mobility in bigger CZs. The opposite is true for JUJ events. In bigger CZs they are less common, but lead to larger earnings losses on impact. Overall this leads to smaller earning losses in bigger CZs due to JUJ events.

The reduced form results, already point towards thick market effects related to reduced unemployment in bigger labor markets. While the correlation of city size and JJ mobility is interesting, it will become clear that the effects do not persist, once I control for worker characteristics more rigorously.

5.2 Estimation Strategy

I follow a similar two step estimation strategy as laid out in section (4.2) above, but use quarterly *changes* in log earnings conditioned on the source of change: earnings change due to JJ, JUJ, JOJ and OTJ events. Notice the shift in variable from earnings level to change. The shift is necessary,

Table 5.1: Decomposing contributors to earnings change across CZ size groups; stars indicate that the difference to lowest population group is statistically significant (p<0.05).

		quarterly earnings change	JJ change	JUJ change	OTJ change	JOJ change
smallest CZs		0.0024	0.0011	−0.0007	0.0019	0.0001
second CZ group		0.0021*	0.0008*	−0.0009	0.0021	0.0001
third CZ group		0.0025	0.0010	−0.0006	0.0019	0.0002
biggest CZs		0.0028*	0.0012	−0.0006	0.0020	0.0002

(a) average quarterly log earnings changes, by source of change stars indicate that the difference to lowest population group is statistically significant (p<0.05).

		quarterly JJ rate	quarterly JUJ rate
smallest CZs		0.0233	0.0227
second CZ group		0.0245*	0.0232*
third CZ group		0.0254*	0.0180*
biggest CZs		0.0300*	0.0150*

(b) average quarterly incidence rate of events

		quarterly JJ increment	quarterly JUJ increment
smallest CZs		0.0451	−0.0292
second CZ group		0.0325*	−0.0357*
third CZ group		0.0355*	−0.0309*
biggest CZs		0.0388*	−0.0360*

(c) average increment conditional on event

because I measure the sources of exposure effects as changes. In the first step, I estimate average changes in log earnings by CZ. In the second stage, I estimate the sensitivity of earnings change with respect to population size and core CZ density.

To get from equation (4) above for earnings levels of person i living in CZ c at time t, to earnings changes for JJ, JUJ, OTJ and JOJ events, I difference a version of the equation for earnings levels, over time. Since I am interested in more disaggregated changes, I make more assumptions to estimate the effect of city size on earnings changes. First, since establishments barely move in the sample, any effect they have on earnings becomes part of the CZ fixed effects. To better understand the equation I use for changes I include establishment effects explicitly and assume establishments do not move.

Then, I assume that establishments have two effects on earnings. Assigning each worker i and establishment $f_t(i)$, where the assignment function depends on t reflecting that the match changes over time, I assume that

1. each establishment has a time invariant effect $\alpha^{OTJ}_{f_t(i)}$ on earnings.
2. earnings change on the job occurs probabilistically; in case of an OTJ earnings shock, earnings change by an establishment specific growth factor $\kappa^{OTJ}_{f_t(i)}$.

Therefore, the earnings equation looks as follows:

$$\begin{aligned}w_{if_t(i)t} &= (\hat{\alpha}_{f_t(i)} + D_{it}^{OTJ} \cdot \hat{\kappa}_{f_t(i)}^{OTJ}) + \gamma_i + X_{it}\widehat{\beta} + \\ &+ \sum_f \left[\delta_{1K(f_t(i))} \times e_{iK(f_t(i))t} + \delta_{2K(f_t(i))} \times e_{iK(f_t(i))t}^2\right] + \hat{\epsilon}_{ift}\end{aligned}$$

$K(f) \in \{1,2,3,4\}$ is the CZ quartile of establishment f and D^{OTJ} is an indicator for OTJ earnings change. Differencing the equation yields:

$$\begin{aligned}\Delta w_{if_t(i)f_{t-1}(i)t} &= \left(\widehat{\alpha}_{f_t(i)} - \widehat{\alpha}_{f_{t-1}(i)}\right) + D_{it}^{OTJ} \cdot \widehat{\kappa}_{f_t(i)}^{OTJ} + \Delta\widehat{X}_{it}\beta + \\ &+ \left[\delta_{1K(f_{t-1}(i))} \times \Delta e_{iK(f_{t-1}(i))t} + \delta_{2K(f_{t-1}(i))} \times \Delta e_{iK(f_{t-1}(i))t}^2\right] + \vartheta_{it(t-1)}\end{aligned} \tag{5}$$

$\Delta w_{if_t(i)f_{t-1}(i)t} = w_{if_t(i)} - w_{if_{t-1}(i)}$ is the quarterly log earnings change of person i who worked for $f_{t-1}(i)$ at time $(t-1)$ and for $f_t(i)$ at time t. The change in experience is gained in the CZ of $f_{t-1}(i)$. $\widehat{X}_{it}$ is the set of controls, as above, excluding the time invariant education controls; since I difference over time, any time invariant variables would drop out. In particular, differencing eliminates the worker's innate productivity term.

Equation (5) depends on establishments and earnings in periods $t-1$ and t. To simplify the estimation, I aggregate the equation up to the city level, using the city of establishment $f_t(i)$ at time t, called $c(f_t(i))$:

$$\begin{aligned}\Delta w_{ic(f_t(i))c(f_{t-1}(i))t} &= \left(\widehat{\alpha}_{c(f_t(i))} - \widehat{\alpha}_{c(f_{t-1}(i))}\right) + D_{it}^{OTJ} \cdot \widehat{\kappa}_{c(f_t(i))}^{OTJ} + \Delta\widehat{X}_{it}\beta + \\ &+ \left[\delta_{1K(c(f_{t-1}(i)))} \times \Delta e_{iK(c(f_{t-1}(i)))t} + \delta_{2K(c(f_{t-1}(i)))} \times \Delta e_{iK(c(f_{t-1}(i)))t}^2\right] + \vartheta_{it(t-1)}\end{aligned}$$

Average over all possible $f_{t-1}(i)$, to get

$$\begin{aligned}\Delta w_{ic(f_t(i))t} &= \kappa^{c(f_t(i))} + D_{it}^{OTJ} \cdot \widehat{\kappa}_{c(f_t(i))}^{OTJ} + \Delta\widehat{X}_{it}\beta + \\ &+ \sum_{c(f_{t-1}(i))} \left[\delta_{1K(c(f_{t-1}(i)))} \times \Delta e_{iK(c(f_{t-1}(i)))t} + \delta_{2K(c(f_{t-1}(i)))} \times \Delta e_{iK(c(f_{t-1}(i)))t}^2\right] + \widehat{\vartheta_{it}}\end{aligned}$$

where $\kappa^{c(f_t(i))} = \sum_{c(f_{t-1}(i))} \left(\widehat{\alpha}_{c(f_t(i))} - \widehat{\alpha}_{c(f_{t-1}(i))}\right)$ is the average earnings change due to job mobility that end with employment in c. Simplifying notation yields the estimation equation

$$\begin{aligned}\Delta w_{ict} &= \kappa^c + D_{it}^{OTJ} \cdot \widehat{\kappa}_c^{OTJ} + \Delta\widehat{X}_{it}\beta + \\ &+ \sum_c \left[\delta_{1K(c)} \times \Delta e_{iK(c)t} + \delta_{2K(c)} \times \Delta e_{iK(c)t}^2\right] + \widehat{\vartheta_{it}}\end{aligned} \tag{6}$$

To see that this equation can be estimated separately for each of the earnings change types use the following notation to aggregate:

- $\alpha_{f_t(i)} - \alpha_{f_{t-1}(i)} = \kappa_{JJ}^{c(f_t(i))}$ if $f_t(i) \neq f_{t-1}(i)$ and no unemployment spell
- $\alpha_{f_t(i)} - \alpha_{f_{t-1}(i)} = \kappa_{JUJ}^{c(f_t(i))}$ if $f_t(i) \neq f_{t-1}(i)$ and there was an unemployment spell
- $\alpha_{f_t(i)} - \alpha_{f_{t-1}(i)} = \kappa_{JOJ}^{c(f_t(i))}$ if $f_t(i) \neq f_{t-1}(i)$ and worker was out of sample
- $\widehat{\kappa}_{f_t(i)}^{OTJ} = \kappa_{OTJ}^{c(f_t(i))}$

The average earnings change in $c(f_t(i))$ $\kappa^{c(f_t(i))}$ is then$\kappa^{c(f_t(i))} = \kappa_{JJ}^{c(f_t(i))} + \kappa_{JUJ}^{c(f_t(i))} + \kappa_{OTJ}^{c(f_t(i))} + \kappa_{JOJ}^{c(f_t(i))}$. Since the left hand side can be separated into its components in a similar way, I can estimate each kind of earnings change separately.

When I estimate the equation for each of the source categories E, where $E \in \{$JJ, JUJ, OTJ, JOJ$\}$, the κ_E^c measure average earnings *change* at event type E in CZ c, and are not to be directly compared to the μ_c estimated before. They are also not identified by CZ movers: there are no worker fixed effects since innate worker productivity is already eliminated by difference earnings to get growth.

As before, in the second stage, I regress the CZ fixed effects on log CZ population and and log CZ core populations density as follows:

$$\kappa^c \quad = \quad \alpha + \varphi_1 \, p_c + \varphi_2 \, d_c + \eta_c \tag{7}$$

φ_1 and φ_2 are the semi-elasticities of earnings growth with respect to population and core population density, respectively. Note again, the left hand side of the equation is additive. Therefore, I can separate out the estimation by source and the semi-elasticities for the sources add up to the total. Since the κ cannot be directly compared to the μ_c estimated in section (4), the elasticities have a different interpretation, too. Hence, the estimation does not reproduce the exact coefficients from before. Nevertheless, I show that the estimation produces elasticities consistent with the previous estimation.

5.3 Earnings Growth Differences by City Size and Source

The results of the second stage regressions from equation (7) are shown in Table (5.2). The columns show results by source of earnings growth. Column (1) shows the second stage regression for overall earnings growth. The reduced form result survives: CZs with, both, higher population and higher core density, yield higher earnings growth. To understand the magnitude, remember the comparison from the previous section, relating earnings is areas like Straubing, to earnings in medium sized cities like Dortmund, Munster or Aachen. As before, I compare a worker who spends working 30 years in the biggest CZs to a worker with the same work experience, but in the smallest CZs. The difference consists of two parts. First the differences in average earnings, reflected in the CZ fixed effects. Second, in the first stage regression (not reported) there are diminishing returns to earnings growth in bigger cities. Taken together, the estimated differences in earnings growth suggest earnings in

Table 5.2: Decomposition of earnings growth by source of growth

	(1) CZ premium all	(2) CZ premium JJ	(3) CZ premium JUJ	(4) CZ premium OTJ	(5) CZ premium JOJ
Log population	0.0004	0.0001	0.0009	-0.0009	0.0003
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Log core density	0.0001	-0.0001	-0.0002	0.0003	0.0001
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
N	113	113	113	113	113
R^2	0.0672	0.0032	0.3917	0.4105	0.3137

Standard errors in parentheses

Germany's medium sized cities after 30 years of work there, are higher by 14.0 percent, assuming no initial differences in earnings. The estimate higher than the 9.2 percent estimated in the previous section; but since the descriptive results suggest slightly lower early career earnings levels in bigger cities, the initial assumption might be driving the difference.[13]

Focus next on the results for the decomposition into sources of earnings growth, JJ, JUJ, OTJ and JOJ, presented in columns (2)-(5). The coefficients of the four conditional earnings growth measures add up to the overall earnings growth measure. Two main findings emerge.

First, after controlling for work force composition, JJ mobility and JOJ events have only mild explanatory power for higher earnings in bigger or more densely populated CZs. The most important differences between CZs are the differences in OTJ and JUJ events. OTJ earnings differences are actually smaller in bigger CZs; this difference is fully compensated by lower losses due to JUJ mobility. Translating the semi-elasticities into earnings differences by comparing the differences if a worker works 30 years in Dortmund vs. Straubing helps understand the magnitudes. [14] If JUJ events were the only source of earnings change, the worker's earnings would be 84.1 percent higher in Dortmund after 30 years. If OTJ events were the only source of earnings change, the worker's earnings would be 53.5 percent lower in Dortmund.

Second, core density and population size affect different aspects of earnings change. Population size affects OTJ gains negatively and JUJ changes positively; core population density has the opposite effect (positive for OTJ and negative for JUJ). Higher core population density reduces the average distance between workers. When more workers interact closely to each other, they lose more due to unemployment, but gain more due to OTJ gains. Both findings are consistent with theories that emphasize worker competition as a mechanism leading to agglomeration economies. Higher overall population, on the other hand, lends support to theories of agglomeration focusing on labor market pooling. I discuss the relevance of these findings for the different theories in section (5.5).

5.4 The impact of turnover

Next consider the likelihood of JJ and JUJ mobility. Until now, I have analyzed the total earnings change due to the different events. For the variables associated with mobility events that can be

[13] see Appendix A for details.

[14] Again, see Appendix A for details.

interpreted (JJ and JUJ) it is insightful to consider the event rate separately. The total change is the product of the likelihood of the event and the increment conditional on the event. For JUJ events that means:

$$E\left[\Delta w_{\mathrm{JUJ}}\right] \quad = \quad \Pr\left(\mathrm{JUJ}\right) \times E\left[\Delta w_{\mathrm{JUJ}} | \mathrm{JUJ}\right]$$

Differentiating between the likelihood and increment can shed light on the mechanisms leading to the different labor market outcomes for workers across CZs. Focus on JUJ mobility for instance. On average, JUJ mobility is associated with earnings drops. It is possible that workers in bigger CZs have lower quarterly earnings losses in bigger CZs because they experience fewer JUJ events, or because, conditional on experiencing JUJ mobility, the earnings losses are lower. The economic mechanisms associated with each case may be different. While in the first case the findings point to labor pooling mechanisms and employment insuring effects of bigger markets, in the second case mechanisms related to reduced skill loss or specialization are more likely to be important.

Similar arguments can be made for JJ mobility. The lack of difference in earnings growth due to JJ mobility might be due to similar frequency and increments across CZs; it might be due lower frequency and higher increment or the opposite. Frequency and increments can be informative about returns to tenure: if workers are willing to change jobs frequently, even though change in not compensated, returns to tenure are likely low. Investments in specialized human capital might be less relevant in such a market.

Without specifying a more structural model, I cannot identify the different mechanisms cleanly, but measuring mobility is still insightful. Therefore, I focus on the differences in job mobility patterns between CZs.

5.4.1 Estimation strategy

The estimation is carried out in two steps. I estimate the following first stage equation with and without worker fixed effects:

$$R^{\bar{E}}_{ict} \quad = \quad \rho_c + X_{ict}\widehat{\beta} + \sum_{k=2}^{4}\left[\delta_{1k} \times e_{ikt} + \delta_{2k} \times e^2_{ikt}\right] + \widehat{\epsilon}_{ict}$$

where $\bar{E} \in \{\mathrm{JJ}, \mathrm{JUJ}\}$ and for each $\bar{E}$, $R^{\bar{E}} = 1$ if an event took place and $R^{\bar{E}} = 0$ otherwise. Subsequently, I run a second stage equation of the CZ averages ρ_c on log population size and core density. The estimation is similar to the dynamic estimation in equations (4). Similar to earnings, incidences of mobility and earnings change are determined by the experience, job tenure and education of workers.

Including person fixed effects in a earnings regression deals with the issue of unobserved worker productivity. I include worker fixed effects in the mobility regressions to account for a distinct dimension of unobserved worker ability: workers can have different preferences for mobility; including fixed effects in the mobility regression allows to understand the importance of sorting of intrinsically more mobile workers into certain CZs. But the worker effects also control for any other time invariant, innate worker characteristics that lead to differences in mobility rates. For instance, it is plausible to think that JUJ rates are intrinsically lower for more productive workers, since they tend to experience fewer unemployment spells. Hence the JJ and JUJ type control for a bundle of distinct, unobserved and time invariant characteristics. I assess the correlation between the different worker fixed effects I have used in section (5.4.3) below.

When worker fixed effects are included, the CZ averages are identified by movers. If movers and non-movers between CZs have similar mobility preferences, then two things are true. First, sorting of workers with higher job mobility preferences into bigger CZs is less likely. Second, since we control for observed and unobserved characteristics, the premium in the second stage is likely not only reflective of the CZ movers, but can also be applied to the stayers. I analyze the person fixed effects distribution next.

5.4.2 Results

Tables (5.3) and (5.4), respectively, show the estimation results for the JJ and JUJ rate estimation. The estimation shows two interesting results.

First, when I control for unobserved JJ type, the positive estimates for the CZ mobility premium (in Table (5.3) compare column (4) to (2)) turn negative and become statistically indistinguishable from zero. Hence, unobserved JJ type is positively correlated with working in a bigger city; high mobility workers select into bigger CZs. Once I control for that, mobility is statistically the same across cities of different size.

Second, there are negative exposure effects of JJ mobility after controlling for the JJ type. The more experience a worker gathers in big CZs, the less likely he experiences JJ mobility. The effect of exposure is largest for the biggest CZ group and declines to zero for the second lowest group. Therefore, even though moving to a bigger CZ has almost no effect on JJ mobility, over time mobility declines slightly with exposure to the market. Previously, I estimated the earnings growth semi-elasticity due to JJ mobility to be indistinguishable from zero, meaning that earnings growth due to JJ mobility is unrelated to city size. Since the JJ rate is not different between CZs, the increment (i.e. earnings growth conditional on JJ mobility) are unrelated to city size, as well.[15]

The findings confirm sorting of more mobile workers into bigger cities and they can only partly speak to thick labor market mechanisms related to higher job mobility in bigger cities. While average rates of JJ mobility are higher in bigger cities, it cannot be concluded that the market makes workers more mobile: workers that are intrinsically more mobile, as demonstrated by their job mobility rates in other cities, tend to move to bigger cities. This result does not explain, however, *why* more mobile workers move to bigger cities. In particular, the result does not stand in the way of mechanisms similar to those discussed in Bleakley & Lin (2012), where younger workers in bigger cities move jobs more frequently than in smaller cities, to find better matches or to find which occupation is the best fit for them. If such search is productive, bigger cities favor it, and the

[15] I confirm this separately; results available upon request.

personal cost of job mobility varies among workers, one would expect sorting of more mobile workers to bigger cities, as observed in the data. In particular, Bleakley & Lin (2012) also report that job mobility declines more with experience, in bigger cities in the US, consistent with my finding of exposure effects in the JJ rate.

A similar argument can be made for "stepping-stone mobility" as discussed in Jovanovic & Nyarko (1997). The lack of difference in the JJ rate by city size does not indicate that labor markets in bigger and smaller cities are similarly beneficial for learning at lower rungs of the job ladder. The fact that more mobile workers tend to sort into bigger cities suggests that there are benefits of mobility, but they likely come at a personal cost; therefore, they are reaped only by workers with low costs of mobility.

Turning to JUJ events, I find that going through unemployment is less likely in bigger CZs. On impact of moving from the rural cities in the lowest CZ population quartile to the medium sized cities of the highest, the likelihood of going through unemployment is reduced by 1.76 percentage points. The magnitude is economically meaningful, considering the unconditional, average JUJ rate in the small CZs is 2.27.

There are also exposure effects for the JUJ rate. More work experience in bigger CZs reduces JUJ rates over time.[16] The effects is declines with population size.

5.4.3 Fixed effects comparison

The worker fixed effects in the JJ and the JUJ estimations use variation among movers. Therefore, similar concerns apply here, as did in the estimation of the city size earning premium. Furthermore, the worker fixed effects represent distinct worker characteristics in the mobility estimation. In the earnings regressions, person fixed effects measure unobserved worker productivity. In the mobility estimation, the fixed effects sum up several time invariant worker characteristics that lead to higher mobility. One characteristic I emphasized above is the preference to move between jobs. It is also possible that the labor market leads to more mobility for certain kinds of workers. For instance, workers in jobs requiring specialized knowledge might be less mobile than other workers. In the JJ estimation the worker fixed effects reflect innate preference to move between jobs, market possibilities for mobility of a worker, after we have controlled for observed characteristics like firm tenure, work experience and education level. In particular, if innate worker productivity is related to the choices and opportunities related to mobility, the JJ type can also reflect that. In the JUJ estimation the type reflects the tendency of workers to go through unemployment. The tendency to be unemployed correlates with preference for mobility, but also with worker skill. Therefore, the JUJ effects are likely to measure both.

To show the link between innate worker mobility and productivity, I use the unobserved productivity estimates from section (4) and correlate them with the JJ and JUJ types (each person is one observation). JJ and JUJ types are strongly negatively correlated with worker productivity, with

[16] I can divide JUJ event up further into job separation (JU) and job finding rates, but find that both margins are similarly affected by CZ size. Results not reported. Available if deemed necessary.

Table 5.3: Turnover regressions: job-job mobility, with and without worker fixed effects

	(1) JJ	(2) CZ premium (1)	(3) JJ	(4) CZ premium (3)
Log population		0.0015		-0.0003
		(0.0003)		(0.0009)
Log core density		-0.0002		-0.0013
		(0.0003)		(0.0008)
Experience	0.0030		0.0080	
	(0.0001)		(0.0003)	
Experience^2	-0.0001		- 0.0002	
	(0.0000)		(0.0000)	
Experience top quartile CZs (pop)	-0.0003		-0.0016	
	(0.0001)		(0.0002)	
Experience top quartile CZs (pop) × experience	0.0000		0.0000	
	(0.0000)		(0.0000)	
Experience 3rd quartile CZs (pop)	0.0000		-0.0007	
	(0.0001)		(0.0002)	
Experience 3rd quartile CZs (pop) × experience	-0.0000		0.0000	
	(0.0000)		(0.0000)	
Experience 2nd quartile CZs (pop)	-0.0000		-0.0004	
	(0.0001)		(0.0003)	
Experience 2nd quartile CZs (pop) × experience	0.0000		0.0000	
	(0.0000)		(0.0000)	
tenure	0.0137		-0.0150	
	(0.0000)		(0.0001)	
tenure^2	0.0005		0.0006	
	(0.0000)		(0.0000)	
University	0.0034			
	(0.0003)			
Vocational	0.0037			
	(0.0002)			
Constant	0.0220	-0.0145	-0.0631	0.0201
	(0.0025)	(0.0032)	(0.0047)	(0.0091)
worker fixed eff.	no	N/A	yes	N/A
CZ fixed effects	yes	N/A	yes	N/A
N	9.3e+06	113	9.3e+06	113
R^2	0.0422	0.1638	0.0269	0.0690

Standard errors in parentheses

coefficients of -0.4625 and -0.7777, respectively. This confirms that more productive workers are less mobile. The finding may be due to choice: the returns to tenure and firm specific specialization might be higher for more productive workers, reducing their willingness to move jobs. The finding may also be due to differences in labor market opportunities and match quality: match quality may be more important for higher productivity workers. This incentivizes search and higher match quality for more productive workers, thereby reducing mobility. Still, the JJ type also captures

Table 5.4: Turnover regressions: job-unemployment-job mobility, with and without worker fixed effects

	(1) JUJ	(2) CZ premium (1)	(3) JUJ	(4) CZ premium (3)
Log population		-0.0032		-0.0020
		(0.0008)		(0.0007)
Log core density		0.0008		0.0006
		(0.0008)		(0.0007)
Experience	0.0023		-0.0144	
	(0.0001)		(0.0003)	
Experience^2	-0.0001		-0.0001	
	(0.0000)		(0.0000)	
Experience top quartile CZs (pop)	-0.0010		-0.0006	
	(0.0001)		(0.0002)	
Experience top quartile CZs (pop) × experience	0.0000		0.0000	
	(0.0000)		(0.0000)	
Experience 3rd quartile CZs (pop)	-0.0008		-0.0007	
	(0.0001)		(0.0002)	
Experience 3rd quartile CZs (pop) × experience	0.0000		0.0000	
	(0.0000)		(0.0000)	
Experience 2nd quartile CZs (pop)	-0.0004		-0.0004	
	(0.0002)		(0.0002)	
Experience 2nd quartile CZs (pop) × experience	0.0000		0.0000	
	(0.0000)		(0.0000)	
tenure	-0.0101		-0.0083	
	(0.0000)		(0.0000)	
tenure^2	0.0004		0.0003	
	(0.0000)		(0.0000)	
University	-0.0036			
	(0.0003)			
Vocational	0.0016			
	(0.0002)			
Constant	0.1018	0.0173	0.3185	0.0266
	(0.0044)	(0.0075)	(0.0050)	(0.0069)
worker fixed eff.	no	N/A	yes	N/A
CZ fixed effects	yes	N/A	yes	N/A
N	9.3e+06	113	9.3e+06	113
R^2	0.0457	0.1163	0.0189	0.0723

Standard errors in parentheses

other elements, besides worker productivity. The JUJ type on the other hand is largely aligned with productivity.

As before, I use the share of experience gained in the quartile of biggest CZs to understand the importance of selection on unobserved worker type. The correlations of JJ and JUJ type with the experience share are -0.0150 and -0.0361, respectively. Once more, there is little evidence of selection of a certain worker type into bigger CZs. The small difference after including worker fixed effects is consistent with the finding.

Table (5.5) shows the worker fixed effects of the JJ and JUJ estimations by education group and by mover status.[17] As expected from the preceding discussion mobility is negatively related to education levels. Focusing on differences between movers and stayers, CZ movers tend to be more mobile between jobs. This finding holds for medium and high education levels and does not hold for the lowest level. Table (4.4) shows that movers are slightly more productive than stayers; they are also slightly more mobile between jobs. Hence the differences in JJ and JUJ type do not reflect productivity, despite the relatively high correlation of the type and productivity. A distinct dimension of unobserved characteristic differs between movers and stayers, such as the preference for mobility.

Table 5.5: Analysis of worker fixed effects by education and mover status

(a) JJ

	Worker Fixed Effects – JJ rate	
	stayer	mover
≤ high-school		
mean	-0.0250	-0.0125
N	42,411	26,062
vocational		
mean	-0.0076	0.0099
N	147,699	111,322
≥ college		
mean	-0.0591	-0.0305
N	17,297	23,470

(b) JUJ

	Worker Fixed Effects – JUJ rate	
	stayer	mover
≤ high-school		
mean	-0.1035	-0.1373
N	42,411	26,062
vocational		
mean	-0.0123	0.0168
N	147,699	111,322
≥ college		
mean	-0.1006	-0.0750
N	17,297	23,470

To summarize, unobserved characteristics are likely to play a subordinate role in the JJ and JUJ rate estimations. While there are differences between movers and stayers, worker types are not strongly selected into labor markets.

5.5 Discussion

To summarize, I find three sources for earnings differences by city size in Germany. First, earnings losses due to unemployment are lower in bigger CZs. Lower unemployment rates in bigger cities contribute to the smaller losses. Second, there are no statistically significant differences between big and small German cities regarding patterns of job mobility without intermittent unemployment. Neither the rate of job-job mobility, nor the earnings change conditional on it differ by city size. Finally, earnings grow less on the job in bigger cities. The findings can inform several mechanisms leading to agglomeration economies, suggested in the literature. I begin by discussing four mechanisms which assume a fixed relationship between productivity and earnings, and end by addressing the possibility for bargaining mechanisms altering this relationship.

First, consider thick labor market effects that lead to higher quality of the matches formed in bigger cities. Better matches might be a consequence of reduced search frictions in bigger markets,

[17] Remember, mover status refers to moving between CZs, not between jobs.

allowing employers and employees with different specializations to find each other more easily. Better matches may also be a result of higher levels of specialization of workers and firms. Further, better matches can be a consequence of more intense searching by workers, due to higher returns to search in bigger markets.

For either of the cases, higher match qualities lead to higher earnings, since better matches produce more output. The finding of lower unemployment rates are also consistent with higher match qualities. On the one hand, it may be the case that workers become more choosy about the establishment for which they start working; this would lead to higher unemployment rates in bigger cities. On the other hand, higher expected match quality can incentivize higher contact rates in bigger cities, countering this negative effect. Further, higher match qualities also lower separation rates. The net of these effects is consistent with lower unemployment rates in bigger cities. The finding of similar Job-Job rates by city size is also consistent with this result since higher equilibrium match qualities result in lower incentives for workers to move.

Second, thicker labor markets might increase productivity levels by reducing unemployment. Employers with unfilled vacancies lose profits. As documented, for instance, by Bagger *et al.* (2011), workers lose human capital while unemployed. If thick labor markets lead to lower incidence of unemployment, they increase market productivity.

I find evidence in support of the mechanism. First, workers go through unemployment less frequently in bigger cities. Second, they lose less earnings due to unemployment spells. Since unemployment leads to human capital depreciation, in markets with more unemployment, the same amount of work experience translates to lower levels of human capital. Controlling for the origin of experience accounts for the human capital depreciation in periods of unemployment. Lower levels of earnings growth on the job in bigger cities, may then be related to decreasing returns effects of human capital: at lower human capital level, human capital grows more steeply.

Labor pooling has been suggested as a reason for reduced unemployment. Therefore, I consider labor poling mechanisms next. Labor pooling refers to the possibility that firms in bigger labor markets might find it easier to adjust employment in response to idiosyncratic shocks. Since the labor needs of a single firms will be less important to the market's labor demand curve the bigger the market, idiosyncratic firm shocks do not affect labor costs of firms. In case of a positive shock, wages do not increase by much, allowing the firm to hire workers more cheaply; in case of a negative shock, labor costs do not fall much and rehiring is anticipated to be easy, such that firms reduce the work force more easily. This reduces labor costs for firms and allows them to produce more.

Labor pooling, however, is not consistent with my empirical findings. If negative idiosyncratic shocks are similarly common in big and small markets, labor pooling mechanisms would lead to more mobility, and not as observed in the data to less mobility in bigger markets. Labor pooling can lead to less unemployment for workers, which is consistent with the findings. But changing employers should be more common in bigger cities, which is inconsistent with the findings.

Fourth, consider learning mechanisms. It has often been suggested that there are knowledge spillovers in thicker labor markets, facilitating learning and human capital accumulation in bigger cities, thereby increasing productivity in bigger markets. I use OTJ earnings growth to assess a particular type of human capital accumulation. Suppose that learning increases the marginal productivity of workers and that employers compensate workers for increased productivity in similar ways across local labor markets. Steeper human capital growth in that case should be measurable as steeper earnings growth on the job. OTJ earnings growth eliminates differences in earnings due to potential productivity differences between employers or differences in match quality. It allows focusing on changes in productivity between the same worker and firm, and is commonly interpreted as a measure for human capital growth due to learning on the job. As such, if there were differences in learning on the job between big and small cities, the differences should be reflected in OTJ earnings change.

The value of the worker's added human capital for other employers matters, when assessing how learning might affect mobility. Learning on the job can be very specific to the employer. In that case, higher learning binds the worker to the employer and reduces mobility, since the worker will not be compensated as much by other employers. We would expect lower JJ mobility, and higher susceptibility to unemployment. If learning is less employer specific, higher human capital levels might increase the worker's attractiveness for potential employers and increase JJ mobility, while reducing the likelihood of unemployment.

Increased learning on the job, as a source of agglomeration economies is not fully consistent with my findings. The lower JJ mobility in bigger cities is consistent with learning mechanisms, where the type of added human capital is firm specific. Higher specialized knowledge can make the worker more vulnerable to idiosyncratic employer shock; this effect is not observed. Further, earnings growth on the job is lower in bigger cities. Unless there is a reason why workers are compensated less for gained human capital levels in bigger cities, explanations relying on learning mechanisms seem inconsistent with the data.

Finally, monopsony power of employers plausibly changes with city size, affecting how much a worker is compensated for his marginal productivity. Differing monopsony power adds another dimension to the interpretation of the impact and exposure effects I estimate. To understand the influence of monopsony power on the impact effect, notice that in a monopsonic labor market, firms and workers produce a surplus, which is shared between them. The more bargaining power workers have, the more of the surplus they can secure. Further, monopsony power is lower the more available jobs there are for a given worker. A thicker labor market, supposedly, has more such opportunities and workers there are more likely to receive job offers from other firms. Using such outside poaching offers, it is easier for workers in bigger cities to bid up their earnings. In equilibrium, less monopsony power of firms in bigger cities, therefore, should lead to a higher share of the match surplus for the worker (compare to predictions by Postel-Vinay & Robin (2002) and Cahuc *et al.* (2006)).

Monopsony power by firms can also affect the interpretation of the exposure effect. Presumably, the monopsony power of firms is also impacted by the levels of worker specialization. There are fewer attractive job opportunities for a more specialized worker. Hence, higher levels of specialization tie a worker to the firm he is currently working for, increasing the monopsony power the firm has over the worker. A thicker labor market, especially if it has more specialized firms, is likely to dampen the impact of worker specialization on the firm's monopsony power. Hence, the exposure effect measures how workers in smaller labor markets have fewer options for job change, resulting in fewer poaching offers that could be used to bid up earnings. The interaction of location and experience in this case captures how increased human capital is compensated differently across cities of differing size.

Notice that the higher monopsony power of firms in smaller cities, after investing in specialization, might lead to different choices by workers regarding how much to specialize. If the specificity of human capital investment is the same across cities, workers in smaller cities would invest less in it, ex ante, anticipating that they will be able to bargain a lower share of the investment's surplus ex post. This would compound the effect of monopsony power.

The empirical findings are partially consistent with mechanisms related to increasing monopsony power with city size. Even if productivity were the same across cities, a higher bargaining position of workers would lead to higher earnings in bigger cities. As explained, the exposure effect can also be explained: in smaller cities, monopsony power might rise more steeply with worker experience.

Mobility patterns are only partially consistent with higher monopsony power in bigger cities. First, higher monopsony power can either increase or decrease unemployment. On the one hand workers are more vulnerable to unemployment when they are involuntarily separated from a job; this could lead to higher unemployment rates in smaller cities. On the other hand, it might lead to less voluntary unemployment because the probability of finding a new match is lower. I observe less unemployment in bigger cities. Second, the effect of monopsony power on JJ mobility is also ambiguous. Higher levels of specialization in bigger cities, as a results of lower monopsony power of firms, would reduce mobility, both JJ (and JUJ), in bigger cities. On the other hand, due to lower monopsony power and more options of firms, I would expect JJ mobility to be higher in bigger cities. I observe that JJ mobility rates are similar across cities. Lastly, the combination effect of higher monopsony power and experience is inconsistent with my findings. The effect monopsony and experience effect should lead to a larger reduction of JJ and JUJ mobility with experience, the smaller the city. I observe the opposite: the larger the city, the more experience reduces mobility. In conclusion, while the observed earnings patterns are consistent with higher monopsony power of firms in smaller cities, the observed mobility patterns are only partially so.

6 Conclusions

There are several key takeaways of the analysis in this paper. First, workers gain earnings on impact of moving to a bigger city. Moving from rural parts of Germany to medium sized cities increases earnings by 6.1 percent. This finding underestimates the effect of of bigger labor markets, however, since working in bigger cities further increases earnings. A worker who works 30 years in a city like Dortmund, as opposed to Straubing, can expect his earnings to be higher by an extra 3.1 percent.

To understand possible reasons for the emergence of earnings differences, I decompose earnings growth due to job change and on the job. Earnings change when workers change jobs can be categorized into three sub-cases: job mobility without going through unemployment (JJ mobility), job mobility with intermittent unemployment (JUJ mobility) and job mobility without further information since the worker moves out of the sample (JOJ mobility). I show that mobility events are crucial for understanding the emergence of the earnings premium in bigger cities. While on-the-job earnings gains are smaller in bigger cities, workers lose fewer earnings when going through unemployment, and gain slightly at JJ and JOJ events. The JUJ finding is partly driven by the lower incidence of unemployment in bigger cities. But it is also true that mobility, both JJ and JUJ, decline as workers gain more work experience in bigger cities.

The findings support the view that bigger markets increase, both, the likelihood of matches in the labor market as well as their quality. Lower mobility and higher earnings levels are both consistent with higher match qualities in bigger markets. Lower unemployment is consistent with higher likelihood of matching. I interpret the findings as evidence for the importance of matching mechanisms in the labor market for understanding agglomeration economies in Germany.

References

Bagger, Jesper, Fontaine, François, Postel-Vinay, Fabien, & Robin, Jean-Marc. 2011. Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics.

Baum-Snow, Nathaniel, & Pavan, Ronni. 2011. Understanding the city size wage gap. *The Review of economic studies*, **79**(1), 88–127.

Behrens, Kristian, Duranton, Gilles, & Robert-Nicoud, Frédéric. 2014. Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, **122**(3), 507–553.

Bleakley, Hoyt, & Lin, Jeffrey. 2012. Thick-market effects and churning in the labor market: Evidence from US cities. *Journal of urban economics*, **72**(2), 87–103.

Blossfeld, Hans-Peter. 1983. *Die Entwicklung der qualifikationsspezifischen Verdienstrelationen von Berufsanfängern zwischen 1970 und 1982: regressionsanalytische Untersuchungen auf der Basis von Volkszählungs-und Mikrozensusdaten*. VASMA-Projekt, Institut für Sozialwissenschaften, Universität Mannheim.

Cahuc, Pierre, Postel-Vinay, Fabien, & Robin, Jean-Marc. 2006. Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, **74**(2), 323–364.

Carlino, Gerald A, Chatterjee, Satyajit, & Hunt, Robert M. 2007. Urban density and the rate of invention. *Journal of Urban Economics*, **61**(3), 389–419.

Ciccone, Antonio, & Hall, Robert E. 1996. Productivity and the Density of Economic Activity. *The American Economic Review*, **86**(1), 54–70.

Combes, Pierre-Philippe, Duranton, Gilles, & Gobillon, Laurent. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, **63**(2), 723–742.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, & Roux, Sébastien. 2010a. Estimating agglomeration economies with history, geology, and worker effects. *Pages 15–66 of: Agglomeration economics*. University of Chicago Press.

Combes, Pierre-Philippe, Duranton, Gilles, & Gobillon, Laurent. 2010b. The identification of agglomeration economies. *Journal of Economic Geography*, **11**(2), 253–266.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, Puga, Diego, & Roux, Sébastien. 2012. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, **80**(6), 2543–2594.

Duranton, Gilles, & Puga, Diego. 2004. Micro-foundations of urban agglomeration economies. *Handbook of regional and urban economics*, **4**, 2063–2117.

Eckey, Hans-Friedrich, Schwengler, Barbara, & Türck, Matthias. 2007. *Vergleich von deutschen Arbeitsmarktregionen*. Tech. rept. IAB Discussion paper.

Eckey, HF, Kosfeld, R, & Türck, M. 2006. Delineation of German labour market regions. *Raumforschung und Raumordnung*, **64**(4), 299–309.

Eeckhout, Jan, Pinheiro, Roberto, & Schmidheiny, Kurt. 2010. Spatial sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled.

Eeckhout, Jan, Pinheiro, Roberto, & Schmidheiny, Kurt. 2014. Spatial sorting. *Journal of Political Economy*, **122**(3), 554–620.

Glaeser, Edward L, & Mare, David C. 2001. Cities and skills. *Journal of labor economics*, **19**(2), 316–342.

Glaeser, Edward L, & Resseger, Matthew G. 2010. The complementarity between cities and skills. *Journal of Regional Science*, **50**(1), 221–244.

Glaeser, Edward L, Gottlieb, Joshua D, *et al.* 2008. The Economics of Place-Making Policies. *Brookings Papers on Economic Activity*, **39**(1 (Spring)), 155–253.

Hagedorn, Marcus, & Manovskii, Iourii. 2013. Job Selection and Wages over the Business Cycle. *American Economic Review*, **103**(2), 771–803.

Hout, Michael, & DiPrete, Thomas A. 2006. What we have learned: RC28's contributions to knowledge about social stratification. *Research in social stratification and mobility*, **24**(1), 1–20.

Jolivet, Gregory, Postel-Vinay, Fabien, & Robin, Jean-Marc. 2006. The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US. *European Economic Review*, **50**(4), 877–907.

Jovanovic, Boyan, & Nyarko, Yaw. 1997. Stepping-stone mobility. *Pages 289–325 of: Carnegie-Rochester Conference Series on Public Policy*, vol. 46. Elsevier.

Krugman, Paul R. 1991. *Geography and trade*. MIT press.

Marshall, Alfred. 1890. Principles of political economy. *Maxmillan, New York*.

Mion, Giordano, & Naticchioni, Paolo. 2009. The spatial sorting and matching of skills and firms. *Canadian Journal of Economics/Revue canadienne d'économique*, **42**(1), 28–55.

Moretti, Enrico. 2004. Workers' education, spillovers, and productivity: evidence from plant-level production functions. *The American Economic Review*, **94**(3), 656–690.

Moretti, Enrico. 2011. Local labor markets. *Handbook of labor economics*, **4**, 1237–1313.

Ngai, L Rachel, & Tenreyro, Silvana. 2014. Hot and Cold Seasons in the Housing Market. *The American Economic Review*, **104**(12), 3991–4026.

Petrongolo, Barbara, & Pissarides, Christopher. 2006. Scale effects in markets with search. *The Economic Journal*, **116**(508), 21–44.

Petrongolo, Barbara, & Pissarides, Christopher A. 2001. Looking into the black box: A survey of the matching function. *Journal of Economic literature*, **39**(2), 390–431.

Postel-Vinay, Fabien, & Robin, Jean-Marc. 2002. Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, **70**(6), 2295–2350.

Roca, Jorge De La, & Puga, Diego. 2017. Learning by working in big cities. *The Review of Economic Studies*, **84**(1), 106–142.

Rosenthal, Stuart S, & Strange, William C. 2004. Evidence on the nature and sources of agglomeration economies. *Handbook of regional and urban economics*, **4**, 2119–2171.

A Appendix: interpreting the estimate

In section (5) I relate the estimated growth differences to the estimated differences in earnings levels. In particular, I compare the differences in earnings of a worker who works in the top quartile of the CZ population distribution, to his earnings if he works in the bottom quartile, using both methods. To understand how this happens, let experience be measured in quarters; further let $X^T\,(exp = 120)$ denote the earnings in the top and $X^B\,(exp = 120)$ the earnings in the bottom after 30 years of work experience. The average quarterly earnings growth in CZ c is estimated via the fixed effects κ_c. I call the growth rate in the top quartile g^T and in the bottom g^B.

First, I make assumptions about the earnings growth process. In particular, assume, as suggested by the descriptive findings in Figure (4.4), that $X^T\,(exp = 0) = X^B\,(exp = 0) = X_0$. With a quarterly growth rate of g^T and g^B in the top and bottom quartiles, respectively, we have

$$\begin{aligned} X^T\,(exp = 120) &= X_0 \left(1 + g^T\right)^{120} \\ X^B\,(exp = 120) &= X_0 \left(1 + g^B\right)^{120} \end{aligned}$$

To be able to relate the findings in section (5) to differences in earnings levels, as in section (4) I am interested in Π:

$$\begin{aligned} \Pi &= \frac{X^T\,(exp = 120) - X^B\,(exp = 120)}{X^B\,(exp = 120)} \\ &= \frac{\left(1 + g^T\right)^{120}}{\left(1 + g^B\right)^{120}} - 1 \end{aligned}$$

Using that $\log\,(1 + x) \approx x$ for small x, I take logs on both sides of the equation to get

$$\Pi \approx \exp\left\{120\left(g^T - g^B\right)\right\} - 1 \tag{8}$$

Hence after determining the difference in growth rates, I can use Equation (A) to translate the growth difference into the earnings difference after 30 years.

To get the quarterly growth difference, suppose I compare a CZ in the middle of the top quartile to a CZ in the bottom. From the estimation equation (5) I get

$$\begin{aligned} \left(g^T - g^B\right) &= E\left[\Delta w | c = T\right] - E\left[\Delta w | c = B\right] \\ &= \kappa^T - \kappa^B + \delta_{1,T} \times \Delta\bar{e}_T + \delta_{2,T} \times \Delta\bar{e}_T^2 \end{aligned}$$

The first part of the equation, $\kappa^T - \kappa^B$ can be evaluated using the second stage regression: $\kappa^T - \kappa^B = \varphi \times \Delta pop$. For overall earnings growth, that would be 0.0004×6.86 for the difference in population between the top and the bottom. The second part of the equation, $\delta_{1,T} \times \Delta\bar{e}_T + \delta_{2,T} \times \Delta\bar{e}_T^2$, can be evaluated using the δ coefficients from the first stage regression together with the average gained experience level, leading to a wage change event in the top CZs, $\Delta\bar{e}_T$.