MATCHING IN CITIES

Wolfgang Dauth
Institute for Employment Research (IAB) and University of Bamberg, Germany

Sebastian Findeisen
University of Konstanz, Germany

Enrico Moretti
University of California, Berkeley, USA

Jens Suedekum
DICE University of Düsseldorf, Germany

Abstract
Using administrative German data, we show that large cities allow for a more efficient matching between workers and firms and this has important consequences for geographical inequality. Specifically, the match between high-quality workers and high-quality plants is significantly tighter in large cities relative to small cities. Wages in large cities are higher not only because of the higher worker quality but also because of a stronger assortative matching. Strong assortative matching in large cities magnifies wage differences caused by worker sorting, and is a key factor in explaining the growth of geographical wage disparities over the last three decades. (JEL: R11, R12)

1. Introduction

In most countries around the world, there are large wage and income disparities between cities and regions. In the United States, the 2014 average hourly wage of a worker in Stamford, CT was twice that of a worker with the same education and

The editor in charge of this paper was Giovanni Peri.

Acknowledgments: We thank our editor Giovanni Peri, four referees, David Card, Gilles Duranton, Laurent Gobillon, Simon Jaeger, Fabian Lange, Attila Lindner, Elena Manresa, Giordano Mion, Michel Serafinelli, Raffaele Saggio, and seminar participants at the ASSA Meetings, Barcelona, UC Berkeley, Cambridge, UC Davis, RWI Essen, Essex, EUI, Haas, Halle, Hamburg, Helsinki, Heidelberg, Mannheim, Munich, Nuremberg, St. Gallen, Trier, Utrecht, and Warwick for helpful comments and suggestions. Linda Borrs, Florian Knauth, and Evan Rose provided excellent research assistance. We thank the department DIM at the IAB and Hans Ludsteck for their support and suggestions with the dataset. We received financial support from the Deutsche Forschungsgemeinschaft (DFG) within the Priority Program 1764 "The German Labour Market in a Globalised World—Challenges through Trade, Technology, and Demographics". Grant numbers: DA 1576/1-2, FI 1939/1-2, SU 413/3-2.

E-mail: wolfgang.dauth@iab.de (Dauth); sebastian.findeisen@uni-konstanz.de (Findeisen); moretti@berkeley.edu (Moretti); suedekum@dice.hhu.de (Suedekum)

Journal of the European Economic Association 2022 20(4):1478–1521
https://doi.org/10.1093/jeea/jvac004

demographics in Flint, MI—a difference significantly larger than in 1980. In Germany, after conditioning on the same variables, the 2014 average wage in Munich was 43% higher than in Uelzen, a small city at the bottom of the wage distribution. This difference is significantly larger today than it was in 1985.[1]

These large and growing wage disparities between communities have become an important source of policy concern, but the exact reasons for their existence are still debated. Geographical wage disparities appear to be associated, at least in part, with city size. In most countries, larger cities tend to enjoy higher wages than medium-sized cities, and medium-sized cities have higher wages than small cities. In Germany, doubling city size is associated with 3.7% higher conditional wages. A similar relationship has been documented, among others, in the United States (Glaeser and Mare 2001), France (Combes, Duranton, and Gobillon 2008), Spain (De La Roca and Puga 2017), the United Kingdom (Rice, Venables, and Patacchini 2006), or Japan (Keisuke 2017).

Which economic forces are responsible for generating such geographical disparities? One possibility is that they reflect sorting of workers with different unobserved skills. Residents of high-wage cities like Munich or Stuttgart may potentially have higher ability than residents of low-wage cities such as Uelzen or Hof.[2] Yet, as we will show, wage disparities remain significant, and city size remains highly correlated with wages even after controlling for worker fixed effects. Hence, there appears to be something else beyond worker quality that systematically affects wages and depends on a worker's location. What causes the link between wages and locations is a fundamental question in labor and urban economics, and while much progress has been made, the exact answer is arguably still unclear.[3]

In this paper, we empirically identify an important mechanism behind geographical wage disparities—namely, spatial variation in the quality of the match between workers and firms. We find that, compared to small cities, large cities allow for a more efficient matching between workers and firms, and this has important consequences for geographical inequality. Large labor markets have long been hypothesized to produce more productive matches between workers and firms than small markets. In many urban economics models, labor pooling is an important advantage of large cities,[4] but direct evidence is scant.

We study the role played by assortative matching between workers and plants in explaining wage differences between German cities. When worker quality and plant quality are gross complements in production, average productivity, and wages are

1. Over this period, the standard deviation of the average conditional wage distribution across metropolitan areas in the United States and Germany grew by 48% and 42%, respectively, suggesting a growing *divergence* in the fortunes of local communities (Glaeser and Kohlhase 2004; Moretti 2011, 2012).

2. The importance of sorting has been documented for the United Kingdom by D'Costa and Overman (2014), for France by Combes, Duranton, and Gobillon (2008), and for Italy by Mion and Naticchioni (2009).

3. For a general discussion, see Duranton and Puga (2014) or Rosenthal and Strange (2004).

4. See Helsley and Strange (1990), Acemoglu (1997), Rotemberg and Saloner (2000), and Krugman (1991).

higher with assortative matching—that is, when high-quality workers are matched with high-quality plants. We show that the match between high-quality workers and high-quality plants is significantly tighter in large cities relative to small cities. Wages in large cities are higher, not only because of the higher quality of their labor force but also because of a stronger assortative matching. Strong assortative matching in large cities magnifies wage differences caused by worker sorting, and is a key factor in explaining the growing disparities between communities over the last three decades. Our findings empirically validate the intuition behind many urban economics models of labor pooling.[5]

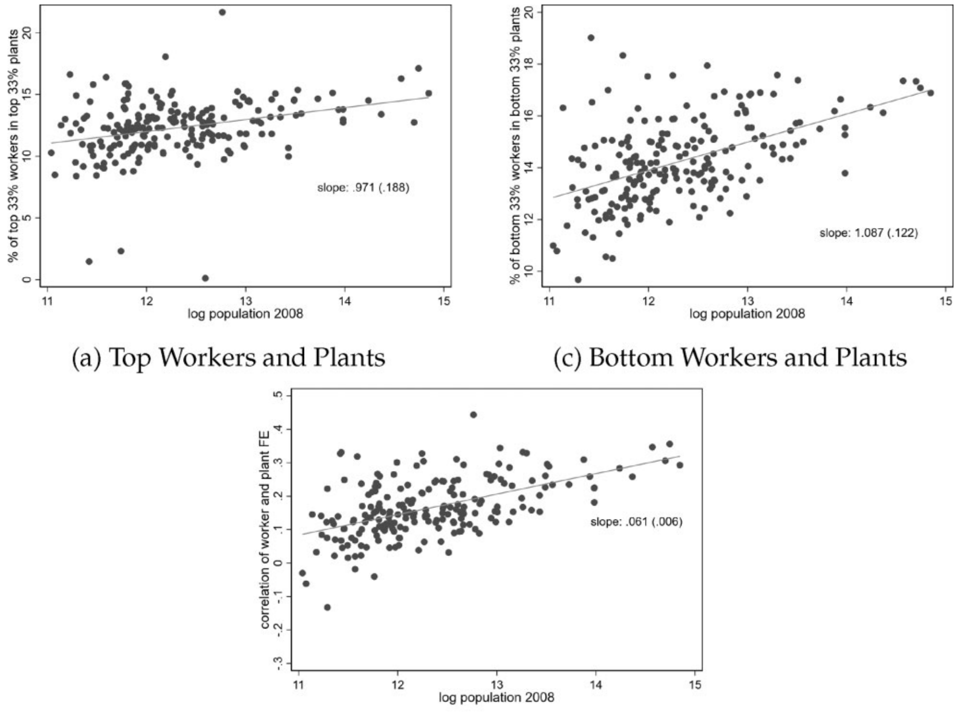
Our analysis is based on a detailed administrative dataset that covers the full job history of the universe of private sector workers in Germany from 1985 to 2014—about 30 million individuals per year—and links them to their plant of work. For each worker and plant, we estimate a fixed effect, which measures worker and plant quality, respectively. The correlation between those two sets of fixed effects (FE) is then used to measure assortative matching.

There are two ways in which assortative matching can arise in practice: between and within local labor markets. *Between city* assortative matching—which we refer to as *co-location*—is the tendency of high-quality workers to locate in cities with many high-quality plants. Empirically, mean worker and mean plant effects display a strong covariance across cities, driven in large part by the fact that both high-quality workers and high-quality plants tend to be over-represented in large cities. The second and perhaps more interesting source of assortative matching then takes place *within cities* for a given spatial distribution of workers and plants across cities. Of particular interest is the relationship with labor market size. We find that larger and denser cities display significantly higher correlations of worker and firm fixed effects. This means that high-quality workers are significantly more likely to be employed in high-quality plants within large cities than within small cities.

Compare Munich (2,531,068 residents) with a medium-sized city like Balingen (190,291 residents) and a small city like Cochem (64,689 residents). The share of workers with fixed effects in the top 33% of the respective city's distribution, who work in plants with fixed effects in the top 33% of the distribution in the three cities, is 17.1%, 12.3%, and 8.5%, respectively. Similarly, the share of workers with fixed effects in the bottom 33% of the distribution who, work in plants with fixed effects in the bottom 33% is 17.1%, 13.5%, and 10.9% in the three cities, respectively. Overall, the correlation between worker and plant effects clearly increases with city size—and is 0.356 in Munich, 0.138 in Balingen, and −0.062 in Cochem.

This pattern holds more broadly across all German cities, as shown in Figure 1, which illustrates our main finding. The top left (right) panel shows the correlation between city size and the share of top (bottom) tercile workers matched with top (bottom) tercile plants in their respective local fixed effects distributions. Both shares

5. Assortative matching has also been shown to be an important driver of changes in the nationwide wage distribution in Germany and other countries (see Card, Heining, and Kline 2013; Card et al. 2018; Torres et al. 2018; Jaeger and Heining 2020), but without reference to the local dimension.



(a) Top Workers and Plants

(c) Bottom Workers and Plants

(b) Correlation of Worker and Plant Effects

FIGURE 1. City size and assortative matching—2008–2014. The vertical axes of this figure stem from an individual level AKM estimation of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies for the period 2008–2014. The figures visualize the correlation between the share of workers in the upper (lower) tercile of the worker fixed effect distribution employed in the plants in the upper (lower) tercile of the plant fixed effect distribution, relative to all workers in a city, and the log population. The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

tend to be significantly higher in larger cities. In the bottom panel, we plot the strength of assortative matching—measured by the within-city correlation of plant and worker effects—against city size. When the population doubles, this correlation increases by 6.1 percentage points.[6]

Thus, our evidence supports the notion that larger cities allow for a more efficient matching between workers and firms. We find an even stronger association when we define a local labor market to be a city-occupation pair, rather than the city as a whole. This is important because workers looking for a job, and plants looking for an employee, are likely to search within specific occupations. Figure 2 shows this relationship for

6. This finding remains highly robust in many empirical specifications. In particular, using the approaches proposed by Bonhomme, Lamadon, and Manresa (2019, 2021) and by Kline, Saggio, and Solvsten (2020), we find no evidence that this finding is driven by limited mobility bias .

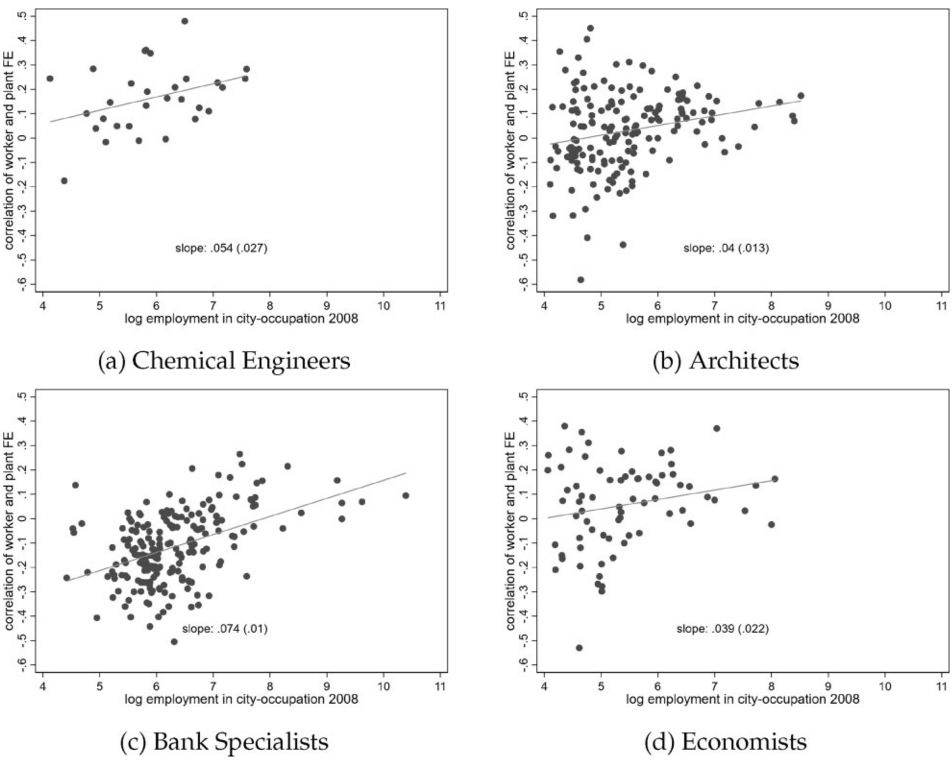


FIGURE 2. Occupation-city size and strength of assortative matching in individual occupations—2008–2014. This figure visualizes the bivariate correlation of assortative matching and log employment across 204 cities for four selected occupations. Assortative matching is defined as the occupation-city level correlation coefficient of worker and plant effects. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

four examples: chemical engineers, architects, bank specialists, and economists. The first panel shows that the probability that a high-quality chemical engineer is matched with a high-quality chemical plant is significantly higher in cities where the stock of chemical engineers is larger. Intuitively, a high-quality plant in Munich looking for a high-quality chemical engineer is more likely to find one because, at any moment in time, there are many more chemical engineers of all qualities looking for jobs in Munich. By contrast, a high-quality plant in Balingen may have to settle for a low-quality engineer, simply because there are not that many candidates around looking for jobs in the city at any given moment in time. In fact, in an average year, there are 106 matches of chemical engineers of any quality to plants of any quality in Munich, and only two matches in Balingen. Across all German cities, we find that doubling the size of a city-occupation increases assortative matching by 5.4 percentage points for chemical engineers.[7] When bunching all occupations together, this elasticity even increases to 6.5 percentage points.

Overall, our findings indicate that wages are higher in larger cities not only because they host more high-quality workers and firms but also because the matching of workers to firms is more efficient. Better matching in large cities is a key explanation for spatial wage disparities in Germany. We show that geographical inequality would decrease significantly if the strength of *within-city* matching was the same in all cities, irrespective of their size.

The relationship between size and assortativeness has been growing stronger over time—it is 75% higher now than in the period 1985–1991. Our simulations indicate that such increase in assortativeness has substantially increased inter-city wage inequality.

At the same time, the increase in assortativeness also had a positive effect on aggregate earnings in Germany. When worker and firm quality are complement, more assortativeness means higher productivity, output, and aggregate earnings. We estimate that the increase in within-city assortative matching observed between 1985 and 2014 increased aggregate labor earnings in Germany by 31 billion euros. Hence, we conclude that assortative matching not only increases earnings inequality across communities, but it also generates important efficiency gains for the German economy as a whole.

The empirical literature on worker–firm matching in local labor markets is still in its infancy. Petrongolo and Pissarides (2006) quantify scale effects in job search by comparing the number of job matches in labor markets of different sizes. Wheeler (2008) and Bleakley and Lin (2012) find that the probability of changing occupation or industry is positively correlated with city size for young workers and negatively associated with city size for older workers. Andersson, Burgess, and Lane (2007) find stronger assortative matching in denser counties in Florida and California. By contrast, Figueiredo, Guimaraes, and Woodward (2014) find limited support for stronger assortative matching in Portugal. Mion and Naticchioni (2009) find a negative correlation between assortativeness and area density. Andini et al. (2013) use survey questions, which measure workers' assessments of match quality. Orefice and Peri (2020) also find evidence for assortativess matching across French regions, which is strengthened by inward migration. Our paper is the first to uncover direct evidence on how the quality of the match between workers and firms varies across cities, and to quantify its effects on geographical inequality and aggregate earnings.

The remainder of this paper is organized as follows. In Section 2, we present the data and some key stylized facts. In Section 3, we describe our empirical approach. In Section 4, we present our empirical results on co-location, while Section 5 shows the results on assortative matching within cities. In Section 6, we conduct some counterfactual experiments. Section 7 concludes.

7. The remaining three panels of Figure 2 show that the same is true for the other examples. In an average year, there are 861 matches of architects, 2,143 matches of bank specialists, and 794 matches of economists of any quality to plants of any quality in Munich, but only 21 matches of architects, 50 matches of bank specialists, and 211 matches of economists in Balingen. The elasticities of assortative matching with regard to city size range from 3.9 to 7.4 percentage points.

2. Geographical Wages Differences in Germany

In this section, we introduce our data, we provide some background about major developments in the German labor market, and present some initial descriptive facts about spatial wage disparities.

Data. We use data from the Employee History of the Institute for Employment Research (IAB).[8] It follows the full job history of the universe of private workers from 1985 to 2014, excluding the self employed. It includes 298,565,604 worker-year-observations and a total of 29,187,865 individuals and 3,252,487 plants. We focus on male full-time workers aged 20–60 in West Germany for the main part of the paper, because data for East Germany became available only after 1991. In the robustness checks, we also report results that encompass East Germany. Worker wages are defined as gross labor earnings per day.[9] Means and standard deviations for all our variables are reported in Appendix Table A.1.

Our main geographical unit of analysis are 204 consistently defined travel-to-work areas (*Arbeitsmarktregionen*), which are similar to US commuting zones and the closest approximation to local labor markets. We refer to those units as *cities*, and in additional specifications, we also use alternative spatial units to address the robustness of our results.[10]

Background. Our empirical analysis covers a 30-year period during which the German labor market underwent major changes. Most prominently, the German reunification in 1990 integrated two hugely different market systems and led to a net-migration of more than 1.5 million people from East to West. Still, even at its peak, this inflow comprised less than 1% of total employment in the relatively larger West German labor market, on which we focus in this paper.[11]

During the observation period, Germany also saw massively increasing trade flows (mainly with Eastern Europe and China) and the surge of new technologies such as robots. Both shocks differentially affected local labor markets, mostly depending on the precise specialization patterns of their manufacturing sectors (Dauth, Findeisen,

8. Specifically, the data is called Beschäftigtenhistorik—BEH, Version V10.01.00–160816. See Oberschachtsiek et al. (2009) for a detailed description of an earlier version of this dataset.

9. One well-known problem in this data is the top-coding of wages at the social security contribution ceiling, which was around 140 euros per day in 2010. We deal with this problem in two ways. First, we follow the imputation methodology proposed by Card, Heining, and Kline (2013). All our main tables are based on this approach. In addition, we re-estimate equation (1) using top-coded data and find that our results are not sensitive. The detailed results tables for that approach are available upon request.

10. More specifically, there we distinguish 325 administrative counties (*Landkreise und kreisfreie Städte, without Berlin)* (which are NUTS-3 regions comparable to US counties), 108 larger commuting zones, and 8,277 small-scale municipalities (*Gemeinden*).

11. Findeisen et al. (2021) argue that the patterns in the micro data make it unlikely that the absorption of these workers had a big impact on the West German wage structure and the allocation of workers.

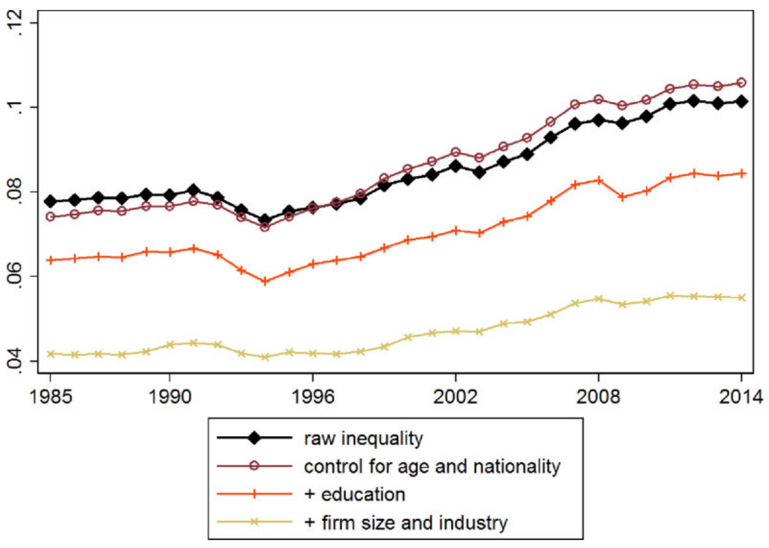


FIGURE 3. Spatial wage dispersion—1985–2014. The figure shows the standard deviation of city-level average log daily wages by year. The black line represents the standard deviation of the observed wages, while the others represent the standard deviations of residuals from regressions controlling for age, nationality, education, plant size, and two-digit industry.

and Suedekum 2014; Dauth et al. 2021). Systematic differences between larger and smaller cities—the key theme of this paper—played a lesser role, however.

Finally, the observation period involved the Hartz labor market reforms, and a general trend away from collective bargaining agreements (which for a long time set industry-level wage floors) toward decentralized wage setting schemes at the firm-level. This led to a decline in the share of workers covered by industry-wide agreements from around 75% in the mid-1990s to 55% in 2008, while the share of employees in firm-level agreements has remained flat over time. Those institutional changes are one key explanation for the strong decline in aggregate unemployment, which is often paraphrased as the transition of Germany from the "sick man of Europe" to becoming its "economic superstar" (Dustmann et al. 2014). That discussion has been largely orthogonal to the causes of rising between-city wage inequality, however, on which we focus in this paper.

Spatial Wage Inequality. Our focus in this paper is on spatial inequality. Like in most countries, Germany exhibits vast differences in mean wages across cities. For instance, among the 204 local labor markets in West Germany, Munich has an average wage roughly 62% higher than the cities at the bottom of the distribution.

Those differences have grown over time. The black line in Figure 3 depicts the evolution of the standard deviation in raw average wages across cities. This measure of spatial wage inequality has been increasing over most of the observation period, with some flattening of the trend since 2008. Controlling for various observable characteristics of the local workforce, such as age, nationality, education, plant size, and industry composition reduces the *level* of wage inequality across regions, but the three curves at the bottom of Figure 3 reveal that the time trends of conditional wage inequality remain very similar.[12]

Wages and City Size. In most countries, wages tend to be positively correlated with city size.[13] Germany is no different in that respect. The two maps in Figure 4 show average daily gross wages and population sizes in 2014 for all 204 German local labor markets. A visual comparison shows a strong correlation. Table 1 shows the average 2008–2014 wage in the five largest cities in terms of population, the median city, and the five smallest cities. Differences in observed wage levels between large and small cities are substantial. For example, the 90−10 difference and 99−1 difference are 0.25 and 0.43 log points, respectively. Column (4) shows the conditional average log wage, after controlling for a cubic in age, education levels, gender, and nationality of the worker. As expected, spatial differences drop significantly, but remain economically important. For example, 90−10 difference and 99−1 difference are 0.20 and 0.33 log points, respectively.

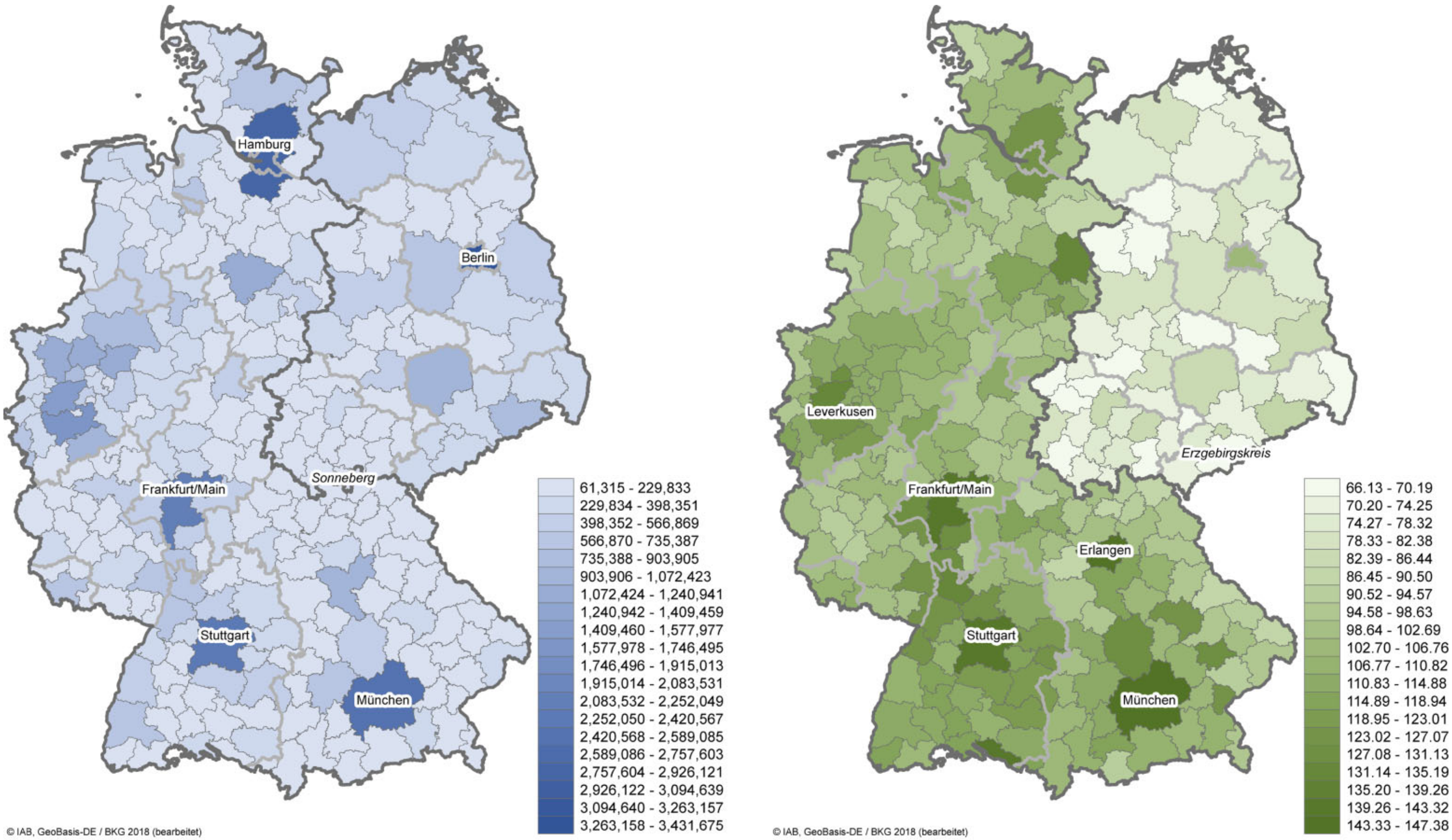
The left panel in Figure 5 plots the mean log wage after controlling for workers' education, demographic variables, time effects, and industry in each of the 204 metropolitan areas in West Germany between 2000 and 2014 against log population.[14] The estimated elasticity is 0.037, indicating that a 10% increase in population is associated with a 3.7% increase in average wages, holding constant workers' observables.

It is, of course, possible that workers with high *unobserved* ability sort into cities with higher wages. For example, Combes et al. (2012) have shown that the sorting of high-ability individuals plays a major role in explaining spatial wage

12. The literatures on geographical sorting and agglomeration economies is very large. See, among others, Glaeser and Mare (2001), Yankow (2006), Gould (2007), Combes, Duranton, and Gobillon (2008), Baum-Snow and Pavan (2012), Eeckhout, Pinheiro, and Schmidheiny (2014) or De La Roca and Puga (2017), who describe different forms of sorting mechanisms and agglomeration effects, and provide evidence at the worker-level; or Henderson (2003), Moretti (2004), Combes et al. (2012), Gaubert (2018) who analyze agglomeration effects from the perspective of firms. Behrens, Duranton, and Robert-Nicoud (2014) present a tractable framework in which founders and workers with heterogenous talents can self-select into cities. Consistent with the predictions of their model, we find positive selection of workers and plants in cities.

13. To give a few examples, this is true for the United States, where Glaeser and Mare (2001) find an urban wage premium of 24.5% in metropolitan areas in cities with at least half a million citizens after controlling for experience, education, race, tenure, and occupation. It is also true for France, where Combes, Duranton, and Gobillon (2008) find an elasticity of unconditional wages with respect to city size of 5.15, and for Spain, where De La Roca and Puga (2017) find an elasticity of wages with respect to city size of 4.6 after controlling for similar characteristics.

14. In practice, we follow Combes, Duranton, and Gobillon (2008) and use a two-stage procedure. In the first step, we regress log individual wage on a city fixed-effect for the current location of a worker and a vector of standard individual-level control variables. We use the same control variables in this model as in the AKM specification later on, namely education-specific age profiles with a cubic functional form and a set of year dummies. In the second step, we regress the estimated city fixed-effects on an intercept and log population.



(a) Population

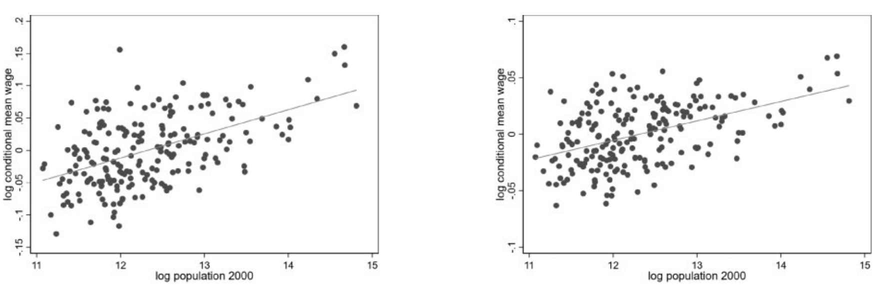
(b) Average Wage

FIGURE 4. Population and average wage by city, 2014. The maps visualize the spatial distribution of the population size (left) and average wage (right).

TABLE 1. Average wage in the largest and smallest cities.

Rank	City	(1) Population	(2) Daily wage	(3) Log daily wage	(4) Res. log daily wage
1	Hamburg	2,803,463	12349	4.676	4.608
2	München	2,531,068	14738	4.845	4.732
3	Stuttgart	2,419,694	14250	4.834	4.758
4	Frankfurt/Main	2,124,514	14312	4.815	4.712
5	Köln	1,737,116	12703	4.710	4.639
102	Balingen	190,294	11206	4.632	4.648
200	Holzminden	75,092	10094	4.530	4.532
201	Kronach	71,609	9184	4.440	4.464
202	Lichtenfels	68,617	8866	4.409	4.438
203	Cochem	64,489	9483	4.480	4.502
204	Daun	62,201	9646	4.500	4.527
	Standard deviation	397,441	1230	0.098	0.080
	75–25	186,275	1536	0.132	0.110
	90–10	532,636	3046	0.250	0.204
	99–01	2,351,077	5596	0.430	0.334

Notes: Columns (2) and (3) report (log) wage levels by city for all West German cities in the period 2008–2014. Column (4) reports residualized log wages. Controls for the residual wages: educational attainment, nationality, gender, and cubic terms in age.



(a) Controlling for Worker Observables (b) Controlling for Worker Observables and Fixed Effects

FIGURE 5. City size and average conditional wages—2000–2014. Average wages are conditional on worker characteristics. The upper panel controls for education, and education-specific cubic age profiles, education-specific time effects, and two-digit industry fixed effects. Slope: 0.037 (s.e. = 0.004). The lower panel controls for the same variables and also individual fixed effects. Slope 0.017 (s.e. = 0.002).

differences between French cities. D'Costa and Overman (2014) show that workers in the United Kingdom sort into bigger cities on the basis of observable and unobservable characteristics. Once this sorting is accounted for, they find no further wage growth in larger cities. Mion and Naticchioni (2009) show that three quarters of spatial wage differences in Italy can be explained by the sorting of more able workers.

In Germany, sorting also matters. But after controlling for worker fixed effects, there continues to be a positive relationship between average wages and city size.

This is shown in the right panel of Figure 5. The elasticity drops to 1.7%, but remains economically meaningful and statistically significant. Combes et al. (2012) use historical population data as instruments for current city size. If we use the 1952 population as an instrument, then the estimated two-stage least squares (2SLS) elasticity is 3.8%, virtually unchanged relative to the ordinary least squares (OLS) elasticity in the left panel of Figure 5. Controlling for individual fixed effects yields an elasticity of 1.8%.

Overall, the relationship between wages and city size in Germany is similar as in other countries. The results so far do not take into account the matching of workers and firms within and between local labor markets, however, which is the key aspect of the this paper.

3. The Geography of Assortative Matching: Conceptual Framework and Estimation

To empirically measure worker and plant quality, and their matching, we follow the approach by Abowd, Kramarz, and Margolis (1999), henceforth labeled AKM. This approach has been widely used in the growing literature on assortative matching, though not yet in the context of local labor markets. The paper in this literature, that is, closest to ours is the study by Card, Heining, and Kline (2013), who explore the contributions of assortative matching to the rise in nationwide wage inequality in Germany, but do not focus on intra-national spatial disparities in assortative matching across and within cities.

Abowd, Kramarz, and Margolis (1999) decompose variation in wages into worker- and plant-specific pay components by assuming that the log wage of a worker i can be written as

$$\ln(\text{wage}_{it}) = \mu_i + \Psi_{\mathrm{J}(i,t)} + X'_{it}\gamma + \varepsilon_{it}, \tag{1}$$

where μ_i are worker effects, $\Psi_{\mathrm{J}(i,t)}$ are plant effects[15], and X_{it} is a vector of observable worker characteristics, which in our application includes education-specific age profiles—in the form of a quadratic and a cubic term in age interacted with four dummies for educational attainment—and year fixed effects.

In our empirical analysis, we focus on the covariance $\mathrm{Cov}(\mu_i, \Psi_{\mathrm{J}(i,t)})$. We define (positive) assortative matching as a (positive) correlation between the worker effects μ_i and the plant effects $\Psi_{\mathrm{J}(i,t)}$ in a local labor market.

The AKM model is not a structural model of the labor market, and the identified worker and plant effects do not necessarily measure true ability or productivity (Abowd et al. 2004, 2018; Eeckhout and Kircher 2011). It is consistent, however, with a

15. The plant effects are identified by the mobility of workers between plants during the 7-year period. This means that plants are disregarded if they are not connected to other plants by plant-movers. While workers may have a spell of non-employment between working at two plants, entries, and exits in and out of the labor force do not contribute to the identification of the plant effects.

variety of production functions where the plant effect captures plant-specific total factor productivity. Consider, for example, the case where the output generated by the match between worker i and plant J is given by

$$Y_J \propto \tilde{\mu}_i^a \tilde{\Psi}_{\mathrm{J}(i,t)}^b. \tag{2}$$

With such a production function, log wages are additive in worker and plant effects and, thus, have no match-specific component. But wages are multiplicative in levels, and thus, "good" workers (with high μ_i) earn relatively more than "bad" workers (with low μ_i) when working for a "good" plant (with high $\Psi_{\mathrm{J}(i,t)}$). Therefore, if worker quality and plant quality are complements in production, then positive assortative matching might arise, whereby high (low) quality workers tend to be matched with high (low) quality plants.

3.1. Assortative Matching Between and Within Cities

Assortative matching has potentially important consequences for economic geography. If worker quality and plant quality are complements, then assortative matching magnifies wage differences across cities for two reasons. First, wages in cities with more good workers and good plants are higher than wages in cities with fewer good workers and good plants—not only because of the difference in quality but also because of this match component. In addition, if the strength of assortative matching increases with city size, then geographical wage differences will be further magnified.

More specifically, we can decompose the overall covariance between worker and plant effects into the part of assortative matching that takes place between cities and the part that takes place within them

$$\mathrm{Cov}(\mu_i, \Psi_{\mathrm{J}(i,t)}) = \underbrace{\mathrm{Cov}\left(E_c[\Psi_{\mathrm{J}(i,t)}], E_c[\mu_i]\right)}_{\text{Between}} + \underbrace{E\left[\mathrm{Cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})\right]}_{\text{Within}}, \tag{3}$$

where c indicates a city and E_c and Cov_c are the respective moments at the city level. The first term describes the covariance between average worker and plant quality across cities and captures the *between part*. This term measures the degree to which high-quality workers and plants co-locate in the same cities. The second term captures the formation of matches *within* cities.

Between-City Matching. We might see positive assortative matching between cities if high-quality workers tend to locate in cities where high-quality plants are also located. In this case, the correlation of average worker and plant effects should be positive, indicating that cities with an above-average share of high-quality plants—which in Germany tend to locate in larger cities—will also have a higher than average share of high-quality workers. We will refer to this form of assortative matching as *co-location*.

In practice, there are a variety of reasons why we might observe co-location occurring in the data. If worker quality and plant quality are complements in production, then good workers and good plants have an incentive to co-locate. Alternatively, it is also possible that larger cities offer workers better or more varied consumption amenities (Glaeser, Kolko, and Saiz 2001; Diamond 2016) and also offer productivity advantages to plants in the form of productive amenities-for example, transportation infrastructure or other locational advantages. If good workers have stronger tastes for consumption amenities found in large cities and good plants have a higher return to productive amenities found in large cities, then co-location might arise even in the absence of complementarities. For the purpose of our analysis, what matters is the degree of co-location and not the specific mechanism that generates it.

Within-City Matching. For a given spatial distribution of worker and plant effects across cities, assortative matching might take place within each city. If worker and plant quality are complements in production, then there is an incentive for good workers in a city to be matched with good plants in that city.

Of particular interest for us is the relationship between the strength of assortative matching and labor market size. The degree of assortative matching within cities does not have to be geographically uniform, but may increase on city size. Large labor markets have long been hypothesized to produce better matches between workers and plants than small markets.

The intuition was first provided in the barter model by Diamond (1982), where the probability of finding a trading partner depends on the number of potential partners available, so that an increase in the size of the market makes trade easier. In urban economics models, labor pooling has long been assumed to be a potentially important advantage of large cities. In Helsley and Strange (1990), a worker–plant match is more productive in areas where there are many plants offering jobs and many workers looking for jobs.[16]

As a concrete example of Diamond (1982)s intuition, consider again the market for chemical engineers in a large city like Munich (2,531,068 residents) and in a small city like Balingen (190,291 residents), mentioned earlier in the introduction. Our data indicate that in the period 2008–2014, there were a total of 633 matches of chemical engineers (of all qualities) with 361 plants (of all qualities) in Munich but only 12 matches of chemical engineers with 10 plants in Balingen. The barter model suggests that one reason one might expect stronger assortative matching in larger cities is that a high-quality firm looking for a high-quality chemical engineer in a small city like Balingen has a lower probability of finding a high-quality chemical engineer, simply because there are not that many chemical engineers looking for jobs at any given moment in time, and it may need to settle for a worse match. By contrast, the probability that a high-quality firm finds a high-quality chemical engineer should be higher in a large city like Munich, because there are many chemical engineers of all qualities looking for jobs at any given moment in time.

16. Acemoglu (1997) and Rotemberg and Saloner (2000) propose alternative mechanisms. See Moretti (2011) for a survey. See Chade, Eeckhout, and Smith (2016) for a precise characterization of the conditions for assortative matching between heterogeneous workers and firms.

A separate reason for why the degree of assortative matching may increase in city size is that under production complementarities, the incentive of forming better worker—plant matches is stronger for "good workers" and "good plants", when large cities have more "good workers" and "good plants" than small cities.

This point can be illustrated with a simple toy model. Suppose there are two cities: C (large) and R (small). Assume that there are two equally large groups of workers in city C with ability $\mu_i^C = \{11, 9\}$ and two equally large groups of plants with productivity $\nu_J^C = \{110, 90\}$. In R, we also have two equally sized groups, but their quality is lower: $\mu_i^R = \{6, 4\}$ and $\nu_J^R = \{60, 40\}$. Suppose that a worker–plant match generates revenue $\mu_i \times \nu_J$—as in the production function in equation (2)—which is equally split. With initial random matching within every city, the average wage (and profit) is thus $(10 \times 100)/2 = 500$ euros in city C and only $(5 \times 50)/2 = 125$ euros in city R. Assume that workers can search only locally and that the plants can switch their single employee at a fixed cost F. If the costs F are similar in the two regions, then it is easy to see that good plants in the larger city C have a strongest incentive to re-match than good plants in the smaller city R and this ultimately ends up increasing assorative matching in C and raising wages there. A good plant in C, that is, initially matched with a bad worker can gain $(110 \times 11 - 110 \times 9)/2 = 110$ euros when switching to a good worker. This gain is larger than the one for a good plant in R (60 euros). For the range $60 < F < 110$, only plants in C have turnover in equilibrium and eventually all good plants end up employing good workers. In R, by contrast, there is no turnover. In this case, the average wage increases to 505 euros in C, and it remains unchanged in R. Thus, the increased assortative matching in C raises the mean wage in C and the wage inequality between cities.[17]

Ultimately, the relationship between assortative matching and city size is an empirical question, one that we address in our empirical analysis below.

3.2. Estimation Issues

The worker and plant effects in equation (1) are identified by individuals moving across plants. Thus, the estimation of AKM models requires a large longitudinal data set that ideally covers the country's entire workforce and all of its plants. Our data is well suited because it allows us to follow the entire job history of all private sector worker from 1985 to 2014. We split the sample into five 7-year time intervals: (i) 1985–1991, (ii) 1990–1996, (iii) 1996–2002, (iv) 2002–2008, and (v) 2008–2014 and estimate the AKM model (1) separately for every interval.[18]

17. Of course, some bad workers in C lose. With perfect assortative matching, they all end up with a wage of 405 euros, lower than the 495 euros for those who initially happened to be in good plants under random matching. But this loss is smaller than the gain for re-matched good workers, whose wage increases from 495 euros to 605 euros.

18. For every individual worker, we record the main job held on June 30 in every year and compute the correlations of μ_i and $\Psi_{J(i,t)}$ pertaining to the job held in the first year of the respective time interval.

Estimation of the AKM model (1) hinges on several assumptions. Card, Heining, and Kline (2013) provide a detailed discussion and a series of empirical tests.[19] We successfully replicated their validity tests for our version of the data set. The log additive structure of the AKM model provides a good approximation to the German wage distribution.[20]

Limited Mobility Bias. An important issue in the estimation of the AKM model (1) is the presence of limited mobility bias, which could potentially lead to a downward bias in the estimated covariance of worker and plant effects (Abowd et al. 2004). In our context, we are particularly concerned that this bias may vary systemically with local labor market size. Larger markets tend to have more worker mobility, and this could drive some our results in Section 5 on assortative matching within cities.

We address this concern in two ways. The first way is to employ the leave-out estimation by Kline, Saggio, and Solvsten (2020). This procedure extends the AKM framework by unbiased estimates of the variance components. In order to obtain one correlation of worker and plant effects for each city, we run this procedure city-by-city.[21]

Our second way to address the limited mobility bias is to follow the discussion in Bonhomme, Lamadon, and Manresa (2019, 2021) and estimate a set of grouped-fixed effects models. Instead of obtaining a fixed effect for each plant, we allocate all plants in our sample into $k = 10, 15, 20$ groups with similar wage structures using a k-means cluster analysis. We then measure the distribution of wages in each plant by $m = 20, 40$ wage perceniles.[22] Since there is much more mobility of workers

Notice that both μ_i and $\Psi_{J(i,t)}$ may vary across time intervals if the same worker or plant is observed in the data in more than one interval.

19. They find that workers of different skill groups receive approximately the same proportional wage premiums at a given plant—consistent with the simple additive structure of equation (1). Second, a fully saturated model with job-specific fixed-effects only yields a marginal improvement in terms of data fit. Finally, the match-specific component of the residual is uncorrelated with the direction of job switches between high- and low-paying plants.

20. Bonhomme, Lamadon, and Manresa (2019) find in Swedish data that the log additive structure of the AKM model provides a good approximation to the wage distribution. The results and tests in Macis and Schivardi (2016) also suggest a good fit of the AKM model in Italy. On the identification of positive assortative matching in AKM models, see also Eeckhout and Kircher (2011), Abowd et al. (2018), Chade, Eeckhout, and Smith (2016). In Online Appendix B.1, we demonstrate that the assumption of exogenous mobility also holds when we distinguish between plant movers that switch either between or within cities. We further show that there is enough spatial worker mobility to ensure a well-connected set of cities. Both suggest that the AKM model can be applied to our spatial setting.

21. This implies that the plant effects are now only identified from within-city mobility. Their first moments are therefore not comparable across cities. This prevents us from using this procedure as our baseline model. The procedure is computationally very demanding, both in terms of speed and memory. To implement it on our 8-core, 64 gigabyte server, we follow the suggestions for large datasets in the Computational Appendix of Kline, Saggio, and Solvsten (2020). First, we residualize the outcome $\ln(\text{wage}_{it})$ from the observable characteristics X_{it}. Second, we collapse the dataset to the match level. Third, we use the Johnson–Lindenstrauss approximation.

22. The plant clusters represent the heterogeneity of wages across German plants remarkably well. Depending on the interval and the choice of k and m, they capture between 94.48% and 97.98% of the variation of average wages across all German plants.

between these plant clusters than between individual plants, any bias should be mitigated.

Residual Plant Effects. Another issue arises in interpreting the plant effects in equation (1). Since plants virtually never change their location, all time-invariant characteristics of the plant—in particular, its industry affiliation—are captured by the $\Psi_{J(i,t)}$.[23] We construct a set of plant effects that are purged of those influences. In particular, we de-mean the estimated plant effects at the industry level (two-digit) and label those *residual plant effects*. These estimates are orthogonal to local industry structures by construction. In what follows, we present our key results for both the unconditional and the residual plant fixed effects.

Nominal Versus Real Wages. Finally, we estimate equation (1) using nominal wages as the outcome variable. The reason is that they are costs from the perspective of plants. The variation of nominal wages therefore reflects productivity differences, while real wages reflect workers' utility (Moretti 2011). Nonetheless, we use real wages, deflated by a regional price index, in a robustness check. Since log real wages are the sum of log nominal wages and the log price deflator, the latter will simply be a constant added to the fixed effects of all plants in the same city, as long as plants do not move while keeping their plant-id, which is very uncommon. Using real wages instead of nominal wages therefore has implications only for co-location but not for our main results on within-city matching, as adding a constant does not affect the covariance.

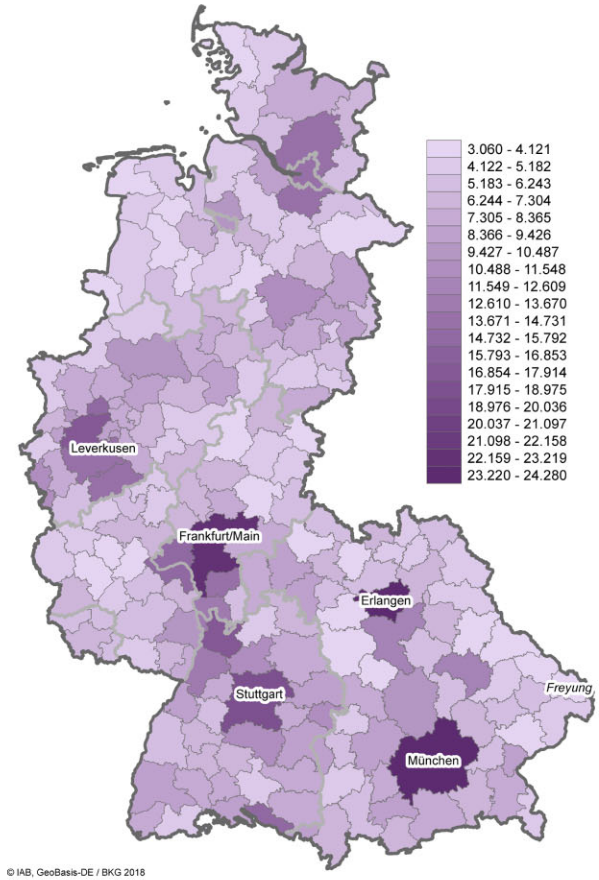
4. Empirical Results on Between-City Assortative Matching

In this section, we quantify the degree of assortative matching between German cities (co-location) and the role played by city size in co-location patterns. In the next section, we study the link between the strength of within-city assortative matching and city size.

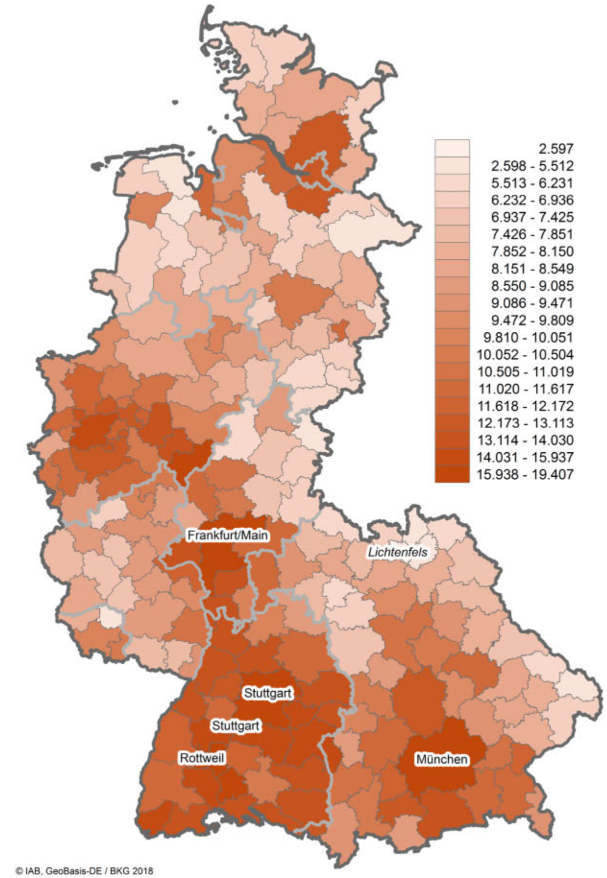
4.1. Co-Location and City Size

The left panel in Figure 6 shows for each city the fraction of workers in the top 10% of the national distribution of fixed effects. The top three cities according to this measure in 2008–2014 are Munich, Erlangen, and Frankfurt. Stuttgart, Duesseldorf, and most of the largest and densest cities are all among the top 20. The bottom three

23. To quantify how much of the geographical variation in plant effects is driven by nationwide differences across industries and how much is due to spatial differences orthogonal to industry effects, we regress the plant effects identified in equation (1) on a vector of city identifiers, a vector of two-digit industry identifiers, and a quadratic in initial employment. Following Huettner and Sunder (2012), we decompose the overall fit of the model into the contributions of the each variable group. We find that at most 17% of the variation in plant effects can be explained by the combination of plant size, industry, and geography (Appendix Table A.2). The bulk of the variation seems to stem from genuine plant-specific differences orthogonal to those dimensions.

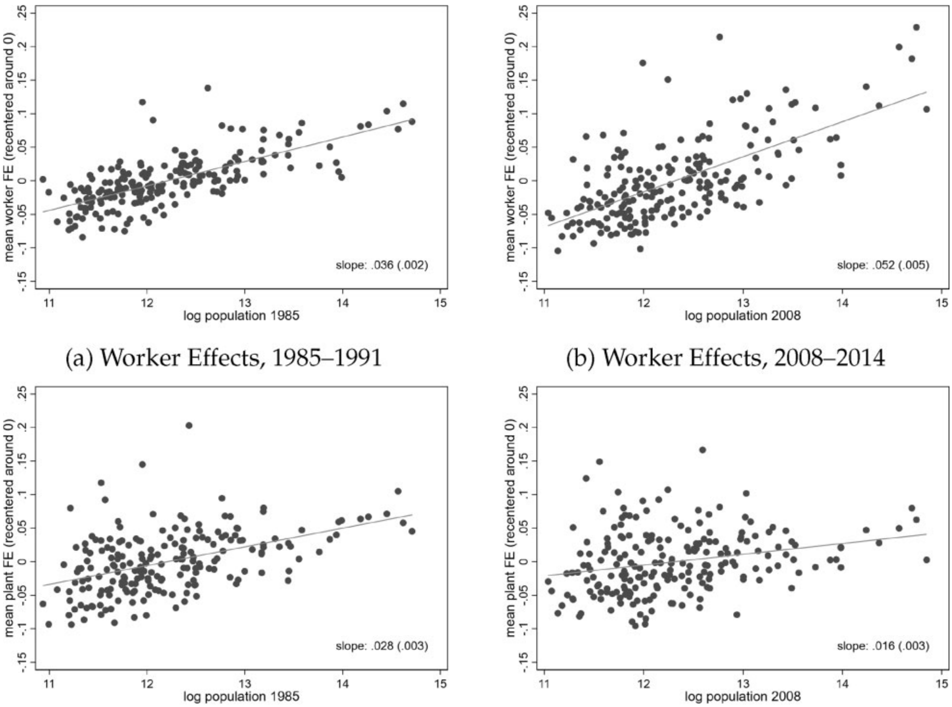


(a) Top decile workers



(b) Top decile plants

FIGURE 6. Percentages of the national top decile workers and plants in the local labor market, 2008–2014. The maps visualize the spatial distribution of the share of a city's workers with a worker effect in the top decile of the national distribution (left) and the share of a city's plants with a plant effect in the top decile of the national distribution (right). The worker and plant effects stem from an individual level AKM estimation of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies for the period 2008–2014.



(a) Worker Effects, 1985–1991

(b) Worker Effects, 2008–2014

(c) Plant Effects, 1985–1991

(d) Plant Effects, 2008–2014

FIGURE 7. Correlation of worker and plant fixed effects and population across 204 labor markets. This figure visualizes the bivariate correlation of log population and average worker effects (top) or average plant effects (bottom) across 204 cities. The worker and plant effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

cities are Lichtenfels, Cloppenburg, and Neustadt/Aisch. The right panel shows the fraction of plants in the top 10% of the national fixed effects distribution. The top three cities in 2008–2014 are Wolfsburg, Salzgitter, and Dingolfing, which are all prominent locations of car manufacturers. The bottom three cities are Hof, Leer, and Uelzen.

A comparison of the two maps points to a positive association across German cities between worker and plant effects. Some cities—mostly large cities—appear to have a disproportionate share of both workers with high fixed effects and plants with high fixed effects, while other cities—mostly small cities—appear to have a lower share of both.

Panels (a) and (b) of Figure 7 depict the relationship between worker quality and market size. In particular, it plots initial (log) population size against the mean worker effect in the respective city. The corresponding elasticities for the early and latest periods are 0.036 (s.e. = 0.002) and 0.052 (s.e. = 0.005) , respectively. Hence, doubling the local population is associated with a 5.2% higher mean worker effect in

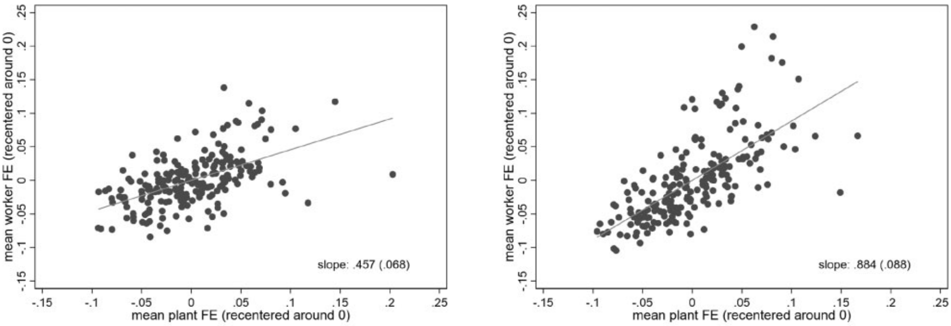


FIGURE 8. Correlation of worker FE and plant FE. This figure visualizes the bivariate correlation of average worker effects and average plant effects across 204 cities. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

2008–2014, a relationship that has become more pronounced over time. Put differently, "good workers" with high individual-specific wage components are mostly found in large, dense cities, and particularly so in more recent years. If we instrument the current population with the 1952 population, then the estimated elasticity for the most recent period increases from 0.052 to 0.056 (s.e. = 0.005).

Panels (c) and (d) of Figure 7 depict the relationship between plant effects and population. The elasticity of mean regional plant effects with respect to (log) population is 0.028 (s.e. = 0.003) in the first period and 0.016 (s.e. = 0.003) in the last period. Unlike for worker effects, this elasticity has thus decreased over time.[24]

Overall, it appears that in Germany, "good workers" and "good plants" are concentrated in larger cities. The concentration of "good workers" has become more important over time, while the concentration of "good plants" is stable. These two facts together imply that worker and plant effects display a strong degree of correlation across German cities and that this correlation is growing over time.

Figure 8 plots mean plant effects (x-axis) against worker effects (y-axis) in 1985–1991 and 2008–2014. The figure confirms that the already strong correlation in the early period has become even stronger more recently. Table 2 shows that the slope in a regression of mean plant on mean worker effects across cities is 0.61 (s.e. = 0.091)

24. A regression of the residual plant effects, which we have demeaned at the two-digit industry level, on the (log) population yields almost identical elasticities. If we instrument the current population with the 1952 population, then the estimated elasticity is 0.031 (s.e. = 0.004) in the first period and 0.017 (s.e. = 0.004) in the last period.

TABLE 2. Co-location: regression of mean worker fixed effects on mean plant fixed effects.

	(1) 1985–1991	(2) 1990–1996	(3) 1996–2002	(4) 2002–2008	(5) 2008–2014
Panel A: unweighted					
Average plant FE	0.4574***	0.4319***	0.7378***	0.9069***	0.8836***
	(0.068)	(0.075)	(0.086)	(0.076)	(0.088)
R^2	0.291	0.255	0.405	0.469	0.462
Panel B: weighted by lagged population					
Average plant FE	0.6027***	0.6418***	1.0166***	1.2401***	1.2993***
	(0.091)	(0.106)	(0.130)	(0.157)	(0.184)
R^2	0.364	0.332	0.474	0.485	0.478

Notes: City-level regressions. The dependent variable is the city level average worker effect from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. Robust standard errors in parentheses. Levels of significance: ***1%.

in 1985–1991 and 1.30 (s.e. = 0.184) in 2008–2014 when we weight by lagged city size.[25]

Summing up, this suggests that "good workers" and "good plants" tend to co-locate mainly in larger cities.

4.2. Variance Decomposition

In order to assess quantitatively the importance of co-location in explaining wage differences across German cities, we use equation (1) to decompose the between-city variance in mean wages into the following components:

$$\begin{aligned}\mathrm{Var}(E_c[\ln\ \mathrm{wage}_{it}]) = {} & \mathrm{Var}(E_c[\mu_i]) + \mathrm{Var}(E_c[\Psi_{\mathrm{J}(i,t)}]) + \mathrm{Var}(E_c[X'_{it}\gamma]) \\ & + 2\,\mathrm{Cov}(E_c[\Psi_{\mathrm{J}(i,t)}], E_c[\mu_i]) + 2\,\mathrm{Cov}(E_c[\mu_i], E_c[X'_{it}\gamma]) \\ & + 2\,\mathrm{Cov}(E_c[\Psi_{\mathrm{J}(i,t)}], E_c[X'_{it}\gamma]).\end{aligned}$$

The decomposition indicates that geographical wage differentials depend on geographical differences in mean worker quality $\mathrm{Var}(E_c[\mu_i])$ and mean plant quality $\mathrm{Var}(E_c[\Psi_{\mathrm{J}(i,t)}])$. Crucially, it also depends on assortative matching between cities (co-location), measured here by the covariance of mean worker quality and mean plant quality: $\mathrm{Cov}(E_c[\Psi_{\mathrm{J}(i,t)}], E_c[\mu_i])$. Thus, for a given amount of spatial differences in mean worker and plant quality, a stronger degree of assortative matching across cities results in larger spatial wage differences across cities.[26]

25. The corresponding slopes in an unweighted regression are 0.46 in the first period and 0.88 in the last period. These results also hold for the other spatial units: 325 administrative counties, 108 larger commuting zones, and 8277 small-scale municipalities (*Gemeinden*).

26. This decomposition is analogous to the one proposed by Card, Heining, and Kline (2013) for the decomposition of national wage inequality between occupation and education groups. Note that for

TABLE 3. Decomposition of across-city variation in average wages.

	(1) 1985–1991	(2) 1990–1996	(3) 1996–2002	(4) 2002–2008	(5) 2008–2014	(6) %	(7) Δ 2014–1985	(8) %
Var mean log wages	61.7	58.2	64.0	79.4	94.6	100	32.9	100
Var mean worker effects	15.2	15.6	22.8	32.1	37.7	39.8	22.5	68.4
Var mean plant effects	21.1	21.3	17.0	18.3	22.3	23.6	1.2	3.6
Var mean Xb	0.8	0.4	0.3	0.4	0.6	0.6	−0.2	−0.7
2 Cov(worker, plant)	19.3	18.4	25.0	33.2	39.4	41.7	20.1	61
2 Cov(worker, Xb)	2.8	1.5	−0.8	−3.0	−3.5	−3.7	−6.3	−19.1
2 Cov(plant, Xb)	2.4	1.1	−0.2	−1.5	−1.9	−2	−4.3	−13.1

Notes: Variance decomposition of between-city wage inequality across 204 labor markets. The parts of the decomposition stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. The city average log wages are city-level averages of the fitted values. Column (6) is the percentage contribution of each term in the last cross-section 2008– 2014. Column (8) is the percentage contribution of each term to the change in the between-region variance between the first and last time interval.

Table 3 shows the results of this decomposition exercise. Columns (5) and (6) indicate that 40% and 24% of the cross-sectional differences in average wages across German cities in 2008–2014 reflect variation in worker and plant effects, respectively. Differences in workers' observable characteristics—namely, education-specific age profiles—play a negligible role after conditioning on fixed effects. The finding that worker effects are an important determinant of spatial disparities is not new. Indeed, it is consistent with findings established in Combes, Duranton, and Gobillon (2008) for France, who show that sorting of high-quality workers is an important explanation for wage differences across French cities.

The key finding in Table 3, however, is that 42% of the cross-sectional differences in average wage across German cities in 2008–2014 are due to co-location (column (6)). Put differently, between-city matching across German cities is the most important part in this decomposition, accounting for almost half of the cross-sectional variation.

Turning to the evolution over time, we find that the importance of assortative matching across cities is not fading but getting stronger. Columns (1)–(5) show that, after staying constant from the mid-1980s to the mid-1990s, the covariance of worker and plant effects has been steadily increasing. Co-location of "good workers" and "good plants" accounts for as much as 61% of the increase in spatial wage disparities over the entire observation period (column (8)). A more unequal distribution for worker effects across space is the other important driver, accounting for 68% of the increased variance. By contrast, changes in the dispersion of the plant effects (i.e., rising workplace heterogeneity) have not contributed significantly to changes in geographical inequality between high-wage and low-wage cities.

simplicity here, we are focusing on the variance in log wages rather than wages in levels. This means that we are abstracting from the role played by within-city assortative matching and focusing only on the role played by co-location.

5. Within-City Assortative Matching and City Size

We now turn to assortative matching *within* cities and its relationship with city size. We first estimate the degree of assortativeness of worker–plant matching for every city in our sample. We then estimate how the degree of assortativeness varies as a function of city size.

5.1. Baseline Estimates

There are large differences in the strength of assortative matching across space. Figure 9 shows geographical differences in the strength of assortative matching. Specifically, for each city, the map shows the correlation between worker and plant effects in the period 2008–2014. The map clearly shows that assortative matching varies significantly across areas. The three cities with the highest degree of assortative matching are Erlangen (0.44), Munich (0.36), and Frankfurt (0.35). The difference between the cities with the largest and smallest correlations is 0.58. The 75−25 quartile spread is 0.12 (Appendix Table A.3).

In Figure 10, we plot the city-specific correlation coefficients against (log) population across all 204 local labor markets in 1985–1991 (panel (a)) and in 2008–2014 (panel (b)). We find a positive and statistically significant relationship, with an estimated elasticity of 0.038 (s.e. = 0.005) in the early period and 0.061 (s.e. = 0.006) in the most recent period.

This means that larger cities are characterized by a substantially stronger degree of assortativeness.[27] Notably, the relationship between city size and the degree of assortativeness has become stronger over time. Larger cities were already characterized by a strong degree of assortative matching in their local labor market in 1985–1991, but this relationship is much stronger is the most recent period.

The corresponding estimates are reported in Table 4. Panel A reports OLS estimates of a regression of the degree of assortativeness on the current (log) population. The coefficients of columns (1) and (5) correspond to those shown in Figure 10. It is possible that city size in a given year may be driven, at least in part, by the expectation of better matching within the local labor market. In panel B, we show 2SLS estimates if we instrument the current population with the 1952 population. The estimated elasticity is 0.042 in the first and 0.070 in the last time period, which are slightly larger than the corresponding OLS coefficients. In panel C, we measure city size by the log employment in a city, instead of the log population. We find that the elasticity of

27. Notice that the *level* of the correlation coefficient on the vertical axis is not particularly high. For some regions, it is even negative. This is a common pattern in applications of the AKM approach; however, that is scrutinized in a substantial literature which argues that the measured correlation understates the true degree of assortative matching (e.g., Andrews et al. 2008; Eeckhout and Kircher 2011; Abowd et al. 2018). Also see the applications by Card, Heining, and Kline (2013) and Andersson, Burgess, and Lane (2007), who find similar magnitudes and domains for the correlations as in our Figure 10. For this paper, we are less interested in the *level* of the correlation than how it varies with population size and over time. This inference is unaffected by those identification problems when the bias is constant.

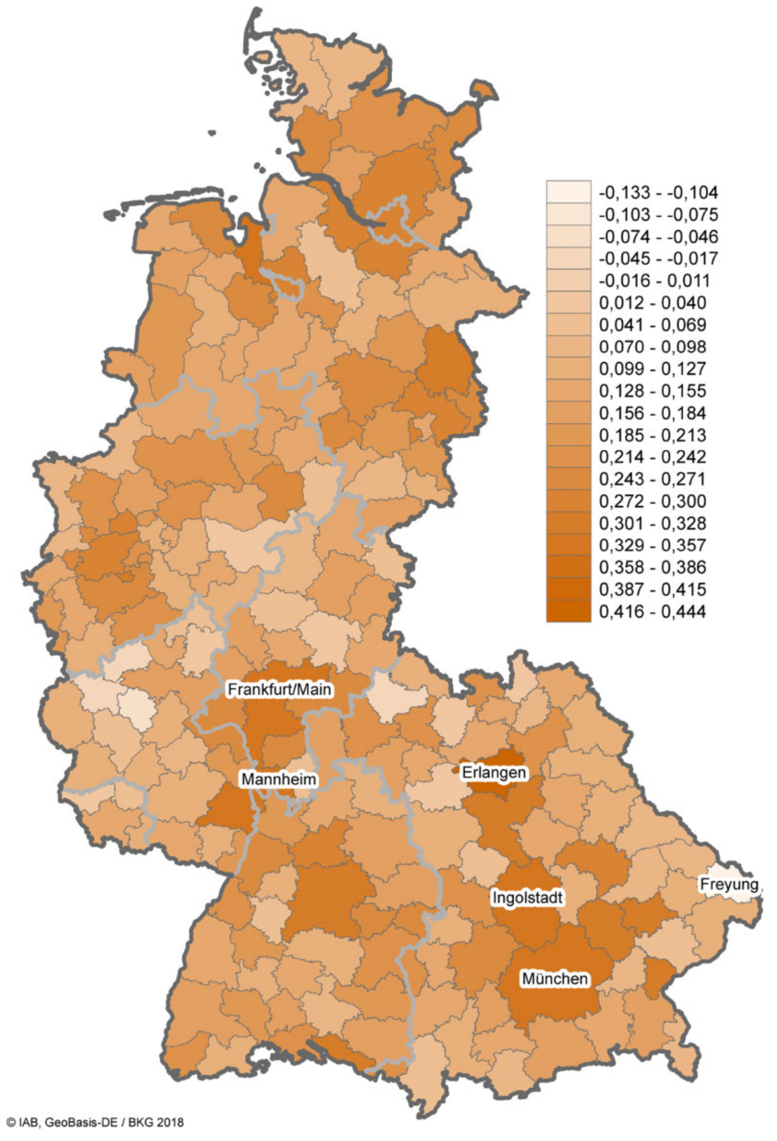
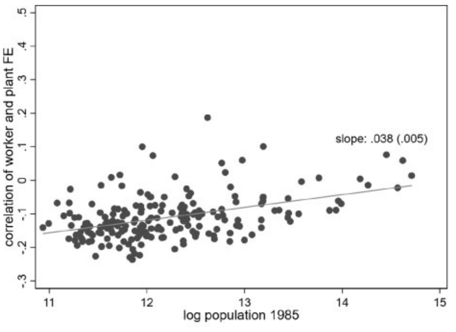


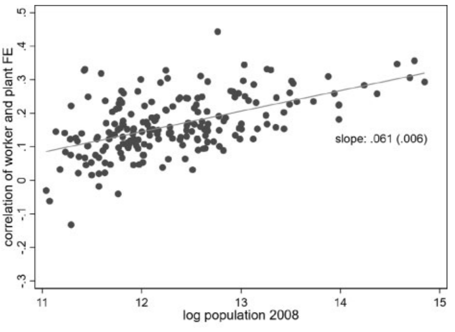
FIGURE 9. Degree of assortative matching by city, 2008–2014. The map visualizes the spatial distribution of the city level correlation coefficients of worker and plant effects. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies.

assortativeness with respect to the total number of workers in the local labor market is very similar to the elasticity with respect to the population.

Overall, Table 4 indicates that the degree of positive assortative matching between workers and plants is stronger in larger cities and that the relationship with city size has grown over time. When we re-estimate our models using residual plant effects, we find similar results (panels D and E), suggesting that the stronger assortative matching



(a) 1985–1991



(b) 2008–2014

FIGURE 10. City size and strength of assortative matching. This figure visualizes the bivariate correlation of assortative matching and log population across 204 cities. Assortative matching is defined as the city level correlation coefficient of worker and plant effects. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

in larger cities is not driven by local industry mix. Rather, our findings indicate that larger cities exhibit better worker–plant matching independently of the local industry mix.

To better understand where the association between assortative matching and city size is coming from, we consider the largest and smallest West German cities as examples. Panel (a) of Figure 11 plots the joint density of plant effects and worker effects measured in 2008–2014 in Hamburg, and panel (b) shows a similar graph for Daun. Both figures show assortativeness, but the degree of assortativeness is much stronger in the large cities. In Hamburg, the higher degree of assortative matching seems to be driven mostly by the matching of the highest two deciles of workers with the highest paying plants. But there are also comparatively high densities of the other combinations among the main diagonal, whereas the probabilities of matching top plants and bottom workers and vice versa are considerably smaller. In Daun, the picture is somewhat different. Here, we find a lot of matches of top plants with workers from the bottom of the distribution of fixed effects, and the degree of assortativeness in worker–plant matching appears much less tight.

5.2. Occupation-Specific Local Labor Markets

So far, we have measured the size of cities by the overall population. However, it is possible that localized increasing returns in the matching technology may be specific to narrower definition of local labor markets (Moretti 2011, 2012). For example, the assortativeness of matching bioengineering workers with bioengineering firms may be better in local environments where many such bioengineers are present, and only weakly better in larger cities with more inhabitants or workers in general.

TABLE 4. Correlation between city size and strength of assortative matching.

	(1) 1985–1991	(2) 1990–1996	(3) 1996–2002	(4) 2002–2008	(5) 2008–2014
Panel A: OLS. Dependent variable: correlation of worker and plant FE					
Log population	0.0380***	0.0380***	0.0579***	0.0610***	0.0613***
	(0.005)	(0.005)	(0.005)	(0.007)	(0.006)
R^2	0.188	0.190	0.290	0.265	0.286
Panel B: 2SLS. Dependent variable: correlation of worker and plant FE					
Log population	0.0415***	0.0438***	0.0664***	0.0701***	0.0708***
	(0.006)	(0.005)	(0.006)	(0.007)	(0.007)
R^2	0.186	0.184	0.282	0.259	0.275
First-stage F	294.140	291.613	288.923	280.681	276.485
Panel C: OLS. Dependent variable: correlation of worker and plant FE					
Log employment	0.0390***	0.0361***	0.0568***	0.0624***	0.0612***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
R^2	0.260	0.218	0.341	0.346	0.350
Panel D: OLS. Dependent variable: correlation of worker and residual plant FE					
Log population	0.0412***	0.0384***	0.0601***	0.0590***	0.0618***
	(0.005)	(0.004)	(0.005)	(0.006)	(0.006)
R^2	0.248	0.233	0.317	0.263	0.280
Panel E: 2SLS. Dependent variable: correlation of worker and residual plant FE					
Log population	0.0422***	0.0425***	0.0664***	0.0650***	0.0695***
	(0.005)	(0.005)	(0.006)	(0.007)	(0.008)
R^2	0.248	0.230	0.309	0.261	0.271
First-stage F	294.140	291.613	288.923	280.681	276.485

Notes: City-level regressions. $N = 204$ (Panels A, C, D) and $N = 200$ (Panels B, E). The dependent variables stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. The residual plant effects in Panels D and E are the residuals from a regression of the plant effects on two-digit industry effects. Instrument variable for log population is log population in 1952. (Since Saarland has not been part of Germany in 1952, the respective regions are omitted in Panels B and E.) The first-stage coefficient is 0.7726 (s.e. = 0.045) in the first period and 0.7574 (s.e. = 0.046) in the last period. Robust standard errors in parentheses. Levels of significance: *** 1%.

To explore this, we now turn to a more fine-grained definition of local labor markets using cells for particular occupations. We distinguish 89 different two-digit occupations and compute the correlation of worker and plant effects separately for each occupation within each city.[28]

A doubling of city-occupation cell size results in a 6.4% increase in assortative matching, as shown in Figure 12. Table 5 reports more detailed results. In column (1), we regress the degree of occupation-specific assortative matching on overall city employment only. In column (2), we use the total number of workers in the relevant city-occupation cell.

28. Since in small cities, some occupational cells are empty or very small, we restrict our analysis to cells with at least 50 workers and 5 plants.

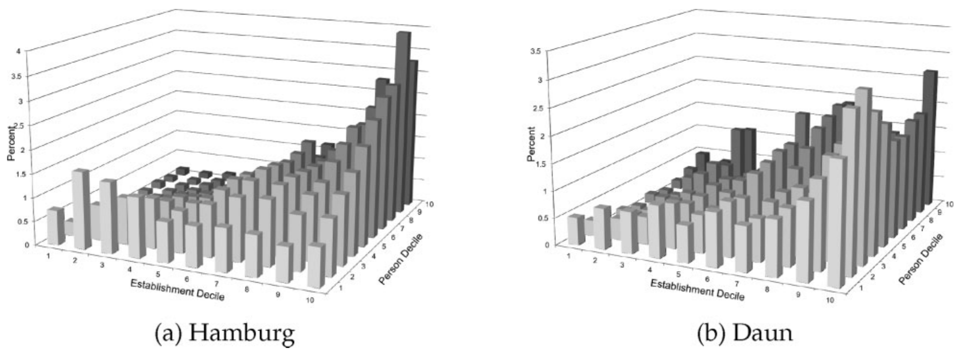


FIGURE 11. Joint density of plant effects and worker effects in the biggest and smallest cities, 2008–2014. This figure visualizes the joint distribution of worker and plant effects in West Germany's largest and smallest city, respectively. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. Each bar represents the share of workers with an individual fixed effect at the decile depicted at the right axis who are employed at a plant with a plant effect at the decile depicted at the horizontal axis.

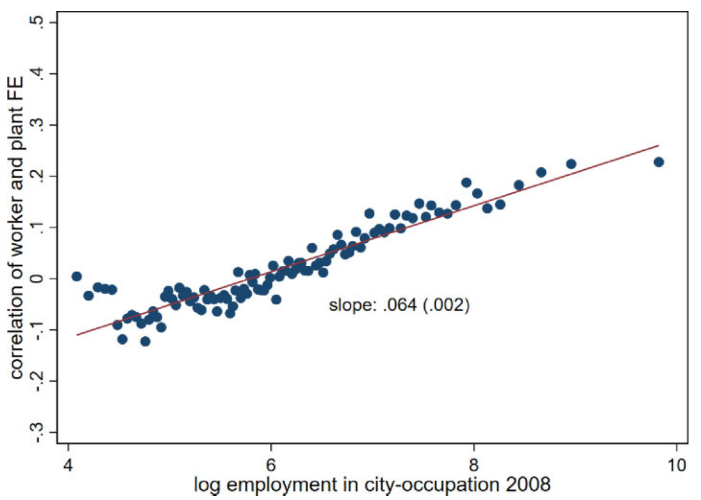


FIGURE 12. Occupation-city size and strength of assortative matching, 1985–2014. This figure visualizes the bivariate correlation of assortative matching and log employment across 10,293 occupation-city pairs. Assortative matching is defined as the occupation-city level correlation coefficient of worker and plant effects. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. The figure is a binned scatter plot: All 10,293 occupation-city pairs are grouped into 100 percentiles according to their employment. The dots represent the average values of the correlation coefficient of worker and plant effects (y-axis) plotted against the average log employment (x-axis) in each percentile category. The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

TABLE 5. Correlation between city-occupation cell size and strength of assortative matching.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: correlation of worker and plant FE							
1985–1991							
Employment in city	0.0229***		−0.0000				
	(0.002)		(0.002)				
Employment in city-occupation		0.0359***	0.0359***	0.0356***	0.0264***	0.0277***	0.0381***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
× % college degree							−0.0001
							(0.000)
% college degree							0.0010**
							(0.000)
City FE	–	–	–	yes	–	yes	yes
Occupation FE	–	–	–	–	yes	yes	–
N	10,321	10,321	10,321	10,321	10,321	10,320	10,321
R^2	0.011	0.044	0.044	0.076	0.395	0.427	0.077
2008–2014							
Employment in city	0.0313***		−0.0102***				
	(0.002)		(0.003)				
Employment in city-occupation		0.0645***	0.0681***	0.0682***	0.0388***	0.0305***	0.0794***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
× % college degree							−0.0006***
							(0.000)
% college degree							0.0041***
							(0.000)
City FE	–	–	–	yes	–	yes	yes
Occupation FE	–	–	–	–	yes	yes	–
N	10,293	10,293	10,293	10,293	10,293	10,293	10,293
R^2	0.016	0.116	0.117	0.160	0.410	0.452	0.173

Notes: Regressions at the level of occupation-specific local labor markets, defined as all combinations of 204 cities and 89 two-digit occupations with at least 5 plants and 50 workers. The dependent variables stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies.

Employment is measured as log number of employees. Robust standard errors in parentheses. Levels of significance: ** 5% and *** 1%.

Notably, we find a stronger elasticity than we do at the city level. Including both measures at the same time in column (3), we find that only the local-occupation size measure remains positive and significant. In column (4), we include city fixed effects and thus focus on the variation in assortative matching across occupations within cities. We obtain coefficients similar to column (2). In column (5), we include occupation fixed effects and in column (6), we include both city and occupation fixed effects simultaneously. This cuts the correlation between the size of the specific local labor market and matching in half. Apparently, this correlation holds more strongly across than within occupations. Finally, in column (7), we include the percentage of college-educated workers in each cell as well as an interaction term with the size of the cell. We find that while more skill-intensive city-occupations have a higher correlation of worker and plant effects per se, the elasticity with respect to the size of the local labor market actually declines with skill intensity. Taken together, the results show that the benefits of size for assortativeness seem to be strongly confined to the specific local labor markets for particular occupations.

Comparing the upper and the lower panel of Table 5, we find higher elasticities for 2008–2014 than for 1985–1991.[29] This mirrors the pattern found before that the elasticity of assortative matching with respect to labor market size has become stronger over time.[30]

5.3. Robustness and Limited Mobility Bias

In Online Appendix B.2, we present several robustness checks and alternative specifications. In particular, we address two concerns regarding the interpretation and estimation of the AKM model. First, the theoretical literature on the interpretation of assortative matching has pointed out that if wages are non-monotonous with respect to plant-productivity, then firm effects from the AKM model do not capture productivity and, hence, the correlation of worker and firm effects does not reflect the strength assortative matching (Lentz and Mortensen 2010; Eeckhout and Kircher 2011; Bartolucci, Devicienti, and Monzón 2018). Lopes de Melo (2018) proposes to measure the intensity of assortative matching by the correlation of worker effects with the average worker effects of their respective co-workers. We find that this measure of sorting is also significantly related to city size with an elasticity of 3.1% in the first and 4.7% in the last period (the baseline elasticities were 3.8% and 6.1%, respectively).

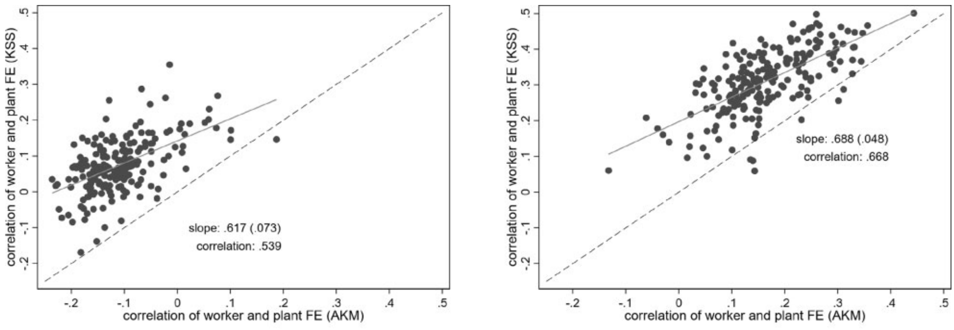
Second, we obtain our worker and plant effects from one unique AKM estimation over the entire period 1985–2014 instead of splitting it into five shorter sub-periods. This exercise yields elasticities of 2.8% in the first and 5.5% in the last period, close to the baseline values. Since this estimation prevents worker and plant effects to change over time, those results are exclusively driven by the allocation of workers and plants, which appears to improve over time.

Further robustness checks include cities from East Germany, address regional price differences, the definition of geographical units, and measure density instead of population size. In Online Appendix B.3, we also account for the possibility of a spurious correlation caused by an omitted variable by including control variables or federal state fixed-effects and by stacking our periods and including city fixed-effects. Our key conclusions remain fairly robust across all those specifications.

Limited Mobility Bias. One particularly important issue to consider is the presence of limited mobility bias, which may generate a downward bias in the estimated covariance of worker and plant effects. We start by running the leave-out estimation by Kline, Saggio, and Solvsten (2020, henceforth KSS) to obtain unbiased estimates of the

29. Note that the number of occupation-city cells decreased by 0.3% from the first to the last period, which induces a slight imprecission to this comparison.

30. Since occupation-specific local labor markets are considerably smaller than whole local labor markets, those results might be driven by limited mobility bias. In the Online Appendix, we therefore replicate Table 5 using plant effects from plants grouped into $k = 20$ groups. This does not change the overall result.



(a) 1985–1991

(b) 2008–2014

FIGURE 13. Assortative matching from AKM versus KSS. This figure visualizes the correlation of the city level correlation coefficients of worker and plant effects computed by either individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies or the unbiased estimates of the variance components of the corresponding leave-out estimations by Kline, Saggio, and Solvsten (2020). The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

variance components of the AKM model. We run this procedure city-by-city and obtain one correlation of worker and plant effects for each city and period. Figure 13 plots those unbiased correlation coefficients against our baseline estimates and reveals a positive but moderate correlation. The KSS correlations are markedly higher than the AKM counterparts, which implies that AKM was indeed affected by limited mobility bias.

Next, we regress the correlations of worker and firm effects on the log city population and compare the elasticities between our baseline results from the AKM model from the KSS procedure. The results are reported in Figure 14 and reveal that, while the intercepts are larger for the KSS results (in line with Figure 13), the slope coefficients are virtually identical. This is remarkable since the correlations stem from different estimations, carried out either nationwide (AKM) or city-by-city (KSS).[31] This result strongly suggests that limited mobility bias reduces the correlations of worker and firm effects in all cities, but not systematically more strongly in smaller cities.

In Online Appendix B.2.3, we discuss two additional ways to address the issue of limited mobility bias. First, following the idea behind Bonhomme, Lamadon, and Manresa (2019, 2021), we group plants into $k = 10, 15, 20$ groups of plants with similar wage structures and run the AKM procedure with fixed effects for clusters, which are connected by much more worker mobility than individual plants. A second and simpler approach is to re-estimate our model after dropping the largest or smallest cities, or after dropping the cities with the smallest or highest turnover. Grouping

31. Repeating this for 106 more aggregated local labor markets yields slopes that are similar between AKM and KSS but not identical any more.

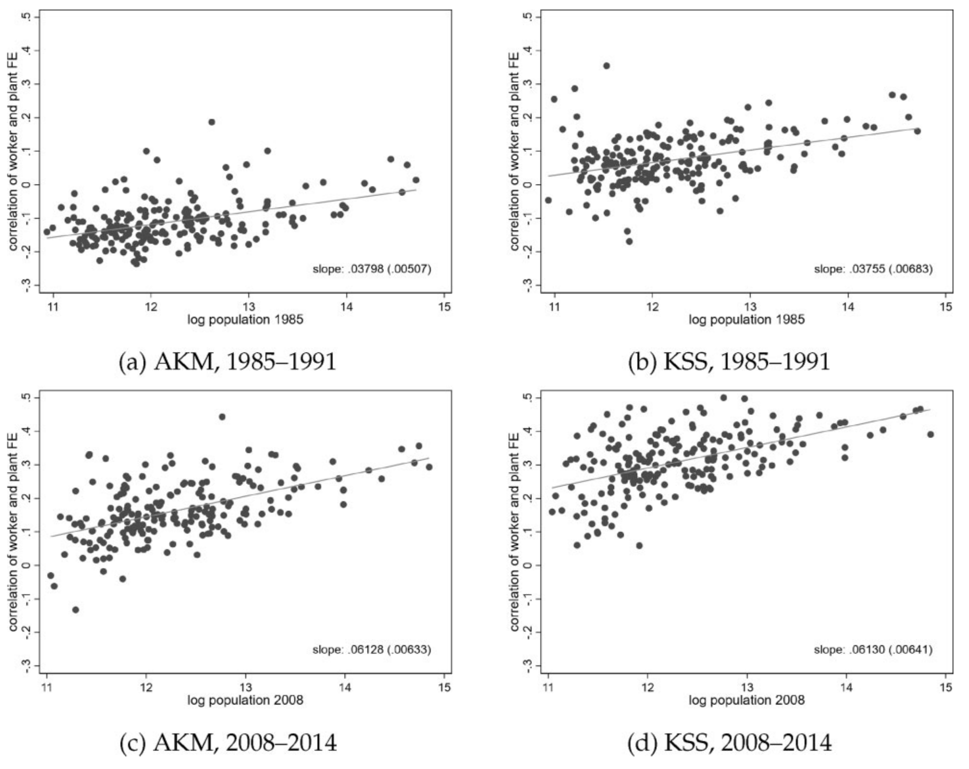


FIGURE 14. City size and strength of assortative matching—AKM versus KSS. This figure visualizes the bivariate correlation of assortative matching and log population across 204 cities. Assortative matching is defined as the city level correlation coefficient of worker and plant effects. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies (panels (a) and (c)) or the unbiased estimates of the variance components of the corresponding leave-out estimations by Kline, Saggio, and Solvsten (2020) (panels (b) and (d)). The solid line represents the regression coefficient of a bivariate regression. The numbers in parentheses are robust standard errors.

plants yields smaller elasticities in the last period compared to our baseline results, while dropping the smallest cities that comprise half of the total population raises the elasticity in the first period. This suggests that the elasticity remained roughly constant over time, but by-and-large, our conclusions remain unchanged.

5.4. Discussion

Overall, the evidence in this section indicates that assortative matching is significantly stronger in larger cities. The elasticity of assortative matching with respect to city size implies large and growing differences in the degree of assortative matching between large and small cities. We find an even stronger association when we define a local labor market as a city-occupation pair, rather than the city as a whole.

We conclude that larger labor markets—whether cities or city-occupation pairs—allow for a more efficient matching between workers and firms. Our findings are consistent with previous theoretical papers, where labor market size is an important agglomeration advantage (Helsley and Strange 1990; Acemoglu 1997; Rotemberg and Saloner 2000; Lazear 2009). They are also consistent with stronger incentives for assortative matching in larger cities, as we argued by our toy model in Section 4.

Could our evidence be explained by learning externalities as opposed to matching? Imagine, for example, that workers in bigger cities learn faster and become more productive over time (De La Roca and Puga 2017). In this case, our model may be mispecified, and it might generate high fixed-effects for workers in large cities due to faster wage growth. It is also possible that if the gains to assortative matching are higher for good firms in big cities workers in big cities churn between firms more often, as firms re-optimize, so that the wage growth is faster in large cities as workers ascend the "firm quality" ladder.[32]

To investigate this alternative interpretation, we divide our sample based on job order. In particular, after running AKM, we group together all workers who are in their first job, those who are in their second job, those in their third job, and so on. We then estimate the correlation of worker and plant effects for each of those groups. We find that this correlation increases monotonously as workers climb up the job ladder. However, we do not find evidence that the elasticity of assortative matching with respect to city size increases as workers climb up the job ladder (see Online Appendix B.3). In other words, workers do not seem to learn faster in cities by churning through different jobs at increasingly productive firms. We conclude that two explanations for the urban wage premium, matching and learning, seem to be complementary.

Finally, it is in principle plausible that denser and larger cities reduce search costs, facilitating PAM in those cities. In Online Appendix B.3, we show that, in the cross-section and conditional on industry-fixed effects, city size is negatively correlated with the probability that a plant faces difficulties in filling a vacancy. This also offers an explanation why the elasticity of assortative matching with respect to city size has increased over time.

6. Consequences of Assortative Matching

Assortative matching matters for at least two reasons. First, it can potentially magnify wage differences across communities. Wages in cities with more good workers and good plants are higher than wages in cities with fewer good workers and good plants. This is true not only because of the difference in quality but also because of the

32. A more sophisticated variant of this hypothesis is that these channels reinforce each other. Big cities provide learning opportunities, talented workers move to big cities to learn and workers learn by churning through different jobs. In this version, learning occurs through better matching—so that matching and learning mechanisms are hard to separate.

match component. If the strength of assortative matching increases with city size, then geographical wage differences are magnified even further.

Second, from a macroeconomic point of view, assortative matching can potentially increase aggregate output and earnings. Intuitively, for a given distribution of worker and plant quality, a country can produce more output if good workers and good plants are tightly matched than under random matching of workers and firms.

In this final section, we use a simple approach to quantify the importance of assortative matching for geographical wage differences and aggregate earnings in Germany. Based on equation (1), the mean log wage in city c can be written as

$$E_c[\ln \mathrm{wage}_{it}] = E_c[\mu_i + \Psi_{\mathrm{J}(i,t)} + X'_{it}\gamma].$$

If wages at the city level are log-normally distributed so that μ_i, $\Psi_{\mathrm{J}(i,t)}$, and X'_{it} are jointly normally distributed in each city, we can write the mean wage in city c as

$$\begin{aligned} E_c[w] = {} & \exp\left(\bar{\mu}_c + \bar{\Psi}_c + \bar{X}_c\gamma\right) \\ & \times \exp\Bigg[\left(\frac{1}{2}\left(\sigma^2_{\mu(c)} + \sigma^2_{\Psi(c)} + \sigma^2_{X\gamma(c)}\right)\right) \\ & + \Bigg(\underbrace{\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})}_{\text{Assortative Matching}} + \mathrm{cov}_c(\mu_i, X'_{it}\gamma) + \mathrm{cov}_c(\Psi_{\mathrm{J}(i,t)}, X'_{it}\gamma)\Bigg)\Bigg], \end{aligned} \tag{4}$$

where $\bar{\mu}_c$, Ψ_c, and $\bar{X}_c$ are means, and $\sigma^2_{\mu(c)}$, $\sigma^2_{\Psi(c)}$, and $\sigma^2_{X\gamma(c)}$ are the corresponding variances at the city level c.

Equation (4) shows that average wages are higher in cities with larger average plant and worker effects. The average wage is also higher in cities with a higher variance of the fixed effects and worker observables due to the skewness of the log-normal distribution. Importantly for us, the expression indicates that stronger assortative matching—namely, a higher covariance term $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})$—increases the mean wage level in a city. This makes intuitive sense. If worker and plant quality are gross complements, cities where workers are more likely to be matched to plants of higher quality enjoy higher average wages.

We use equation (4) to estimate what would happen to wage differences across cities if the distribution of assortative matching—measured by $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})$—differed from the observed amount.[33]

33. We note that the log-normality assumption is convenient, but not crucial. This assumption creates a simple and transparent covariance structure given by the sum of the three respective covariances. To probe the robustness of our findings to this assumption, we re-estimated all the counterfactuals without the log-normality assumption under the assumption of no correlation between the X's and the worker/plant effects. This assumption is consistent with the empirical finding that the fixed effects show a near zero correlation with the $X_c\gamma$ parts (see also Table 3). If for simplicity of exposition, we assume that $\bar{X}'_c\gamma = 0$, we can write average wages at the city level as

$$E_c[w] = E_c[\exp(\mu_i + \Psi_{J(i,t)})] = E_c[\mu_i^* \Psi^*_{J(i,t)}] = \mathrm{Cov}_c[\mu_i^*, \Psi^*_{J(i,t)}] + E_c[\mu_i^*]E_c[\Psi^*_{J(i,t)}],$$

TABLE 6. Effects on geographical wage differences.

Differences of ln wage	(1) Across cities 90–10	(2) Across cities 75–25	(3) Across cities s.d.	(4) Within cities s.d.
(1) Observed across city dispersion of average wages	0.28	0.15	0.11	0.42
	% difference to observed dispersion:			
(2) Random sorting of workers within cities (corr = 0)	−5.24	−5.60	−4.75	−3.06
(3) Sorting of workers within cities as in median city	−4.62	−3.83	−3.98	−0.28
(4) Zero elasticity of city size and sorting	−2.21	−1.21	−1.75	−0.08
(5) 1985 sorting of workers within cities	−5.26	−4.73	−2.44	−5.10
(6) No spatial sorting of workers across cities	−55.27	−54.68	−54.56	–
(7) No mobility of workers across cities	−6.93	0.55	−3.89	–

Notes: Counterfactual exercises for inequality of average log wages across cities, measured by the 90−10 quantile difference (column (1)), the 75−25 quartile spread (column (2)), and the standard deviation (column (3)) and for the standard deviation of wages within cities (column (4)).
Row 1 reports the inequality measures of city level average log wages, defined as the average fitted values from an individual level AKM estimation of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies for the period 2008–2014.
Counterfactual exercises: We change the means, variances, and covariances of the estimated worker and plant effects, as described in the main text. We assume that the distribution of X's remains unaffected. Also changing the distribution of the X's in accordance with the counterfactuals gives almost identical results.

6.1. Geographical Differences in Wages

Our findings are summarized in Table 6. For reference, the first row shows three measures of spatial wage variation in our most recent period by inserting the observed moments of the within-city distributions of μ_i, $\Psi_{\mathrm{J}(i,t)}$, and X_i into equation (4). To examine the effect of assortative matching within cities (rows 1–4), we then replace the observed values of $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})$ with counterfactual values. Alternatively, we replace the observed values of $\bar{\mu}_c$ and $\sigma^2_{\mu(c)}$ to examine the effect of spatial sorting across cities (rows 5–6).

Homogenizing Within-City Matching. We start with a set of counterfactuals that equalize the degree of within-city matching. This describes how much of regional inequalities are driven by the advantage of some labor markets in generating positive assortative matching conditional on the spatial distribution of workers and plants. In row 2, we randomly assign workers to plants within each city but keep the location of workers and plants equal to that observed in the data. In practice, we set $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)}) = 0$. The 90−10 difference, the 75−25 difference, and the standard deviation decline by 5.2%, 5.6%, and 4.8%, respectively. In row 3, we equate $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})$ across cities by setting it to the median value (0.152, found in the city

where * to denote exponentiated variables; and $\mathrm{Cov}_c[\mu_i^*, \Psi^*_{J(i,t)}]$ captures within-city matching and $E_c[\mu_i^*]E_c[\Psi^*_{J(i,t)}]$ the cross-city component. Empirically, using this assumption we find counterfactuals that are very similar to the ones based on the log-normality. The results are available on request.

of Steinfurt). The results show that the 90−10 difference, the 75−25 difference, and the standard deviation decline by 4.6%, 3.8%, and 4.0%, respectively. This shows that the homogenizing matching across cities has similar effects regardless of the precise value $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})$ is set to.

Next, we ask what fraction of the urban wage premium can be explained by better assortative matching. To obtain this counterfactual, we exchange $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})$ with the residuals from a regression of this covariance on log city size. We re-center these residuals to have the same median as the original covariances, which means that the result can be compared to the previous counterfactual. Row 4 of Table 6 reveals that between 1.2% and 2.2% of geographical wage differences can be attributed to the positive elasticity of sorting and city size. This is between one third and one half of the reduction in row 3, where sorting was restricted to be equal across cities. For comparison, 27.6% of geographical wage differences can be attributed to city size. This implies that between 4% and 8% of the unconditional urban wage premium stem from stronger assortative matching in bigger cities.

We found that the correlation between assortative matching and city population has increased by 75% between 1985–1991 and 2008–2014. This increase has contributed to the observed increase in spatial wage inequality. In row 5, we use the geographical distribution of workers and plants observed in 2008–2014, but assign to each labor market the degree of assortative matching it had in 1985–1991. This experiment should be interpreted as a lower bound, since geographical sorting might have increased because of these trends in matching. Results show that the 90−10 difference, the 75−25 difference, and the standard deviation declines by 5.3%, 4.7%, and 2.4%, respectively, which is quantitatively similar to the effect of making $\mathrm{cov}_c(\mu_i, \Psi_{\mathrm{J}(i,t)})$ homogenous across cities.

Geographical Worker Sorting. In the next scenario, we consider workers randomly sorting themselves into cities. Specifically, we equate $\bar{\mu}_c$ and $\sigma^2_{\mu(c)}$ across cities, setting it to the moments of the national distribution of worker effects, but keep the location of plants and city population as well as assortative matching as observed in the data. Row 6 shows that in this counterfactual, the 90−10 difference, the 75−25 difference, and the standard deviation declines by 55.3%, 54.7%, and 54.6%, respectively. This is a remarkably large effect, indicating that co-location plays a major role in explaining geographical differences in wages: Geographical sorting of workers to labor markets accounts for more than half of observed spatial inequalities.[34]

Worker Mobility. In row 7 of Table 6, we quantify how much spatial inequality is driven by geographical worker mobility after entering the labor market. We allocate workers to cities based on their origin labor market, but keep the locations of plants and the degree of assortative matching as observed in the

34. This finding is related to our earlier finding of a large contribution of the covariance term in Table 3.

data.[35] This scenario sheds light on the question how much of the total geographical worker sorting effect is due to labor market mobility over the career versus at labor market entry. The answer is not by much, as row 6 shows, at least when compared to the sorting estimate from row 5. Surprisingly, sorting into cities after labor market entry therefore only seems to play a relatively small role for spatial wage disparities. The standard deviation of log wages and the 90−10 spread move by about 4% and 7%, respectively.[36]

Within-City Inequality. In the next exercise, we quantify the contribution of assortative matching to within-city wage dispersion, which we measure by the standard deviation of the expected log wage

$$\sigma_c[\ln \text{wage}_{it}] = \Big[\sigma^2_{\mu(c)} + \sigma^2_{\Psi(c)} + \sigma^2_{X\gamma(c)} + \text{cov}_c(\mu_i, \Psi_{\text{J}(i,t)}) + \text{cov}_c(\mu_i, X'_{it}\gamma) + \text{cov}_c(\Psi_{\text{J}(i,t)}, X'_{it}\gamma)\Big]^{0.5}.$$

As before, we replace the observed values of $\text{cov}_c(\mu_i, \Psi_{\text{J}(i,t)})$ with the same counterfactuals for within labor market sorting. We average this measure across cities and report the percentage change in each counterfactual. The numbers are reported in column (4) of Table 6.

Turning off assortative matching by setting $\text{cov}_c(\mu_i, \Psi_{\text{J}(i,t)}) = 0$ reveals a decline in the (average) within-city inequality measure of 3.1%. Homogenizing matching across cities to its median value or switching off the elasticity of matching, and city size has a negligible impact on average within-city inequality, as inequality increases in some cities are cancelled out by reductions in other cities. This is in contrast to the results on between-city inequality in columns (1)–(3). Strikingly, row 5 shows that the observed increases in assortative matching since 1985 have caused average within-city inequality to grow by 5.1%, which is significantly more than in the no-sorting counterfactual from row 2. This echoes again the importance of increases in assortative matching observed in the 30-year period.

6.2. *Aggregate Earnings*

In Table 7, we quantify how much differential assortative matching across cities contributes to aggregate earnings at the national level. The aggregate wage at the national level is the weighted average of city-level wages according to equation (4) across all cities, with the weights reflecting the number of workers in each city N_c: $\sum_c N_c E_c[w]$. The first row reports the observed average wage in West Germany in 2014. In column (1), we report the counterfactual daily wages and in column (2) their

35. The origin labor market each worker is assigned to is the worker's location when he first appears in the social security data, that is, the location of his first job. For the majority of individuals, this will be the place of their apprenticeship.

36. The 75−25 spread is almost unchanged and implies slightly lower inequality. This is because we observe moves mostly from very low ranked labor markets to very high ranked ones, which drive the larger effect on the 90−10 spread.

TABLE 7. Effects on aggregate earnings.

	(1) Average daily wage	(2) %-diff.	(3) Δ billions
(1) Observed	117.30		
(2) Random sorting of workers within cities (corr = 0)	115.25	−1.75	−25.97
(3) Perfect sorting of workers within cities (corr = 1)	123.97	5.68	84.37
(4) Reverse sorting of workers within cities (corr = −1)	107.16	−8.65	−128.39
(5) Sorting of workers within cities as in median city	116.54	−0.65	−9.70
(6) Zero elasticity of city size and sorting	116.70	−0.51	−7.62
(7) 1985 sorting of workers within cities	114.83	−2.11	−31.32
(8) No spatial sorting of workers across cities	116.78	−0.44	−6.59

Notes: Counterfactual exercises for national average wage levels. Row 1 reports the national average log wage, defined as the average fitted value from an individual level AKM estimation of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies for the period 2008–2014.
Counterfactual exercises: We change the means, variances and covariances of the estimated person- and plant fixed-effects, as described in the main text. We assume that the distribution of X's remains unaffected. Also changing the distribution of the X's in accordance with the counterfactuals gives almost identical results.
The last column reports the implied absolute change in the total compensation of employees according to the national accounts in 2014 in billion euros.

percentage difference to the observed wage. In column (3), we report the implied absolute change in the total compensation of employees according to the national accounts in 2014 in billion euros.

Homogenizing Within-City Matching. Switching off assortative matching in row 2 implies an aggregate earnings loss of around 1.8%. Row 3 shows the consequences of perfect assortative matching, which would increase aggregate earnings by 5.7%. Conversely, a perfect negative correlation would reduce aggregate earnings by 8.7% (row 4). While these numbers should not be taken literally, they can be seen as the upper and lower bounds of the effect of assortative matching. In the next three counterfactuals, we assume more realistic matching levels that, in the light of the previous result, yield quite significant effects.

In row 5, we see that if the level of matching was homogenous across cities and equal to the observed median value, aggregate earnings would decrease by −0.7%. Cities with a large positive assortative matching level lose more than cities with a small level gain. This is because cities with a high level of assortative matching also attract better workers and plants, and the gains from assortative matching are increasing in the means of the local worker and firm distributions. This result directly stems from the supermodularity in the model as illustrated in Section 3. Overall, this implies a loss of 9.7 billion euros in aggregate earnings.

Row 6 reports an even more conservative counterfactual where sorting is allowed to differ across cities, but not systematically with size. This still yields an income loss of 0.5% or 7.6 billion euros. Enforcing random matching within labor markets in row 2 leads to losses around 3.4 times as large.

Assigning each labor market its 1985–1991 matching level in row 7 would reduce aggregate labor earnings by 2.1%, a loss of 31.32 billion euros. Note that these estimated gains in earnings did not come from new investment in physical or human capital, or any geographical change on the part of workers or firms. This rise of average labor earnings came exclusively from improved matching of workers to plants within each city.

Geographical Worker Sorting. In contrast, the counterfactual scenario changing the allocation of workers across space in row 8 shows only very small effects on aggregate earnings. Hence, while the geographical sorting of workers of different ability is very important to explain spatial inequalities, it matters less for aggregates. This result follows from the fact that the distribution of plant effects is relatively homogeneous across labor markets.[37] In the simple counterfactuals, high-quality workers are able to find high-quality firms in all locations. Consequently, geographical worker sorting has a first-order effect on regional inequality but only a second-order effect on aggregate earnings.

7. Conclusion

Geographical disparities in labor market outcomes across cities are large and persistent, generating concerns about rising inequality and the appropriate policy response. However, crafting appropriate place-based policies crucially depends on understanding the ultimate economic sources of those disparities.

We show that assortative matching plays an important and growing role in explaining differences in labor market outcomes across German cities. In particular, we show that larger and denser cities display significantly more within-city assortative matching, both across and within specific occupations. This finding empirically validates the intuition behind many urban economics models of labor pooling.

One key implication is that assortative matching plays an important role in explaining geographical wage differences across German cities. Our results suggest that wages are higher in larger cities not only because larger cities have more high-quality workers but also because high-quality workers are significantly more likely to work in high-quality plants. We estimate that geographical inequality would decrease significantly if small cities had the same degree of assortative matching.

While stronger assortativeness in large cities increases geographical inequalities, it also has a positive effect on aggregate earnings in Germany and its growth over time. We estimate that the increase in within-city assortative matching observed between 1985 and 2014 increased aggregate labor earnings in Germany by roughly 31 billion euros.

The growing geographical disparities between dynamic large metro areas and struggling small cities have generated a wealth of place-based policy proposals both in Europe and the United States. They are typically aimed at transferring resources from the most productive areas to the least productive areas in order to offset some of the economic disparities. In Germany, more than 1 billion euros have been spent annually

37. We have seen this already in our findings from Section 4, where we found the elasticity of plant effects with respect to population was comparably small and from our decomposition in Table 3, where the variance of plant effects explained much less than worker and match effects.

since reunification to support economically struggling communities plagued by weak labor market outcomes.

Our findings indicate that policies designed to subsidize labor demand or labor supply in small struggling communities create a previously unrecognized efficiency loss due to a decline in match quality. Thus, redistributing economic resources and population from large productive cities to smaller cities might help reduce earning differences, but it also has aggregate costs, creating a classic case of an equity-efficiency trade-off (Gaubert, Kline, and Yagan 2021; Moretti 2021). An alternative strategy could be one that seeks to improve the matching efficiency within small cities through the creation of search platforms to facilitate job search specifically in small labor markets. Our findings suggest that if small cities could achieve the degree of assortative matching of large cities, spatial wage disparities would decline and aggregate earnings would increase.

Appendix

TABLE A.1. Descriptive statistics.

	(1) Mean	(2) s.d.	(3) Min	(4) p10	(5) p25	(6) p50	(7) p75	(8) p90	(9) Max
Panel A: worker level									
Period 1985–1991									
Log wage	4.532	0.349	2.346	4.167	4.326	4.495	4.718	5.004	6.312
Wage	98.921	37.569	10.440	64.515	75.646	89.579	111.968	148.952	551.174
Age	39.072	10.999	20	24	30	39	48	54	60
Missing educ	0.008	0.088	0	0	0	0	0	0	1
No degree	0.107	0.309	0	0	0	0	0	1	1
Vocc	0.737	0.440	0	0	0	1	1	1	1
College	0.048	0.214	0	0	0	0	0	0	1
University	0.100	0.300	0	0	0	0	0	0	1
Worker FE	0.000	0.295	−5.244	−0.317	−0.197	−0.048	0.160	0.432	2.899
Plant FE	0.000	0.153	−3.130	−0.164	−0.079	0.015	0.093	0.161	4.945
Corr(worker, plant)	−0.037								
Workers	10,405,925								
Plants	673,359								
Period 2008–2014									
Log wage	4.586	0.507	2.588	3.960	4.274	4.570	4.896	5.252	6.728
Wage	111.746	61.435	13.306	52.475	71.815	96.582	133.773	190.877	835.531
Age	40.945	10.155	20	26	33	42	49	55	60
Missing educ	0.006	0.079	0	0	0	0	0	0	1
No degree	0.044	0.204	0	0	0	0	0	0	1
Vocc	0.658	0.474	0	0	0	1	1	1	1
College	0.104	0.305	0	0	0	0	0	1	1
University	0.188	0.391	0	0	0	0	0	1	1
Worker FE	0.000	0.381	−5.087	−0.416	−0.265	−0.068	0.226	0.574	6.415
Plant FE	0.000	0.218	−5.499	−0.288	−0.110	0.025	0.159	0.240	5.900
Corr(worker, plant)	0.293								
Workers	12,105,712								
Plants	880,656								

TABLE A.1. Continued

	(1) Mean	(2) s.d.	(3) Min	(4) p10	(5) p25	(6) p50	(7) p75	(8) p90	(9) Max
Panel B: city level									
Period 1985–1991									
Log wage	4.465	0.076	4.293	4.367	4.417	4.458	4.515	4.561	4.731
Wage	91.833	7.676	76.132	82.497	86.779	90.702	96.535	101.498	119.819
Worker FE	0.000	0.039	−0.084	−0.049	−0.024	−0.003	0.021	0.046	0.138
Plant FE	0.000	0.046	−0.094	−0.058	−0.032	−0.003	0.031	0.058	0.203
Corr(worker FE, plant FE)	−0.111	0.066	−0.236	−0.180	−0.155	−0.123	−0.083	−0.023	0.187
Population	290,033	357,809	56,092	86,745	115,037	168,943	283,096	535,544	2,447,241
Area (km)	1,218	670	75	627	794	1,129	1,442	1,998	4,735
Density (1000/km)	0.292	0.394	0.054	0.083	0.111	0.167	0.260	0.618	2.627
Obs	204								
Period 2008–2014									
Log wage	4.575	0.098	4.394	4.462	4.501	4.562	4.633	4.712	4.847
Wage	107.534	12.300	87.639	94.056	98.490	105.298	113.847	124.516	147.379
Worker FE	0.000	0.061	−0.105	−0.065	−0.044	−0.010	0.030	0.081	0.229
Plant FE	0.000	0.047	−0.096	−0.057	−0.035	−0.005	0.030	0.064	0.167
Corr(worker FE, plant FE)	0.164	0.087	−0.133	0.064	0.108	0.152	0.228	0.289	0.444
Population	321281	397,449	62,201	94,828	130,873	189,795	317,148	627,464	2,803,463
Area (km)	1,218	670	75	627	794	1,129	1,442	1,998	4,735
Density (1000/km)	0.314	0.392	0.054	0.094	0.120	0.184	0.294	0.665	2.603
Obs	204								

Notes: Panel A: descriptive statistics at the level of individual workers. This data is used to carry out the AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. Worker and plant effects are re-centered around zero. Panel B: city level descriptive statistics for 204 local labor markets. The worker and plant fixed effects stem from AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies. Worker and plant effects are re-centered around zero.

TABLE A.2. Decomposition of the AKM plant effects.

	(1) 1985–1991	(2) 1990–1996	(3) 1996–2002	(4) 2002–2006	(5) 2008–2014
Number of plants	648,695	692,098	728,426	709,176	698,979
R^2	0.1456	0.17137	0.17088	0.14875	0.1405
% Contribution of variable groups to R^2					
Plant size	2.37	2.71	2.94	3.58	5.28
Local labor market	10.83	8.09	5.39	4.90	6.56
Industry	86.79	89.20	91.67	91.52	88.16

Notes: Decomposition of the R^2 of a regression of pre-estimated industry dummies. The plant effects stem from AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies.

TABLE A.3. Differentials in assortative matching by city.

Rank	City	(1) Assortative matching	(2) Population
1	Erlangen	0.444	349,366
2	Munich	0.356	2,531,068
3	Frankfurt/Main	0.347	2,124,514
4	Ingolstadt	0.344	456,651
5	Mannheim	0.332	574,807
102	Steinfurt	0.152	444,399
200	Bad Kissingen	−0.018	105,770
201	Daun	−0.030	62,201
202	Ahrweiler	−0.040	128,509
203	Cochem	−0.062	64,489
204	Freyung	−0.133	80,044
	s.d.	0.087	397,441
	75−25	0.120	186,275
	90−10	0.225	532,636
	99−01	0.387	2,351,077

Notes: The table reports differentials of assortative matching across 204 cities in the period 2008–2014. Assortative matching is defined as the city level correlation coefficient of worker and plant effects. These effects stem from individual level AKM estimations of the log wage on worker effects, plant effects, skill-specific cubic age profiles, and year dummies.

References

Abowd, John M., Francis Kramarz, and David N. Margolis (1999). "High Wage Workers and High Wage Firms." *Econometrica*, 67, 251–333.

Abowd, John M., Francis Kramarz, Paul Lengermann, and Sébastien Pérez-Duarte (2004). "Are Good Workers Employed by Good Firms? A Test of A Simple Assortative Matching Model for France and the United States." Centre de Recherche en Economie et en Statistique, Working Paper.

Abowd, John M., Francis Kramarz, Sébastien Pérez-Duarte, and Ian M. Schmutte (2018). "Sorting between and within Industries: A Testable Model of Assortative Matching." *Annals of Economics and Statistics*, 1–32.

Acemoglu, Daron (1997). "Training and Innovation in an Imperfect Labour Market." *Review of Economic Studies*, 64, 445–464.

Andersson, Fredrik, Simon Burgess, and Julia I. Lane (2007). "Cities, Matching and the Productivity Gains of Agglomeration." *Journal of Urban Economics*, 61, 112–128.

Andini, Monica, Guido De Blasio, Gilles Duranton, and William C. Strange (2013). "Marshallian Labour Market Pooling: Evidence from Italy." *Regional Science and Urban Economics*, 43, 1008–1022.

Andrews, Martyn J., Len Gill, Thorsten Schank, and Richard Upward (2008). "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171, 673–697.

Bartolucci, Cristian, Francesco Devicienti, and Ignacio Monzón (2018). "Identifying Sorting in Practice." *American Economic Journal: Applied Economics*, 10, 408–438.

Baum-Snow, Nathaniel and Ronni Pavan (2012). "Understanding the City Size Wage Gap." *Review of Economic Studies*, 79, 88–127.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud (2014). "Productive Cities: Sorting, Selection, and Agglomeration." *Journal of Political Economy*, 122, 507–553.

Bleakley, Hoyt and Jeffrey Lin (2012). "Thick-Market Effects and Churning in the Labor Market: Evidence from US Cities." *Journal of Urban Economics*, 72, 87–103.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2019). "A Distributional Framework for Matched Employer Employee Data." *Econometrica*, 87, 699–738.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa (2021). "Discretizing Unobserved Heterogeneity." Working Paper No. 2019-16, University of Chicago, Becker Friedman Institute for Economics.

Card, David, Ana Rute Cardoso, Jörg Heining, and Patrick Kline (2018). "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36, S13–S70.

Card, David, Jörg Heining, and Patrick Kline (2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics*, 128, 967–1015.

Chade, Hector, Jan Eeckhout, and Lones Smith (2016). "Sorting Through Search and Matching Models in Economics." *Journal of Economic Literature*, 55, 493–544.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon (2008). "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics*, 63, 723–742.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux (2012). "The Productivity Advantages of Large Cities: Distinguishing Agglomeration from firm Selection." *Econometrica*, 80, 2543–2594.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum (2014). "The Rise of the East and the Far East: German Labor Markets and Trade Integration." *Journal of the European Economic Association*, 12, 1643–1675.

Dauth, Wolfgang, Sebastian Findeisen, Jens Suedekum, and Nicole Woessner (2021). "The Adjustment of Labor Markets to Robots." *Journal of the European Economic Association*, 19, 3104–3153.

D'Costa, Sabine and Henry G. Overman (2014). "The Urban Wage Growth Premium: Sorting or Learning?" *Regional Science and Urban Economics*, 48, 168–179.

De La, Roca Jorge, and Diego Puga (2017). "Learning by Working in Big Cities." *Review of Economic Studies*, 84, 106–142.

Diamond, Peter A. (1982). "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy*, 90, 881–894.

Diamond, Rebecca (2016). "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000." *American Economic Review*, 106(6), 479–524.

Duranton, Gilles and Diego Puga (2014). "The Growth of Cities." In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf, vol. 2, chap. 5. Elsevier, pp. 781–853.

Dustmann, Christian, Bernd Fitzenberger, Uta Schönberg, and Alexandra Spitz-Oener (2014). "From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy." *Journal of Economic Perspectives*, 28(1), 167–88.

Eeckhout, Jan and Philipp Kircher (2011). "Identifying Sorting—in Theory." *Review of Economic Studies*, 78, 872–906.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny (2014). "Spatial Sorting." *Journal of Political Economy*, 122, 554–620.

Figueiredo, Octávio, Paulo Guimaraes, and Douglas Woodward (2014). "Firm–Worker Matching in Industrial Clusters." *Journal of Economic Geography*, 14, 1–19.

Findeisen, Sebastian, Sang Yoon (Tim) Lee, Tommaso Porzio, and Wolfgang Dauth (2021). "Transforming Institutions Labor Reallocation and Wage Growth in a Reunified Germany." Working Paper, University of Konstanz.

Gaubert, Cecile (2018). "Firm Sorting and Agglomeration." *American Economic Review*, 108(11), 3117–3153.

Gaubert, Cecile, Patrick M. Kline, and Danny Yagan (2021). "Place-Based Redistribution." NBER Working Paper No. 28337.

Glaeser, Edward L. and Janet E. Kohlhase (2004). "Cities, Regions and the Decline of Transport Costs." In *Fifty Years of Regional Science. Advances in Spatial Science*, edited by Raymond J.G.M. Florax and David A. Plane. Springer, Berlin, pp. 197–228.

Glaeser, Edward L., Jed Kolko, and Albert Saiz (2001). "Consumer City." *Journal of Economic Geography*, 1, 27–50.

Glaeser, Edward L. and David C. Mare (2001). "Cities and Skills." *Journal of Labor Economics*, 19, 316–342.

Gould, Eric D. (2007). "Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium." *Review of Economic Studies*, 74, 477–506.

Helsley, Robert W. and William C. Strange (1990). "Matching and Agglomeration Economies in a System of Cities." *Regional Science and Urban Economics*, 20, 189–212.

Henderson, Vernon J. (2003). "Marshall's Scale Economies." *Journal of Urban Economics*, 53, 1–28.

Huettner, Frank and Marco Sunder (2012). "Axiomatic Arguments for Decomposing Goodness of Fit According to Shapley and Owen values." *Electronic Journal of Statistics*, 6, 1239–1250.

Jaeger, Simon and Joerg Heining (2020). "How Substitutable Are Workers? Evidence from Worker Deaths." Working Paper, MIT Press.

Keisuke, Kondo (2017). "Urban Wage Premium Revisited: Evidence from Japanese Matched Employer-employee Data." RIETI Working Paper No. 17-E-047.

Kline, Patrick, Raffaele Saggio, and Mikkel Solvsten (2020). "Leave-Out Estimation of Variance Components." *Econometrica*, 88, 1859–1898.

Krugman, Paul R. (1991). *Geography and Trade*. MIT Press.

Lazear, Edward P. (2009). "Firm-Specific Human Capital: A Skill-Weights Approach." *Journal of Political Economy*, 117, 914–940.

Lentz, Rasmus and Dale Mortensen (2010). "Labor Market Models of Worker and Firm Heterogeneity." *Annual Review of Economics*, 2, 577–602.

Lopes de Melo, Rafael (2018). "Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence." *Journal of Political Economy*, 126, 313–346.

Macis, Mario and Fabiano Schivardi (2016). "Exports and Wages: Rent Sharing, Workforce Composition, or Returns to skills?" *Journal of Labor Economics*, 34, 945–978.

Mion, Giordano and Paolo Naticchioni (2009). "The Spatial Sorting and Matching of Skills and Firms." *Canadian Journal of Economics/Revue canadienne d'économique*, 42, 28–55.

Moretti, Enrico (2004). "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review*, 94(3), 656–690.

Moretti, Enrico (2011). "Local Labor Markets." *Handbook of Labor Economics*, edited by O. Ashenfelter and D. Card, 1st ed., vol. 4, chap 14. Elsevier, pp. 1237–1313.

Moretti, Enrico (2012). *The New Geography of Jobs*. Houghton Mifflin Harcourt.

Moretti, Enrico (2021). "The Effect of High-Tech Clusters on the Productivity of Top Inventors." *American Economic Review*, 111(10), 3328–3375.

Oberschachtsiek, Dirk, Patrycja Scioch, Christian Seysen, and Joerg Heining (2009). "Integrated Employment Biographies Sample IEBS—Handbook For the IEBS in the 2008 Version." FDZ-Datenreport Working Paper No. 03/2009.

Orefice, Gianluca and Giovanni Peri (2020). "Immigration and Worker-Firm Matching." NBER Working Paper No. 26860.

Petrongolo, Barbara and Christopher Pissarides (2006). "Scale Effects in Markets with Search." *Economic Journal*, 116, 21–44.

Rice, Patricia, Anthony J. Venables, and Eleonora Patacchini (2006). "Spatial Determinants of Productivity: Analysis for the Regions of Great Britain." *Regional Science and Urban Economics*, 36, 727–752.

Rosenthal, Stuart S. and William C. Strange (2004). "Evidence on the Nature and Sources of Agglomeration Economies." *Handbook of Regional and Urban Economics*, vol. 4. Elsevier, pp. 2119–2171.

Rotemberg, Julio J. and Garth Saloner (2000). "Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade." *Regional Science and Urban Economics*, 30, 373–404.

Torres, Sonia, Pedro Portugal, John T. Addison, and Paulo Guimarães (2018). "The Sources of Wage Variation and the Direction of Assortative Matching: Evidence from a Three-Way High-Dimensional Fixed Effects Regression Model." *Labour Economics*, 54, 47–60.

Wheeler, Christopher H. (2008). "Local Market Scale and the Pattern of Job Changes Among Young Men." *Regional Science and Urban Economics*, 38, 101–118.
Yankow, Jeffrey J. (2006). "Why do Cities Pay More? An Empirical Examination of Some Competing Theories of the Urban Wage Premium." *Journal of Urban Economics*, 60, 139–161.

Supplementary data

Supplementary data are available at *JEEA* online.