Regional Science and Urban Economics 96 (2022) 103806

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

Assortative labor matching, city size, and the education level of workers☆

Stefan Leknes [a], Jørn Rattsø [b,*], Hildegunn E. Stokke [b]

[a] Research Department, Statistics Norway, Norway
[b] Department of Economics, Norwegian University of Science and Technology, Norway

ARTICLE INFO

JEL classification:
J24
J31
R23

Keywords:
Assortative matching
Education-specific plant fixed effects
City size
Limited mobility bias
Tasks
Register data

ABSTRACT

We investigate the heterogeneity of assortative labor matching with respect to geography, skills, and tasks. Our contribution is to separate plant quality by education level and occupation tasks using the AKM-model. We introduce a geology-related instrument to analyze the city effect and address limited mobility bias. Using rich administrative worker-plant dataset for Norway, we show that matching of the college educated have a strong city effect. The IV estimates indicate that a doubling of city size increases the correlation between worker and plant quality by 9 percentage points. A wage decomposition shows that matching accounts for 22% of the urban wage premium adjusted for sorting. In terms of occupations, better matching in cities is observed only for non-routine abstract tasks.

1. Introduction

The urban wage premium is found to be increasing with the level of education of workers. Bacolod et al. (2009) offer an early overview of the allocation of skills across cities and the impact of agglomeration on wages dependent on skill. Recent studies improve the identification strategies. Carlsen et al. (2016) expand the analysis of De la Roca and Puga (2017) of dynamic urban wage gaps by allowing for heterogeneity with respect to education. They show that college-educated workers are positively selected into cities and benefit more than the low educated from working in cities even when sorting is accounted for. Baum-Snow and Pavan (2012) develop and estimate a structural model and find advantages for college-educated workers living in large cities. The result that urban productivity is higher for workers with higher education is convincing, but there is scarce evidence of the mechanisms. We pursue the importance of the education level of workers and occupation tasks for matching of workers and firms in local labor markets – a potential explanation of the productivity effect of urban scale.

The role of assortative matching for regional wage differences is recently analyzed by Dauth et al. (2022) based on data for private sector workers in Germany. They find that assortative matching matters – wages are higher in large cities because they attract high-quality workers, but also because high-quality workers are likely to be better matched to high-quality firms. The method applied was innovated by Abowd et al. (1999), called the AKM-model. It estimates wage components attributable to workers and firms and addresses unobservable characteristics. Card et al. (2013, 2018) have developed the empirical methodology and shown that assortative matching affects the wage distribution. Earlier studies related to urbanization inclu de Andersson et al. (2007), Melo and Graham (2014) and Figueiredo et al. (2014). The finding that assortative matching is stronger in large cities is in accordance with the standard understanding that large labor markets facilitate more productive matching between workers and firms. Our results show that the matching effect strongly depends on the allocation of workers across firms by skill.

Our main contribution is to extend the model formulation to analyze the role of skill heterogeneity for the geographic variation of assortative matching. The standard AKM-model assumes identical firm fixed effects for all workers in a firm independent of their time-invariant characteristics. However, when technologies of firms are affected by skill-replacing and skill-biased technological change, the quality of the firm is likely to be different for workers with different skill levels. Key references regarding the understanding of skill-bias are Acemoglu (2007) and

☆ We appreciate discussions at the 2020 EALE/SOLE/AASLE Conference, the 2020 Virtual Meeting of the Urban Economics Association, and the department seminar at Umeå University, and in particular comments from the editor and referees, Manudeep Bhuller, Gonzalo Dona, Magne Mogstad, Paolo Naticchioni, Kjell G. Salvanes and Bjarne Strøm. The project is funded by the Research Council of Norway (grant number 255509).

* Corresponding author.
E-mail addresses: sfl@ssb.no (S. Leknes), jorn.rattso@ntnu.no (J. Rattsø), hildegunn.e.stokke@ntnu.no (H.E. Stokke).

https://doi.org/10.1016/j.regsciurbeco.2022.103806
Received 11 October 2021; Received in revised form 31 May 2022; Accepted 1 June 2022
Available online 7 June 2022

Acemoglu and Autor (2011). We suggest a more flexible formulation of the AKM-model by estimating the fixed effects in separate regressions for three worker groups defined by their level of education (primary, high school and college). This allows the plant fixed effects to differ by workers' skill levels within the same plant. Giannone (2019) offers an analysis of the role of skill-bias in regional convergence separating between education groups. An alternative approach is to apply productivity data for the assessment of assortative matching, recently discussed by Mendes et al. (2010). In this literature, firm-specific productivity terms for each firm are quantified and related to the skill distribution of workers in the firm. However, this method does not allow for a measure of firm quality related to different skill levels of workers. The recent literature has emphasized routine-biased technological change with Autor et al. (2003) as an early contribution. We extend the analysis of heterogeneity to include a separation of workers based on occupation tasks.

A concern with the AKM-analysis is the dominance of negative correlations between worker and firm fixed effects. A background understanding of the limited mobility bias problem is offered by Andrews et al. (2012). We show how the correlations vary with changes of the sample selection and that negative correlations originate from a small part of the labor market – firms in the bottom of the employment size distribution. This finding sheds light on a limitation of the standard AKM-model, and enables us to test several amendments to the basic estimation strategy to mitigate limited mobility bias.

A second contribution of the paper is the handling of endogeneity and omitted variables affecting regional population size. Dauth et al. (2022) recognize identification challenges related to labor market size and use historical population data to instrument present populations in city regions. We introduce historical mines as an alternative instrument for population size, which yields point estimates that are about 50% higher than with OLS. The difference may reflect seemingly bad matches in cities due to urban amenities compensating wages, specialized industries in the periphery with good matches, and college-cities with bad matches. The strengths of the mining instrument are that the exclusion restriction is clearly stated, the historical mining activity has ceased, and that the impact on present city size can be understood through path dependence. Another relevant topic in the empirics of agglomeration economies is selection. Sorting of workers into populous places have been found to exacerbate regional wage disparities (Ahlfeldt and Pietrostefani, 2019). To gauge the importance of spatial selection processes for our results, contrafactual decomposition analysis is offered.

Using rich administrative data for Norway linking workers and firm plants, we investigate whether the overall positive relationship between city size and matching varies by the education level of workers. We find that the positive relationship between city size and assortative matching is driven mainly by college-educated workers. The better matching of the college educated in cities is consistent with studies showing that agglomeration effects increase with the level of education. The decomposition of spatial wage differences shows that for the college educated, assortative matching accounts for 22% of the urban wage premium adjusted for spatial worker sorting. Furthermore, better matching in cities for the college educated explains about 1/3 of the difference in the adjusted urban wage premium between non-college and college-educated workers. In an extension, we analyze the differences according to occupation tasks instead of skills. The results show that the importance of city size for matching is concentrated to workers conducting non-routine abstract tasks, corroborating the results for workers with higher education.

The heterogeneity with respect to gender and age distribution of workers is also studied. The relationship between city size and matching is confirmed for both male and female college-educated workers. We explore the age gradient with respect to assortative matching in regions of different sizes. College-educated workers are again better matched in more populous areas, and in particular among workers in the middle of the age distribution.

The econometric strategy separating plant fixed effects by education groups is discussed in section 2. Econometric challenges related to limited mobility bias and endogenous population size are discussed in section 3. The main results are reported in section 4, including the relationship between city size and matching using education-specific measures of firm quality, a discussion of the identification challenges, and a comparison of the results for education groups versus occupation task groups. Section 5 discusses the importance of the findings for urban wage premiums by decomposing spatial wage differences. Section 6 investigates the robustness of results, primarily for limited mobility bias. Further heterogeneity with respect to gender and age distribution is pursued in section 7. Concluding remarks are given in section 8.

2. Estimating assortative matching separating between education groups

To study if there is a skill-heterogeneous relationship between assortative matching and city size, we employ the method innovated by Abowd et al. (1999) estimating two-way worker and plant fixed effects.[1] In this framework, plant effects are identified by workers switching plants. We use a longitudinal dataset of employer-employee register data for the entire Norwegian labor force on hourly wages and worker characteristics during 2003–2014. We concentrate on full-time workers aged 25–65 employed in the private sector. The dataset includes about 8.1 million observations, covering 1.2 million workers and 118,000 plants. As explained below, in the main analysis we exclude plants with five or fewer workers to mitigate limited mobility bias resulting from plants with few job shifts.[2]

The AKM-strategy entails estimation of individual-level wage equations with both worker and plant fixed effects. We start out with the following aggregate specification:

$$\ln w_{it} = \mu_i + \varphi_{J(i,t)} + X_{it}\beta + \varepsilon_{it} \tag{1}$$

where w_{it} is the hourly wage income for worker i in year t. Worker fixed effects are represented by μ_i, and $\varphi_{J(i,t)}$ captures plant fixed effects of all employees of plant J. The vector of time-varying worker characteristics (X_{it}) includes job tenure (measured as the length of tenure for a worker in a specific year and a specific plant, given information on work contracts back to 1993), education-specific cubic age profiles (quadratic and cubic age terms interacted with dummies for three levels of education), and year dummies. β is a vector of parameters and ε_{it} is the error term.

The additive formulation of the AKM-model has been questioned, notably by Eeckhout and Kircher (2011), Lise and Robin (2017) and Lopes de Melo (2018). Macis and Schivardi (2016) and Bonhomme et al. (2019) provide empirical support for the additive structure of the AKM-model based on Italian and Swedish data, respectively. To test whether the AKM-model is a good approximation of the wage structure in Norway, we compare the fit with that of a match effect model in Appendix Table B.1. The table shows that the adjusted R-squared from the match effect model is only slightly higher relative to the AKM-model and root MSE is only slightly lower independent of sample restrictions. The improvement in fit is extremely modest in line with the findings of Card et al. (2013) and Macis and Schivardi (2016).

The correlation between the estimated worker and plant fixed effects gives a measure of assortative matching. To retrieve the strength of assortative matching within a city region, we calculate the correlation between fixed effects within each local labor market, r. The geographic units used in the analysis are based on information about commuting flows between municipalities. They are constructed by Statistics Norway,

[1] In the analysis, we estimate plant fixed effects, but use the terms *plant* and *firm* interchangeably throughout the paper.

[2] Further details on the dataset are given in Appendix A with descriptive statistics in Table A.1.

which divides Norway into 89 economic regions (broadly consistent with NUTS-4 regions in the EU standard). To study the relationship between assortative matching and city size, we regress the correlation between worker and plant effects within each labor market ($CorrFE_r$) on regional population size (Pop_r):

$$CorrFE_r = \alpha_0 + \alpha_1 \ln Pop_r + \varepsilon_r \tag{2}$$

There are large differences in population size across regions with an average of about 50,000 inhabitants and a standard deviation of 75,000. About 40% of the regions have population below 20,000. As an alternative to the relationship between continuous labor market size and strength of assortative matching in equation (2), we study how the matching varies between regions grouped based on population size. Consistent with earlier studies of Norwegian regions, we separate between 7 large cities with population above 150,000, 13 small cities with population in the range 65,000–150,000, and the remaining regions with less than 65,000 inhabitants.

We proceed with three subgroups of workers according to level of education: primary, high school, and college.[3] In the dataset applied in the main analysis, about 18% of the workers have no more than compulsory schooling, while workers with high school and college education account for 50% and 32% of the sample, respectively. As discussed in the introduction, we allow plant fixed effects to differ by level of education within the plant by estimating the fixed effects separately for the three education groups. This formulation takes into account possible complementarity between skill level of workers and plant-specific productivity. We compare the plant fixed effects distributions for the three education groups to study if the quality of plants varies by the skill level of workers (documented in Appendix Table B.2). The rank correlation between the plant FE distributions of primary and high school educated and between high school and college educated are both 0.26 (in about 24,000 plants). As expected, the rank correlation between plant fixed effects related to primary and college educated is lower, 0.15 (in about 19,000 plants). The plant fixed effects estimated separately for education groups are quite different. To extend the documentation that the plant fixed effects differ across education groups, we calculate the quintiles of the plant fixed effect distributions and identify the overlap of plants in each quintile. Comparing the top 20% of plants for each education group, we find 36–37% overlap between primary and high school educated and between high school and college educated, while the overlap for primary and college educated is 29%. The degree of overlap is in the order of 20–30% for comparisons in other quintiles. The low degree of overlap in the distributions of plant fixed effects suggests that the assumption of skill-independent plant effects is strong and that the plant fixed effects should be allowed to differ between education levels. Arguably, this will contribute to better measurement of the true strength of assortative matching.[4] In a further study of how the differences relate to firm-level and region-level characteristics, we show that both sector and city size matter.[5]

In the analysis, we demonstrate how allowing for separate plant fixed effects for workers with different education level matters for the relationship between matching and city size (in section 4) and for the decomposition of spatial wage differences (in section 5). The flexible specification with education-specific plant FEs is compared with the aggregate AKM-model where the fixed effects are estimated in one common regression under the assumption that plant fixed effects are identical for all workers in each plant.

We apply the same approach when extending the evidence to occupation tasks, separating between non-routine manual, routine and non-routine abstract occupations (as described by Acemoglu and Autor, 2011). Occupational data are available for the period 2003–2010 and include 345 occupations (see descriptions in Appendix A). The fixed effects are estimated separately for three subgroups of workers defined by the main task content of their occupation. The models estimated for studying differences across occupation task groups are identical to those used for skill groups.

3. Econometric challenges: limited mobility bias and endogenous population size

Limited mobility bias is a potential problem with the AKM-model. Fig. 1 illustrates the relationship between estimated plant fixed effects and small plant size when using all data available. The figure shows large variation in the plant fixed effect among small plants, while the variation decreases with plant size. The dispersion in plant quality is expected to be larger in the lower end of the size distribution, as plant size may signal success, small plants can be highly specialized and productive, and the strength of applicant screening may be lower for small plants. However, the marked pattern shown in the figure is likely to also reflect measurement problems. Our interpretation is that the estimated fixed effects are noisy in small plants with few job switches. This may in turn affect the calculated correlation between worker and plant fixed effects measuring the strength of assortative matching.

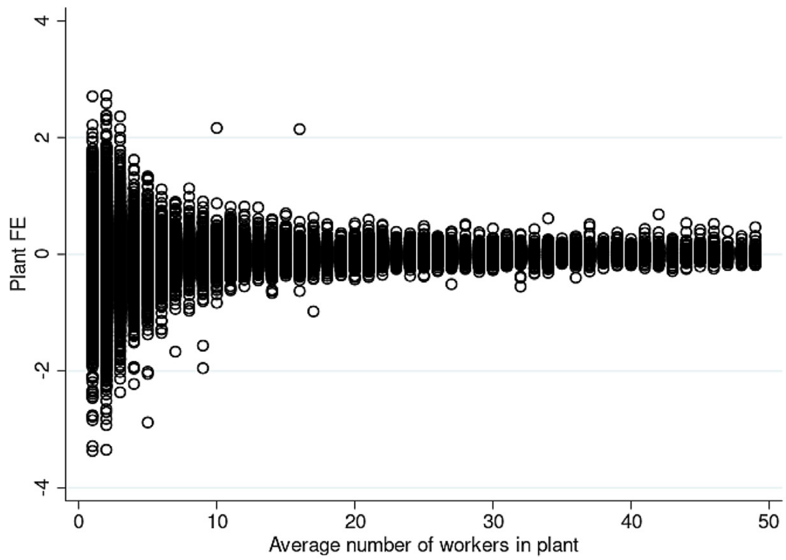
Using all observations in our dataset, the correlation between worker and plant qualities is negative and equals −0.038. Andrews et al. (2012) study the negative correlation between worker and plant fixed effects based on German register data. They show that the correlation between fixed effects is negative in limited samples where worker mobility between firms is low, while it turns positive for larger samples with more inter-firm mobility. The correlation between fixed effects is increasing and concave in the number of movers per establishment and approaches the true correlation asymptotically. In Fig. 2, we illustrate a similar pattern in the Norwegian data. As we exclude small plants with few job switches from the dataset, the overall correlation between worker and plant fixed effects increases gradually from −0.038 in the full dataset towards 0.187 when plants with ten or fewer workers are excluded. At each point in the figure, the fixed effects are re-estimated given the new sample. The relationship between minimum plant size and the education-specific fixed effect correlations is illustrated in Figures C.6 – C.8 in the separate online appendix. Due to less labor mobility within education groups, the education-specific correlations are lower than the aggregate correlation, especially for workers with only primary education. In the full dataset, the correlation ranges from −0.4 for the primary educated to −0.163 and −0.097 for workers with high school and college education, respectively. Importantly, excluding small plants with few job switches increases the education-specific correlations between worker and plant fixed effects for all education groups. For workers with high school or college education, the correlations turn positive as minimum plant size increases.

We explore the importance of limited mobility bias for the city size effect on matching by excluding small plants. We start out by excluding plants with five or fewer workers on average. In sensitivity analyses, we compare results with alternative cutoff levels for the exclusion of small

[3] Among workers with high school education, about 80% have a degree based on 3–4 years of schooling, while the remaining 20% have 1–2 years of high school education. Among the college educated, 87% either have a graduate degree (3–4 years of higher education, equivalent to a bachelor degree) or a postgraduate degree (more than 4 years of higher education), while workers with 1 or 2 years of higher education (but no degree) account for 3% and 10%, respectively.

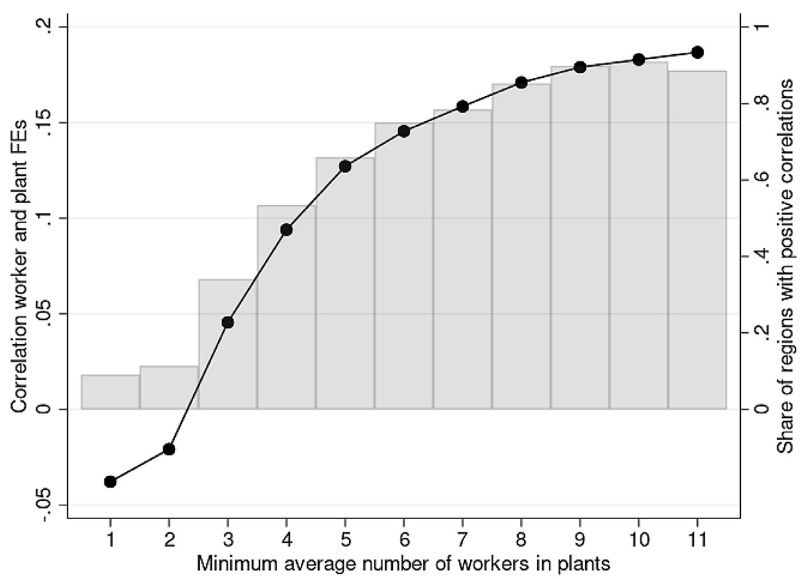
[4] The distributions of the education-specific plant fixed effects are shown in Figure C.1 in a separate online appendix available from the authors. The shape of the distributions is similar, but as shown above, the ranking of plants within the distributions differs substantially across education groups.

[5] Table C.1 in the online appendix studies whether the differences in a plant's position (percentile) in two respective education-specific plant fixed effect distributions vary systematically with plant-level and region-level characteristics like sector and population size. Defining the dependent variable as the absolute value of the difference in percentiles, the findings show that the difference in a plant's position in the fixed effect distributions decreases with the regional population size and is higher in services than in manufacturing.



Notes: The figure illustrates the relationship between estimated plant fixed effects and plant size for plants with up to 50 employees. Fixed effects are centered around zero. In the separate online appendix, we present the same figure for all plants, as well as a more detailed view of the relationship for plants with less than 20 employees (Figure C.2). The respective relationships for the three education levels are given in Figures C.3 – C.5 of the online appendix.

Fig. 1. Estimated plant fixed effects and plant size.



Notes: The overall correlation between worker and plant fixed effects is measured on the left axis and illustrated by the dotted line. The share of regions with positive correlation between fixed effects is measured on the right axis and illustrated by histograms.

Fig. 2. Correlation between FEs and minimum plant size.

plants and also show the results when all plants are included. The main finding is that the exclusion of small plants matters for the estimated correlation between worker and plant fixed effects (as discussed above), but the relationship between the strength of assortative matching and city size remains robust to variation in sample restrictions. This indicates that the degree of limited mobility bias does not vary systematically with regional population size.

In the basic dataset, we exclude plants with five or fewer workers on average. This includes 500,000–600,000 workers every year during the period 2003–2014, and a total of about 6.5 million worker-year observations and 32,760 plants. Workers can enter and leave the labor market during the twelve-year period, and in total about one million different workers are included.[6] It should be noted that the geographic dispersion of the excluded plants does not differ much from the plants remaining in the dataset. Small plants (no. of workers $\leq$5) are well represented in both cities and in more peripheral regions. The share of plants located in one of the seven large cities equals 41% and 43% for small and large plants,

[6] The geographic pattern of key variables is displayed in Figure B.1 (regional correlations between worker and plant fixed effects) and in Figures C.9 – C.11 in the online appendix (population size of regions, mean worker fixed effects, and mean plant fixed effects).

respectively. When it comes to industry structure, small plants are somewhat overrepresented in retail trade, hotels/restaurants and personal service activities. Workers employed in the small excluded plants (about 200,000 workers) consist of 24% primary educated, 54% with high school education and 22% college-educated workers. Compared to workers in large plants, the college educated are somewhat underrepresented.

The strength of assortative matching is measured as the correlation between estimated worker and plant fixed effects (aggregate and separately for each education group). With minimum plant size of six workers, the overall correlation equals 0.146 (as seen from Fig. 2). Descriptive statistics on the regional correlations are shown in Table 1, covering all regions and three groups of regions based on population size (large cities, small cities, periphery). The average correlation across regions equals 0.061. The education-specific correlations are negative on average across regions, but with important geographic heterogeneity. In the periphery and in small cities, the high school educated have best matching, while the college educated is best matched in large cities. For college-educated workers, the strength of assortative matching varies by population size (stronger in large cities than in periphery regions and small cities), while there is not much variation by population size for primary and high school educated workers.

Identification of the role of labor market size for the strength of assortative matching is challenging because workers and firms are drawn into urban areas motivated by superior labor matching opportunities. Although this may be an important productive advantage of cities, it bedevils interpretation of the population scale coefficient. An equally important challenge to identification is omitted variables – missing local variables that affect both matching opportunities and population size. To handle these possible sources of bias, we apply an instrument for current population size using information about the geographic distribution of historical mines introduced by Leknes (2015). The mining industry was one of the first industries in Norway. In the same way as the locations of mineral resources are random, it can be argued that the geographic distribution of the mining industry is random. Also, all the historical mines were exhausted before the period of our study.[7] The discoveries of valuable mining resources incited economic activity that spurred local population growth, which is traceable in the population patterns of today.

Norway has a long history of mining. The written source “Historia Norvegia” from 1170 mentions a silver mine in Oslo. Later the mining industry gained momentum, and in the 18th century, mining was one of the largest national industries. In the period of our study, however, the traditional mining industry is of marginal importance and all the historical mines are closed. Our argument is that historical mines predict present day city size, but do not impact city size other than through path dependence. The understanding resembles that of Bleakley and Lin (2012), which use natural features related to rivers, portage sites, and their importance in historical times for local economic activity and population growth. They argue that the natural advantages of these places can be considered forfeited today and find that they can still partially explain contemporaneous population patterns. The fact that the historical mining activity has ceased sets the instrument apart from another instrument in the urban economics literature, historical population size. Using historical population size as instrument, identification hinges on different drivers of regional population growth historically compared to today. In this case, the mechanisms causing historical population growth is not explicitly stated and it is therefore more demanding to justify that they are not important today.

It can be argued that the instrument is not solely based on the existence of mineral deposits, but also on the decision to mine these resources. If labor was an important factor in production, a relationship may follow whereby number of mines in a region is driven by historical population size. We are not able to test this empirically because of the lack of suitable historical population data. However, the scarce historical sources we have found typically suggest discovery by chance or by foreign professional miners traveling by the king's decree in search of valuable metals.[8] We acknowledge that if there were cases where early population size influenced the decision to mine, the instrument can be interpreted as a proxy for very early population. As larger lags of population size should be more reliable, we still prefer such a proxy over measured historical population with a smaller lag. In the analysis below, we compare our suggested mining instrument to instruments based on historical population size and the findings suggest that the bias compared to standard OLS depends on the chosen year of historical population. In further investigation of the instrumentation, we back up that the presence of mines is not much correlated with observed geographic features that may have a direct impact on wage outcomes.

A detailed description of the data on historical mines can be found in Leknes (2015). We define historical mines as mines that opened sometime between the 12th and 19th century. As in the applications by Leknes (2015) and Carlsen et al. (2016), we use the number of historical mines in each region. The historical mining activity was reasonably spread out across the country. The number of mines in a region ranges from zero to six with a mean number of 0.7 per region (s.e. = 1.28).[9]

Further analyses are made to control for omitted variables. As the plant fixed effects also include time-invariant characteristics, a concern is that the estimated relationship with urban scale reflects industry differences. We purge the plant fixed effects of industry influences by regressing them on 2-digit industry fixed effects and use the residual plant fixed effects in the correlation with worker FE. Another potential confounder to population size is how regions may vary in industry specialization. In sensitivity analysis, we include a Herfindahl index of industry concentration as control variable.

Table 1
Correlation between worker and plant fixed effects (mean values across regions).

	Aggregate (1)	Primary (2)	High school (3)	College (4)
Overall	0.061	−0.353	−0.058	−0.143
Large cities	0.143	−0.285	−0.031	0.017
Small cities	0.072	−0.358	−0.039	−0.131
Periphery	0.05	−0.36	−0.064	−0.161

Notes: Large cities are regions with population above 150,000 in 2003 (7 regions), small cities are regions with population in the range 65,000–150,000 (13 regions), while the remaining regions constitute the periphery. The worker and plant fixed effects are estimated separately for three subgroups of workers defined by the level of education.

4. Assortative matching, skills, and tasks

Our main hypothesis is that the estimated relationship between strength of matching and labor market size is positive and increasing in the formal skill level. The aggregate OLS estimated effect of population size on assortative matching equals 0.025 and is statistically significant at 5% level. Panel A of Table 2 compares the OLS estimates with the second stage IV estimates using an instrument for population size based on historical mining. The first stage estimate is documented in the first

[7] We omit one region where there is mining activity in the vicinity of a historical mine.

[8] See for instance Moen (1978) discussing chance discovery of silver at Kongsberg, Steen (1986) references discoveries in the north by the Sami during seasonal reindeer migration, and Thuesen (1979) and Øisang (1942) treating the importance of foreign miners for discoveries in the southern and middle part of the country.

[9] Table C2 in the online appendix provides more descriptive detail on the mining locations and years of operation.

Table 2
City size and strength of assortative matching: Heterogeneity across education levels.

	Dependent variable: Correlation of worker and plant FE			
	All (1)	Primary (2)	High school (3)	College (4)
Panel A: Group-specific fixed effects estimated separately for each education level				
OLS:				
Log population	0.025**	0.022*	0.011	0.048***
	(0.01)	(0.012)	(0.014)	(0.012)
R^2	0.054	0.03	0.006	0.14
IV-2SLS:				
Log population	0.039*	0.018	0.02	0.09***
	(0.023)	(0.029)	(0.033)	(0.027)
Exogeneity test (p-value)	0.499	0.885	0.778	0.054
Panel B: Fixed effects estimated from one common regression for all workers				
OLS:				
Log population		0.007	0.015	0.022**
		(0.008)	(0.011)	(0.011)
R^2		0.005	0.017	0.035
IV-2SLS:				
Log population		0.013	0.017	0.042*
		(0.021)	(0.026)	(0.023)

Notes: The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88). In columns (2)–(4) of panel A, the fixed effects are estimated separately for three subgroups of workers defined by the level of education, while in panel B, the correlation is calculated based on fixed effects estimated from one common regression for all workers. The fixed effects follow from individual level AKM estimations during 2003–2014 of the log hourly wage on worker effects, plant effects, education-specific cubic age profiles, job tenure, and year dummies. The regional population level is measured in 2003. In the IV-estimations, the instrument for log population is the number of historical mines opened before the 19th century. The first stage estimation is given in column (1) of Table 3. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

column of Table 3 and shows a strong relationship between number of historical mines opened before the 19th century and current regional population size (the first stage F statistic is about 18). IV-estimation with historical mines as instrument gives a coefficient of 0.039, significant at 10% level. The interpretation of the result is that a doubling of the regional population size increases the correlation between worker and plant fixed effects by 3.9 percentage points.

Our main innovation is to study how assortative matching differs between education groups by taking into account that plant fixed effects may vary dependent on education. In the analysis presented in columns (2)–(4) in panel A of Table 2, the worker and plant fixed effects are estimated separately for primary-, high school-, and college-educated workers. The correlation of worker and plant fixed effects for each education group is related to population size. In the IV model, the strength of matching among the college educated increases with city size, while there is no significant relationship between city size and assortative matching for low-educated workers.[10] The coefficient for the college educated equals 0.09 and is significant at 1% level. The implication is that doubling the regional population size leads to an increase in assortative matching for this group by 9 percentage points, as measured by the correlation coefficient of worker and plant fixed effects.[11] Comparing the OLS and IV point estimates for the college educated, the difference indicates 47% downward OLS-bias.[12] We acknowledge that we have not shown that the difference is statistically significant, and it is well known that the degree of imprecision increases when the instrument variable approach is used. To back up our understanding of bias, we show the rejection of exogeneity in a Wu-Hausmann test for the college educated.[13] The importance of separating the plant fixed effects for the education groups is shown by comparing the results to the case where FEs are estimated jointly for all workers, given in panel B of Table 2. When plant fixed effects are not allowed to differ by level of education, the positive effect of city size on matching for the college educated is halved and is only significant at 10% level. The difference between education groups is less clear in the aggregate model.

The analysis above investigates the relationship between continuous labor market size and strength of assortative matching. In an alternative formulation, we study how the matching varies across groups of cities dependent of population size – large cities, small cities, and periphery. We instrument the large city group with the historical mines, and the first stage estimates are reported in column (2) of Table 3. The mines are strongly related to the large city group. The results of the OLS and IV estimates of the labor market size effect in Table 4 show strong assortative matching for the college educated concentrated to large cities, while there is no significant difference between small cities and the periphery.

We investigate historical population size as an alternative instrument, as reported in Appendix Table B.3. Three alternative years are reported, 1875, 1910 and 1952. The first stage regressions are shown in columns 5–7 in Table 3, and we see that all coefficients are close to and somewhat above 1. The F-values are definitely large, increasing from 148 using 1875 to 873 using 1952. The 1952-estimates in Table B.3 are all below the OLS estimates shown in Panel A of Table 2. Using this instrument, the OLS bias is positive. When the historical year is moved backwards, to 1910 and 1875, the estimates for all workers and the college educated are increasing, and they indicate negative OLS bias using 1875. The 1875 estimates are still lower than the estimates we get when using the mining instrument.

A pertinent question is whether historical mines pick up the effects of fundamentals. Combes et al. (2010) and Rosenthal and Strange (2008) exploit geology to generate IVs for fundamentals. They emphasize how geology acts through a supply side channel, as discussed by Ahlfeldt and Barr (2022) – stable soil types help reduce the cost of floor space supply. We investigate the relationship between mining and natural features by including geographic and weather controls in the first stage estimation of the relationship between historical mines and population size. The covariates include geographic variables (region size in square km, average slope, km of coastline and mountainous area share) and climate variables (January temperatures, wind speed, amount of precipitation).

[10] It should be noticed that recent research in the US, notably Autor (2019), has shown that the effect of city size on wages has had different development over time between skill groups. It follows that our results may be specific to the period covered.

[11] Our main result that assortative matching concentrates on college-educated workers in large cities apparently is at odds with a finding of Dauth et al. (2022). They study the importance of skill-intensive city-occupations and estimate that the elasticity with respect to the size of the local labor market declines with skill intensity (measured at the regional level). To compare the results, we replicate their analysis based on occupation data available for the period 2003–2010, as documented in Table C.3 in the online appendix. We find a strong effect of employment in city-occupations on matching, but the interaction term with the share of college educated is not significant indicating that the effect of size on matching does not vary with skill intensity. However, the analysis of skill-intensive city-occupations is not directly comparable to our main analysis since it is done at the regional level. In our understanding, a proper investigation of the skill dimension requires individual observation of the education level of workers and education-specific plant fixed effects.

[12] Scatterplots of the OLS estimation, the first stage, the second stage, and the reduced form estimation of the main specification are given in Figures C.12 – C.21 in the separate online appendix. When we assess the indicated OLS bias based on the confidence intervals of the parameters, we find that 95 percent confidence intervals are overlapping and that the confidence interval of the IV estimate encompasses the OLS parameter of population size.

[13] The Wu-Hausmann test of exogeneity is reported in Table 2, and we can reject the null hypothesis of exogeneity at less than 6% level for the college educated.

Table 3
First stage IV estimation.

	Log population (1)	Large city dummy (2)	Log population (3)	Log population (4)	Log population (5)	Log population (6)	Log population (7)
Historical mines	0.288***	0.104***	0.232***	0.287***			
	(0.068)	(0.03)	(0.065)	(0.064)			
Small city dummy		−0.064					
		(0.042)					
Log population 1875					1.095***		
					(0.09)		
Log population 1910						1.132***	
						(0.047)	
Log population 1952							1.147***
							(0.039)
Herfindahl index			−10.1***				
			(2.159)				
Geographic and weather controls				Y			
Observations	88	88	88	88	88	88	88
R^2	0.148	0.255	0.349	0.355	0.707	0.816	0.917
F statistic	18.04	12.45	12.91	20.08	148.29	581.87	873.69

Notes: In column (2), the dependent variable is the large city dummy defined in the notes to Table 1, while in the other columns, the dependent variable is regional population size in 2003 (log form). In columns (1)–(4), the instrument is the number of historical mines opened before the 19th century, while historical population size (measured in 1875, 1910 or 1952) is the instrument in columns (5)–(7). In column (3), a Herfindahl index for industry concentration is included as control variable. Column (4) includes covariates on geographic and weather fundamentals: geographic variables (region size, average slope, km of coastline and mountainous area share) and climate variables (January temperature, wind speed, amount of precipitation). Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

Table 4
Alternative model specification: City dummies.

	Dependent variable: Correlation of worker and plant FE			
	All (1)	Primary (2)	High school (3)	College (4)
OLS:				
Large cities	0.094***	0.075**	0.031	0.179***
	(0.032)	(0.037)	(0.041)	(0.03)
Small cities	0.023	0.002	0.025	0.031
	(0.024)	(0.025)	(0.036)	(0.024)
R^2	0.061	0.028	0.007	0.157
IV-2SLS:				
Large cities	0.112	0.048	0.059	0.254***
	(0.073)	(0.084)	(0.096)	(0.096)
Small cities	0.025	−0.001	0.028	0.038
	(0.026)	(0.026)	(0.038)	(0.025)

Notes: The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88). In columns (2)–(4), the fixed effects are estimated separately for three subgroups of workers defined by the level of education. Large cities are regions with population above 150,000 in 2003 (7 regions), while small cities are regions with population in the range 65,000–150,000 (13 regions). The reference category is the remaining 68 regions. In the IV estimations, only the large city indicator is instrumented. The first stage estimation is given in column (2) of Table 3. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

Table 5
City size and strength of assortative matching: Heterogeneity across occupation tasks.

	Dependent variable: Correlation of worker and plant FE		
	Non-routine manual (1)	Routine (2)	Non-routine abstract (3)
OLS:			
Log population	−0.001	−0.011	0.043***
	(0.022)	(0.017)	(0.014)
R^2	0.000	0.004	0.085
IV-2SLS:			
Log population	0.013	−0.031	0.056**
	(0.035)	(0.034)	(0.023)

Notes: The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88). The fixed effects follow from individual level AKM estimations during 2003–2010 of the log hourly wage on worker effects, plant effects, education-specific cubic age profiles, job tenure, and year dummies. The fixed effects are estimated separately for three subgroups of workers defined by the task content of their occupation. We separate between non-routine manual, routine and non-routine abstract occupations. The regional population level is measured in 2003. In the IV-estimations, the instrument for log population is the number of historical mines opened before the 19th century. The first stage estimation is given in column (1) of Table 3. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

The first stage coefficient of the historical mines instrument is not much affected, as shown in Table 3 column 4, and the instrumented effect of city size including these covariates on fundamentals does not affect the matching results (documented in Table B.3, to be compared with Table 2). The inclusion of geographic fundamentals does not change the relationship between instrumented city size and matching. This result strengthens the interpretation that the mining instrument captures path dependence consistent with Bleakley and Lin (2012). In a broader context, we refer to the recent reviews of Combes and Gobillon (2015) and Ahlfeldt and Pietrostefani (2019). They conclude that sorting is the more pressing issue when it comes to estimating agglomeration effects than unobserved locational fundamentals. The importance of sorting for geographic wage disparities are examined in a decomposition analysis below.

While we concentrate on skill differences in accordance with the agglomeration literature, the heterogeneity of occupation tasks is of interest as part of the understanding of polarization of the labor market. Fredriksson et al. (2018) apply a modified AKM model to study matching of workers and jobs, where the latter is understood as tasks performed within firms. Our more limited ambition is to study whether a differentiation with respect to tasks produces results consistent with the formal skill dimension. Education levels and occupation tasks measure different, but related, aspects of skill heterogeneity. We follow the convention of separating between three groups: non-routine manual, routine, and non-routine abstract occupations. The overlap between education level groups and occupation task content varies, but is particularly strong for the college educated. As reported in Table A.2, about 80% of college-educated workers have non-routine abstract work and the college educated account for 57% of workers in this task group. The estimates of the city size effect on assortative matching for workers by occupations are reported in Table 5 and show that it is significant only for workers with non-routine abstract tasks. The importance of city size for matching is consistent between the two classifications of worker groups (by skill level or occupation tasks). Doubling city size increases the correlation between worker and plant quality by 9 percentage points for the college educated (as shown in Table 2). When workers are separated by occupation tasks, the corresponding effect of city size on matching for non-routine abstract tasks equals 5.6 percentage points. Since 57% of workers in non-routine abstract occupations are college educated, we expect a lower effect of city size for non-routine abstract occupation if the effect is driven by the college educated. Matching for workers with non-routine manual and routine tasks do not vary by city size, and these worker groups have much lower share of college educated.

We have pursued the identification issues to learn more about the underlying adjustment mechanisms. In a discussion of the endogeneity of labor market scale, Combes et al. (2008, 2010) emphasize two-way causality. Larger labor markets have been found to be more productive and are therefore more likely to attract firms and workers, which in turn increase the labor market scale. Reverse causality suggests an upward bias of OLS estimates. But bias can also be related to omitted variables. There may be unobserved traits of the region that have a positive relationship to population size and negative relations to degree of assortative matching (or vice versa). In this case, the OLS estimate would be downward biased. Our OLS and IV point estimates indicate such a negative bias. Broadly, urban amenities that compensate for wages may deflate the measured fixed effect. Better amenities in larger cities may reduce wages of high quality workers and look like bad matches with high quality firms. Industry composition may be another source of omitted variable bias. Small regions with specialized high-return industries may exhibit high assortative matching and attenuate the scale effect. Inspection of the regional data shows quite a few regions with small population that are dominated by a single industry (typically resource based, many of them linked to waterfalls and electricity production in fjords and valleys) and have a specialized and well-matched workforce. Another factor that may dilute the scale effect is the structure of the education market. The larger universities are in urban areas and provide specialized educations that match well with specific jobs/occupations. Specialization of jobs have been found to be higher in cities (Duranton and Jayet, 2011). Mismatch may then occur when persons graduating with these educations settle temporarily for sub-optimal matches waiting for better job opportunities to arise. Inspection of data for the regions again indicates that this may be important for some larger regional 'college cities'. More broadly, there are some large city regions with heterogenous private industries and many public sector jobs (not included in the analysis) that have less correlation between worker and firm quality. The larger gap between IV and OLS point estimates among college-educated workers compared to the aggregate results may follow from the 'college city' effect discussed above.

5. The importance of assortative matching for urban wage premium: decomposition

Labor markets differ with respect to skill composition of workers and consequently we expect the matching between workers and plants to reflect differences both at the supply and the demand side. Our main result implies that the market for the college educated takes advantage of larger city regions to achieve better quality matches between workers and plants. The result is consistent with the agglomeration literature discussed in the introduction – the college educated are overrepresented in larger cities (Bacolod et al., 2009) and the agglomeration effect is larger for college-educated workers (Carlsen et al., 2016). The urban wage premium has been shown to be smaller for workers with lower education. For primary- and high school-educated workers, we find that the strength of assortative matching does not vary with regional population size, and this may contribute to the explanation of the lower agglomeration effect for the low educated shown in the literature.

We offer a decomposition analysis of the spatial wage dispersion allowing for separation of the contributions from assortative matching, worker sorting, and plant sorting. The mean log wage in city region r follows from equation (1) as:

$$E_r[\ln w_{it}] = E_r\left[\mu_i + \varphi_{J(i,t)} + X_{it}\beta\right] \tag{3}$$

As in Dauth et al. (2022), the mean wage in city region r can then be written as:

$$\begin{aligned} E_r\left[w\right] = \exp(\overline{\mu}_r + \overline{\varphi}_r + \overline{X}_r\beta)\ \cdot \exp\Bigg[\left(\frac{1}{2}\left(\sigma^2_{\mu(r)} + \sigma^2_{\varphi(r)} + \sigma^2_{X\beta(r)}\right)\right) \\ +\ \left(cov_r\left(\mu_i, \varphi_{J(i,t)}\right) + cov_r\left(\mu_i, X'_{it}\beta\right) + cov_r\left(\varphi_{J(i,t)}, X'_{it}\beta\right)\right)\Bigg] \end{aligned} \tag{4}$$

where $\overline{\mu}_r$, $\overline{\varphi}_r$ and $\overline{X}_r$ are mean values of worker fixed effects, plant fixed effects and worker characteristics in region r, while $\sigma^2_{\mu(r)}$, $\sigma^2_{\varphi(r)}$ and $\sigma^2_{X\beta(r)}$ are the corresponding variances. The average observed wage level in each region is calculated from equation (4) and geographic wage differences are described by five measures: 90–10 quantile difference, 75–25 quartile spread, the standard deviation, and the urban wage premium estimated given population size in 2003 using both OLS and IV. Broadly, the observed spatial wage dispersion in Norway shown in row 1 of Table 6 is similar to the observations for Germany given by Dauth et al. (2022). The 90–10 quantile difference equals 0.24, while the urban wage premium elasticity equals 0.055 (OLS estimation).

Counterfactual experiments are made to detect the relative importance of assortative matching, worker sorting and plant sorting. The importance of matching is analyzed using two experiments (documented in rows 2 and 3 of Table 6). First, we assume homogenous matching across regions by setting the covariance between worker and plant fixed

Table 6
Decomposition of spatial wage differences: The role of assortative matching, worker sorting, and plant sorting (aggregate).

	90-10 (1)	75-25 (2)	s.d. (3)	UWP OLS (4)	UWP IV (5)
Observed spatial wage dispersion	0.24	0.153	0.1	0.055	0.075
	% difference to the observed wage dispersion:				
No assortative matching:					
a) Homogeneous matching across regions, $\mathrm{cov}_r(\mu_i, \varphi_{J(i,t)}) = 0$	−2.9	−1.3	−3.2	−2.2	−2.3
b) Zero elasticity of city size and assortative matching	−0.4	−3	−0.8	−3.1	−2.3
No spatial sorting of workers: $\overline{\mu}_r$ and $\sigma^2_{\mu(r)}$ equal across regions	−57.4	−58	−58.8	−70.7	−73.8
No spatial sorting of plants: $\overline{\varphi}_r$ and $\sigma^2_{\varphi(r)}$ equal across regions	−33.1	−37.7	−30.4	−28.8	−28.2
	% difference to wage dispersion corrected for spatial sorting of workers:				
No spatial sorting of workers and homogeneous matching across regions	−5.8	−3.8	−8	−7.1	−8.2

Notes: Geographic wage differences are measured by the 90–10 quantile difference (column 1), the 75–25 quartile spread (column 2), the standard deviation (column 3), and the urban wage premium (using log population in 2003) estimated by OLS and IV (columns 4 and 5, respectively). In the IV estimations, the instrument for log population is the number of historical mines opened before the 19th century. The first row reports the observed spatial wage dispersion, calculated by inserting regional averages into equation (4). Based on counterfactual exercises (described in the text), the following rows document the contribution from assortative matching, worker sorting, and plant sorting to the spatial wage dispersion.

Table 7
Decomposition of spatial wage differences by level of education: The role of assortative matching, worker sorting, and plant sorting (non-college vs. college-educated workers).

	90–10	75–25	s.d.	UWP OLS	UWP IV
Panel A: Workers without college education					
Observed spatial wage dispersion	0.227	0.13	0.084	0.033	0.037
	% difference to the observed wage dispersion:				
No assortative matching:					
a) Homogeneous matching across regions, $\mathrm{cov}_r(\mu_i, \varphi_{J(i,t)}) = 0$	−2.9	−4.5	−4.5	−2.5	−2.4
b) Zero elasticity of city size and assortative matching	−0.7	0.6	−0.4	−2.8	−2.5
No spatial sorting of workers: $\overline{\mu}_r$ and $\sigma^2_{\mu(r)}$ equal across regions	−48.9	−51.3	−48.8	−62.6	−70.2
No spatial sorting of plants: $\overline{\varphi}_r$ and $\sigma^2_{\varphi(r)}$ equal across regions	−36.8	−40	−36.3	−36	−36.7
	% difference to wage dispersion corrected for spatial sorting of workers:				
No spatial sorting of workers and homogeneous matching across regions	−9.7	−5.7	−9.5	−6.4	−8.1
Panel B: College-educated workers					
Observed spatial wage dispersion	0.27	0.16	0.105	0.062	0.088
	% difference to the observed wage dispersion:				
No assortative matching:					
a) Homogeneous matching across regions, $\mathrm{cov}_r(\mu_i, \varphi_{J(i,t)}) = 0$	−7	−5.4	−5.5	−5	−6.8
b) Zero elasticity of city size and assortative matching	−4	−1.1	−3.1	−9.8	−6.9
No spatial sorting of workers: $\overline{\mu}_r$ and $\sigma^2_{\mu(r)}$ equal across regions	−63.3	−65.8	−60.5	−74.9	−69.2
No spatial sorting of plants: $\overline{\varphi}_r$ and $\sigma^2_{\varphi(r)}$ equal across regions	−17.7	−30.6	−18.6	−20.8	−23.5
	% difference to wage dispersion corrected for spatial sorting of workers:				
No spatial sorting of workers and homogeneous matching across regions	−16.8	−13.3	−13	−20.5	−22.4

Notes: We separate between college-educated workers and workers without college education (primary or high school). For further descriptions, see the notes to Table 6.

effects equal to zero in all regions ($cov_r(\mu_i, \varphi_{J(i,t)}) = 0$). Second, we impose zero elasticity of matching on city size by replacing the covariance between fixed effects with the residual from a regression of the covariance against population size using IV estimation. Recalculating regional wages and the different measures of spatial wage dispersion in these two counterfactual alternatives, we show that assortative matching explains only 2–3% of the observed geographic wage differences. The importance of worker sorting is studied by assuming that the worker fixed effects have the same distribution in all regions ($\overline{\mu}_r$ and $\sigma^2_{\mu(r)}$ set equal across regions). The same is assumed for plant fixed effects to study the relative importance of spatial sorting of plants. The results are presented in rows 4 and 5 of Table 6 and show that the main explanatory factors behind observed regional wage differences are spatial sorting of workers and plants, which accounts for 60–70% and 30–35% of wage differences, respectively. In the final row of Table 6, we calculate the relative importance of matching for the urban wage premium corrected for worker sorting. We compare the spatial wage dispersion in the case of no sorting of workers (similar distribution of worker fixed effects across regions) to the case where we assume away both spatial sorting of workers and assortative matching (covariance between fixed effects set equal to zero in all regions). The agglomeration elasticity adjusted for worker sorting and estimated with IV equals 0.0196 (consistent with findings in the literature). Assuming homogeneous matching across regions reduces the elasticity to 0.018, which implies that assortative matching accounts for 8% of the corrected urban wage premium.

Investigating the role of education in this setting, we simplify the analysis by comparing workers without college education and college-educated workers. The results are reported in Table 7, where panel A represents workers without college education and panel B refers to the college educated. The agglomeration elasticity adjusted for worker sorting and estimated with IV equals 0.027 and 0.011 for high- and low-educated workers, respectively. Assuming homogeneous matching across regions reduces these elasticities to 0.021 and 0.01, respectively. This implies that assortative matching accounts for 22% of the urban wage premium for the college educated compared to only 8% of the urban wage premium for workers without college education. Assortative matching matters more for the college educated. The corresponding difference in agglomeration elasticities between the two education groups decreases from 0.016 to 0.011, which implies that better assortative matching in cities for the college educated compared to the non-college educated explains about 1/3 of the difference in urban wage premia between the two education groups. When the decomposition of spatial wage differences is based on common estimation of plant fixed effects rather than education-specific estimation, assortative matching cannot explain the difference in urban wage premium between the two education groups. This shows that the plant fixed effects should be allowed to vary between education levels to capture important heterogeneity.

6. Robustness: limited mobility bias

The AKM-model applied to estimate assortative matching has challenges related to limited mobility bias and interpretation of the plant effects. The plant fixed effects are identified based on job switchers, and the estimates may be noisy when mobility is low. Thus, the relationship between the strength of matching and local labor market scale may potentially be biased. We investigate the importance of small plants and few switchers to throw more light on the limited mobility problem. Several methods are applied to test the sensitivity of results to limited worker mobility, including adjustments to the specification and the sample. The analyses are based on the model assuming separate plant fixed effects for the different levels of education and with continuous labor market size effect.

We start the investigation by studying alternative cutoffs for the exclusion of small plants. In the main analyses, we exclude plants with five or fewer workers. Table 8 reports four alternatives. Panel A includes all plants in the dataset, panel B excludes plants with 2 or fewer workers, panel C excludes plants with 7 or fewer workers, and panel D excludes plants with 10 or fewer workers. As shown in the table, the number of plants is strongly reduced when the cutoff is increased. The total dataset has about 118,000 plants, but this is reduced to 16,000 when plants with 10 or fewer workers are excluded. The number of workers is much less affected. As discussed in section 2, the cutoff is important for the share of regions with positive correlation between worker and plant fixed effects. A higher cutoff means a higher share of positive correlations, which is considered more economically plausible. The effect of population size for the strength of matching is very similar across alternative cutoffs with IV-estimates for the college educated that vary between 0.075 and 0.095.

Table 8
Alternative cutoff level for exclusion of small plants.

	Dependent variable: Correlation of worker and plant FE			
	All (1)	Primary (2)	High school (3)	College (4)
Panel A: Including all plants independent of size (117,684 plants and 1,218,862 workers)				
Log population	0.053**	0.033	0.042	0.092***
(IV-2SLS)	(0.022)	(0.024)	(0.028)	(0.028)
Panel B: Excluding plants with 2 or fewer workers (72,566 plants and 1,157,256 workers)				
Log population	0.049**	0.035	0.035	0.088***
(IV-2SLS)	(0.022)	(0.027)	(0.03)	(0.029)
Panel C: Excluding plants with 7 or fewer workers (23,154 plants and 937,777 workers)				
Log population	0.042*	0.02	0.012	0.095***
(IV-2SLS)	(0.024)	(0.035)	(0.035)	(0.029)
Panel D: Excluding plants with 10 or fewer workers (15,676 plants and 856,743 workers)				
Log population	0.019	0.026	−0.011	0.075***
(IV-2SLS)	(0.022)	(0.033)	(0.035)	(0.025)

Notes: The table presents robustness checks on the cutoff level for exclusion of small plants. The main regressions in Table 1 exclude plants with five or fewer workers, which leaves 32,760 plants and 1,008,170 workers. Panels A–D estimate the effect of city size on the strength of assortative matching for alternative cutoff levels with respect to plant size. The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88). In columns (2)–(4), the fixed effects are estimated separately for each education level. In each panel, the fixed effects are re-estimated given the new sample. The regional population level in 2003 is instrumented with the number of historical mines opened before the 19th century. The first stage estimation is given in column (1) of Table 3. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

This indicates that the limited mobility bias does not vary systematically with population size.[14]

An alternative strategy is to exclude plants based on the education-specific plant size (number of workers in the respective education group).[15] There are two disadvantages of such an approach. First, the common set of plants across the three education groups is reduced. While a general cutoff level based on the total number of workers excludes the same plants for all education groups, an education-specific cutoff gives separate plant samples for each education group (and could make results less comparable across groups). Second, compared to a general cutoff level, the sample is reduced much more when using an education-specific cutoff. Figure B.2 illustrates the loss of observations for the college educated for alternative general and education-specific cutoff levels. A general cutoff level at 5 (excluding plants with 5 or fewer workers) leaves 88% and 45% of the worker and plant sample, respectively. The same cutoff level for the number of college-educated workers in plants implies that only 75% and 15% of the worker and plant samples remain. An education-specific cutoff at 2 would entail about the same loss of observations as the general cutoff at 5. The two other education groups have similar patterns, as documented in Figures C.25 – C.26 in the online appendix. The consequences of different assumptions about the education-specific cutoff for the city-size effects on matching are shown in Appendix Table B.5. The main conclusion of the paper, the positive and statistically significant association between regional population size and the degree of assortative matching for the college educated is robust across cutoff values. With an education-specific cutoff at 2, the size of the effect is the same and a similar number of observations are excluded as with a general cutoff at 5. When the education-specific cutoff level is set higher, the size of the effect goes down and eventually becomes insignificant when a sizeable share of the sample is dropped.

Further robustness checks are reported in Table 9 (starting out from the main dataset). In panel A, the dependent variable is the correlation between worker fixed effects plus observable worker characteristics and plant fixed effects. Using a broader measure of worker quality (taking observable worker characteristics into account) does not alter the baseline findings. In panel B, we apply a more direct approach to address noisy estimates that originate from limited mobility by excluding the top and bottom percentile of the fixed effects distributions. The results do not change noteworthy and the city region size effect is still significant only for the college educated. As the plant fixed effects also include time-invariant characteristics, a concern is that the estimated relationship with urban scale reflects industry differences. In Panel C, we purge the plant FE of industry influences by regressing them on 2-digit industry fixed effects and use the residual plant fixed effects in the correlation with worker FE. The results are similar to the main analysis with an effect concentrated to the college educated. Industry concentration may also influence both matching opportunities and population size. In panel D, we include a Herfindahl index of industry concentration as control variable. The additional control does not affect the estimated effect for college-educated workers.

Table 9
Robustness checks.

	Dependent variable: Correlation of worker and plant FE			
	All (1)	Primary (2)	High school (3)	College (4)
Panel A: Correlation between $\mu_i + X_{it}\beta$ and plant FEs				
Log population	0.033	0.012	0.003	0.097***
(IV-2SLS)	(0.023)	(0.028)	(0.03)	(0.03)
Panel B: Excluding outliers in the FE distributions				
Log population	0.03*	0.015	0.004	0.069***
(IV-2SLS)	(0.017)	(0.022)	(0.019)	(0.022)
Panel C: Plant FEs purged for industry fixed effects				
Log population	0.055**	0.031	0.044	0.103***
(IV-2SLS)	(0.023)	(0.027)	(0.03)	(0.029)
Panel D: Herfindahl index for industry concentration included as control variable				
Log population	0.039	0.023	0.025	0.099***
(IV-2SLS)	(0.029)	(0.037)	(0.041)	(0.036)
Panel E: Plant FEs purged for average plant size				
Log population	0.025 (0.02)	0.016	0.014	0.09***
(IV-2SLS)		(0.027)	(0.027)	(0.026)
Panel F: Worker FEs purged for occupation fixed effects				
Log population	0.037	0.018	0.01	0.099***
(IV-2SLS)	(0.024)	(0.03)	(0.035)	(0.029)
Panel G: Largest connected group				
Log population	0.038*	0.017	0.018	0.088***
(IV-2SLS)	(0.023)	(0.029)	(0.033)	(0.028)

Notes: The table presents robustness checks on the regressions in panel A of Table 2. In columns (2)–(4), the fixed effects are estimated separately for each education level. In panel A, the dependent variable is the correlation between worker fixed effects plus observable worker characteristics and plant fixed effects. In panel B, we exclude workers that are in the top or bottom 1% of the distribution of worker and/or plant fixed effects. In panel C, the dependent variable is the correlation between worker and residual plant fixed effects at the regional level, where the residual plant effects are the residuals from a regression of the plant effects on 2-digit industry effects. In panel D, a Herfindahl index for industry concentration is included as control variable in the regional level regressions. In panel E, the dependent variable is the correlation between worker and residual plant fixed effects at the regional level, where the residual plant effects are the residuals from a regression of the plant effects on average plant employment size. In panel F, we purge 2-digit occupation fixed effects from the worker fixed effects and use the correlation between residual worker fixed effects and plant fixed effects as the dependent variable. In panel G, only the largest connected group is included in the estimation of fixed effects. In all panels, the instrument for log population is the number of historical mines opened before the 19th century. For panel D, the first stage estimation is given in column (3) of Table 3, while for all other panels, the first stage estimation is given in column (1) of Table 3. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

[14] An alternative exclusion criterium is investigated in Table B.4 in Appendix B, where the cutoff is based on few mobility events rather than the size of the plant. The main conclusion holds in all four alternatives investigated in the table (from 2 or fewer job mobility events to 10 or fewer) – matching of college-educated workers is increasing with population size.

[15] Figures C.22 – C.24 in the online appendix illustrate a positive relationship between education-specific correlations of fixed effects and the minimum education-specific plant size.

It is of interest to investigate whether the large city assortative matching result is the effect of large plants. Given the estimated fixed effects from the AKM estimation, we purge average plant employment size from the plant fixed effects. This robustness test is documented in panel E of Table 9. The average plant size explains little of the variation in plant fixed effect (with R-squared in the range 0.04–0.12). Another issue investigated using the same method concerns occupations. We purge 2-digit occupations from the worker fixed effects, and the new estimates are shown in panel F of Table 9. The 2-digit occupation explains little of the variation in worker fixed effects (the R-squared in the range from 0.09 to 0.27). For the college educated, occupation group explains 20% of worker fixed effects. We acknowledge that this approach only captures the role of occupation within each level of education, while it is more challenging to identify the importance of occupations for the difference between educational groups (since not many occupations overlap between the three levels of education). Finally, in panel G of Table 9, we include only the largest connected group in the analysis. This gives a more correct comparison of fixed effects, but since very few observations are unconnected, the main results are not affected.[16]

Next, we pursue an alternative approach to cutting the sample based on clustering of plants. In this analysis, we therefore apply all plant observations in the data (including plants with five or fewer workers). Bonhomme et al. (2019, 2022) suggest to cluster firms to increase statistical support. Following this reasoning, we estimate a fixed-effect model where all plants are allocated to clusters with similar wage structures and take benefit of the higher mobility of workers between these clusters. As in the application of Dauth et al. (2022), we characterize the wage structure in each plant by 20 and 40 wage quantiles and use a k-means clustering technique to classify plants into 10, 15 and 20 clusters. Ongoing research develops methods of bias correction to mitigate possible limited mobility bias, notably Bonhomme et al. (2020). This is especially important when the role of firm effects for wage inequality is evaluated. Interestingly, the covariance results resemble those of Bonhomme et al. (2020) with bias correction, easing the concern of downward bias caused by insufficient worker mobility.[17]

The clustered plant estimation washes out much of the plant heterogeneity and is therefore a harsh test of result robustness to limited mobility bias. Nonetheless, as can be seen from panels A–F of Appendix Table B.6, the overall conclusions hold for most specifications. There is a positive urban gradient in assortative matching that seems to originate from the outcomes of college-educated workers. However, the estimated coefficients are generally lower (about a half of the baseline for the full sample and a fourth for the college sample). The reduction in coefficients suggests that the limited worker mobility may cause an upward bias in the relationship between city size and assortative matching. In panel G of Table B.6, the clustering is concentrated to small plants (10 or fewer workers) and based on industry and location. The rationale being that plants operating within the same activity in the same local market should have comparable productivity. In this case, the effect of population size for the college educated is somewhat larger.

The leave-out estimator developed by Kline et al. (2020) is a major contribution to the estimation of two-way fixed effects for the aggregate economy. Unfortunately, the estimator is not developed for backing out correlations between worker and plant fixed effects for separate regions with inter-regional mobility. Noting that the samples and estimates are not directly comparable, Dauth et al. (2022) run the estimator for each region in Germany (with plant effects only identified from within-region mobility) and find fairly similar results as when using the AKM model. The general pattern is that the correlations between fixed effects shifts upwards, but the population gradient is close to unchanged. However, the observations across regions are then not linked and the estimated fixed effects are not comparable. We suggest mitigating limited mobility bias by excluding plants in the bottom of the employment size distribution. In this way, fixed effects are still estimated based on inter-regional mobility. Using different plant size cutoffs, we show that correlation between fixed effects increase in plant size cutoff value while the population gradient on assortative matching is more or less unchanged (as documented in Fig. 2 and Table 8). The similar pattern as described by Dauth et al. (2022) comparing the leave-out and AKM estimators may be incidental, but indicates that the gradient with respect to urban scale is robust across settings, samples and estimators.

7. Assortative matching heterogeneity: gender and age

In the following, we explore if there are other traits of workers in conjunction with education that may explain the results. We focus on worker heterogeneity with respect to gender and age. Gender differences at the labor market have been studied in the context of urbanization and regional size. Hirsch et al. (2013) deal with regional differences in the gender pay gap, while Phimister (2005) analyzes the urban wage premium by gender. They find that the gender pay gap is lower in cities and the urban wage premium is larger for women, respectively. Based on worker and plant fixed effects estimated separately for the three education groups, we find that both male and female college-educated workers take advantage of larger labor markets, as documented in panel A of Table 10. The point estimate of city size on assortative matching is somewhat higher for college-educated men than for college-educated women, but the difference is not statistically significant. The main gender difference is that assortative matching among female workers with high school education also is increasing with regional population size. Consequently, a larger share of the female worker population in cities may display better matching compared to men. While not conclusive, these results are in line with higher urban wage premium for women found in Phimister (2005).

As a further analysis of the gender heterogeneity, we estimate group-specific fixed effects at a more detailed level. We allow for different assortative matching for each education-gender group, where the fixed effects are estimated separately for six worker groups combining gender and three levels of education. The results confirm the gender pattern – the population size matters for the matching of male and female college-educated workers and also for high school educated women.[18]

Following the literature on job search, we expect workers early in their career to have imperfect information about their own preferences and abilities. They are therefore expected to experience worse matches early in their career relative to later. Yankow (2009) shows that workers tend to find more productive job matches over time. Because of data censoring (first observations are in 2003), we have limited knowledge about individual work histories. However, we can infer that younger workers have come shorter in the search process for better job matches. In panel B of Table 10, we split the workforce into four age categories: 25–34, 35–44, 45–54 and 55–65. We obtain the same conclusion as from the first part of the paper, positive urban assortative matching is concentrated to college-educated workers. Among these, in line with our expectations, the youngest age group has the lowest score and the magnitude of effect is largest (and most significant) for workers of intermediate ages (35–54 years of age).

[16] Unconnected worker-year observations that are removed from the analysis account for only 0.17% of the aggregate sample. When fixed effects are estimated separately for each education group, the share of unconnected observations equals 4.79%, 0.97% and 1.77% for primary, high school and college-educated workers, respectively.

[17] The correlation between worker and plant cluster fixed effects is 39 percent for the full sample and ranges from 27 to 36 percent for the education groups. The correlation increases in education level suggesting stronger sorting effects for highly educated workers.

[18] The point estimate of city size on assortative matching is about 0.08 for both male and female college-educated workers (significant at 1% level) and equals 0.064 for high school educated women (significant at 5% level).

Table 10
City size and strength of assortative matching: Heterogeneity by gender and age.

	Dependent variable: Correlation of worker and plant FE					
	Primary (1)	High school (2)	College (3)	Primary (4)	High school (5)	College (6)
Panel A: Gender heterogeneity						
	MEN			WOMEN		
Log population (IV-2SLS)	0.017	0.029	0.091***	0.024	0.032**	0.055***
	(0.029)	(0.032)	(0.029)	(0.019)	(0.016)	(0.016)
Panel B: Age heterogeneity						
	AGE 25–34			AGE 35–44		
Log population (IV-2SLS)	−0.015	−0.007	0.056**	0.054	0.049	0.127***
	(0.029)	(0.028)	(0.024)	(0.049)	(0.041)	(0.038)
	AGE 45–54			AGE 55–65		
Log population (IV-2SLS)	0.007	0.01	0.095***	−0.052	−0.001	0.08**
	(0.035)	(0.03)	(0.033)	(0.039)	(0.052)	(0.038)

Notes: The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88) for subgroups of workers defined based on level of education, gender, and age. The fixed effects are estimated separately for each education level and follow from individual level AKM estimations during 2003–2014 of the log hourly wage on worker effects, plant effects, education-specific cubic age profiles, job tenure, and year dummies. The regional population level in 2003 is instrumented with the number of historical mines opened before the 19th century. The first stage estimation is given in column (1) of Table 3. Robust standard errors are in parentheses. *** indicates significance at the 1 percent level.

8. Concluding remarks

The education level of workers has been found to impact the urban wage premium. Recent evidence shows that superior labor matching may be a potential source of the premium, but has not been put in relation to formal skills. Using rich administrative data for Norway linking workers and plants, we study if there is a skill-heterogeneous relationship between assortative matching and city size. Workers are separated by education level – primary, high school and college. The method innovated by Abowd et al. (1999), the AKM-model, is employed to estimate two-way worker and plant fixed effects, from which we derive measures of the strength of assortative matching. The analysis addresses the concern of possible bias related to endogeneity and omitted variables. We instrument population size using historical mines before the 19th century that are obsolete today. Further methodological issues are handled related to heterogeneity of firm fixed effects and robustness for limited mobility bias.

The analysis shows that the overall positive relationship between city size and assortative matching is driven by college-educated workers. This conclusion is corroborated by results for occupation-task groups as workers conducting high-skilled work, non-routine abstract tasks, display a similar pattern. Using a wage decomposition method, we find that, net of worker sorting, assortative matching explains 22% of the urban wage premium found for higher skilled workers. The analysis is extended to address how the role of education varies across age and gender. The relationship between city size and matching is confirmed for both male and female college-educated workers, but female workers with high school education also benefit from city size. Finally, we explore the age gradient with respect to assortative matching in regions of different sizes. College-educated workers of all ages are better matched in more populous areas, and in particular among workers in the middle of the age distribution.

Methodological issues remain, in particular estimating the quality of firms. The plant fixed effects incorporate time-invariant characteristics that may be related to labor market size. It is of interest to develop characterizations of firms with more data about their structure and performance.

Author statement

The authors share responsibility for the paper.

Declaration of competing interest

The authors have no conflicts of interest regarding this paper.

Appendix C. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.regsciurbeco.2022.103806.

Appendix A. Additional details on the dataset

The individual level dataset is computed from three administrative registers: employment, tax, and education. The employment register links workers and plants and gives information on work contracts for all employees. It includes the duration of the contract, the type of contract, and the exact number of hours worked per week. We calculate the number of hours worked per year, which is combined with data on annual wage income from the tax register to give a measure of hourly wages for all employees. The education register covers the entire adult population and gives information about workers' level of education. We also have information on the age, gender, industry affiliation, plant affiliation, and home region of all individuals.

We exclude workers in the primary industries (agriculture, fishing, and forestry), as well as public sector workers. The employment register separates between three contract types: full-time contracts with at least 30 h of work per week, part-time contracts with 20–29 h of work per week, and part-time contracts with fewer than 20 h of work per week. We concentrate on full-time workers aged 25–65. This gives a dataset of about 12 million worker-year observations during 2003–2014. The tax register gives information on total annual earnings, rather than separate earnings for each work contract. Workers with more than two contracts during a year, as well as workers with one full-time and one part-time contract, are excluded. For workers with two full-time contracts, we allow for a maximum of three months of overlap between the contracts. We also exclude workers whose contract length is less than one month during a year. These restrictions reduce the dataset by about 1.3 million observations. Missing data on hours worked, annual earnings, level of education, or industry affiliation, together with exclusion of workers that change education level after entering the labor market as full-time employees, further excludes approximately 1.9 million observations. To avoid extreme observations, we exclude the top and bottom 1% of the wage distribution. Finally, we exclude observations that do not contribute to the estimation of plant fixed effects, workers that are observed in a single year and plants without any job switches (amounts to about 0.5 million observations). This leaves a dataset with about 8.1 million observations covering 1.2 million workers and 118,000 plants. To mitigate limited mobility bias, we exclude plants with five or fewer workers, which gives the main dataset of about 6.5 million observations during 2003–2014 covering 1,008,170 workers and 32,760 plants.

Table A.1 below reports descriptive statistics of the individual level data, both aggregate and for three levels of education. The average hourly wage in the dataset is 314 NOK (log wage of 5.65) and wages of primary educated and high school educated workers are, respectively, 40% and 26% below wages of the college educated. The average age is 43 years and decreases with the level of education, from 44 years among the primary educated to 41 years among college-educated workers. Overall, 72% of workers are male, but the share is lower among the college educated. Separating between manufacturing industries and services, about 55% of primary and high school educated workers are employed in services, increasing to 70% for the college educated. The mean value of worker fixed effects increases with the level of education, from −0.17 among the primary educated to 0.18 among college-educated workers (fixed effects are centered around zero overall).

We define worker groups based on occupation tasks and separate between non-routine manual, routine and non-routine abstract occupations (as described in Acemoglu and Autor, 2011). Occupational data is available for the period 2003–2010 and includes 345 occupations. We use the occupational crosswalk created by Hoen (2016) to link the Norwegian occupation codes to the occupational characteristics from the O*NET database (version 9.0, December 2005). Task characteristics included in the three occupation task groups are given below. For each occupation, the task intensity of manual, routine and abstract tasks is calculated as an average of the respective importance scales (range 1–5) and the occupation is included in the task group with the highest intensity.

Non-routine manual tasks:

- 4.A.3. a.4 Operating vehicles, mechanized devices, or equipment
- 4.C.2. d.1. g Spend time using hands to handle, control or feel objects, tools or controls
- 1.A.2. a.2 Manual dexterity
- 1.A.1. f.1 Spatial orientation

Routine tasks:

- 4.C.3. b.7 Importance of repeating the same tasks
- 4.C.3. b.4 Importance of being exact or accurate
- 4.C.3. d.3 Pace determined by speed of equipment
- 4.A.3. a.3 Controlling machines and processes
- 4.C.2. d.1. i Spend time making repetitive motions

Non-routine abstract tasks:

- 4.A.2. a.4 Analyzing data/information
- 4.A.2. b.2 Thinking creatively
- 4.A.4. a.1 Interpreting information for others
- 4.A.4. a.4 Establishing and maintaining personal relationships
- 4.A.4. b.4 Guiding, directing and motivating subordinates
- 4.A.4. b.5 Coaching/developing others

Table A.2 shows the degree of overlap between education levels and occupation task content. Overall, the share of workers employed in non-routine manual, routine and non-routine abstract occupations equal 24%, 32% and 44%, respectively. Separating by level of education, the broad picture is that college-educated workers are overrepresented in non-routine abstract occupations (80% of the college educated belong to this group), while workers with high school or primary education are more likely to be employed in non-routine manual or routine occupations.

Table A.1
Descriptive statistics (mean values)

	All	Primary	High school	College
Log hourly wage	5.65	5.45	5.59	5.85
Hourly wage (in NOK)	314.2	251.2	291.8	383.1
Age	42.9	44.0	43.7	40.9
Male	0.72	0.74	0.76	0.65
Manufacturing	0.41	0.44	0.48	0.30
Services	0.59	0.56	0.52	0.70
Job tenure	5.27	5.38	5.76	4.43
Worker fixed effect	0.00	−0.17	−0.05	0.18
Plant fixed effect	0.00	−0.02	−0.01	0.02
No. of workers	1,008,170	179,732	509,352	319,086
Share of workers	1.00	0.18	0.50	0.32

Notes: The values for hourly wage, age and job tenure refer to the average value across the period 2003–2014. Job tenure is calculated based on actual days worked in the worker's present plant from 1993 onwards and is expressed in years.

Table A.2
Overlap between education level and occupation task content

Panel A: Share of workers overall and within each education level by occupation task group				
	All	Primary	High school	College
Non-routine manual	0.237	0.35	0.316	0.042
Routine	0.325	0.471	0.377	0.154
Non-routine abstract	0.438	0.179	0.307	0.804
Panel B: Share of workers overall and within each occupation task group by level of education				
	All	Non-routine manual	Routine	Non-routine abstract
Primary	0.177	0.261	0.257	0.072
High school	0.514	0.684	0.597	0.361
College	0.309	0.055	0.146	0.567

Notes: Based on data for the period 2003–2010.

Appendix B. Additional tables and figures

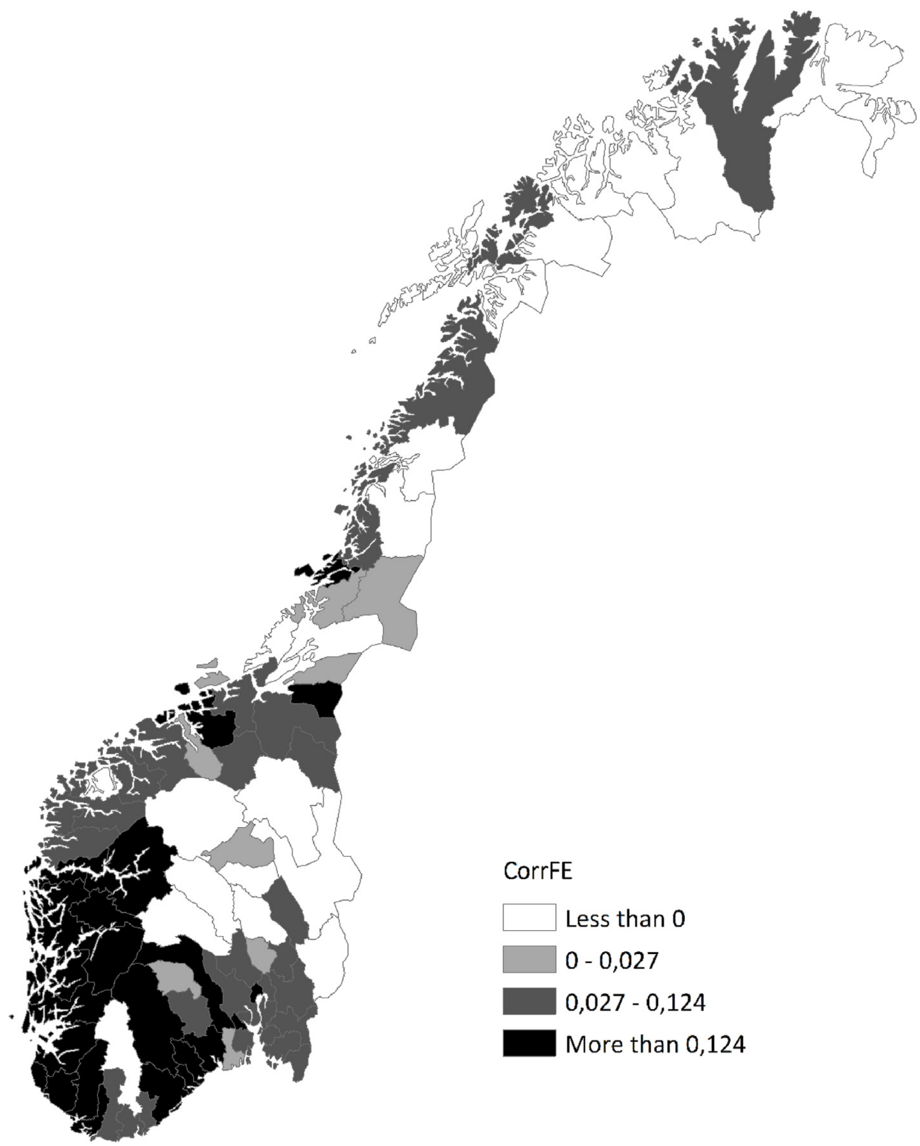


Fig. B.1. Regional distribution of the correlation between worker and plant FEs.

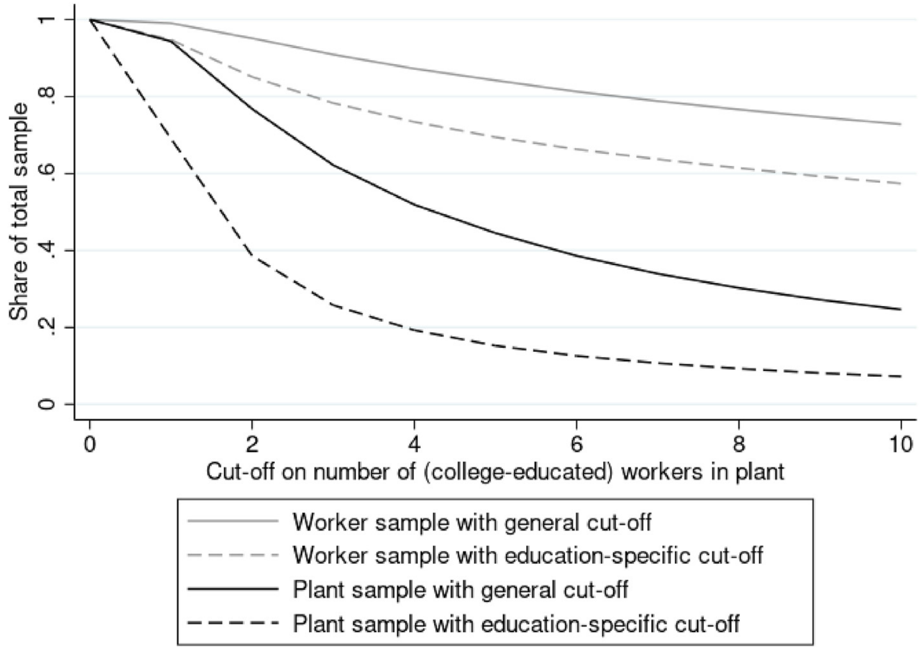


Fig. B.2. Loss of observations for the college educated for alternative (general and education-specific) cutoff levels.

Table B.1
AKM model versus model with full set of firm-worker fixed effects

	Baseline AKM-model (1)	Saturated match-model (2)
Panel A: All workers		
Adjusted R-squared	0.7897	0.8216
Root MSE	0.1986	0.1814
Panel B: Primary education		
Adjusted R-squared	0.6816	0.7166
Root MSE	0.2180	0.2029
Panel C: High school		
Adjusted R-squared	0.7680	0.7992
Root MSE	0.1916	0.1767
Panel D: College		
Adjusted R-squared	0.7992	0.8301
Root MSE	0.1936	0.1763

Notes: The table shows goodness of fit by adjusted R-squared and root mean square error (RMSE) for two specifications of the wage variation in Norway: the baseline AKM model specified by equation (1) and a corresponding specification where the firm and worker fixed effects are replaced by firm-worker fixed effects.

Table B.2
Comparison of the distributions of education-specific plant FEs

	Primary vs. High school (1)	High school vs. College (2)	Primary vs. College (3)
Correlation between plant FEs distributions	0.19	0.19	0.10
Rank correlation between plant FEs distributions	0.26	0.26	0.15
Percent overlap of plants within each quintile:			
Quintile 1	31.0	31.1	25.4
Quintile 2	24.9	24.0	23.0
Quintile 3	18.2	21.5	22.7
Quintile 4	26.1	24.8	24.0
Quintile 5	36.1	37.1	29.1
No. of plants	24,439	24,027	19,065

Notes: The table compares plant fixed effects distributions of two education groups at the time by calculating the percent overlap of plants within each quintile of the distribution. The overall correlation and rank correlation between the two distributions is also reported. In the comparison, we only consider plants that employ workers from both education groups.

Table B.3
Robustness of IV estimation: Additional covariates on fundamentals and historical population size as instrument

	Dependent variable: Correlation of worker and plant FE			
	All (1)	Primary (2)	High school (3)	College (4)
Instrument:				
Historical mines (including covariates on fundamentals)	0.046*	0.049	0.053	0.079***
	(0.025)	(0.033)	(0.038)	(0.029)
Log population 1952	0.018*	0.018*	0.006	0.039***
	(0.01)	(0.011)	(0.014)	(0.012)
Log population 1910	0.026**	0.018	0.019	0.041***
	(0.01)	(0.011)	(0.015)	(0.012)
Log population 1875	0.035***	0.017	0.033*	0.05***
	(0.012)	(0.013)	(0.017)	(0.013)

Notes: The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88). In columns (2)–(4), the fixed effects are estimated separately for three subgroups of workers defined by the level of education. In the first row, we add additional covariates on fundamentals to our main instrument for log population (historical mines opened before the 19th century). The covariates comprise geographic variables (region size in square kilometers, average slope, km of coastline and mountainous area share) and climate variables (January temperature, wind speed, amount of precipitation). Descriptive statistics of the variables are found in Table C4 in the separate online appendix. In the last three rows, the instrument for log population is historical population level in 1952, 1910 and 1875, respectively. The first stage estimations are given in columns (4)–(7) of Table 2. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

Table B.4
Alternative exclusion criterium: excluding plants with low job mobility

	Dependent variable: Correlation of worker and plant FE			
	All (1)	Primary (2)	High school (3)	College (4)
Panel A: Excluding plants with 2 or fewer job mobility events (89,803 plants and 1,306,309 workers)				
Log population (IV-2SLS)	0.050**	0.038	0.019	0.103***
	(0.023)	(0.026)	(0.027)	(0.031)
Panel B: Excluding plants with 5 or fewer job mobility events (54,427 plants and 1,220,085 workers)				
Log population (IV-2SLS)	0.043*	0.037	−0.003	0.093***
	(0.022)	(0.025)	(0.028)	(0.029)
Panel C: Excluding plants with 7 or fewer job mobility events (42,850 plants and 1,173,246 workers)				
Log population (IV-2SLS)	0.043*	0.049	−0.012	0.104***
	(0.023)	(0.030)	(0.029)	(0.031)
Panel D: Excluding plants with 10 or fewer job mobility events (32,230 plants and 1,113,673 workers)				
Log population (IV-2SLS)	0.020	0.004	−0.032	0.071***
	(0.019)	(0.028)	(0.029)	(0.024)

Notes: The table presents robustness checks of the results using alternative cutoff levels for exclusion of plants with low job mobility. Job mobility events are measured as the sum of switches in and out of the plant. The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88). In columns (2)–(4), the fixed effects are estimated separately for each education level. In each panel, the fixed effects are re-estimated given the new sample. The regional population level in 2003 is instrumented with the number of historical mines opened before the 19th century. The first stage estimation is given in column (1) of Table 2. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

Table B.5
Robustness: Excluding small plants based on the education-specific plant size

	Dependent variable: Correlation of worker and plant FE		
	Primary (1)	High school (2)	College (3)
Panel A: Excluding plants with 2 or fewer workers in the respective education group			
Log population (IV-2SLS)	0.017	0.039	0.084***
	(0.036)	(0.034)	(0.027)
Number of workers	174,507	560,601	321,767
Number of plants	19,060	45,867	21,887
Panel B: Excluding plants with 4 or fewer workers in the respective education group			
Log population (IV-2SLS)	−0.022	0.015	0.048**
	(0.031)	(0.035)	(0.021)
Number of workers	131,452	482,774	285,627
Number of plants	7,570	22,827	10,890
Panel C: Excluding plants with 6 or fewer workers in the respective education group			
Log population (IV-2SLS)	0.033	0.000	0.042*
	(0.05)	(0.033)	(0.023)
Number of workers	107,462	429,576	263,007
Number of plants	4,267	14,361	7,085

(*continued on next page*)

Table B.5 (*continued*)

	Dependent variable: Correlation of worker and plant FE		
	Primary (1)	High school (2)	College (3)
Panel D: Excluding plants with 8 or fewer workers in the respective education group			
Log population (IV-2SLS)	−0.024	−0.02	0.043*
	(0.034)	(0.033)	(0.025)
Number of workers	91,252	390,940	246,903
Number of plants	2,804	10,237	5,214
Panel E: Excluding plants with 10 or fewer workers in the respective education group			
Log population (IV-2SLS)	−0.026	−0.016	0.012
	(0.033)	(0.027)	(0.022)
Number of workers	79,495	360,140	233,244
Number of plants	2,016	7,744	4,068

Notes: The table presents robustness checks on the education-specific cutoff level for exclusion of small plants. The dependent variable is the correlation between worker and plant fixed effects at the regional level (N = 88). The fixed effects are estimated separately for three subgroups of workers defined by the level of education. For each education group, plants with few workers in the respective education group are excluded. In each panel, the fixed effects are re-estimated given the new sample. The regional population level in 2003 is instrumented with the number of historical mines opened before the 19th century. The first stage estimation is given in column (1) of Table 3. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

Table B.6
Clustering of plants

	Dependent variable: Correlation of worker and clustered plant FE			
	All (1)	Primary (2)	High school (3)	College (4)
Clustering all plants based on wage structures				
Panel A: 10 clusters defined by 20 wage quantiles				
Log population (IV-2SLS)	0.023*	0.001	−0.012	0.019*
	(0.012)	(0.011)	(0.015)	(0.010)
Panel B: 15 clusters defined by 20 wage quantiles				
Log population (IV-2SLS)	0.026**	0.003	−0.008	0.012
	(0.012)	(0.010)	(0.014)	(0.010)
Panel C: 20 clusters defined by 20 wage quantiles				
Log population (IV-2SLS)	0.028**	0.001	−0.006	0.023**
	(0.012)	(0.011)	(0.014)	(0.011)
Panel D: 10 clusters defined by 40 wage quantiles				
Log population (IV-2SLS)	0.026**	0.002	−0.006	0.015*
	(0.012)	(0.011)	(0.015)	(0.009)
Panel E: 15 clusters defined by 40 wage quantiles				
Log population (IV-2SLS)	0.020*	0.008	−0.009	0.013
	(0.011)	(0.011)	(0.014)	(0.010)
Panel F: 20 clusters defined by 40 wage quantiles				
Log population (IV-2SLS)	0.028***	−0.000	−0.004	0.023**
	(0.012)	(0.011)	(0.014)	(0.011)
Panel G: Clustering small plants (10 or fewer workers) based on industry and geography				
Log population (IV-2SLS)	0.031*	−0.001	−0.019	0.059**
	(0.018)	(0.019)	(0.03)	(0.028)

Notes: Panels A–F classify plants into k = 10, 15, 20 clusters with similar wage structures measured by 20 and 40 quantiles (following discussion in Bonhomme et al., 2019, 2022). In panel G, small plants with 10 or fewer workers on average are clustered based on industry and geography in two steps: first, we cluster based on 2-digit industry code and labor market affiliation. Second, we cluster the remaining small plants based on aggregate codes (1-digit industry and county). Worker fixed effects are defined at the education level. Robust standard errors are in parentheses. ***, ** and * indicate significance at the 1 percent, 5 percent and 10 percent level, respectively.

References

Abowd, J., Kramarz, F., Margolis, D., 1999. High-wage workers and high-wage firms. Econometrica 67 (2), 251–333.

Acemoglu, D., 2007. Equilibrium bias of technology. Econometrica 75 (5), 1371–1409.

Acemoglu, D., Autor, D., 2011. Skills, tasks and technologies: implications for employment and earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 4b. Elsevier, North-Holland, pp. 1043–1171.

Ahlfeldt, G., Barr, J., 2022. The economics of skyscrapers: a synthesis. J. Urban Econ. 129, 103419.

Ahlfeldt, G., Pietrostefani, E., 2019. The economic effects of density: a synthesis. J. Urban Econ. 111, 93–107.

Andersson, F., Burgess, S., Lane, J., 2007. Cities, matching and the productivity gains from agglomeration. J. Urban Econ. 61, 112–128.

Andrews, M., Gill, L., Schank, T., Upward, R., 2012. High wage workers match with high wage firms: clear evidence of the effects of limited mobility bias. Econ. Lett. 117, 824–827.

Autor, D., 2019. Work of the past, work of the future. American Economic Review Papers and Proceedings 109, 1–32.

Autor, D., Levy, F., Murnane, R., 2003. The skill content of recent technological change: an empirical exploration. Q. J. Econ. 118 (4), 1279–1333.

Bacolod, M., Blum, B., Strange, W., 2009. Skills in the city. J. Urban Econ. 65, 136–153.

Baum-Snow, N., Pavan, R., 2012. Understanding the city size wag gap. Rev. Econ. Stud. 79 (1), 88–127.

Bleakley, H., Lin, J., 2012. Portage and path dependence. Q. J. Econ. 127 (2), 587–644.

Bonhomme, S., Lamadon, T., Manresa, E., 2019. A distributional framework for matched employer employee data. Econometrica 87, 699–739.

Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M., Setzler, B., 2020. How Much Should We Trust Estimates of Firm Effects and Worker Sorting? NBER Working Paper, p. 27368.

Bonhomme, S., Lamadon, T., Manresa, E., 2022. Discretizing Unobserved Heterogeneity. Econometrica, forthcoming.

Card, D., Henning, J., Kline, P., 2013. Workplace heterogeneity and the rise of West German wage inequality. Q. J. Econ. 128 (3), 967–1015.

Card, D., Cardoso, A., Heining, J., Kline, P., 2018. Firms and labor market inequality: evidence and some theory. J. Labor Econ. 36 (51), S13–S70.

Carlsen, F., Rattsø, J., Stokke, H., 2016. Education, experience, and urban wage premium. Reg. Sci. Urban Econ. 60, 39–49.

Combes, P.P., Gobillon, L., 2015. The empirics of agglomeration economies. In: Duranton, G., Henderson, V., Strange, W. (Eds.), Handbook of Urban and Regional Economics, vol. V. Elsevier, North-Holland, pp. 247–348.

Combes, P.P., Duranton, G., Gobillon, L., 2008. Spatial wage disparities: sorting matters. J. Urban Econ. 63, 723–742.
Combes, P.P., Duranton, G., Gobillon, L., Roux, S., 2010. Estimating agglomeration economies with history, geology, and worker effects. In: Glaeser, E. (Ed.), Agglomeration Economics. The University of Chicago Press, pp. 15–66.
Dauth, W., Findeisen, S., Moretti, E., Suedekum, J., 2022. Matching in Cities. Journal of the European Economic Association, forthcoming.
De la Roca, J., Puga, D., 2017. Learning by working in big cities. Rev. Econ. Stud. 84, 106–142.
Duranton, G., Jayet, H., 2011. Is the division of labour limited by the extent of the market? Evidence from French cities. J. Urban Econ. 69 (1), 56–71.
Eeckhout, J., Kircher, P., 2011. Identifying sorting – in theory. Rev. Econ. Stud. 78, 872–906.
Figueiredo, O., Guimaraes, P., Woolward, D., 2014. Firm-worker matching in industrial clusters. J. Econ. Geogr. 14, 1–19.
Fredriksson, P., Hensvik, L., Nordström Skans, O., 2018. Mismatch of talent: evidence on match quality, entry wages, and job mobility. Am. Econ. Rev. 108 (11), 3303–3338.
Giannone, E., 2019. Skill-biased Technical Change and Regional Convergence. University of Chicago working paper.
Hirsch, B., König, M., Möller, J., 2013. Is there a gap in the gap? Regional differences in the gender pay gap. Scot. J. Polit. Econ. 60 (4), 412–439.
Hoen, M., 2016. Occupational Crosswalk, Data and Language Requirements. Ragnar Frisch Centre for Economic Research. Working Paper 1/2016.
Kline, P., Saggio, R., Sølvsten, M., 2020. Leave-out estimation of variance components. Econometrica 88, 1859–1898.
Leknes, S., 2015. The more the merrier? Evidence on quality of life and population size using historical mines. Reg. Sci. Urban Econ. 54, 1–17.
Lise, J., Robin, J., 2017. The macrodynamics of sorting between workers and firms. Am. Econ. Rev. 107 (4), 1104–1135.
Lopes de Melo, R., 2018. Firm wage differentials and labor market sorting: reconciling theory and evidence. J. Polit. Econ. 126 (1), 313–346.
Macis, M., Schivardi, F., 2016. Exports and wages: rent sharing, workforce composition, or returns to skills? J. Labor Econ. 34 (4), 945–978.
Melo, P., Graham, D., 2014. Testing for labour pooling as a source of agglomeration economies: evidence for labour markets in England and Wales. Pap. Reg. Sci. 93 (1), 31–52.
Mendes, R., van den Berg, G., Lindeboom, M., 2010. An empirical assessment of assortative matching in the labor market. Lab. Econ. 17, 919–929.
Moen, K., 1978. *Kongsberg Sølvverk 1623-1957*. Kongsberg: Sølvverksmuseets Venner.
Øisang, O., 1942. Rørosboka: Vol. 2. *Røros Kobberverks Historie*, Trondheim. Nidaros boktrykkeri AS.
Phimister, E., 2005. Urban effects on participation and wages: are there gender differences? J. Urban Econ. 58 (3), 513–536.
Rosenthal, S., Strange, W., 2008. The attenuation of human capital spillovers. J. Urban Econ. 64 (2), 373–389.
Steen, S., 1986. Bals gruve: landsdelens eldste gruve. Ottar 3, 12–16.
Thuesen, G., 1979. Den Første Dokumenterte Bergverksdrift I Norge, *Volund*. Norsk Teknisk Museum, pp. 7–60.
Yankow, J., 2009. Some empirical evidence of the efficacy of job matching in urban labor markets. Int. Adv. Econ. Res. 3 (2), 233–244.