

CEP Discussion Paper No 1608

March 2019

The Urban Wage Premium in Imperfect Labour Markets

Boris Hirsch

Elke J. Jahn

Alan Manning

Michael Oberfichtner

Abstract

Using administrative data for West Germany, this paper investigates whether part of the urban wage premium stems from fierce competition in thick labour markets. We first establish that employers possess less wage-setting power in denser markets. Local differences in wage-setting power predict 1.8–2.1% higher wages from a 100 log points increase in population density. We further document that the observed urban wage premium from such an increase drops by 1.5–1.9pp once conditioning on local search frictions. Our results therefore suggest that a substantial part of the urban wage premium roots in differential imperfections across local labour markets

Key words: urban wage premium, imperfect labour markets, monopsony, search frictions
JEL Codes: R23; J42; J31

This paper was produced as part of the Centre's Labour Markets Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

We thank William Boal, Michael Burda, Angus Deaton, Albrecht Glitz, Dan Hamermesh, Peter Kuhn, Ismir Mulalic, Michael Pflüger, Claus Schnabel, Uta Schönberg, Jens Südekum, Gerard van den Berg, Rudolf Winter-Ebmer, and Thomas Zwick for very useful suggestions. We further appreciate comments by participants of the EALE 2016, RES 2016, UEA 2015, VfS 2015, and WEAI 2017 conferences, the Industrial Relations Section, Princeton University, 2018 Monopsony Conference in Sundance, UT, the BGPE 2015 and DFG SPP 1764 workshops, as well as seminar participants in Berlin, Göttingen, Halle, Hamburg, Lüneburg, Mainz, Nuremberg, and Munich. We also gratefully acknowledge financial support from the Fritz Thyssen Foundation.

Boris Hirsch, Leuphana University and IZA. Elke J. Jahn, Institute for Employment Research, University of Bayreuth and IZA. Alan Manning, London School of Economics and Centre for Economic Performance, London School of Economics. Michael Oberfichtner, Institute for Employment Research.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

The urban wage premium in imperfect labour markets^{*}

Boris Hirsch[†], Elke J. Jahn[‡], Alan Manning[§], and Michael Oberfichtner^{**}

Abstract: Using administrative data for West Germany, this paper investigates whether part of the urban wage premium stems from fierce competition in thick labour markets. We first establish that employers possess less wage-setting power in denser markets. Local differences in wage-setting power predict 1.8–2.1% higher wages from a 100 log points increase in population density. We further document that the observed urban wage premium from such an increase drops by 1.5–1.9pp once conditioning on local search frictions. Our results therefore suggest that a substantial part of the urban wage premium roots in differential imperfections across local labour markets.

Keywords: urban wage premium, imperfect labour markets, monopsony, search frictions

JEL classification: R23, J42, J31

^{*} We thank William Boal, Michael Burda, Angus Deaton, Albrecht Glitz, Dan Hamermesh, Peter Kuhn, Ismir Mulalic, Michael Pflüger, Claus Schnabel, Uta Schönberg, Jens Südekum, Gerard van den Berg, Rudolf Winter-Ebmer, and Thomas Zwick for very useful suggestions. We further appreciate comments by participants of the EALE 2016, RES 2016, UEA 2015, VfS 2015, and WEAI 2017 conferences, the Industrial Relations Section, Princeton University, 2018 Monopsony Conference in Sundance, UT, the BGPE 2015 and DFG SPP 1764 workshops, as well as seminar participants in Berlin, Göttingen, Halle, Hamburg, Lüneburg, Mainz, Nuremberg, and Munich. We also gratefully acknowledge financial support from the Fritz Thyssen Foundation.

[†] Boris Hirsch, Leuphana University of Lüneburg and IZA, Scharnhorststr. 1, 21335 Lüneburg, Germany, email: hirsch@leuphana.de

[‡] Elke J. Jahn, Institute for Employment Research, University of Bayreuth, and IZA, Regensburger Str. 104, 90478 Nuremberg, Germany, email: elke.jahn@iab.de

[§] Alan Manning, London School of Economics and Centre for Economic Performance, Houghton Street, London WC2A 2AE, United Kingdom, a.manning@lse.ac.uk

^{**} Michael Oberfichtner, Institute for Employment Research, Regensburger Str. 104, 90478 Nuremberg, Germany, email: michael.oberfichtner2@iab.de

1 Introduction

Following Glaeser and Maré (2001), a large empirical literature has investigated differences in wages across labour markets of different sizes. The general finding of this literature is that a significant urban wage premium exists and that this premium consists both of a level effect that accrues directly upon arrival in an urban labour market and a growth effect that arises as workers gain urban work experience (Heuermann *et al.*, 2010; De la Roca and Puga, 2017). The conventional interpretation of this evidence is that the urban wage premium stems from higher worker productivity in thick labour markets rooted in agglomeration economies (Puga, 2010; Moretti, 2011). The wage level effect reflects a higher level of worker productivity in denser markets, and higher urban wage growth mirrors that worker productivity is also growing at higher pace in thick markets.

This conventional interpretation attributes the urban wage premium entirely to differences in workers' marginal productivity, it implicitly views labour markets as perfectly competitive. In contrast, we argue that part of the urban wage premium is the result of denser urban labour markets being thicker and more competitive than non-urban labour markets. In imperfect labour markets workers receive a share of their marginal product of labour, and the share is higher in urban areas if thick labour markets are more competitive, perhaps because search frictions are lower. If this view is correct, prior estimates of the urban wage premium may exaggerate the part of the urban wage premium that is due to higher worker productivity.

To support this view, we present evidence from German administrative data that workers find it easier to move to better-paying employers in denser labour markets: quit rates are more sensitive to wages in denser markets and the share of hires from non-employment is lower (both commonly used measures of the degree of competition in labour markets—see, for example, Manning, 2003). We also show that the urban wage (growth) premium is considerably lower once we condition on our measures of labour market competitiveness. Consequently, our

findings suggest that a substantial part of the urban wage premium roots in competition effects rather than merely reflecting productivity effects. Our estimates suggest that about half of the urban wage level premium is due to differences in labour market competition and about two thirds of the urban wage growth premium.

The paper is organised as follows: Section 2 briefly reviews the relevant literature and describes our hypotheses. Section 3 describes the data. Section 4 provides a descriptive analysis of urban labour markets, presenting several empirical observations consistent with the view that denser labour markets are more competitive. Section 5 explains our empirical approach and presents results. Section 6 considers issues of robustness, and Section 7 concludes.

2 Review of the literature and theoretical considerations

An increasing body of international evidence has established that workers earn significantly higher wages in urban than in rural labour markets. This urban wage premium has proven to be robust to controlling for unobserved worker heterogeneity by means of fixed-effects techniques (e.g. Glaeser and Maré, 2001; D’Costa and Overman, 2014) and to endogenising workers’ location decision in structural approaches (e.g. Gould, 2007; Baum-Snow and Pavan, 2012). The premium is thus unlikely to reflect mere worker sorting. As a general finding, the literature has documented that the urban wage premium stems both from a wage level and a wage growth effect (see, e.g., the survey by Heuermann *et al.*, 2010). In other words, urban experience–wage profiles have been found to possess both a larger intercept and a larger slope than rural profiles.

The standard explanation for these findings is that agglomeration economies raise the level and growth in worker productivity in thick markets (Duranton and Puga, 2004; Puga, 2010; Moretti, 2011; De la Roca and Puga, 2017). There is good empirical evidence that agglomeration economies exist (see, e.g., the surveys by Rosenthal and Strange, 2004, or Combes and Gobillon, 2015) but the precise mechanisms are less clear e.g. knowledge spill-overs, faster

learning or a more efficient matching process have been proposed. In many papers, the underlying model of the labour market is not made explicit but implicitly would seem to be a perfectly competitive model in which wages are equal to marginal products.

If labour markets are imperfectly competitive and employers possess some wage-setting power over their workers, “wages are ... only proportional and not equal to labour productivity by a factor that depends on the local monopsony power of the firm” (Combes and Gobillon, 2015: 283). And local monopsony power may depend upon market density. If thick labour markets are more competitive, as suggested by Manning (2010) and Hirsch *et al.* (2013), workers in denser markets obtain a larger share of their marginal product, and we might observe an urban wage premium even if agglomeration economies were completely absent.

If employers have some market power this is of interest in its own right but also has the potential to alter our understanding of why agglomeration exists. If labour markets in all areas are perfectly competitive with higher wages in urban than non-urban labour markets, then one can only explain why employers locate in urban areas if there are productivity gains from doing so i.e. if there are agglomeration effects on productivity. If labour markets are imperfectly competitive, agglomeration equilibria may exist without any variation of productivity with location. Manning (2010) shows that an agglomeration equilibrium in which there are no incentives for workers or firms to move areas could result if labour markets are more competitive in agglomerations and firms are heterogeneous in productivity. High productivity firms choose to locate in agglomerations because they want to be large and the wage they would have to pay to be large is greater outside the agglomeration (though these wages are not observed as they do not locate there in equilibrium).

Our estimates imply that differences in competition can explain some but not all of the urban wage premium so it is likely that there are agglomeration economies, though smaller than implied by other estimates. And many of the proposed mechanisms for the urban wage premium are consistent with our model: if there is more efficient matching then it is easier for workers

to change employers which might be expected not just to raise the average quality of a match but to give workers more market power in any particular match. And if labour markets are imperfectly competitive, one would expect high productivity employers to pay higher wages, consistent with the evidence in Dauth *et al.* (2018).

Why should denser local labour markets be more competitive? In the last two decades, a growing literature has investigated the prevalence and causes of imperfect competition in the labour market (for recent surveys, see Ashenfelter *et al.*, 2010, or Manning, 2011). Employer market power derives from search frictions, mobility costs, or job differentiation. All these factors are likely to impede workers' responsiveness to wages causing the labour supply curve to a single firm to be upward-sloping, rather than being horizontal as under perfect competition. In line with this prediction, numerous studies have found that the wage elasticity of the labour supply to the firm is limited (see Manning, 2011), suggesting that employers possess substantial wage-setting power and pay workers only part of the marginal product of labour. What is more, all three sources of employer market power are likely to play less of a role in thick labour markets with many competing employers that we therefore expect to be more competitive.

3 Data

We combine two administrative German data sets for the period 1985–2010: the Integrated Employment Biographies (IEB) and a quarterly version of the Establishment History Panel (BHP), which are both provided by the Institute for Employment Research (IAB). Since the information contained in these data is used to calculate social security contributions, it is highly reliable and especially suited for analyses on wages and job durations.

The data on job durations (at daily frequency), wages, and worker characteristics (education, experience, occupation, and nationality) come from a 5% random sample of the IEB (for details on the IEB, see Jacobebbinghaus and Seth, 2007). The IEB comprises all wage and

salary employees registered with the German social security system, where about 80% of all people employed in Germany are part of the system. Note that the IEB dates back until 1975, so that we have information on workers' employment biographies from 1975 onwards. Note, however, that we will not use pre-1985 wage information in our analysis because of changes in the wage variable, which does not include bonus payments before 1985 but contains these from 1985 onwards. In the following, we will further restrict our sample to workers born no earlier than 1960, i.e. workers who were at maximum 15 years old in 1975 and for whom we thus have complete information on their work experience.

The data on employers come from a quarterly version of the BHP which also consists of data from the German social insurance system aggregated at the level of the plant at the end of each quarter (for details on the BHP, see Spengler, 2008). It contains information on plants' workforce composition, industry, size, and on plant location at the NUTS 3 level. We use this latter information to assign workers and their jobs to 103 local labour markets in West Germany identified by Kosfeld and Werner (2012) based on commuting links (rather than on mere administrative boundaries). Figure 1 depicts these local labour markets and their time-averaged population density (i.e. population per square kilometre) by quintile along with large cities of more than 500,000 inhabitants.

Although our data contain observations for East German workers from 1992 onwards, restricting our analysis to the post-unification period would markedly reduce our period of observation and thus the scope of our investigation. We will thus focus our analysis throughout on workers in West Germany (excluding Berlin) during the period 1985–2010, and we further restrict to males to circumvent selectivity issues regarding female employment and because female and male workers have been shown to differ significantly in their firm-level labour supply elasticities (Hirsch *et al.*, 2010).

To calculate the share of hires from non-employment at the local labour market level, we distinguish employment and non-employment as labour market states. Consequently, a new

job may either start after a job-to-job move has taken place (i.e. the new job is with a plant that has a different plant identifier), or following a previous spell in registered unemployment or no spell in the data at all.¹ The latter either means that before starting the new job the individual has been non-employed without receiving unemployment benefits or, for instance, a self-employed worker who is not included in the data. While our data do not enable us to disaggregate this category of unknown origin, information from other German data sets suggests that the vast majority of workers in this category have indeed started new jobs from non-employment.²

Whereas information on job durations and daily gross wages in the data are highly reliable, the data include no detailed information on the number of hours worked. Moreover, wages are top-coded at the social security contribution ceiling, which affects 7.6% of our observations. To deal with the first drawback, we restrict our analysis to full-time workers. To cope with the second, we exclude jobs with wages above the ceiling (though we will also include imputed wage observations in a check of robustness presented in Section 6). In addition, information on workers' education stems from employers and is for this reason inconsistent or missing for some workers. To alleviate this problem, we impute the missing information on education using a procedure proposed by Fitzenberger *et al.* (2006) that allows inconsistent education information to be corrected. After applying this imputation procedure, we have to drop only 2.0% of jobs due to missing or inconsistent information on education.

The merged data for the period 1985–2010 allow us to set up an inflow sample of 1,782,212 jobs held by the 575,014 workers. Out of these 1,782,212 jobs, 246,401 jobs (or 13.8%) have right-censored job durations. In our sample, the number of jobs varies markedly across the 103 local labour markets, with a minimum of 1,401 and a maximum of 98,977. Note

¹ Note that separations to non-employment are ignored if the worker is recalled by the same plant within three months. Similarly, in classifying job-to-job moves we allow a gap of up to three months between two subsequent employment spells with different plants if no other labour market status, like registered unemployment, is recorded in the data.

² See, e.g., Hirsch *et al.* (2018) for a comparison to the Socio-Economic Panel that includes workers who are not registered with the German social security system.

that we observe multiple jobs within a given labour market for the majority (i.e. 56.1%) of workers. For descriptive statistics on our sample, see Table 1.

When estimating the urban wage premium in the second part of our analysis, we will only use wage observations at the 30th of June of a year yielding a panel of 3,702,677 observations at yearly frequency. Again, the number of observations varies considerably across local markets, with a minimum of 3,313 and a maximum of 181,248. Notwithstanding, there are enough observations in every local labour market as well as enough movers across markets to precisely estimate local wage levels.

4 Descriptive analysis of the urban wage premium

This section provides a descriptive analysis of the urban wage premium and its connection to local labour market competition. Based on the 103 local labour markets in our administrative employer–employee data for West Germany over the period 1985–2010, Figure 2 plots workers’ average local log wage against the local time-average of log population density. The resulting regression line has a slope of 0.034, so that an increase in population density by 100 log points is associated with 3.4% higher wages on average.³ As we will see later in Section 5, the density gradient of wages drops slightly when controlling for observable worker, employer, and region characteristics as well as for worker fixed effects.

Apart from higher wage levels, previous research has documented that experience–wage profiles are steeper in denser local labour markets. In particular, De la Roca and Puga (2017) find that additional work experience in denser labour markets leads to a significant urban wage growth premium that adds to the static gains from working in a thick market. To get an impression of this urban wage growth premium in our data, we first fit an extended Mincerian wage

³ Note that the standard deviation in log population density across local labour markets is 0.7, meaning that a rise in density by 100 log points is a reasonable point of departure.

equation controlling for standard worker characteristics and for worker–region fixed effects that includes region-specific coefficients of real work experience and its square.⁴ These coefficients provide us with estimates of experience–wage profiles that are specific to the local labour markets in our data. Since the wage equation comprises worker–region fixed effects, the identifying wage variation stems from workers gaining work experience within a local labour market and thus informs us on workers’ average local wage growth during their careers.⁵

We then regress the region-specific coefficients of experience and its square on log population density and use these estimates to predict the accumulated log wage growth over a worker’s career that is associated with gaining work experience in a 100 log points denser labour market. As is clear from the solid line in Figure 3 that shows this accumulated urban wage growth premium, gaining work experience in denser labour markets is associated with marked additional wage growth. The accumulated urban wage growth premium from entering a 100 log points denser labour market and gaining 20 years of work experience in this market amounts to 32.6 log points.

To scrutinise whether this urban wage growth premium reflects faster within-job wage growth in thick labour markets, we redo our analysis fitting a wage equation that includes job fixed effects rather than region–worker fixed effects and region-specific coefficients of the worker’s job tenure and its square rather than real work experience. Because of the job fixed effects, the identifying wage variation now stems from workers gaining tenure with a specific employer. The dashed line in Figure 3 shows the additional accumulated within-job log wage growth in a 100 log points denser labour market. Within-job wage growth contributes to the

⁴ More specifically, in the wage equation we control for the worker’s education, tenure, one-digit occupation, and nationality, as well as for year dummies.

⁵ Note that the worker–region fixed effects further control for the worker’s previous (time-invariant) work experience gained in other local labour markets. Hence, our approach is very similar in spirit to De la Roca and Puga’s (2017), as is our finding that work experience gained in denser labour markets gives rise to more pronounced wage growth. One possible concern, however, is that workers who repeatedly move between the same regions and who therefore gain work experience within a local labour market at different points of time in their careers may blur our estimates. To rule this out, we redid our analysis for stayers, who do not change regions at all. Reassuringly, this had no impact on our findings.

urban wage growth premium but seems to play less of a role than wage gains from changing jobs.

To get an impression by how much wage gains from job changes vary with labour market density, in a next step we calculate the difference in log wages between two consecutive jobs within a region for every worker and regress this difference on a full set of region dummies controlling for time and worker fixed effects. Then, we regress the coefficients of the region dummies, which reflect a worker's average wage gain between two consecutive jobs, on log population density. As is seen from Model II in Table 2, a rise in population density by 100 log points is associated with on average 0.8 log points larger wage gains from job changes, which is a sizeable effect given that the mean wage gain amounts to 3.5 log points. As Table 2 further makes clear, this positive relationship between wage gains from job changes and density is solely driven by wage gains from direct job-to-job moves. In contrast, wage gains from changing jobs are close to zero and unrelated to labour market density if there is an intervening period of non-employment.⁶

The higher wage gains from job-to-job moves in denser labour markets suggest that workers' on-the-job search is more effective in thick markets. If on-the-job search frictions are lower in denser labour markets, we expect employed workers to receive more job offers and to see more job-to-job moves in thick markets. To check this expectation, Figure 4 plots the average number of job transitions of workers, who stay in the same local labour market, within their first 15 years of real work experience against log population density. As Figure 4 makes clear, the overall number of job transitions does not vary much with population density. This finding, however, masks that the number of job transitions into employment rises with density whereas the number of transitions into non-employment falls.⁷ In other words, in thick labour markets

⁶ These results are rather different from those reported for the US NLSY in Baum-Snow and Pavan (2012).

⁷ Note that we obtain the same pattern for workers' overall job separation rate as well as their separation rates to employment and non-employment when fitting stratified Cox models that control for observable worker, employer, and region characteristics as well as for permanent worker unobservables. Results are available upon request.

workers' inter-employer mobility seems to be larger, suggestive of lower on-the-job search frictions and fiercer competition among employers.

To capture the extent of inter-employer mobility, Manning (2003: 44–49) suggested using the share of hires from non-employment (as opposed to employment).⁸ Intuitively, the higher is this share, the less often are employers hiring workers faced with direct competition with other employers. Figure 5 plots the share of hires from non-employment in local labour markets against log population density and clearly shows that in denser markets new hires less often come from non-employment. This suggests that in thick markets workers find it easier to flee low-paying jobs through job-to-job moves, which erodes employers' wage-setting power. In line with this expectation, plotting average local log wages against the share of hires from non-employment in Figure 6 reveals a tight negative relationship: the regression line has a slope of -0.697 , so that a rise in the share of hires from non-employment by one standard deviation, which amounts to 0.042 across local labour markets, is associated with 3.0 log points higher average local wages. This pattern suggests that part of the urban wage premium may indeed reflect lower on-the-job search frictions and thus fiercer competition among employers in denser labour markets.

This suggestion is further substantiated by the fact that the descriptive urban wage premium falls considerably when conditioning on local search frictions measured by the local share of hires from non-employment. Regressing, respectively, average local log wages and log population density on the share and plotting the wage residuals against the density residuals in Figure 7 more than halves the slope of the regression line *vis-à-vis* Figure 2, which now is 0.015. Hence, when conditioning on search frictions the descriptive urban wage premium from a 100 log points rise in population density just amounts to 1.5%. The corresponding drop in the

⁸ As Manning shows, the share of hires from non-employment has a one-to-one correspondence to the extent of on-the-job search frictions in the Burdett and Mortensen (1998) model and is also likely to be a good proxy for employers' wage-setting power in various other models of imperfect labour markets.

slope from 0.034 in Figure 2 to 0.015 in Figure 7 by 1.9pp represents about half of the descriptive urban wage premium, suggesting that part of the urban wage premium reflects fiercer competition in thick labour markets that are characterised by less on-the-job search frictions.

These stylized facts suggest that it is plausible to think of denser labour markets as being more competitive. The next section provides an assessment of how much more.

5 Econometric approach and results

Section 5.1 outlines our approach to estimating employer market power and Section 5.2 to estimating the urban wage premium.

5.1 Estimating employers' wage-setting power in local labour markets

The first part of our empirical analysis estimates differences in the wage elasticity of the labour supply to a single firm across local labour markets.

To identify the wage elasticity of the labour supply to a single firm, we use the estimation approach by Manning (2003: 96–104) building on search frictions as the source of labour market imperfections. Consider a firm paying some wage w at some point in time.⁹ The change in the labour supply to this firm $L(w)$ can be written as:

$$\dot{L}(w) = R(w) - s(w)L(w), \quad (1)$$

where $R(w) > 0$ denotes the number of recruits arriving at the firm at that point in time with $R' > 0$ and $0 < s(w) < 1$ denotes the separation rate of incumbent workers with $s' < 0$. We thus assume that the firm can increase its labour supply by increasing its wage and that the labour supply adjusts sluggishly over time.

Now consider a steady state with $\dot{L}(w) = 0$. Then, using equation (1) we arrive at

⁹ This assumption of employer wage setting is in line with existing evidence for Germany documenting that wage posting is the predominant form of wage formation (see Brenzel *et al.*, 2014).

$$L(w) = R(w)/s(w) \quad (2)$$

with $L' > 0$.¹⁰ From equation (2) we get the labour supply elasticity to the firm ε_{Lw} as the difference of the wage elasticity of recruitment ε_{Rw} and the wage elasticity of the separation rate ε_{sw}

$$\varepsilon_{Lw} = \varepsilon_{Rw} - \varepsilon_{sw}. \quad (3)$$

Using equation (3) to identify the supply elasticity, however, would require us to estimate the recruitment elasticity ε_{Rw} , which is a hard task given that one typically does not know the firm's recruitment pool.¹¹

To deal with this problem, Manning (2003: 97; 2011) notes that many search models (for example, Burdett and Mortensen, 1998; Bhaskar and To, 1999) imply that the recruitment elasticity is minus the separation elasticity i.e. $\varepsilon_{Rw} = -\varepsilon_{sw}$, so that the labour supply elasticity becomes:

$$\varepsilon_{Lw} = -2\varepsilon_{sw}. \quad (4)$$

Intuitively, this result holds because one firm's wage-related hire is another firm's wage-related quit. Hence, equation (4) allows us to identify the labour supply elasticity to the firm by just estimating the wage elasticity of incumbent workers' job separation rate.¹²

¹⁰ Note that perfect competition is nested as the case with $L' \rightarrow \infty$, i.e. a horizontal labour supply curve to the firm, due to $s' \rightarrow -\infty$ and $R' \rightarrow \infty$ at the competitive wage that equalises supply and demand at the level of the labour market.

¹¹ One rare exception is Falch (2017), who is able to analyse data on employers' recruitment pools for certified teachers in Norway and finds substantial monopsony power in this labour market segment. Reassuringly, his estimates of the firm-level labour supply elasticity are of the same magnitude as in an earlier study (Falch, 2011) that uses the same data but rests on the approach based on workers' job separation rate that we will use, thereby validating the finding of studies lacking such data on employers' recruitment pools.

¹² Some previous studies, e.g. Booth and Katic (2011) and Hirsch *et al.* (2018), applied a more sophisticated estimation approach distinguishing employment and non-employment as distinct labour market states. While our data include information on workers' previous and subsequent labour market states, distinguishing transitions from and to employment from those from and to non-employment is not viable in our application because of the limited number of jobs observed in sparsely populated local labour markets.

To estimate this elasticity, we use a two-step procedure similar in spirit to the approaches by Hirsch and Schumacher (2005), Combes *et al.* (2008), and De la Roca and Puga (2017). In the first step, we will fit individual-level separation equations controlling for several worker and employer characteristics to obtain estimates of the separation elasticity at the local labour market level. In the second step, we will regress these local elasticity estimates on log population density and other local labour market characteristics to assess whether firms' wage-setting power is less pronounced in denser labour markets.

To obtain an estimate of the wage elasticity of incumbent workers' job separation rate for every local labour market, we fit in the first step a stratified Cox model for the separation rate of job m held by worker i at employer j in region r

$$s_m \left(\tau \middle| \log w_m(\tau), \mathbf{x}_i(\tau), \mathbf{z}_j(\tau) \right) = s_{0ir}(\tau) \exp(\theta_r \log w_m(\tau) + \mathbf{x}_i(\tau)' \boldsymbol{\beta} + \mathbf{z}_j(\tau)' \boldsymbol{\gamma}), \quad (5)$$

where τ is the job duration, $\log w_m(\tau)$ is the log wage, $\mathbf{x}_i(\tau)$ is a vector of worker characteristics, $\mathbf{z}_j(\tau)$ is a vector of employer characteristics, $s_{0ir}(\tau)$ is a worker–region-specific baseline hazard, and we treat all covariates as time-varying. In the separation equation (5), the region-specific coefficient of the log wage θ_r provides us with an estimate of the local separation rate elasticity. Furthermore, the baseline hazard $s_{0ir}(\tau)$ in the equation is some arbitrary worker–region-specific function of job duration and thus encompasses permanent unobservables at both the level of the worker and the level of the region.¹³ Controlling for worker unobservables is indispensable in our application because worker sorting on unobservables may simultaneously influence workers' wages, their location, and their job mobility. Furthermore, controlling for region unobservables in the separation equation addresses concerns that quitting for the same wage is not comparable across local labour markets because of regional price or wage level

¹³ Note that by allowing for a worker–region-specific baseline hazard the proportionality assumption inherent to the class of hazard rate models defined by equation (5) needs to hold only for jobs held by the same worker within a particular local labour market, but may well be violated across workers or regions without invalidating identification (see Kalbfleisch and Prentice, 2002: 118/119).

differences, as permanent price and wage level differences are part of the baseline hazard and are thus accounted for.

To estimate the separation equation (5), we adopt the stratified partial likelihood estimator (see Ridder and Tunalı, 1999). This estimator allows us to sweep out the baseline hazard without estimating it directly, similar to the within estimator in linear fixed-effects models. Hence, estimating stratified Cox models with worker–region-specific baseline hazards is viable, and we are able to precisely identify local separation rate elasticities—the θ_r ’s—in the first-step separation equation (5) with our data.

The stratified partial likelihood estimator is identified from within-variation at the worker–region level, e.g. from wage variation occurring in multiple jobs held by the same worker within the same local labour market. In the stratified Cox model, we thus control for workers’ wage relative to the outside offer available to him. Worker controls consist of real experience (linearly and squared) as well as groups of dummies for education (distinguishing low-skilled, medium-skilled, and high-skilled workers¹⁴), one-digit occupation, and non-German nationality. As employer controls we include the shares of part-time, high-skilled, low-skilled, female, and non-German workers among the plant’s workforce as well as groups of dummies for plant size and two-digit industry. We finally add a full set of time dummies. These first-step estimates are available upon request but we do not report them in detail because the coefficients of interest—the region-specific coefficients of the log wage in the separation equation—are hard to summarize.

In the second step, we regress the estimated labour supply elasticity to the firm $\hat{\varepsilon}_{LW,r} = -2\hat{\theta}_r$ on the centred time-average of log population density $\log popdens_r$ and other local labour market characteristics \mathbf{c}_r .

¹⁴ Low-skilled workers are workers with neither a vocational nor an academic degree, while medium-skilled workers possess a vocational degree and high-skilled workers have an academic degree.

$$\hat{\varepsilon}_{Lw,r} = \zeta_0 + \zeta_1 \log popdens_r + \mathbf{c}_r' \boldsymbol{\psi} + v_r \quad (6)$$

with v_r denoting the error term. In this second-step regression (6), we expect ζ_1 to have a positive sign, thereby indicating that employers possess less wage-setting power in denser labour markets with more elastic firm-level labour supply.¹⁵ Our standard errors come from a block bootstrap at the worker level with 400 replications. Other controls are the log employment share of the largest industry in the local labour market (a measure of market specialisation), the log Herfindahl index in industries' local employment levels (a measure of diversification), the share of low- and high-skilled workers among the active working population (to measure skill levels). We present estimates either with or without employer controls in the first-step separation equation because different employer characteristics may themselves root in the agglomeration economies that give rise to regional productivity and wage differences.¹⁶ As the log population density is centred around its mean, the estimated regression constant represents the average elasticity estimate across local labour markets.

Table 3 presents the main results of the second-step regression. In Model I, the average elasticity amounts to 2.43, which is well within the range of previous estimates summarised by Manning (2011). This number implies that employers possess substantial, though not implausibly large wage-setting power over their workers. In line with our expectations, the labour supply elasticity to the firm is significantly larger in denser labour markets. A 100 log points increase in population density comes along with a rise in the elasticity by 0.19 to 2.62.

The positive relationship between the elasticity and density shrinks somewhat when controlling for employer characteristics in Model II. With employer controls in the first-step

¹⁵ Note that our results do not hinge on using the time-average of log population density as agglomeration measure in the second-step regression. We will discuss alternative agglomeration measures in Section 6.

¹⁶ For instance, Manning (2010) shows that larger plant sizes in denser markets, which have been documented to explain part of the urban wage premium in Germany (Lehmer and Möller, 2010), are at odds with canonical models of agglomeration economies, which would predict the opposite to hold. Yet, he also demonstrates that larger plant sizes in denser markets may stem from fiercer competition in these labour markets.

separation equation, the average elasticity amounts to 2.22 and a 100 log points rise in population density is associated with an increase in the elasticity by 0.15 to 2.37.

One concern with our estimates is that the first-step stratified Cox models while controlling for worker and employer observables as well as permanent worker unobservables may still suffer from bias stemming from employer unobservables. As a case in point, compensating wage differentials may result in higher wages for high-turnover employers, thereby contaminating our estimates of the local firm-level labour supply elasticity. Controlling for both permanent worker and employer unobservables by means of a stratified Cox regression, however, is not viable as this would base identification on multiple jobs held by the same worker at the same plant.

In order to alleviate concerns, we estimate the separation equation (5), with another control variable which is the plant wage effect from a two-way fixed effects decomposition of individual wages for our data conducted by Card *et al.* (2013) that builds on Abowd *et al.*'s (1999) methodology. In Abowd *et al.*'s framework, the plant wage effect represents the wage premium enjoyed by every worker employed at a plant and thus comprises all wage components stemming from permanent employer unobservables. As shown by Sorkin (2018), compensating wage differentials account for more than half of the variance in plant wage effects and thus the plant wage effects may serve well as a proxy variable enabling us to control for the non-pecuniary attractiveness of employers.

Card *et al.*'s (2013) plant wage effects are only available for the years 1985–2009 and are missing for some plants in our sample (for details, see Card *et al.*, 2015). Hence, we fit Model III that controls for plant wage effects in the first-step stratified Cox regression on a reduced sample of jobs at plants for which Card *et al.* (2015) provide plant wage effects and where we disregard jobs starting in 2010 and treat jobs ending in 2010 as right-censored. Remarkably, we obtain almost unchanged results when controlling for the plant wage effect. The average local labour supply elasticity amounts to 2.38, and a rise in population density by 100

log points is associated with an increase in the elasticity by 0.15 to 2.53.

These results are about the impact of population density on the wage elasticity of separations. To convert this to an impact, through competition, on wages we use the fact that in monopsonistic labour markets the relationship between wages and marginal products will be given by:

$$w_r = \frac{\varepsilon_{LW,r}}{1 + \varepsilon_{LW,r}} \phi_r, \quad (7)$$

where ϕ_r denotes the marginal product of labour in region r . We next conduct a thought experiment and ask what the urban wage premium would be if we assumed away agglomeration economies that yield different marginal products across regions. Hence, we set $\phi_r \equiv \phi$ in equation (7). Thus, the predicted wage gap across any two local labour markets 1 and 2 gets

$$\frac{w_2 - w_1}{w_1} = \frac{\varepsilon_{LW,2} - \varepsilon_{LW,1}}{\varepsilon_{LW,2}(\varepsilon_{LW,1} + 1)}. \quad (8)$$

Based on the estimated ζ 's from the second-step regression (6) and setting $\varepsilon_{LW,1}$ to the average elasticity across local labour markets, we can calculate the predicted urban wage premium from differential local labour market competition (8) and confront it with estimates of the actual premium.

Our separation elasticity estimates predict an urban wage premium from fiercer competition in thick labour markets of 1.8–2.1% for a 100 log point difference. This prediction comes very close to the drop in the descriptive premium by 1.9pp when conditioning on local search frictions that we found in Section 4. Our results thus suggest that a substantial part of the urban wage premium reflects differences in labour market competition. To estimate how much, we next present estimates of the urban wage premium that condition on worker and employer characteristics as well as on permanent worker unobservables and that thus account for worker sorting on these factors.

5.2 Estimating the urban wage premium

In the second part of our analysis, we compare the predicted urban wage premium due to differential labour market competition from equation (8) to the reduction in the estimated premium that occurs when conditioning on the extent of search frictions in local labour markets measured by the share of hires from non-employment. If these two numbers were of similar magnitude, this would suggest that this part of the urban wage premium reflects fiercer competition in denser labour markets.¹⁷ To estimate the urban wage premium, we will again adopt a two-step procedure. In the first step, we run individual-level wage regressions controlling for several worker and employer characteristics to obtain estimates of local wage levels. In the second step, we regress these wage levels on log population density, the share of hires from non-employment, and other local labour market characteristics to get estimates of the urban wage premium.

To be more precise, the first step consists of running extended Mincerian wage regressions at the level of the individual worker

$$\log w_{ijrt} = \delta_r + \alpha_i + \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + u_{ijrt}, \quad (9)$$

where notation follows the same rules as before, δ_r is a region fixed effect, α_i is a worker fixed effect, and u_{ijrt} is an error term.¹⁸ Our main point of interest in the wage equation (9) are the δ_r 's which provide us with estimates of average local wage levels after controlling for observable worker and employer characteristics and permanent worker unobservables. As made clear by previous studies, such as Glaeser and Maré (2001), Yankow (2006), and De la Roca and

¹⁷ Note that our results do not hinge on using the share of hires from non-employment as measure for workers' on-the-job search frictions. In Section 6, we will demonstrate that we obtain the same results when using an alternative measure of search frictions proposed by van den Berg and van Vuuren (2010).

¹⁸ Note that we do not correct workers' wages for differences in local labour markets' price levels because we are interested in the part of the urban wage premium that reflects workers' marginal productivity rather than differences in local price levels. As stressed by Heuermann *et al.* (2010: 752), "[t]he fundamental point in the debate on whether to use nominal or real wages is that, while spatial differences in nominal wages can be interpreted as productivity differences, regional differences in real wages reflect differences in workers' utility rooted in urban amenities." See Glaeser and Gottlieb (2009), Moretti (2011), and Combes and Gobillon (2015) for similar assessments.

Puga (2017), it is important to include worker fixed effects in the wage equation to tackle the selection bias that would result if workers with higher abilities chose to live in denser labour markets.¹⁹

In the second step, we regress the estimated δ_r 's obtained from the wage regression (9) on the centred time-average of local log population density and other labour market characteristics

$$\hat{\delta}_r = \pi_0 + \pi_1 \log popdens_r + \mathbf{c}_r' \boldsymbol{\eta} + e_r, \quad (10)$$

where e_r denotes an error term and π_1 provides us with an estimate of the urban wage premium. Next, we add our measure of search frictions, the centred time-average of the local share of hires from non-employment $share_{nemp}_r$, to the model

$$\hat{\delta}_r = \tilde{\pi}_0 + \tilde{\pi}_1 \log popdens_r + \tilde{\pi}_2 share_{nemp}_r + \mathbf{c}_r' \tilde{\boldsymbol{\eta}} + e_r. \quad (11)$$

In the second-step regression (11), we expect a negative sign for $\tilde{\pi}_2$ —as in the descriptive analysis in Figure 6—because a higher share of hires from non-employment indicates more search frictions that raise employers' wage-setting power. We will then compare the estimated π_1 from regression (10) and the estimated $\tilde{\pi}_1$ from regression (11) and interpret $\pi_1 - \tilde{\pi}_1$ as an estimate of the part of the urban wage premium reflecting fiercer competition in denser labour markets. Again, we base our inference on standard errors coming from a block bootstrap at the worker level with 400 replications.

We will show results obtained from estimating the first-step wage equation (9) either

¹⁹ Including worker fixed effects, however, means that the identification of local wage levels rests on workers who switch locations, and clearly switching locations may itself be endogenous. Hence, estimated regional wage levels may suffer from bias if worker unobservables and location changes are not orthogonal as is implicitly assumed when applying the fixed-effects approach. While instrumenting workers' location has proven difficult due to the lack of credible, strong instruments (Heuermann *et al.*, 2010) and has, in general, also made no big difference (Melo *et al.*, 2009), another approach chosen in previous studies is to model worker mobility explicitly in a structural setting (Gould, 2007; Baum-Snow and Pavan, 2012). This structural approach, though, comes at the cost of strong functional assumptions and of excluding worker fixed effects from the wage equations.

with or without worker fixed effects. In the wage equation, we include the same worker and employer characteristics as in the separation equation in the previous section and add a group of tenure dummies on top of these. In the second-step regression, we additionally control for the local unemployment rate to avoid that the share of hires from non-employment picks up local differences in unemployment rather than on-the-job search frictions. As before, we will present estimates either without or with employer controls in the first-step wage equation and, for the sake of brevity, we will just show the main results of the second-step regressions (with first-step results being available upon request).

Table 4 summarises our main findings. Panel A presents the second-step regression (10) of local wage levels on log population density and other local labour market characteristics for various specifications of the first-step wage equation (9). When just controlling for observed worker characteristics in the individual-level wage regression (Model I) we arrive at a coefficient of log density of 0.032 that is very similar to the descriptive estimate of 0.034 reported in Section 3. Hence, a 100 log points rise in population density is associated with a rise in local wages by 3.2%. When additionally controlling for employer characteristics in Model II, this number rises somewhat to 3.6%.

One concern with these estimates is that workers in local labour markets of different density may differ in unobservables that affect their marginal productivity and their wages. To account for permanent worker unobservables, we next include worker fixed effects in the first-step wage regressions. In these specifications, identification rests on workers moving across local labour markets. Estimating the first-step regression with worker fixed effects reduces the estimated density coefficient somewhat. In Model III (IV) without (with) employer controls, a 100 log points increase in population density now comes along with a 3.0% (2.8%) increase in wages.

Panel B in Table 3 shows the second-step regression (11) of Models I–IV when adding the local share of hires from non-employment. In line with our expectation and the descriptive

evidence from Figure 6, this measure of workers' on-the-job search frictions has a significantly negative impact on local wages in all specifications. In our preferred Models III and IV, in which the first-step regression includes worker fixed effects, a one standard deviation rise in the share of hires from non-employment, which amounts to 0.042 across local labour markets, is associated with a drop in wages by 1.6–2.0%.

As in the descriptive analysis in Figure 7, conditioning on local search frictions in the second-step regression markedly reduces the estimated urban wage premium, i.e. the coefficient of log population density, by 1.5–2.2 log points, depending on specification. In our preferred Models III and IV, the drop amounts to 1.5–1.9 log points. A 100 log points rise in population density is now only associated with a 1.2–1.4% rise in wages, rather than the 2.8–3.0% found previously when not conditioning on local search frictions (see Panel A). We consider this fall in the urban wage premium by 1.5–1.9pp as a benchmark estimate of the part of the premium that reflects fiercer competition in thick labour markets. Remarkably, this drop is of the same magnitude as the predicted urban wage premium from differential competition across local labour markets from the previous subsection, which amounted to 1.8–2.1%.

To gain further insight into the role of search frictions on the urban wage premium, it is instructive to have a closer look at the static wage gains that accrue directly upon arrival in a denser labour market as well as the additional wage growth over a worker's career there. To estimate the static urban wage premium, we redo our analysis estimating the first-step wage equation with the first-difference rather than the within estimator. Regressing the resulting region fixed effects on log population density while controlling for other local labour market characteristics informs us on the immediate wage premium for workers moving to denser labour markets.

As is seen from Panel A of Table 5, a 100 log points rise in population density is associated with an immediate static urban wage premium of 2.2% (2.4%) when excluding (including) employer controls in the first-step wage equation. Note that these numbers are smaller in

magnitude than in the specifications using the within estimator and thus point at dynamic wage gains over workers' career that add to the static gains from moving to thick labour markets. Once conditioning on search frictions measured by the share of hires from non-employment in Panel B, the static premium drops by 1.4–1.7pp suggesting that this part of the static wage gains reflect fiercer competition in denser labour markets.

To obtain estimates of the urban wage growth premium and its change when conditioning on local search frictions, we re-estimate the first-step wage equation including region-specific coefficients of real experience and its square as well as worker–region fixed effects (as in the descriptive analysis in Section 4). By including worker–region fixed effects (rather than worker fixed effects), we base identification on variation in wages that stems from workers gaining work experience within a local labour market, which is our point of interest. In the second step, we regress the coefficients of experience and its square on the centred time-average of log population density and the other labour market characteristics. Therefore, the estimated regression constants inform us on the average coefficients of experience and its square across local labour markets while the coefficients of log population density inform us on how they vary with density.

As before, Panel A in Table 6 presents the main results of the second-step regression of the local coefficients of experience and its square on log population density and the other labour market characteristics. The density coefficient for the linear experience component is significantly positive and of the same size no matter whether we exclude employer controls in the first-step wage equation (Model I) or include them (Model II). Hence, workers entering denser labour markets have higher wage gains from work experience. In other words, there exists an urban wage growth premium. Further, as the density coefficient for the quadratic experience component is negative (though not statistically significantly so), wage–experience profiles are more concave in denser labour markets.

As an illustration of these findings, Figures 8 and 9 plot the predicted accumulated log

wage growth over a worker's career in a labour market with 100 log points higher population density when, respectively, excluding or including employer controls in the first-step wage equation. For comparison to the descriptive analysis in Section 4, the solid lines show the accumulated urban wage growth premium when not controlling for other labour market characteristics than population density in the second-step regressions. As is clear from the long-dashed lines in Figures 8 and 9, there still exists a substantial urban wage growth premium when conditioning on further labour market characteristics. The accumulated urban wage growth from entering a 100 log points denser labour market and gaining 20 years of work experience there amounts to about 18 log points no matter whether we include or exclude employer controls in the first-step wage equation.

When we condition on search frictions by adding the share of hires from non-employment to the second-step regression, the density coefficients of both the linear and the quadratic experience component fall. As is seen from the short-dashed lines in Figures 8 and 9, this results in a marked drop in the predicted accumulated urban wage growth premium that nonetheless remains non-trivial, with about 6 log points after 20 years of work experience.

Taken together, our findings suggest that the urban wage growth premium stems from two sources. In line with the previous literature, part of workers' dynamic wage gains in denser markets seems to stem from higher wage growth that is most pronounced at the beginning of workers' careers and thus is likely to reflect an acceleration in workers' human capital acquisition due to learning effects. On the other hand, a substantial part of the urban wage growth premium seems to mirror faster search capital growth in more competitive, thick labour markets.

6 Robustness checks

To scrutinise our results further, we perform several checks of robustness along three dimensions. First, we repeat our analysis using different measures of agglomeration in the second-step regressions and, second, including imputed wages for top-coded wage observations. Third, we re-estimate the drop in the observed urban wage premium when conditioning on an alternative measure of local search frictions. Tables 7 and 8 present the key results from these checks when, respectively, excluding or including employer controls in the first-step wage equation and underscore the robustness of our findings.

In the first group of robustness checks, we explore how our results change when using alternative measures of agglomeration than the local time-average of log population density, which we used in our baseline specification. Using the local log population density in 1985, i.e. at the beginning of our period of observation, or in 2010, i.e. at the end of the observational window, rather than its time-average has little impact on our findings. Neither the density gradient of the wage elasticity of the labour supply to the firm, nor the predicted urban wage premium due to differential local labour market competition, nor the drop in the observed urban wage premium when conditioning on local search frictions change in any substantial way. The same holds when including log population and log size as separate regressors in the second-step regression or when using log employment density (i.e. log employment per square kilometre) rather than log population density as agglomeration measure.

In our second check of robustness, we repeat our analysis including top-coded wage observations which we impute using a heteroscedastic single imputation approach developed by Büttner and Rässler (2008) for our data. We do so because top coding occurs at the contribution limit to the German social security system that is the same for all workers and thus independent of job location. As a consequence, top coding has a stronger bite in denser labour markets with higher wage levels, which may arouse some concerns. As Tables 7 and 8 make

clear, our findings do not seem to suffer from this differential bite in top coding across local labour markets and are virtually the same when including imputed wage observations.

In a final group of robustness checks, we re-estimate the drop in the urban wage premium when conditioning on local search frictions using an alternative measure of these suggested by van den Berg and van Vuuren (2010), viz. the local share of job exits into non-employment (as opposed to employment). Like the share of hires from non-employment used in our baseline specification, the share of job exits into non-employment captures how hard it is for workers to move their way up in the local wage distribution by job-to-job moves. As the last column of Tables 7 and 8 makes clear, the drop in the urban wage premium when conditioning on this alternative measure of local search frictions is almost the same as in our baseline specification. What is more, our results based on this alternative measure keep robust when using alternative agglomeration indicators and when including imputed wage observations.

7 Conclusions

Using administrative linked employer–employee data for West Germany comprising the years 1985–2010, we have presented evidence that part of the urban wage premium stems from fiercer competition in thick local labour markets. In the first part of our analysis, we documented that the wage elasticity of the labour supply to the firm, which governs what part of the marginal product of labour accrues to workers in monopsonistic labour markets with employer wage setting, is significantly larger in denser markets. While the average elasticity across local labour markets amounted to 2.22–2.43, depending on specification, an increase in population density by 100 log points came along with an increase in the elasticity by 0.15–0.19. Based on a thought experiment that abstracts from agglomeration economies that cause productivity differences across local markets, our estimates predict workers’ wages to rise by 1.8–2.1%.

In the second part of our analysis, we found that a 100 log points increase in population

density is associated with 2.8–3.0% higher wages when controlling for worker fixed effects and several worker, employer, and local labour market characteristics. However, once we conditioned on the extent of search frictions in local labour markets measured by the share of hires from non-employment, the urban wage premium dropped considerably by 1.5–1.8pp. Remarkably, these numbers are of the same magnitude as the predicted urban wage premium from differential competition in local labour markets obtained in the first part of our analysis.

Thus, our findings are in line with the notion that a substantial, though not all, part of the urban wage premium derives from fiercer competition in thick labour markets. Our results therefore suggest that workers in denser labour markets not only obtain higher wages because worker productivity is greater and grows at higher pace there but because they also receive a larger share of the marginal product of labour.

References

- Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999), “High wage workers and high wage firms”, *Econometrica*, **67**(2): 251–333
- Ashenfelter, O. C., Farber, H., and Ransom, M. R. (2010), “Labor market monopsony”, *Journal of Labor Economics*, **28**(2): 203–210
- Baum-Snow, N. and Pavan, R. (2012), “Understanding the city size wage gap”, *Review of Economic Studies*, **79**(1): 88–127
- Bhaskar, V. and To, T. (1999), “Minimum wages for Ronald McDonald monopsonies: A theory of monopsonistic competition”, *Economic Journal*, **109**(455): 190–203
- Booth, A. L. and Katic, P. (2011), “Estimating the wage elasticity of labour supply to a firm: What evidence is there for monopsony?”, *Economic Record*, **87**(278): 359–369
- Brenzel, H., Gartner, H., and Schnabel C. (2014), “Wage bargaining or wage posting? Evidence from the employers’ side”, *Labour Economics*, **29**(1): 41–48
- Burdett, K. and Mortensen, D. T. (1998), “Wage differentials, employer size, and unemployment”, *International Economic Review*, **39**(2): 257–273
- Büttner, T. and Rässler, S. (2008), Multiple imputation of right-censored wages in the German IAB Employment Sample considering heteroscedasticity, IAB Discussion Paper No. 44/2008, Nuremberg: Institute for Employment Research
- Card, D., Heining, J., and Kline, P. (2013), “Workplace heterogeneity and the rise of West German wage inequality”, *Quarterly Journal of Economics*, **128**(3): 967–1015
- Card, D., Heining, J., and Kline, P. (2015), CHK effects: Version 2, FDZ-Methodenreport 06/2015, Nuremberg: Institute for Employment Research
- Combes, P.-P., Duranton, G., and Gobillon, L. (2008), “Spatial wage disparities: Sorting matters!”, *Journal of Urban Economics*, **63**(2): 723–742

- Combes, P.-P. and Gobillon, L. (2015), “The empirics of agglomeration economics”, in G. Duranton, J. V. Henderson, and W. C. Strange (eds.), *Handbook of Regional and Urban Economics*, vol. 5A, pp. 247–348, Amsterdam: Elsevier
- Dauth, W., Findeisen, S., Moretti, E., and Suedekum, J. (2018), Matching in cities, NBER Working Paper No. 25277, Cambridge, MA: National Bureau of Economic Research
- D’Costa, S. and Overman, H. G. (2014), “The urban wage growth premium: Sorting or learning?”, *Regional Science and Urban Economics*, **48**(1): 168–179
- De la Roca, J. and Puga, D. (2017), “Learning by working in big cities”, *Review of Economic Studies*, **84**(1): 106–142
- Duranton, G. and Puga, D. (2004), “Micro-foundations of urban agglomeration economies”, in J. V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4, pp. 2063–2117, Amsterdam: Elsevier
- Falch, T. (2011), “Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment”, *American Economic Review (Papers and Proceedings)*, **101**(3): 460–465
- Falch, T. (2017), “Wages and recruitment: Evidence from external wage changes”, *Industrial and Labor Relations Review*, **70**(2): 483–518
- Fitzenberger, B., Osikominu, A., and Völter, R. (2006), “Imputation rules to improve the education variable in the IAB employment subsample”, *Schmollers Jahrbuch*, **126**(3): 405–436
- Glaeser, E. L. and Gottlieb, J. D. (2009), “The wealth of cities: Agglomeration economies and spatial equilibrium in the United States”, *Journal of Economic Literature*, **47**(4): 983–1028
- Glaeser, E. L. and Maré, D. C. (2001), “Cities and skills”, *Journal of Labor Economics*, **19**(2): 316–342
- Gould, E. D. (2007), “Cities, workers, and wages: A structural analysis of the urban wage premium”, *Review of Economic Studies*, **74**(2): 477–506
- Heuermann, D., Halfdanarson, B., and Suedekum, J. (2010), “Human capital externalities and the urban wage premium: Two literatures and their interrelations”, *Urban Studies*, **47**(4): 749–767
- Hirsch, B., Jahn, E. J., and Schnabel, C. (2018), “Do employers have more monopsony power in slack labor markets?”, *Industrial and Labor Relations Review*, **71**(3): 676–704
- Hirsch, B., König, M., and Möller, J. (2013), “Is there a gap in the gap? Regional differences in the gender pay gap”, *Scottish Journal of Political Economy*, **60**(4): 412–439
- Hirsch, B., Schank, T., and Schnabel, C. (2010), “Differences in labor supply to monopsonistic firms and the gender pay gap: An empirical analysis using linked employer-employee data from Germany”, *Journal of Labor Economics*, **28**(2): 291–330
- Hirsch, B. T. and Schumacher, E. J. (2005), “Classic or new monopsony? Searching for evidence in nursing labor markets”, *Journal of Health Economics*, **24**(5): 969–989
- Jacobebbinghaus, P. and Seth, S. (2007), “The German Integrated Employment Biographies Sample IEBS”, *Schmollers Jahrbuch*, **127**(2): 335–342
- Kalbfleisch, J. D. and Prentice, R. L. (2002), *The Statistical Analysis of Failure Time Data*, 2nd ed., Hoboken, NJ: John Wiley & Sons
- Kosfeld, R. and Werner, A. (2012), “Deutsche Arbeitsmarkregionen – Neuabgrenzung nach den Kreisgebietsreformen 2007–2011”, *Raumforschung und Raumordnung*, **70**(1): 49–64
- Lehmer, F. and Möller, J. (2010), “Interrelations between the urban wage premium and firm-size wage differentials: A microdata cohort analysis for Germany”, *Annals of Regional Science*, **45**(1): 31–53
- Manning, A. (2003), *Monopsony in Motion: Imperfect Competition in Labor Markets*, Princeton, NJ: Princeton University Press

- Manning, A. (2010), “The plant size–place effect: Agglomeration and monopsony in labour markets”, *Journal of Economic Geography*, **10**(5): 717–744
- Manning, A. (2011), “Imperfect competition in the labor market”, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4B, pp. 973–1041, Amsterdam: Elsevier
- Melo, P. C., Graham, D. J., and Noland, R. B. (2009), “A meta-analysis of estimates of urban agglomeration economics”, *Regional Science and Urban Economics*, **39**(3): 332–342
- Moretti, E. (2011), “Local labor markets”, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 4B, pp. 1237–1313, Amsterdam: Elsevier
- Puga, D. (2010), “The magnitude and causes of agglomeration economies”, *Journal of Regional Science*, **50**(1): 203–219
- Ridder, G. and Tunali, İ (1999), “Stratified partial likelihood estimation”, *Journal of Econometrics*, **92**(2): 193–232
- Rosenthal, S. S. and Strange, W. C. (2004), “Evidence on the nature and sources of agglomeration economies”, in J. V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4, pp. 2119–2171, Amsterdam: Elsevier
- Sorkin, I. (2018), “Ranking firms using revealed preference”, *Quarterly Journal of Economics*, **133**(3): 1331–1393
- Spengler, A. (2008), “The Establishment History Panel”, *Schmollers Jahrbuch*, **128**(3): 501–509
- van den Berg, G. J. and van Vuuren, A. (2010), “The effect of search frictions on wages”, *Labour Economics*, **17**(6): 875–885
- Yankow, J. J. (2006), “Why do cities pay more? An empirical examination of some competing theories of the urban wage premium”, *Journal of Urban Economics*, **60**(2): 139–161

Figures

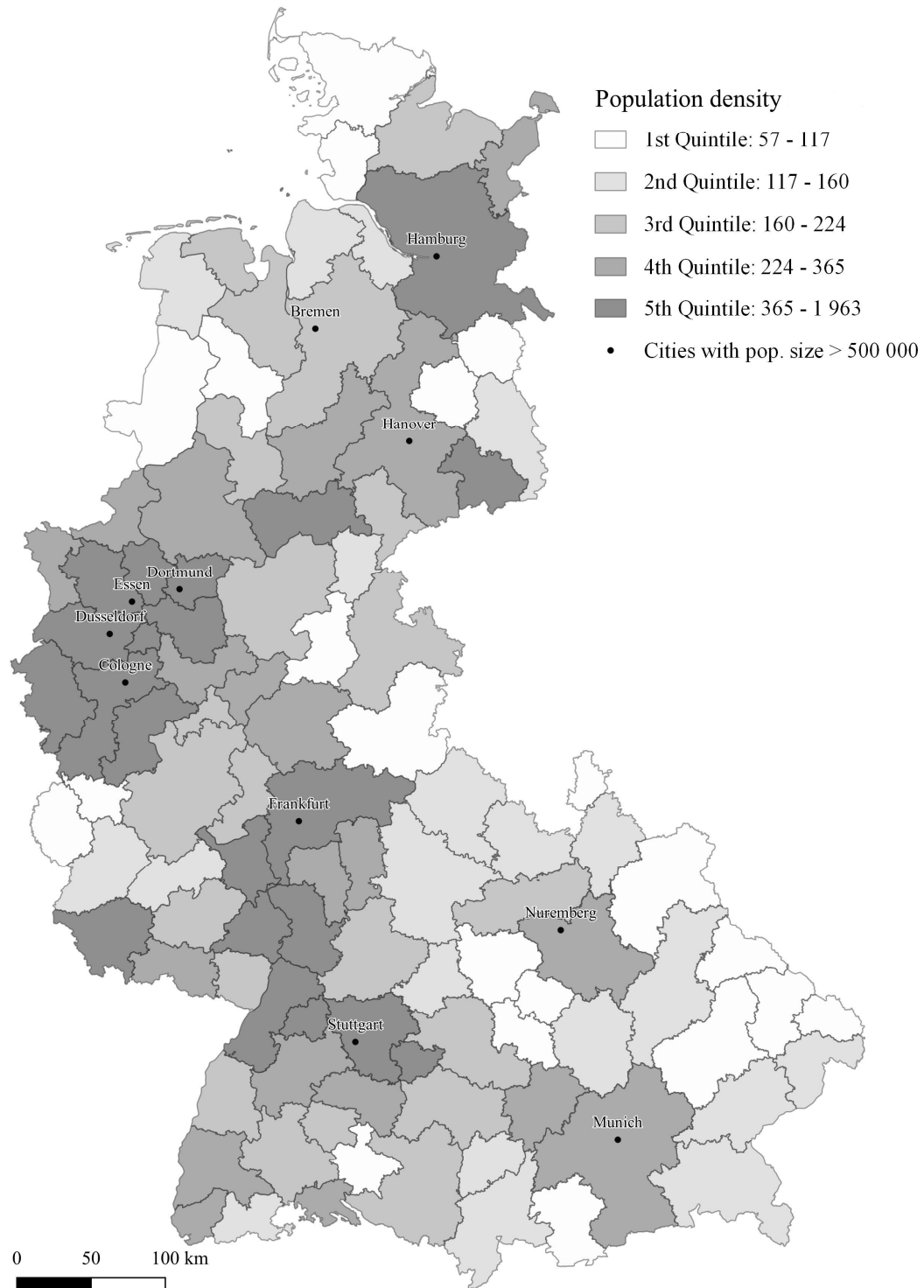


Figure 1: Local labour markets in West Germany and average population density by quintile

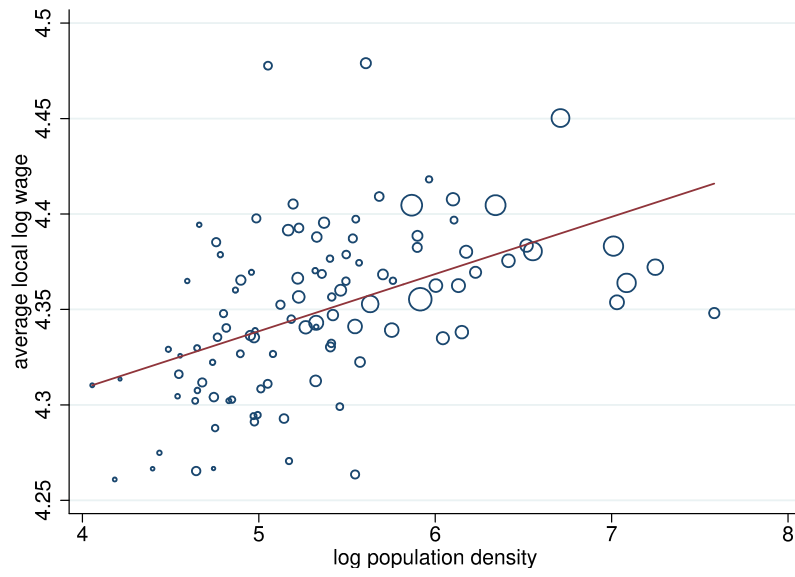


Figure 2: Local average wages and log population density (markers weighted by population size)

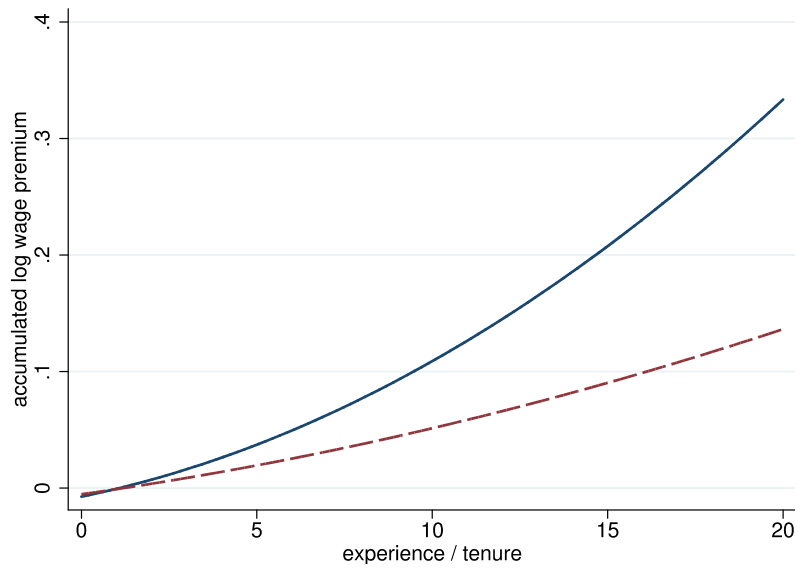


Figure 3: Accumulated additional log wage growth in a 100 log points denser local labour market over workers' real work experience (solid) and tenure (dashed), respectively

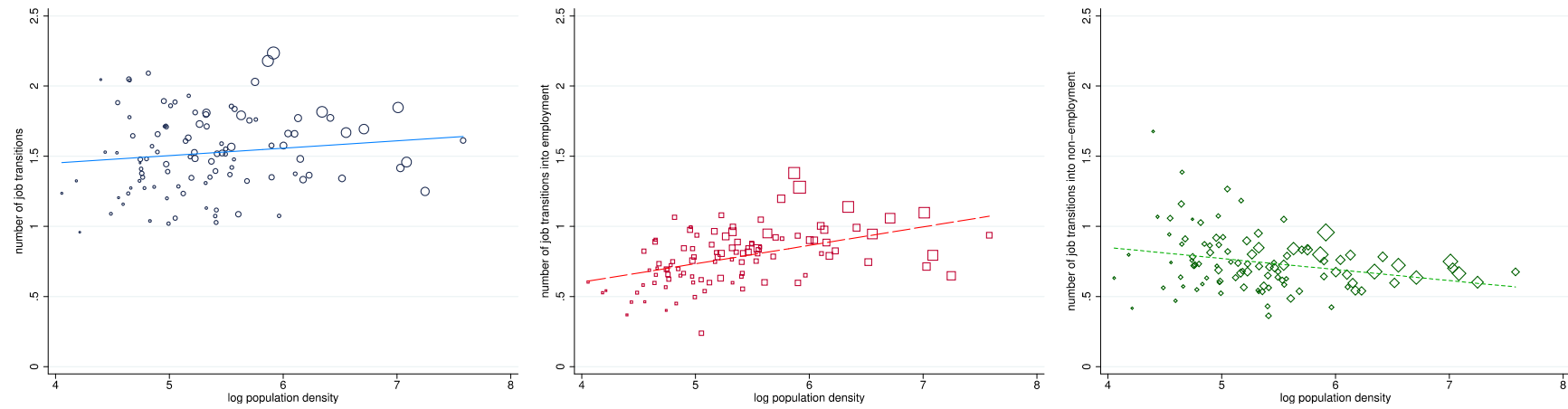


Figure 4: Average number of job transitions of workers, who stay in the same local labour market, within the first 15 years of work experience and log population density (markers weighted by population size)

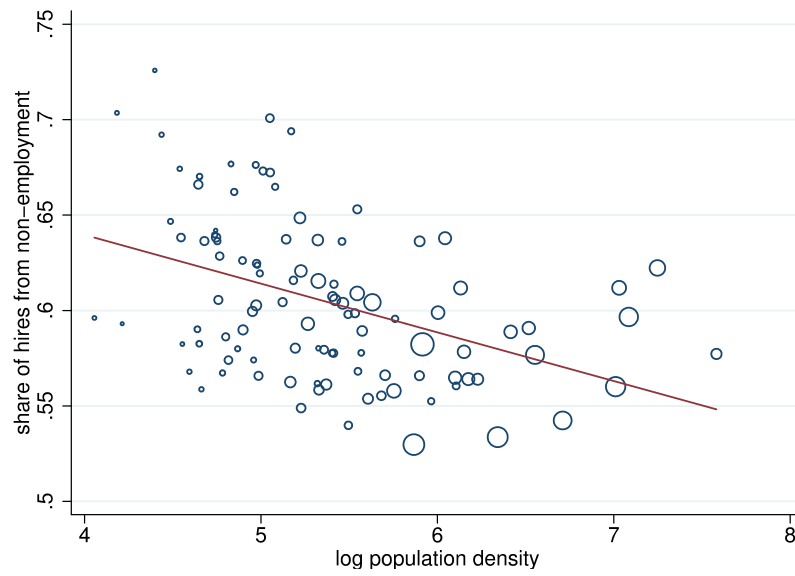


Figure 5: Local share of hires from non-employment and log population density (markers weighted by population size)

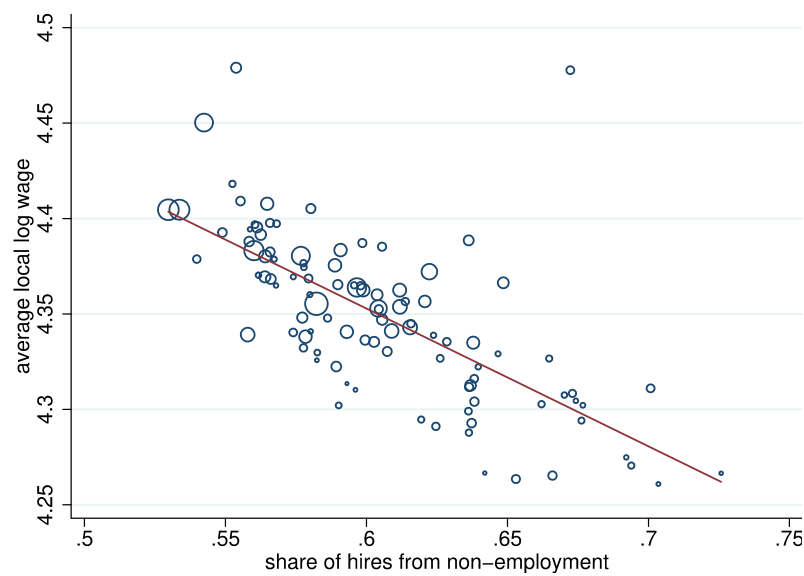


Figure 6: Local average wages and share of hires from non-employment (markers weighted by population size)

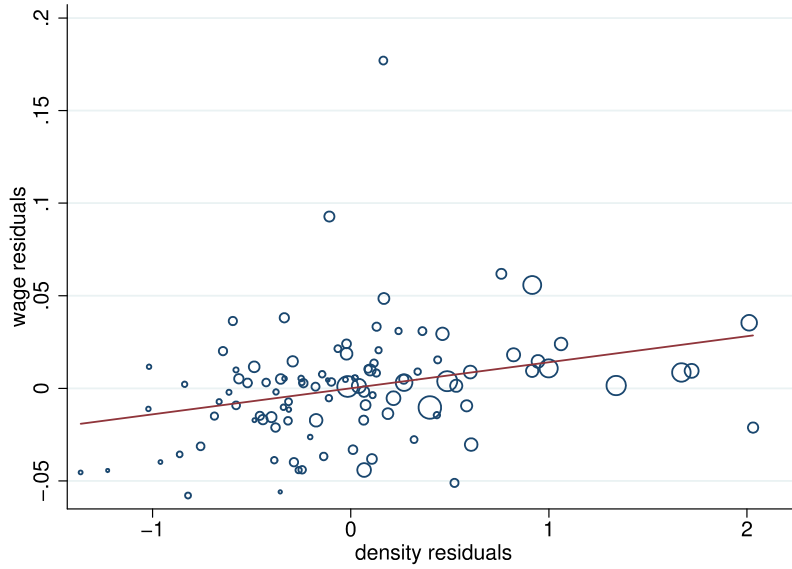


Figure 7: Local average wages and log population density when conditioning on the share of hires from non-employment (markers weighted by population size)

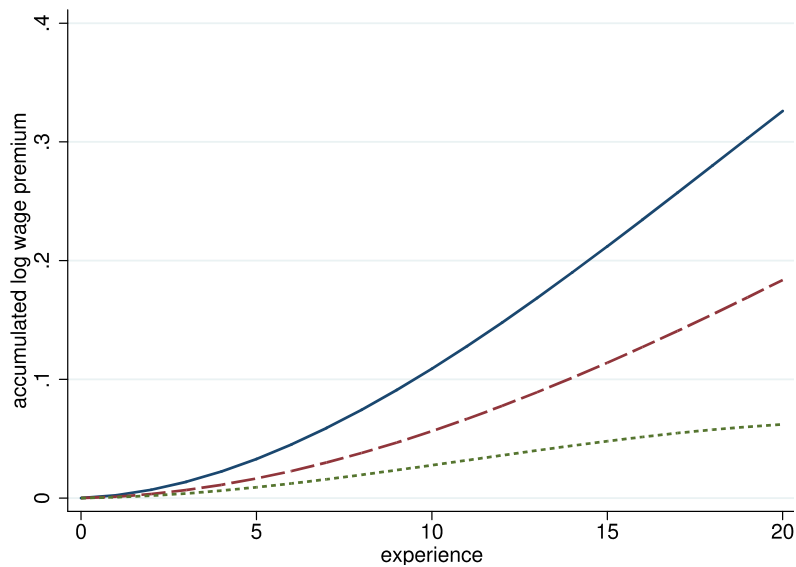


Figure 8: Accumulated additional log wage growth in a 100 log points denser local labour market over workers' real work experience when conditioning on worker characteristics (solid) and additionally on local labour market characteristics (long-dashed) and the share of hires from non-employment (short-dashed)

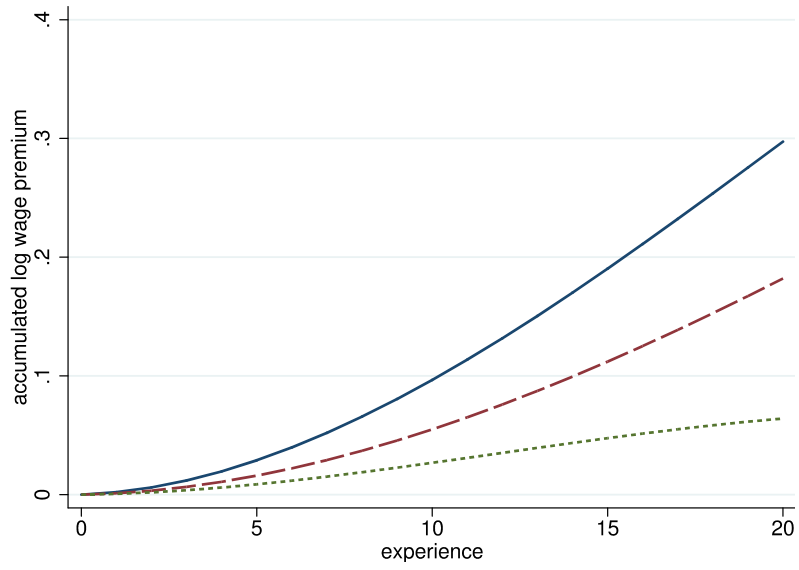


Figure 9: Accumulated additional log wage growth in a 100 log points denser local labour market over workers' real work experience when conditioning on worker and employer characteristics (solid) and additionally on local labour market characteristics (long-dashed) and the share of hires from non-employment (short-dashed)

Tables

Table 1: Descriptive statistics (means)

Log gross daily wage	4.366
Immigrant (dummy)	0.142
Low-skilled (dummy)	0.129
Medium-skilled (dummy)	0.796
High-skilled (dummy)	0.076
Experience (years)	9.506
Tenure (years)	3.534
Plant size below 11 (dummy)	0.156
Plant size 11–50 (dummy)	0.251
Plant size 51–200 (dummy)	0.244
Plant size 201–1000 (dummy)	0.213
Plant size above 1000 (dummy)	0.136
Share of low-skilled workers	0.201
Share of medium-skilled workers	0.613
Share of high-skilled workers	0.059
Share of female workers	0.169
Share of foreign workers	0.098
Share of part-time workers	0.117
Observations	17,010,740

Notes: IEB and BHP, 1985–2010.

Table 2: Wage changes from job changes within regions and log population density

First-step specification	Model I	Model II
	OLS	Worker fixed effects
Second-step results (103 local labour markets)		
Panel A: All transitions		
Log population density	0.0066	0.0083
<i>(Mean log wage change: 0.0351)</i>	<i>(0.0008)</i>	<i>(0.0042)</i>
Panel B: Direct job-to-job moves		
Log population density	0.0057	0.0125
<i>(Mean log wage change: 0.0697)</i>	<i>(0.0013)</i>	<i>(0.0081)</i>
Panel C: Job transitions via non-employment		
Log population density	0.0018	0.0015
<i>(Mean log wage change: 0.0042)</i>	<i>(0.0012)</i>	<i>(0.0074)</i>

Notes: IEB and BHP, 1985–2010. The dependent variable is the region fixed effect from regressing the change in log wages between two consecutive jobs held by a worker within a region on a full set of region and time dummies. Standard errors come from a block bootstrap at worker level with 400 replications.

Table 3: Local differences in the wage elasticity of the labour supply to the firm

First-step specification	Model I	Model II	Model III
	Stratified Cox model with worker but without employer controls	Stratified Cox model with worker and employer controls	Stratified Cox model with worker and employer controls and plant wage effects
Second-step results (103 local labour markets)			
Log population density	0.1879 (0.0352)	0.1486 (0.0351)	0.1490 (0.0386)
Constant	2.4277 (0.0187)	2.2201 (0.0193)	2.3771 (0.0247)
Predicted urban wage premium	2.1%	2.0%	1.8%

Notes: IEB and BHP, 1985–2010. Estimates show the second-step regression (6). The urban wage premium from a 100 log points increase in population density is predicted based on equation (8) that abstracts from agglomeration economies, with $\varepsilon_{LW,1}$ set to the average elasticity across local markets. The dependent variable is the estimated wage elasticity of the labour supply to the firm obtained from the first-step separation equation (5), which we model as a stratified Cox model with a worker–region-specific baseline hazard. Further region controls are the shares of low-skilled and high-skilled workers, the log employment share of the largest two-digit industry, and the log Herfindahl index of employment at industry level, where all second-step regressors are centred around their means. In the stratified Cox regression, worker controls consist of real experience (linearly and squared) as well as groups of dummies for education, one-digit occupation, and non-German nationality. Employer controls are the shares of part-time, high-skilled, low-skilled, female, and non-German workers among the plant’s workforce as well as groups of dummies for plant size and two-digit industry. We also add time dummies. In Model III, we further include the plant wage effect from Card *et al.* (2015) interacted with its reference period. Standard errors come from a block bootstrap at worker level with 400 replications.

Table 4: Estimated urban wage premium

First-step specification	Model I	Model II	Model III	Model IV
	OLS with worker	OLS with worker and	FE with worker	FE with worker and
Second-step results (103 local labour markets)	controls	employer controls	controls	employer controls
Panel A: Estimates of the urban wage premium w/o conditioning on local search frictions				
Log population density	0.0316	0.0360	0.0304	0.0283
	(0.0012)	(0.0011)	(0.0022)	(0.0021)
Panel B: Estimates of the urban wage premium w/ conditioning on local search frictions				
Log population density	0.0097	0.0147	0.0119	0.0138
	(0.0014)	(0.0013)	(0.0024)	(0.0023)
Share of hires from non-employment	-0.5556	-0.5391	-0.4664	-0.3673
	(0.0294)	(0.0265)	(0.0520)	(0.0487)

Notes: IEB and BHP, 1985–2010. Panel A shows estimates for the second-step regression (10) and Panel B for the second-step regression (11). The dependent variable is the local wage level obtained from the first-step wage regression (9). Further region controls are the shares of low-skilled and high-skilled workers, the log employment share of the largest two-digit industry, the log Herfindahl index of employment at industry level, and the unemployment rate, where all second-step regressors are centred around their means. In the first-step wage equation, worker controls consist of real experience (linearly and squared) as well as groups of dummies for education, tenure, one-digit occupation, and non-German nationality. Employer controls are the shares of part-time, high-skilled, low-skilled, female, and non-German workers among the plant's workforce as well as groups of dummies for plant size and two-digit industry. We also add year dummies. Standard errors come from a block bootstrap at worker level with 400 replications.

Table 5: Estimated static urban wage premium

First-step specification	Model I	Model II
Second-step results (103 local labour markets)	FD with worker controls	FD with worker and employer controls
Panel A: Estimates of the static urban wage premium w/o conditioning on local search frictions		
Log population density	0.0224 (0.0025)	0.0239 (0.0023)
Panel B: Estimates of the static urban wage premium w/ conditioning on local search frictions		
Log population density	0.0052 (0.0028)	0.0098 (0.0027)
Share of hires from non-employment	-0.4345 (0.0573)	-0.3571 (0.0540)

Notes: IEB and BHP, 1985–2010. Panel A shows estimates for the second-step regression (10) and Panel B for the second-step regression (11). The dependent variable is the local wage level obtained from the first-step wage regression (9) in first differences. Further region controls are the shares of low-skilled and high-skilled workers, the log employment share of the largest two-digit industry, the log Herfindahl index of employment at industry level, and the unemployment rate, where all second-step regressors are centred around their means. In the first-step wage equation, worker controls consist of real experience (linearly and squared) as well as groups of dummies for education, tenure, one-digit occupation, and non-German nationality. Employer controls are the shares of part-time, high-skilled, low-skilled, female, and non-German workers among the plant's workforce as well as groups of dummies for plant size and two-digit industry. We also add year dummies. Standard errors come from a block bootstrap at worker level with 400 replications.

Table 6: Local differences in experience–wage profiles

Component of wage profile	First-step specification		Model I		Model II	
			FE with worker controls		FE with worker and employer controls	
			Linear	Quadratic ($\times 100$)	Linear	Quadratic ($\times 100$)
Second-step results (103 local labour markets)						
Panel A: Estimates of the urban wage growth premium w/o conditioning on local search frictions						
Log population density			0.0012 (0.0004)	-0.0023 (0.0015)	0.0011 (0.0004)	-0.0020 (0.0014)
Constant			0.0465 (0.0004)	-0.0616 (0.0008)	0.0396 (0.0004)	-0.0561 (0.0008)
Panel B: Estimates of the urban wage growth premium w/ conditioning on local search frictions						
Log population density			0.0007 (0.0004)	-0.0031 (0.0017)	0.0007 (0.0004)	-0.0027 (0.0016)
Share of hires from non-employment			-0.0117 (0.0092)	-0.0219 (0.0355)	-0.0117 (0.0090)	-0.0188 (0.0343)
Constant			0.0465 (0.0004)	-0.0616 (0.0008)	0.0396 (0.0004)	-0.0561 (0.0008)

Notes: IEB and BHP, 1985–2010. The dependent variables are the region-specific coefficients of real experience and its square (times 100), respectively, obtained from a first-step wage regression analogous to (9) including worker–region fixed effects. Panel A shows estimates for the coefficient-specific second-step regression (10) and Panel B for the coefficient-specific second-step regression (11). Further region controls are the shares of low-skilled and high-skilled workers, the log employment share of the largest two-digit industry, the log Herfindahl index of employment at industry level, and the unemployment rate, where all second-step regressors are centred around their means. In the first-step wage equation, worker controls consist of real experience (linearly and squared) as well as groups of dummies for education, tenure, one-digit occupation, and non-German nationality. Employer controls are the shares of part-time, high-skilled, low-skilled, female, and non-German workers among the plant’s workforce as well as groups of dummies for plant size and two-digit industry. We also add year dummies. Standard errors come from a block bootstrap at worker level with 400 replications.

Table 7: Checks of robustness without employer characteristics in the first-step regressions

Estimate	Coefficient of the log of the agglomeration measure in the second-step regression for the labour supply elasticity	Predicted urban wage premium from a 100 log points increase in the agglomeration measure	Drop in the observed urban wage premium when conditioning on the local share of hires from non-employment	Drop in the observed urban wage premium when conditioning on the local share of job exits into non-employment
Check of robustness				
Baseline	0.1879 (0.0352)	2.1%	1.8pp	1.6pp
Alternative measures of agglomeration				
Log population density in 1985	0.1672 (0.0336)	1.9%	1.8pp	1.6pp
Log population density in 2010	0.1950 (0.0371)	2.2%	1.8pp	1.6pp
Log population (controlling for log size separately)	0.1940 (0.0496)	2.2%	1.9pp	1.7pp
Log employment density	0.2085 (0.0376)	2.4%	1.7pp	1.6pp
Including imputed wage observations	0.1861 (0.0356)	2.3%	1.9pp	1.6pp

Notes: IEB and BHP, 1985–2010. The first column shows the coefficient of the agglomeration measure in the second-step regression (6), where the first-step separation equation includes worker controls but no employer controls and a worker–region-specific baseline hazard—as in Model I in Table 3. The second column presents the predicted urban wage premium from a 100 log points increase in the respective agglomeration measure based on equation (8), with $\varepsilon_{LW,1}$ set to the average elasticity. The third column gives the drop in the estimated urban wage premium when conditioning on the share of hires from non-employment, i.e. by moving from the second-step regression (10) to (11), where the first-step wage equation includes worker controls and fixed effects but no employer controls—as in Model III in Table 4. The last column re-estimates the drop in the urban wage premium from the third column using the ratio of job exits into non-employment to job exits into employment as an alternative measure of local search frictions. Standard errors come from a block bootstrap at worker level with 400 replications.

Table 8: Checks of robustness with employer characteristics in the first-step regressions

Check of robustness	Estimate	Coefficient of the log of the agglomeration measure in the second-step regression for the labour supply elasticity	Predicted urban wage premium from an 100 log points increase in the agglomeration measure	Drop in the observed urban wage premium when conditioning on the local share of hires from non-employment	Drop in the observed urban wage premium when conditioning on the local share of job exits into non-employment
Baseline		0.1486 (0.0351)	2.0%	1.5pp	1.3pp
Alternative measures of agglomeration					
Log population density in 1985		0.1325 (0.0335)	1.8%	1.5pp	1.3pp
Log population density in 2010		0.1530 (0.0370)	2.0%	1.4pp	1.2pp
Log population (controlling for log size separately)		0.1397 (0.0494)	1.9%	1.5pp	1.3pp
Log employment density		0.1612 (0.0375)	2.1%	1.4pp	1.3pp
Including imputed wage observations		0.1481 (0.0357)	2.2%	1.5pp	1.3pp

Notes: IEB and BHP, 1985–2010. The first column shows the coefficient of the agglomeration measure in the second-step regression (6), where the first-step separation equation includes worker controls, employer controls and a worker–region-specific baseline hazard—as in Model II in Table 3. The second column presents the predicted urban wage premium from a 100 log points increase in the respective agglomeration measure based on equation (8), with $\varepsilon_{Lw,1}$ set to the average elasticity. The third column gives the drop in the estimated urban wage premium when conditioning on the share of hires from non-employment, i.e. by moving from the second-step regression (10) to (11), where the first-step wage equation includes worker controls and fixed effects as well as employer controls—as in Model IV in Table 4. The last column re-estimates the drop in the urban wage premium from the third column using the ratio of job exits into non-employment to job exits into employment as an alternative measure of local search frictions. Standard errors come from a block bootstrap at worker level with 400 replications.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1607	Max Nathan Anna Rosso	Innovative Events
1606	Christopher T. Stanton Catherine Thomas	Missing Trade in Tasks: Employer Outsourcing in the Gig Economy
1605	Jan-Emmanuel De Neve Christian Krekel George Ward	Employee Wellbeing, Productivity and Firm Performance
1604	Stephen Gibbons Cong Peng Cheng Keat Tang	Valuing the Environmental Benefits of Canals Using House Prices
1603	Mary Amiti Stephen J. Redding David Weinstein	The Impact of the 2018 Trade War on U.S. Prices and Welfare
1602	Greer Gosnell Ralf Martin Mirabelle Muûls Quentin Coutellier Goran Strbac Mingyang Sun Simon Tindermans	Making Smart Meters Smarter the Smart Way
1601	Antoine Dechezleprêtre Caterina Gennaioli Ralf Martin Mirabelle Muûls Thomas Stoerk	Searching for Carbon Leaks in Multinational Companies
1600	Jeremiah Dittmar Skipper Seabold	New Media and Competition: Printing and Europe's Transformation after Gutenberg
1599	Kilian Huber Volker Lindenthal Fabian Waldinger	Discrimination, Managers, and Firm Performance: Evidence from “Aryanizations” in Nazi Germany

1598	Julia Cajal-Grossi Rocco Macchiavello Guillermo Noguera	International Buyers' Sourcing and Suppliers' Markups in Bangladeshi Garments
1597	Alex Bell Raj Chetty Xavier Jaravel Neviana Petkova John Van Reenen	Do Tax Cuts Produce More Einsteins? The Impact of Financial Incentives vs. Exposure to Innovation on the Supply of Inventors
1596	Matthew Baird A.V. Chari Shanthi Nataraj Alexander Rothenberg Shqiponja Telhaj L. Alan Winters	The Public Sector and the Misallocation of Labor: Evidence from a Policy Experiment in India
1595	Stephen J. Redding David E. Weinstein	Aggregation and the Gravity Equation
1594	Andy Feng Anna Valero	Skill Based Management: Evidence from Manufacturing Firms
1593	Leonardo Baccini Giammario Impullitti Edmund J. Malesky	Globalization and State Capitalism: Assessing Vietnam's Accession to the WTO
1592	Andrew Bernard Emmanuel Dhyne Kalina Manova Glenn Magerman Andreas Moxnes	The Origins of Firm Heterogeneity: A Production Network Approach
1591	Brian Bell Jack Blundell Stephen Machin	The Changing Geography of Intergenerational Mobility

The Centre for Economic Performance Publications Unit

Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk

Website: <http://cep.lse.ac.uk> Twitter: @CEP_LSE