

The Urban Wage Premium in Imperfect Labor 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 greater competition in denser labor markets. We show that employers possess less wage-setting power in denser markets. We further document that an important part of the observed urban wage premia can be explained by greater competition in denser labor markets.

Boris Hirsch is Professor of Economics at Leuphana University of Lüneburg and an affiliate of the IWH and IZA. Elke J. Jahn is Professor of Economics at Bayreuth University, Distinguished Researcher at the Institute for Employment Research, and an affiliate of the IZA. Alan Manning is Professor of Economics at the London School of Economics and Director of the Community Programme at the Centre for Economic Performance at the LSE. Michael Oberfichtner is Researcher at the Institute for Employment Research. The authors thank William Boal, Michael Burda, David Card, Angus Deaton, Albrecht Glitz, Peter Haller, Dan Hamermesh, Peter Kuhn, Ismir Mulalic, Michael Pflüger, Claus Schnabel, Uta Schönberg, Jens Südekum, Gerard van den Berg, Rudolf Winter-Ebmer, Thomas Zwick, and two referees for very useful suggestions and 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 by seminar participants in Berlin, Cologne, Göttingen, Halle, Hamburg, Lüneburg, Mainz, Nuremberg, and Munich.

Disclosure statement: This research was supported by the Fritz Thyssen Foundation. The authors use confidential administrative data that are not publicly available due to German data protection laws. Data access for replication is possible through the Research Data Centre of the Institute for Employment Research; see <https://iab.de/en/daten/replikationen.aspx> for further information. The authors are willing to assist (Michael Oberfichtner, michael.oberfichtner@iab.de).

JEL Classification: R23, J42, and J31

I. Introduction

Following Glaeser and Maré (2001), a large empirical literature has investigated differences in wages across labor 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 and a growth effect that arises as workers gain urban work experience (Heuermann, Halfdanarson, and Suedekum 2010; De la Roca and Puga 2017). The conventional interpretation of this evidence is that the urban wage premia result from higher worker productivity in thick labor 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 that worker productivity is also growing faster in thick markets.

This paper presents evidence that part of the urban wage premium is the result of denser labor markets being more competitive. In imperfect labor markets, workers receive a share of their marginal product of labor, and the share is higher in urban areas if denser labor markets are more competitive. If this view is correct, prior estimates of the urban wage premium may exaggerate the part caused by higher worker productivity.

To support this view, we present evidence from German administrative data that separation rates are more sensitive to wages in denser markets, a commonly used measure of the degree of competition in labor markets—see, for example, Manning (2003). We also show that the urban wage and wage growth premium is considerably lower once we condition on our measures of labor market competitiveness. Our estimates suggest that the urban wage level premium falls by about 40 percent when controlling for differences in labor market competition and the cumulated urban wage growth premium after 20 years of experience is about halved.

The paper is organized as follows: Section II describes the data. Section III describes and

estimates our measure of the competitiveness of labor markets and shows that denser markets are more competitive. Section IV describes our estimates of the urban wage and wage growth premium and investigates their relationship with labor market competition. Section V considers issues of robustness, and Section VI concludes.

II. 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 percent 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, in total about 80 percent of employment. 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. We restrict our sample to workers for whom we have complete information on work experience, which implies they were born after 1960.

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 labor 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 labor markets and their time-averaged population density (that is population per square kilometer) by quintile along with large cities of more than 500,000 inhabitants. We use time-averaged population density as our main measure of how dense is a labor market though Section V shows that results are robust to this because population densities are very stable over time—the correlation between population density at the start and end of our sample period is 0.996.

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 therefore focus our analysis throughout on workers in West Germany (excluding Berlin) during the period 1985–2010, and we further restrict our analysis 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 labor supply elasticities (Hirsch, Schank, and Schnabel 2010).

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 percent 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 V). 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, Osikominu, and Völter (2006) that allows inconsistent education information to be corrected. After applying this imputation procedure, we have to drop only 2.0 percent of jobs due to missing or inconsistent information on education.

In part of the paper, we use the share of hires from nonemployment at the local labor market level which requires distinguishing between employment and nonemployment as labor market states. Consequently, a new job may either start after a job-to-job move has taken place (that is 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. The latter either means that before starting the new job the individual has been nonemployed 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 nonemployment.¹ Note that separations to nonemployment 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 labor market status, like registered unemployment, is recorded in the data.

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 the total sample, 246,401 jobs (or 13.8 percent) have right-censored job durations.² In our sample, the number of jobs varies markedly across the 103 local labor markets, with a minimum of 1,401 and a maximum of 98,977. We observe multiple jobs within a given labor market for most (56.1 percent) 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 labor market as well as enough movers across markets to precisely estimate local wage levels.

III. Estimating Competition in Local Labor Markets

As a measure of labor market competitiveness, we use the elasticity of the separation rate with respect to the wage, a measure that has been used extensively in the existing literature following Manning (2003: 96–104). The justification comes from a model where search frictions are the source of labor market imperfections. Consider a firm paying wage w .³ In steady state, the labor supply to this firm $L(w)$ can be written as:

$$(1) \quad L(w) = R(w)/s(w)$$

where $R(w) > 0$ denotes the recruitment rate and $0 < s(w) < 1$ the separation rate (both functions will depend on other factors though this is suppressed in the interest of notational convenience). It is natural to assume that a higher wage makes it easier to recruit workers so that $R' > 0$ and also reduces the separation rate so $s' < 0$. We thus assume that the firm can increase its labor supply by increasing its wage so that it has some market power.

From Equation 1 the labor supply elasticity to the firm ε_{Lw} can be written as the difference of the wage elasticity of recruitment ε_{Rw} and the wage elasticity of the separation rate ε_{sw} :

$$(2) \quad \varepsilon_{LW} = \varepsilon_{RW} - \varepsilon_{SW}$$

The elasticity of the labor supply curve to an individual firm is a natural measure of how much market power the firm has. Using Equation 2 to estimate the labor supply elasticity would seem to require an estimate of both the separation and recruitment elasticity. Whereas the separation elasticity is straightforward to estimate (and later we describe how we do this), estimating the recruitment elasticity in the type of data we have is much harder.⁴

To deal with this problem, Manning (2003: 97, 2011) notes that many models of imperfect competition in labor markets, such as Burdett and Mortensen (1998) and Bhaskar and To (1999), imply that the recruitment elasticity is minus the separation elasticity, that is $\varepsilon_{RW} = -\varepsilon_{SW}$, so that the labor supply elasticity becomes:

$$(3) \quad \varepsilon_{LW} = -2\varepsilon_{SW}$$

Intuitively, this result holds because one firm's wage-related hire is another firm's wage-related quit. Hence, Equation 3 allows us to identify the labor supply elasticity to the firm by just estimating the wage elasticity of incumbent workers' job separation rate.⁵

To estimate the relationship between the labor supply elasticity and population density, we use a two-step procedure similar in spirit to the approaches used for estimating the urban wage premium by Hirsch and Schumacher (2005), Combes, Duranton, and Gobillon (2008), as well as De la Roca and Puga (2017). In the first step, we estimate a model for the separation elasticity in each region. In the second step, we investigate whether these estimated region-specific separation elasticities are correlated with population density controlling for other region characteristics.

To estimate the separation elasticity for each local labor market area in the first step, we fit individual-level separation equations controlling for the log wage and other relevant worker and

employer characteristics. The coefficient on the log wage is interacted with region dummies so that we obtain estimates of the separation elasticity for each local labor market. Specifically, we estimate a Cox model for the separation rate of job m held by worker i at employer j in region r :

$$(4) \quad s_m \left(\tau \mid \log w_m(\tau), \mathbf{x}_i(\tau), \mathbf{z}_j(\tau) \right) = s_{0ir}(\tau) \exp(\theta_r \log w_m(\tau) + \mathbf{x}_i(\tau)' \boldsymbol{\beta} + \mathbf{z}_j(\tau)' \boldsymbol{\gamma})$$

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 Equation 4, 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.⁶ 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 labor markets because of regional price or wage level differences, as permanent price and wage level differences are part of the baseline hazard and are thus accounted for.

To estimate the separation equation, we adopt the stratified partial likelihood estimator (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 Cox models with worker–region-specific baseline hazards is viable, and we are able to identify local separation rate elasticities—the θ_r 's—in the first-step separation equation with our data.

The Cox model is thus identified from within-variation at the worker–region level, for instance from wage variation occurring in multiple jobs held by the same worker within the same local labor market. As the probability of a separation is likely to depend on the wage in this job relative to others in the market, it is important to control for variables likely to influence the general level of wages available to a worker. We estimate a variety of specifications for Equation 4, differing in the controls that are used. The coefficients of interest—the region-specific coefficients of the log wage in the separation equation—are hard to report but available upon request. In Model I, we only include worker controls: real experience (linearly and squared) as well as groups of dummies for education (distinguishing low-skilled, medium-skilled, and high-skilled workers⁷), one-digit occupation, and non-German nationality. In Model II, we additionally include employer controls: 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 to all models a full set of time dummies. The average estimated labor supply elasticity is 2.43 in Model I and 2.22 in Model II, both within the range of previous estimates summarized by Manning (2011) and Sorensen and Sokolova (2020). These numbers imply that employers possess substantial, though not implausibly large wage-setting power over their workers.

We next investigate the robustness of these estimates to different modeling assumptions. One concern with our estimates is that the Cox models though controlling for worker and employer observables as well as permanent worker unobservables may still suffer from bias stemming from employer unobservables. For example, compensating wage differentials may result in higher wages for high-turnover employers, thereby contaminating our estimates of the local firm-level labor supply elasticity. Controlling for both permanent worker and employer unobservables by

means of a Cox model, 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 Model III that adds to the variables included in Model II, the plant wage effect from a two-way fixed-effects decomposition of individual wages for our data conducted by Card, Heining, and Kline (2013) that builds on Abowd, Kramarz, and Margolis's (1999) methodology. In Abowd, Kramarz, and Margolis'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 characteristics. 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 nonpecuniary attractiveness of employers.

Card, Heining, and Kline'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, Heining, and Kline (2015). Hence, we fit Model III that controls for plant wage effects in the Cox model on a reduced sample of jobs at plants for which Card, Heining, and Kline (2015) provide plant wage effects and where we disregard jobs starting in 2010 and treat jobs ending in 2010 as right-censored. We obtain almost unchanged results when controlling for the plant wage effect with an average local labor supply elasticity of 2.38.

We also estimate Model IV that includes as the wage variable the deviation of the individual's log wage from the current plant wage effect. It thus rests identification on wage variation from the combination of any match effect between a specific worker–employer pair and the current wage residual. In this specification, we are seeing whether individuals who are relatively well-paid within a plant are less likely to leave. The average implied labor supply

elasticity is now about 1.5 compared to 2.3 in the other specifications (possibly because of attenuation bias as plant wage effects are estimated with error), suggesting less competition in the labor market.

Having obtained an estimate of the labor supply elasticity of each local labor market, we now investigate the relationship between this measure of labor market competitiveness, that is minus two times the separation elasticity from Equation 4, and population density in the second step. To investigate whether this relationship is robust to controlling for other labor market characteristics, we estimate models of the form:

$$(5) \quad \hat{\varepsilon}_{LW,r} = \zeta_0 + \zeta_1 \log popdens_r + \mathbf{c}'_r \boldsymbol{\psi} + v_r$$

where the dependent variable is the estimated local labor supply elasticity and the regressors are log population density ($\log popdens_r$) and other local labor market characteristics (denoted \mathbf{c}_r). The other controls are the log employment share of the largest industry in the local labor market (a measure of market specialization), 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). Note that we center all regressors around their means so that the regression constant gives the average elasticity across local labor markets.

The coefficients on log population density are shown in Table 2 where the different columns represent the different specifications for the separation equation. For Model I, the labor supply elasticity to the firm is significantly larger in denser labor markets. A 100 log points increase in population density, which is approximately the interquartile range across local labor markets and, to give one example, the difference in log density between Munich (5.87) and Bayreuth (4.85), comes along with a rise in the elasticity by 0.19.

The positive relationship between the elasticity and density shrinks somewhat when

controlling for employer characteristics in Model II and for the plant wage effects in Model III. With employer controls in the first-step 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. In Model IV, the relationship is weaker still though still significantly different from zero. For all four models of the separation elasticity, there is evidence that denser labor markets are more competitive. This is in line with previous work by Manning (2010) and Hirsch, König, and Möller (2013).

Figure 2 plots the estimated elasticity for the preferred specification we use later in the paper (Model II) against population density showing there is a relationship that is not driven by outliers. This section has established that denser labor markets have higher separation elasticities, suggestive of a more competitive labor market. The next section investigates the part of the urban wage premium that can be ascribed to these differences in competition.

IV. The Urban Wage Premium

An increasing body of international evidence has established that workers earn significantly higher wages in denser labor markets. For convenience, we follow convention and refer to the gradient of wages with respect to population density as the urban wage premium though it does not just compare rural and urban areas. The urban wage premium has proved robust to controlling for unobserved worker heterogeneity by means of fixed-effects techniques (Glaeser and Maré 2001; D’Costa and Overman 2014) and to endogenizing workers’ location decision in structural approaches (Gould 2007; Baum-Snow and Pavan 2012). The premium is thus unlikely to reflect mere worker sorting.

The literature has documented that the urban wage premium stems both from a wage level

and a wage growth effect—see, for instance, the survey by Heuermann, Halfdanarson, and Suedekum (2010)—in the form of a higher return to experience in denser labor markets (De la Roca and Puga 2017). We investigate both the level of wages and the returns to experience in this section.

A. The Urban Wage Level Premium

Figure 3 plots the average log wage against log population density for our 103 local labor markets. 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 percent higher wages on average. This estimate may obviously be biased because of a failure to control for other determinants of wages correlated with population density.

To estimate the urban wage premium, we use a two-step procedure similar to the approaches by Hirsch and Schumacher (2005), Combes, Duranton, and Gobillon (2008), as well as De la Roca and Puga (2017) and used by us for the separation elasticity. In the first step, we estimate Mincerian wage equations at the level of the individual worker of the general form:

$$(6) \quad \log w_{ijrt} = \delta_r + \alpha_i + \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{z}'_{jt}\boldsymbol{\gamma} + u_{ijrt}$$

where δ_r is a region fixed effect, α_i is a worker fixed effect, and u_{ijrt} is an error term.⁸ Our main interest in Equation 6 are the δ_r 's which provide us with estimates of average local wage levels after controlling for observable worker and employer characteristics. In the second step, these regional wage premia are regressed on population density and other region characteristics.

We estimate a variety of specifications for the first-step earnings function (results are available upon request). Model I omits individual worker fixed effects and includes only worker controls whereas Model II adds employer controls. In these specifications, the urban wage

premium is identified in part by a comparison of wages for workers with similar characteristics and employers but who always work in different areas. As argued by Glaeser and Maré (2001), Yankow (2006), and De la Roca and Puga (2017), one concern with these estimates is that workers in different local labor markets may differ in permanent unobservables that affect their wages. Models III and IV address these concerns by including worker fixed effects. In these specifications, the urban wage premia are identified from workers who move regions.⁹ Finally, Equation 6 can also be estimated in first-differenced form in which case emphasis is on the wage change in the year when workers move regions; these specifications with and without employer controls are Models V and VI.

These different specifications for the earnings functions all give us estimates of the wage premium for each local labor market. As for the estimates of the separation elasticity, the second step of our estimation process is to regress the estimated wage premia on log population density and other area characteristics, that is a regression like Equation 5, but with the dependent variable now the estimated wage premia:

$$\hat{\delta}_r = \zeta_0 + \zeta_1 \log popdens_r + \mathbf{c}'_r \boldsymbol{\psi} + v_r \quad (7)$$

The coefficients on log population density for the levels equations Models I–IV are reported in Panel A of Table 3. In line with expectations, we find a very significant relationship between all the estimated wage premia and population density: for Model I the coefficient on log density is 0.032, very similar to the descriptive estimate of 0.034 in Figure 2. This implies a 100 log points rise in population density is associated with a rise in local wages by 3.2 percent. When additionally controlling for employer characteristics in Model II, this estimate rises somewhat to 3.6 percent. Including worker fixed effects (Models III and IV) causes the coefficient on log population density to fall a bit but it still remains very significant. 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) percent increase in wages. The estimates for the model in first differences (Models V and VI) are reported in Table 4 and yield a weaker though still significant relationship between the wage premia and log population density.

B. The Link between the Urban Wage Premium and Labor Market Competition

Having shown that both the wage premium and the separation elasticity are significantly related to population density, we now investigate whether the wage premia can be explained by the level of labor market competitiveness. The standard explanation for the wage premia is that agglomeration economies raise worker productivity in thick markets (Duranton and Puga 2004; Puga 2010; Moretti 2011). There is good empirical evidence that agglomeration economies exist—see, for instance, the surveys by Rosenthal and Strange (2004) or Combes and Gobillon (2015)—but the precise mechanisms are less clear, for example knowledge spill-overs, faster learning, or a more efficient matching process have been proposed. In many papers, the underlying model of the labor market is not made explicit but implicitly would seem to be a perfectly competitive model in which wages are equal to marginal products.

However, if labor markets are imperfectly competitive and employers possess some wage-setting power over their workers, “wages are ... only proportional and not equal to labor productivity by a factor that depends on the local monopsony power of the firm” (Combes and Gobillon 2015: 283). Wages will be a mark-down on marginal products with a larger mark-down in less competitive labor markets. If denser labor markets are more competitive, as suggested by Manning (2010), Hirsch, König, and Möller (2013), and our estimates in Section III, workers in these markets obtain a larger share of their marginal product, and we might observe an urban wage premium even if agglomeration economies were completely absent and there were no difference

in marginal products.

To investigate the part of the urban wage premia that can be explained by differences in labor market competitiveness, we modify Equation 7 and include our measure of labor market competitiveness as an additional regressor on the right-hand side, that is we estimate:

$$(8) \quad \hat{\delta}_r = \tilde{\pi}_0 + \tilde{\pi}_1 \log \text{popdens}_r + \tilde{\pi}_2 \hat{\epsilon}_{LW,r} + \mathbf{c}'_r \tilde{\boldsymbol{\eta}} + e_r$$

We are interested in the sign on the estimated labor supply elasticity and how its inclusion affects the coefficient of population density. In these regressions, we use our estimates of the supply elasticity from Model II in Table 2.

Panel B of Table 3 estimates Equation 8 by OLS. Although the coefficient on labor market competitiveness has the expected sign, it is small in magnitude and the inclusion of the labor supply elasticity has little impact on the coefficient on population density. Taken at face value, these results suggest that little or none of the estimated wage premia can be explained by differences in labor market competitiveness.

However, one problem with these OLS regressions is that the labor supply elasticities are estimated with error and are quite noisy. Some indication of this can be seen in Figure 2 where both the largest and smallest elasticities are found for small regions with low population density that one would expect to be estimated with less precision. This problem is likely to lead to an attenuation bias on the labor supply elasticity and a corresponding higher loading on a variable like population density that is correlated with the signal in the elasticity.

To deal with this problem, we seek an alternative measure of labor market competition that can be used as an instrument for the labor supply elasticity. We follow Manning (2003: 44–9) who suggested using the share of hires from nonemployment (as opposed to employment). In the canonical Burdett and Mortensen (1998) model, Manning shows that this variable is a sufficient

statistic for how competitive is the labor market. The intuition is that competition is greater when employers are directly competing against each other for workers which is the case for hires from employment but not for hires from nonemployment. Hence, the higher is the share of hires from nonemployment, the less often are employers hiring workers faced with direct competition with other employers. The outline argument for this is in the Appendix.¹⁰ For our purposes, it does not matter whether this variable is a perfect measure of labor market competition, just that it is correlated with our preferred measure (the labor supply elasticity) and that it is derived from a completely different set of statistics so that any measurement error in the two competition variables is unlikely to be correlated.¹¹

Figure 4 plots the separation elasticity against the share of hires from nonemployment in local labor markets. The first stage suggests the instrument is strong with a first-stage F statistic of 23.38: as it is the same for all models, it is reported in the Appendix. Panel C of Table 3 now presents the IV estimates of the wage premia; the coefficient on the labor supply elasticity is significantly different from zero for all specifications of the urban wage premium. This is evidence that part of the urban wage premium can be explained by more competitive labor markets in more densely populated areas.

The coefficient on log population density is lower in the IV estimates when competition is included as a regressor (Panel C) than when competition is omitted (Panel A) though it remains significantly different from zero in all specifications apart from Model I. The magnitude of the impact of population density falls by about 40 percent, suggesting that there are agglomeration economies, though smaller than implied by other estimates and that differences in competition can explain some but not all of the urban wage premium. Panel D of Table 3 presents the reduced form for our equation. There is a strong relationship between the share of hires from nonemployment

and the wage premia and the impact of population density is much lower than found if the measures of labor market competition are omitted.

Table 4 does a similar exercise for the estimates of the urban wage premium derived from first-differenced Models V and VI. All coefficients are somewhat smaller but the general conclusion remains; more competitive labor markets have higher wage premia and the impact of population density is substantially reduced when controlling for a measure of labor market competition.

It is useful to compare the estimated impact of the labor supply elasticity on the urban wage premium to what would be expected in a simple monopsony model. In that model, the wage is a fraction of the marginal product of labor, that is $w = \frac{\varepsilon_{LW}}{1+\varepsilon_{LW}} MPL$, so that the expected impact of the labor supply elasticity ε_{LW} on log wages is $\frac{\partial \log w}{\partial \varepsilon_{LW}} = \frac{1}{\varepsilon_{LW}(1+\varepsilon_{LW})}$. In our preferred Model II, the mean labor supply elasticity is 2.22. This suggests we would expect to find a coefficient in Table 3 of about 0.14. The estimated impacts are lower than this; one possible explanation is the suggestion in Manning (2011) that employers can affect hiring not just through their choice of the wage but also through their hiring activities. But, our results suggest that the level of wages is higher in more competitive markets and that a sizeable part of the urban wage premium can be explained by this competition effect.

C. The Urban Wage Growth Premium

De la Roca and Puga (2017) find that additional work experience in denser labor markets leads to a significant urban wage growth premium that adds to the static gains from working in a thick market—see also the survey by Heuermann, Halfdanarson, and Suedekum (2010). This is hypothesized to be because of greater human capital accumulation in denser markets but also

because a more efficient matching process causes average match quality to rise faster.

While it is relatively easy to understand why there may be a link between wage levels and market competition, it is perhaps not so obvious why wage growth should also be linked to market competition. But, if it is easier for workers in more competitive markets to find other potential employers, wage growth would be expected to be higher. This is the prediction of the standard canonical Burdett–Mortensen model which assumes identical employers. However, this mechanism may not be so distinct from other hypotheses based on the efficiency of the matching process. In the Burdett–Mortensen model augmented with heterogeneity in productivity across employers, one would also expect the quality of the match to increase faster in more competitive markets. In addition, if there are gains from matching high ability workers with high productivity firms, then denser, more competitive markets will be more effective at generating assortative matching, in line with the evidence in Dauth et al. (2018).

To estimate an urban wage growth premium, we modify the first-step wage equation to include interactions of a quadratic in experience with region-specific dummies. To ensure that we are only capturing returns to experience that occur within a region, we include worker–region fixed effects in this model, meaning that the level of the urban wage premium cannot be identified in these specifications.¹² However, the coefficients on the interactions with experience provide us with estimates of experience–wage profiles that are specific to each local labor market. We then regress these estimated coefficients on population density, our measure of labor market competition, and other area controls as was done earlier for the static urban wage premium.

Table 5 reports the results for two specifications of the earnings function, one with only worker controls, the other also with employer controls. Panel A reports results where the labor market competition variable is omitted. The results are very similar, with a linear term that is

increasing in population density and a quadratic term that is decreasing. As it is not so easy to interpret these coefficients, the solid line in Figure 5 shows the predicted cumulated wage growth at different levels of experience in two labor markets with one being 100 log points denser than the other (corresponding to the estimates in Panel A of Table 5). The accumulated urban wage growth premium from entering a 100 log points denser labor market and gaining 20 years of work experience in this market amounts to about 18 log points.

As for the urban wage level premium, we next investigate whether this urban wage growth premium can be explained by differences in labor market competition. In line with Table 3, Panel B of Table 5 presents an OLS regression of the wage growth premium on our measure of labor market competition, Panel C the IV estimates (the first stage is the same as in Table 3), and Panel D the reduced form. In the IV specification, higher labor market competition is estimated to increase both the linear and quadratic terms though the estimated coefficients are not significantly different from zero. However, the coefficients on log population density in Panels C and D are smaller than those in Panel A. The dashed line in Figure 5 shows that the urban wage growth premium after 20 years is about halved when a control for labor market competition is introduced.

V. Robustness Checks

To scrutinize our results further, we perform several checks of robustness along three dimensions. The results are presented in Table 6 where the first column is the coefficient on log population density in a regression where the dependent variable is the labor supply elasticity (akin to Model II in Table 2), the second column is the coefficient on log population density in a regression where the dependent variable is the urban wage premium and no competition control is included (akin to Model IV in Table 3, Panel A), the third column is the coefficient on log

population density when a competition control is included (akin to Model IV in Table 3, Panel C), and the fourth column is the coefficient on the instrumented labor supply elasticity in the same regression.

The first row in Table 6 shows the baseline estimates as previously reported. The second through fifth rows then repeat our analysis using different measures of agglomeration in the second-step regressions: the log population density at the beginning and end of our sample period, including log population and log area as separate regressors, and using employment rather than population density. The results are very similar for all measures.

The sixth row repeats our analysis including top-coded wage observations which we impute using a heteroskedastic 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 labor markets with higher wage levels, which may arouse some concerns. Results are very similar.

Finally, the seventh row reestimates the model using a different instrument for the estimated labor supply elasticity. We use the measure of search frictions suggested by van den Berg and van Vuuren (2010), namely the local share of job exits into nonemployment (as opposed to employment). Like the share of hires from nonemployment used in our baseline specification, the share of job exits into nonemployment captures how hard it is for workers to move their way up in the local wage distribution by job-to-job moves. The final row of Table 6 shows this has little impact on the results.

VI. 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 level and growth premium stems from greater competition in thick local labor markets. In the first part of our analysis, we documented that the wage elasticity of the labor supply to the firm, which governs what part of the marginal product of labor accrues to workers in monopsonistic labor markets with employer wage setting, is significantly larger in denser markets. The estimated average elasticity across local labor markets in most specifications is about 2.3 with an increase in population density by 100 log points coming along with an increase in the elasticity by about 0.15–0.19.

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 percent higher wages when controlling for worker fixed effects and several worker, employer, and local labor market characteristics. However, once we conditioned on our measure of labor market competition, this effect was reduced by about 40 percent. We also found that more competitive labor markets have higher wages.

Our findings are in line with the notion that a substantial part, though not all, of the urban wage premium derives from fiercer competition in thick labor markets. Our results therefore suggest that workers in denser labor 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 labor.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67(2):251–333.

Baum-Snow, Nathaniel, and Ronni Pavan. 2012. "Understanding the City Size Wage Gap." *Review of Economic Studies* 79(1):88–127.

Bhaskar, Venkataraman, and Ted To. 1999. "Minimum Wages for Ronald McDonald Monopsonies: A Theory of Monopsonistic Competition." *Economic Journal* 109(455):190–203.

Booth, Alison L., and Pamela Katic. 2011. "Estimating the Wage Elasticity of Labour Supply to a Firm: What Evidence is there for Monopsony?" *Economic Record* 87(278):359–69.

Brenzel, Hanna, Hermann Gartner, and Claus Schnabel. 2014. "Wage Bargaining or Wage Posting? Evidence from the Employers' Side." *Labour Economics* 29(1):41–8.

Burdett, Kenneth, and Dale T. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review* 39(2):257–73.

Büttner, Thomas, and Susanne Rässler. 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, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128(3):967–1015.

_____. 2015. "CHK Effects: Version 2." FDZ-Methodenreport 06/2015, Nuremberg: Institute for Employment Research.

Combes, Pierre-Phillipe, Gilles Duranton, and Laurent Gobillon. 2008. "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics* 63(2):723–42.

Combes, Pierre-Phillipe, and Laurent Gobillon. 2015. “The Empirics of Agglomeration Economies.” In *Handbook of Regional and Urban Economics*, Volume 5A, ed. Gilles Duranton, John Vernon Henderson, and William C. Strange, 247–348. Amsterdam: Elsevier.

Dauth, Wolfgang, Sebastian Findeisen, Enrico Moretti, and Jens Suedekum. 2018. “Matching in Cities.” NBER Working Paper No. 25277, Cambridge, MA: National Bureau of Economic Research.

D’Costa, Sabine, and Henry G. Overman. 2014. “The Urban Wage Growth Premium: Sorting or Learning?” *Regional Science and Urban Economics* 48(1):168–79.

De la Roca, Jorge, and Diego Puga. 2017. “Learning by Working in Big Cities.” *Review of Economic Studies* 84(1):106–42.

Duranton, Gilles, and Diego Puga. 2004. “Micro-Foundations of Urban Agglomeration Economies.” In *Handbook of Regional and Urban Economics*, Volume 4, ed. John Vernon Henderson and Jacques-François Thisse, 2063–117. Amsterdam: Elsevier.

Falch, Torberg. 2011. “Teacher Mobility Responses to Wage Changes: Evidence from a Quasi-Natural Experiment.” *American Economic Review (Papers and Proceedings)* 101(3):460–5.

_____. 2017. “Wages and Recruitment: Evidence from External Wage Changes.” *ILR Review* 70(2):483–518.

Fitzenberger, Bernd, Aderonke Osikominu, and Robert Völter. 2006. “Imputation Rules to Improve the Education Variable in the IAB Employment Subsample.” *Schmollers Jahrbuch* 126(3):405–36.

Glaeser, Edward L., and Joshua D. Gottlieb. 2009. “The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States.” *Journal of Economic Literature* 47(4):983–1028.

Glaeser, Edward L., and David C. Maré. 2001. “Cities and Skills.” *Journal of Labor Economics* 19(2):316–42.

Gould, Eric D. 2007. “Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium.” *Review of Economic Studies* 74(2):477–506.

Hethey-Maier, Tanja, and Johannes F. Schmieder. 2013. “Does the Use of Worker Flows Improve the Analysis of Establishment Turnover: Evidence from German Administrative Data.” *Schmollers Jahrbuch* 133(4):477–510.

Heuermann, Daniel, Benedikt Halfdanarson, and Jens Suedekum. 2010. “Human Capital Externalities and the Urban Wage Premium: Two Literatures and their Interrelations.” *Urban Studies* 47(4):749–67.

Hirsch, Barry T., and Edward J. Schumacher. 2005. “Classic or New Monopsony? Searching for Evidence in Nursing Labor Markets.” *Journal of Health Economics* 24(5):969–89.

Hirsch, Boris, Elke J. Jahn, and Claus Schnabel. 2018. “Do Employers Have More Monopsony Power in Slack Labor Markets?” *ILR Review* 71(3):676–704.

Hirsch, Boris, Marion König, and Joachim Möller. 2013. “Is There a Gap in the Gap? Regional Differences in the Gender Pay Gap.” *Scottish Journal of Political Economy* 60(4):412–39.

Hirsch, Boris, Thorsten Schank, Claus Schnabel. 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.

Jacobebbinghaus, Peter, and Stefan Seth. 2007. “The German Integrated Employment Biographies Sample IEBS.” *Schmollers Jahrbuch* 127(2):335–42.

Kalbfleisch, John D., and Ross L. Prentice. 2002. *The Statistical Analysis of Failure Time Data*. Hoboken, NJ: John Wiley & Sons, 2nd edition.

Kosfeld, Reinhold, and Alexander Werner. 2012. “Deutsche Arbeitsmarktregionen – Neuabgrenzung nach den Kreisgebietsreformen 2007–2011.” *Raumforschung und Raumordnung* 70(1):49–64.

Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton, NJ: Princeton University Press.

_____. 2010. “The Plant Size-Place Effect: Agglomeration and Monopsony in Labour Markets.” *Journal of Economic Geography* 10(5):717–44.

_____. 2011. “Imperfect Competition in the Labor Market.” In *Handbook of Labor Economics*, Volume 4B, ed. Orley Ashenfelter and David Card, 973–1041. Amsterdam: Elsevier.

_____. 2020. “Monopsony in Labor Markets: A Review.” *ILR Review*. Forthcoming.

Melo, Patricia C., Daniel C. Graham, and Robert B. Noland. 2009. “A Meta-Analysis of Estimates of Urban Agglomeration Economics.” *Regional Science and Urban Economics* 39(3):332–42.

Moretti, Enrico. 2011. “Local Labor Markets.” In *Handbook of Labor Economics*, Volume 4B, ed. Orley Ashenfelter and David Card, 1237–313. Amsterdam: Elsevier.

Puga, Diego. 2010. “The Magnitude and Causes of Agglomeration Economies.” *Journal of Regional Science* 50(1):203–19.

Ridder, Geert, and İnsan Tunali. 1999. “Stratified Partial Likelihood Estimation.” *Journal of Econometrics* 92(2):193–232.

Rosenthal, Stuart S., and William C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies." In *Handbook of Regional and Urban Economics*, Volume 4, ed. John Vernon Henderson and Jacques-François Thisse, 2119–71. Amsterdam: Elsevier.

Sorensen, Todd A., and Anna Sokolova. 2020. "Monopsony in Labor Markets: A Meta-Analysis." *ILR Review*. Forthcoming.

Sorkin, Isaac. 2018. "Ranking Firms Using Revealed Preference." *Quarterly Journal of Economics* 133(3):1331–93.

Spengler, Anja. 2008. "The Establishment History Panel." *Schmollers Jahrbuch* 128(3):501–9.

Syverson, Chad. 2019. "Macroeconomics and Market Power: Context, Implications, and Open Questions." *Journal of Economic Perspectives* 33(3):23–43.

van den Berg, Gerard J., and Aico van Vuuren. 2010. "The Effect of Search Frictions on Wages." *Labour Economics* 17(6):875–85.

Yankow, Jeffrey J. 2006. "Why Do Cities Pay More? An Empirical Examination of Some Competing Theories of the Urban Wage Premium." *Journal of Urban Economics* 60(2):139–61.

Appendix: The Burdett–Mortensen Model

This appendix briefly summarizes some aspects of the Burdett–Mortensen (1998) model that are used in the main text along the lines of Manning (2003, Chapter 2) where formal derivations can be found.

The marginal product of labor in all firms is p , and the utility from being nonemployed is b . Employers post wages to maximize steady-state profits. Job offers, randomly drawn from employers, arrive at a rate λ for both nonemployed and employed workers. Employed workers lose their jobs at a rate δ . The ratio $k = \lambda/\delta$ is a measure of the competitiveness of the labor market.

Burdett and Mortensen (1998)—see also Manning (2003)—show that the average wage in the labor market can be written as:

$$E(w) = \frac{b + kp}{1 + k}$$

so that more competitive markets (those with a higher value of k) have higher wages conditional on the level of productivity. This also means that, conditional on (b, p) , a higher value of k must be associated with a higher average labor supply elasticity to the firm.

Manning (2003) also shows that the share of hires from nonemployment is given by:

$$\frac{k}{(1 + k) \log(1 + k)}$$

which can be shown to be a decreasing function of the measure of competitiveness, k . This demonstrates that a single measure of competitiveness is associated with the level of wages, the elasticity of labor supply to the firm, and the share of hires from nonemployment.

Figures

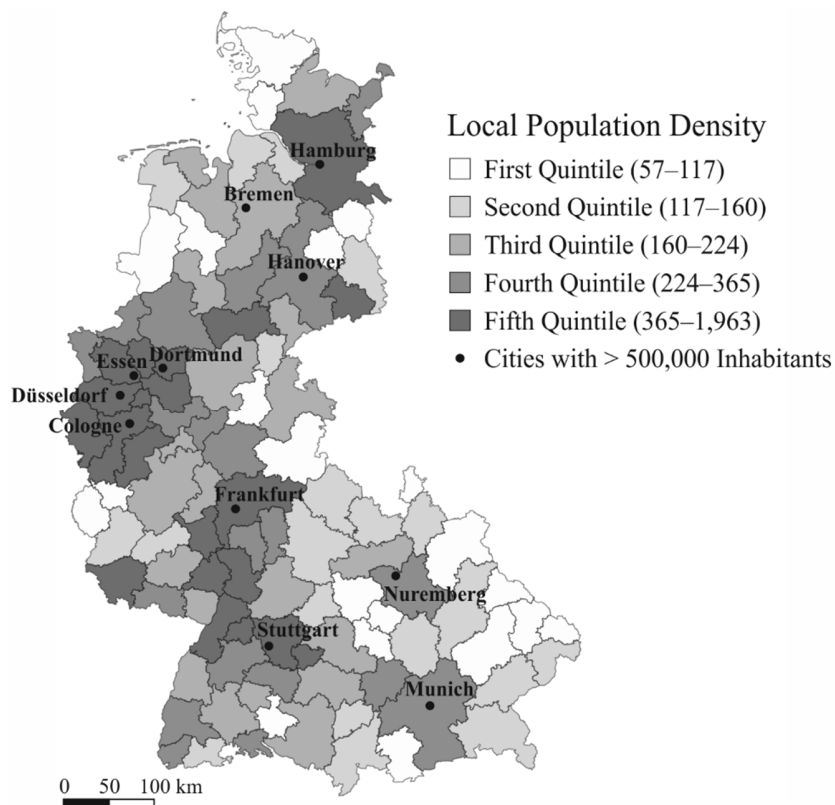


Figure 1

Local Labor Markets in West Germany and Average Population Density by Quintile

Notes: This figure shows the 103 West German local labor markets in our sample and their time-averaged population density (that is population per square kilometer averaged over the years 1985–2010) by quintile along with large cities of more than 500,000 inhabitants.

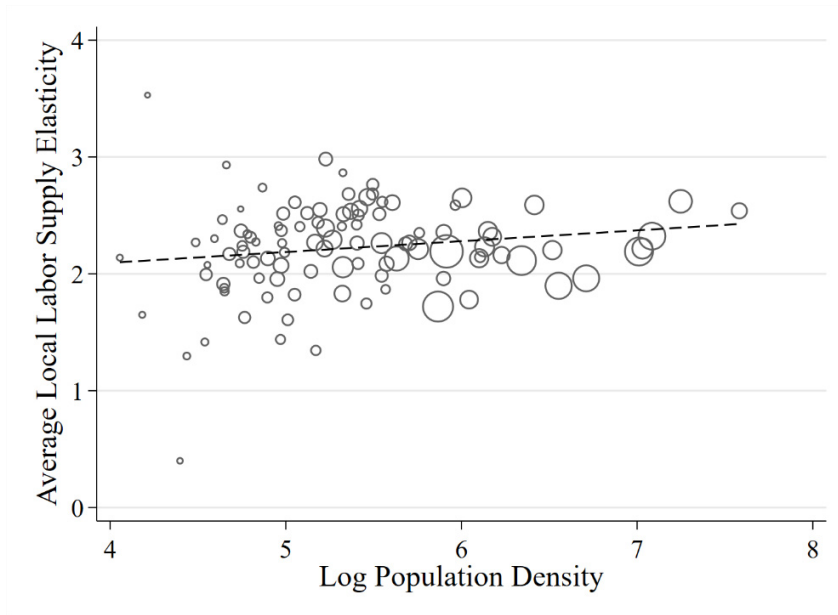


Figure 2

Local Labor Supply Elasticities to the Firm and Log Population Density

Notes: Markers are weighted by population size.

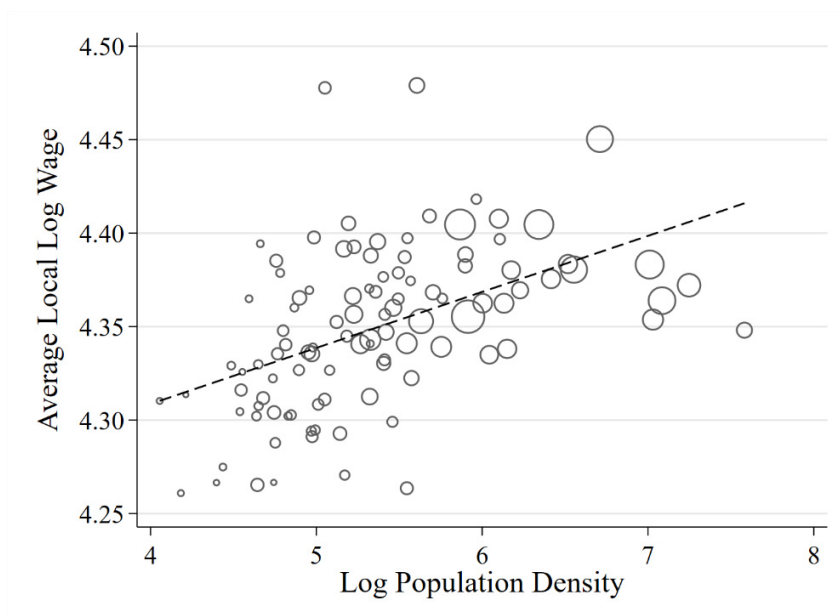


Figure 3

Local Average Wages and Log Population Density

Notes: Markers are weighted by population size.

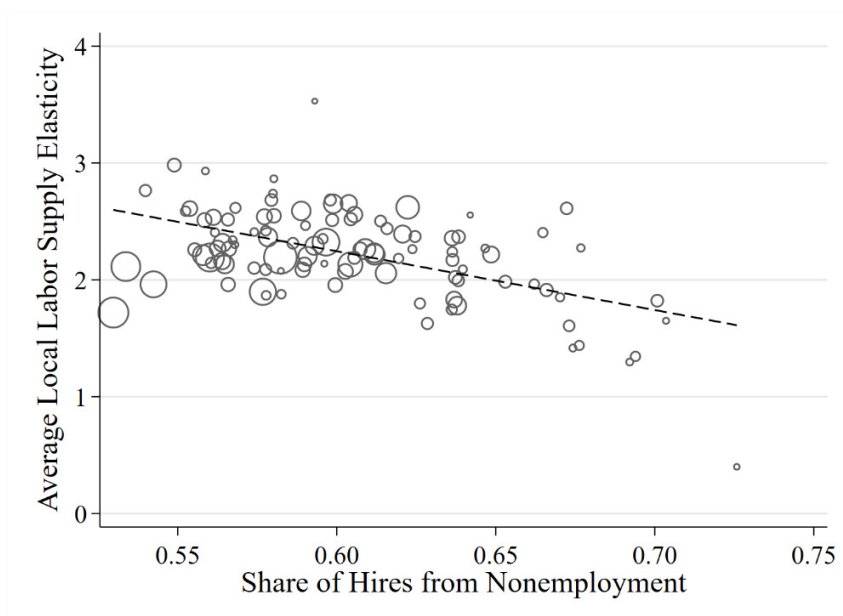


Figure 4

Local Labor Supply Elasticities to the Firm and Share of Hires from Nonemployment

Notes: Markers are weighted by population size.

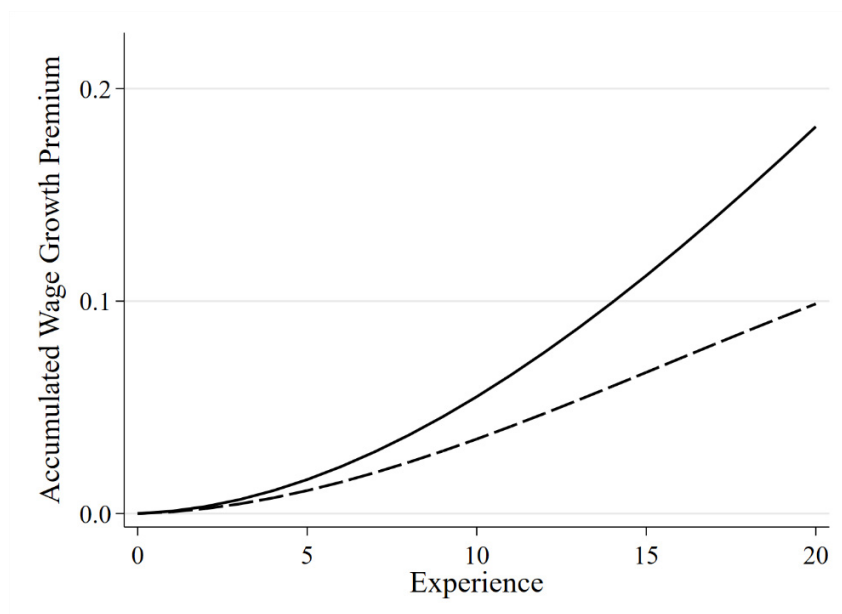


Figure 5

Accumulated Urban Wage Growth Premium over Workers' Real Work Experience

Notes: The figure shows the additional log wage growth in a 100 log points denser local labor market over workers' real work experience (based on the estimates of Model II in Table 5, Panels A and C). The solid line shows the accumulated wage growth when conditioning on worker, employer, and local labor market characteristics (estimates from Panel A). The dashed line shows the accumulated wage growth when additionally conditioning on the labor supply elasticity instrumented with the share of hires from nonemployment (estimates from Panel C).

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 (quarterly job spells)	17,010,740

Notes: IEB and BHP, 1985–2010.

Table 2

Local Differences in the Wage Elasticity of the Labor Supply to the Firm

First-step specification	Model I	Model II	Model III	Model IV
	Cox model with worker but without employer controls	Cox model with worker and employer controls	Cox model with worker and employer controls and plant wage effects	Cox model with worker and employer controls and deviating wage from plant wage effect
Second-step results (103 local labor markets)				
Log population density	0.1879 (0.0694)	0.1486 (0.0658)	0.1490 (0.0663)	0.0942 (0.0241)
Constant	2.4277 (0.0429)	2.2201 (0.0392)	2.3771 (0.0390)	1.4726 (0.0148)

Notes: IEB and BHP, 1985–2010. Estimates show the second-step regression (Equation 5). The dependent variable is the estimated wage elasticity of the labor supply to the firm obtained from the first-step separation equation (Equation 4), which we model as a 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 centered around their means. In the Cox model, 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, Heining, and Kline (2015) interacted with its reference period. In Model IV, the wage regressor is the deviation of the log wage from the plant wage effect from Card, Heining, and Kline (2015). Robust standard errors in parentheses.

Table 3

Estimated Urban Wage Premium from Level Models

First-step specification	Model I	Model II	Model III	Model IV
	OLS with worker controls	OLS with worker and employer controls	FE with worker controls	FE with worker and employer controls
Second-step results (103 local labor markets)				
Panel A: Not conditioning on the labor supply elasticity				
Log population density	0.0316 (0.0081)	0.0360 (0.0073)	0.0304 (0.0053)	0.0283 (0.0050)
Panel B: Conditioning on the labor supply elasticity				
Log population density	0.0240 (0.0096)	0.0319 (0.0084)	0.0254 (0.0062)	0.0257 (0.0058)
Labor supply elasticity	0.0276 (0.0079)	0.0151 (0.0064)	0.0180 (0.0056)	0.0098 (0.0049)
Panel C: Conditioning on the labor supply elasticity instrumented with the share of hires from nonemployment				
Log population density	0.0161 (0.0092)	0.0209 (0.0090)	0.0173 (0.0067)	0.0181 (0.0061)
Labor supply elasticity	0.0566 (0.0155)	0.0549 (0.0158)	0.0475 (0.0109)	0.0374 (0.0103)
Panel D: Conditioning on the share of hires from nonemployment (reduced form)				
Log population density	0.0097	0.0147	0.0119	0.0138

	(0.0074)	(0.0064)	(0.0049)	(0.0047)
Share of hires from nonemployment	-0.5556	-0.5391	-0.4664	-0.3673
	(0.1017)	(0.0885)	(0.0790)	(0.0683)

Notes: IEB and BHP, 1985–2010. Coefficients from second-step regressions as in Equations 7 and 8. The dependent variable is the local wage level obtained from the first-step wage regression (Equation 6). The labor supply elasticity is estimated using Model II from Table 2. The F statistic of the first-stage regression for Panel C is 23.38. 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 centered 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. Robust standard errors in parentheses.

Table 4

Estimated Urban Wage Premium from First-Differenced Models

First-step specification	Model V	Model VI
Second-step results (103 local labor markets)	FD with worker controls	FD with worker and employer controls
Panel A: Not conditioning on the labor supply elasticity		
Log population density	0.0224 (0.0025)	0.0239 (0.0023)
Panel B: Conditioning on the labor supply elasticity		
Log population density	0.0192 (0.0061)	0.0222 (0.0061)
Labor supply elasticity	0.0116 (0.0063)	0.0064 (0.0057)
Panel C: Conditioning on the labor supply elasticity instrumented with the share of hires from nonemployment		
Log population density	0.0103 (0.0067)	0.0140 (0.0062)
Labor supply elasticity	0.0443 (0.0118)	0.0364 (0.0109)
Panel D: Conditioning on the share of hires from nonemployment (reduced form)		
Log population density	0.0052 (0.0028)	0.0098 (0.0027)
Share of hires from nonemployment	-0.4345 (0.0573)	-0.3571 (0.0540)

Notes: IEB and BHP, 1985–2010. Coefficients from second-step regressions as in Equations 7 and 8. The dependent variable is the local wage level obtained from the first-step wage regression (Equation 6) in first differences. The labor supply elasticity is estimated using Model II from Table 2. The F statistic of the first-stage regression for Panel C is 23.38. 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 centered 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. Robust standard errors in parentheses.

Table 5

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	Linear	Quadratic	
Second-step results (103 local labor markets)		($\times 100$)		($\times 100$)	
Panel A: Not conditioning on local search frictions					
Log population density	0.0012 (0.0004)	−0.0023 (0.0015)	0.0011 (0.0004)	−0.0020 (0.0014)	
Panel B: Conditioning on the labor supply elasticity					
Log population density	0.0017 (0.0006)	−0.0042 (0.0018)	0.0017 (0.0006)	−0.0036 (0.0018)	
Labor supply elasticity	−0.0020 (0.0008)	0.0070 (0.0026)	−0.0019 (0.0008)	0.0058 (0.0025)	
Panel C: Conditioning on the labor supply elasticity instrumented with the share of hires from nonemployment					
Log population density	0.0009 (0.0007)	−0.0029 (0.0020)	0.0008 (0.0007)	−0.0025 (0.0021)	
Labor supply elasticity	0.0012 (0.0014)	0.0022 (0.0040)	0.0012 (0.0014)	0.0019 (0.0040)	
Panel D: Conditioning on the share of hires from nonemployment (reduced form)					
Log population density	0.0007 (0.0007)	−0.0031 (0.0025)	0.0007 (0.0007)	−0.0027 (0.0025)	
Share of hires from nonemployment	−0.0117 (0.0132)	−0.0219 (0.0422)	−0.0117 (0.0125)	−0.0188 (0.0422)	

Notes: IEB and BHP, 1985–2010. Coefficients from second-step regressions as in Equations 7 and 8. The dependent variables are the region-specific coefficients of real experience and its square (times 100), respectively, obtained from a first-step wage regression akin to Equation 6 including worker–region fixed effects. The labor supply elasticity is estimated using Model II from Table 2. The F statistic of the first-stage regression for Panel C is 23.38. 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 centered 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. Robust standard errors in parentheses.

Table 6

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

Second-step coefficient of the ...	Log agglomeration measure in regression for the labor supply elasticity	Log agglomeration measure in regression for the urban wage premium when not conditioning on the labor supply elasticity	Log agglomeration measure in regression for the urban wage premium when conditioning on the labor supply elasticity	Labor supply elasticity in regression for the urban wage premium (instrumented)
Check of robustness (103 local labor markets)				
Baseline	0.1486 (0.0658)	0.0283 (0.0050)	0.0181 (0.0061)	0.0374 (0.0103)
Alternative agglomeration measures				
Log population density in 1985	0.1325 (0.0632)	0.0276 (0.0048)	0.0177 (0.0058)	0.0381 (0.0105)
Log population density in 2010	0.1530 (0.0698)	0.0289 (0.0051)	0.0191 (0.0063)	0.0364 (0.0098)
Log population (controlling for log area separately)	0.1397 (0.0895)	0.0308 (0.0056)	0.0213 (0.0070)	0.0357 (0.0094)
Log employment density	0.1612 (0.0745)	0.0266 (0.0044)	0.0164 (0.0057)	0.0383 (0.0109)
Including imputed wage observations	0.1481 (0.0646)	0.0266 (0.0048)	0.0162 (0.0063)	0.0387 (0.111)

Using share of job exits into nonemployment as instrument	0.1486 (0.0658)	0.0283 (0.0050)	0.0185 (0.0059)	0.0358 (0.0100)
--	--------------------	--------------------	--------------------	--------------------

Notes: IEB and BHP, 1985–2010. The first column shows the coefficient of the agglomeration measure in the second-step regression (Equation 5) where the first-step separation equation includes worker controls, employer controls, and a worker–region-specific baseline hazard—akin to Model II in Table 2. The second column gives the estimated urban wage premium when not conditioning on the labor supply elasticity as in Equation 7 where the first-step wage equation includes worker controls and fixed effects as well as employer controls—akin to Model IV in Table 3. The last two columns present the corresponding estimates when conditioning on the labor supply elasticity as in Equation 8 where we instrument the elasticity as in Panel C of Table 3. In the final row, we use the share of job exits into nonemployment as instrumental variable; the F statistic of the first-stage regression underlying this row is 22.77. Robust standard errors in parentheses.

Appendix Table 1

First-Stage Regression for the Labor Supply Elasticity to the Firm

Variable (103 local labor markets)	
Share of hires from nonemployment	−9.813 (2.029)
Log population density	−0.1136 (0.1029)
Share of low-skilled workers	0.5529 (1.574)
Share of high-skilled workers	−4.893 (3.174)
Log Herfindahl index of employment	−0.1766 (0.6505)
Log employment share of the largest industry	0.3067 (0.4490)
Unemployment rate	0.0813 (0.0288)

Notes: IEB, BHP, and official statistics, 1985–2010. Time-averages over years. Results of the first-stage regression underlying the IV estimates that instrument the labor supply elasticity to the firm with the share of hires from nonemployment. Robust standard errors in parentheses.

¹ See, for instance, Hirsch, Jahn, and Schnabel (2018) for a comparison to the Socio-Economic Panel that includes workers who are not registered with the German social security system.

² Note that in our data plant identifiers change in some cases for administrative reasons, such as mergers of plants, spin-offs, etc. To avoid the impact of changing plant identifiers on worker transitions and job durations, we drop all jobs subject to spurious changes in plant identifiers as identified in Hethey-Maier and Schmieder (2013).

³ 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 (Brenzel, Gartner, and Schnabel 2014).

⁴ One rare exception is Falch (2017), who is able to analyze data on employers' recruitment pools for certified teachers in Norway and finds substantial monopsony power in this labor market segment. Reassuringly, his estimates of the firm-level labor 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.

⁵ Some previous studies, such as Booth and Katic (2011) and Hirsch, Jahn, and Schnabel (2018), applied a more sophisticated estimation approach distinguishing employment and nonemployment as distinct labor market states. Although our data include information on workers' previous and subsequent labor market states, distinguishing transitions from and to employment from those from and to nonemployment is not viable in our application because of the limited number of jobs observed in sparsely populated local labor markets.

⁶ 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 4 needs to hold only for jobs held by the same worker within a particular local labor market, but may well be violated across workers or regions without invalidating identification (Kalbfleisch and Prentice 2002: 118–9).

⁷ Low-skilled workers are workers with neither a vocational nor an academic degree whereas medium-skilled workers possess a vocational degree and high-skilled workers have an academic degree.

⁸ Note that we do not correct workers' wages for differences in local labor 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, Halfdanarson, and Suedekum (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.

⁹ One potential problem with this strategy is that 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. Instrumenting workers' location has proven difficult due to the lack of credible, strong instruments (Heuermann, Halfdanarson, and Suedekum 2010) and has, in general, also made no big difference (Melo, Graham, and Noland 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.

¹⁰ Manning (2020) discusses the growing literature on how to measure competition in labor markets with most studies using concentration ratios. However, in the Burdett–Mortensen model, greater competition is associated with higher concentration ratios, a problem also pointed out by Syverson (2019) in the context of product markets.

¹¹ In Section V, we demonstrate that we obtain the same results when using an alternative measure of search frictions proposed by van den Berg and van Vuuren (2010).

¹² Note that the worker–region fixed effects further control for the worker’s previous (time-invariant) work experience gained in other local labor 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 labor 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 labor 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.