City Size, Employer Concentration, and Wage Income Inequality

Daniel Halvarsson, Ratio Institute

(daniel.halvarsson@ratio.se)

&

Martin Korpi, Ratio Institute & EHFF, Stockholm School of Economics

(martin.korpi@ratio.se) [1]

PRELIMINARY DRAFT, 022224

(Please do not quote – comments welcome!)

Abstract: In this paper, we build upon a monopsony framework, suggested by Card et. al. 2016, which links firm level productivity and rent-sharing to wage inequality. Specifically, our research questions address *i)* to which extent labor market concentration across firms (within different types of locally situated industries) affects variation in wages among workers within these firms and industries, and *ii)* how this variation in turn spills over into economy-wide inequality (measured at the level of local labor markets). Using linked employer-employee full population data for Sweden, and an AKM modelling framework to separate between worker- and firm-level heterogeneity, our results suggest that higher firm-level fixed effects (a measure of rent-sharing) is associated with lower labor market employer concentration, something which affects average wage income among firms accordingly. Addressing wage income inequality by applying our model to different segments of the local labor market income distribution, we find that reduced average employer concentration in larger cities accounts for almost all variation in the (positive) link between city size-and wage inequality, except for the largest metropolises where it captures around 30-50 percent of variation depending on the income segment that we focus on.

JEL-codes: D22, J31, J42, R12

Keywords: wage distribution, rent sharing, monopsony, linked employer-employee data, local labor markets

[1] Corresponding author. We wish to thank Erik Lakomaa and seminar participants at EHFF, Stockholm School of Economics, as well as at the Institute for Evaluation of Labor Market and Education Policy (IFAU), Uppsala, for very helpful comments. We are also much obliged for helpful suggestions to earlier versions of this paper from seminar participants at the meetings of the North American Regional Science Council, Vancouver, 2017 (especially Spyridon Stavropoulos, Erasmus University, and Frank Neffke, Harvard University), as well as seminar participants at the Ratio Institute, Stockholm, and the 6th Swedish conference in Economics, Växjö, 2018. Excellent research assistance was provided by Elias Hockerfelt, Pernilla Larsson and Mattias Stålnacke. All potential errors in this work are those of the authors. The research has been supported by a grant from IFAU, Uppsala (research grant no. 184/19).

1. Introduction

In the income inequality literature there has recently been a push to extend the more traditional explanatory approaches (as related to factor input supply- and demand changes, arising from new technology, trade patterns and immigration (Acemoglu & Autor, 2011; Katz, 1999; Wright, Goldin, & Katz, 2009), and highlight how dispersion of firm-level productivity (TFP or output per worker) has been rising over time and how this development closely mirrors the well-known trend of rising wage inequality between workers. For example, Barth, Bryson, Davis, & Freeman (2016) decompose US individual log earnings into dispersion in-between- and within establishments and estimate that in-between dispersion is related to as much as 79 percent of total increase in variance in income among all workers, 1992 to 2007. Similar results are also found in Dunne, Foster, Haltiwanger, and Troske (2004), and Faggio, Salvanes, and Van Reenen (2010).

This extension of the inequality literature rests upon three somewhat different strands of the literature on worker- and firm productivity. Firstly, it is a long since established fact that there is considerable heterogeneity in firm-level TFP, even as measured among observably similar firms and establishments (see Syverson, 2011, for a review). For example, the 90-10 TFP percentile ratio for US manufacturing firms is estimated as being in the order of two, and even larger gaps are found for firms in China and India (Hsieh & Klenow, 2009). Second, this spread in productivity between firms has also been documented as related to differences in wages for workers within these firms (see e.g. Cardoso, 1997; Davis & Haltiwanger, 1991; Skans, Edin, & Holmlund, 2009; Slichter, 1950, among many), but since selection and unobserved heterogeneity among workers is often difficult to capture, researchers have been reluctant to pin-point firm level differences in TFP as causing this variation in wages. Third, the empirical literature on rent-sharing takes a more direct aim on this issue and relates wages for workers within separate firms and industries to various measures of firm profits or rents, and a typical finding in this latter extension of the literature is that a 10 percent increase in value-added per worker leads to somewhere between a 0.5 and 1.5 average percent increase in wages (for a review, see Card, Cardoso, Heining, & Kline, 2018).

However, despite recent progress as to how these wage increases are partitioned between different types of workers, there is still a lack of theoretical work as to how we can explain and model how firm level profits or rents may spill into average firm wages and wages for different worker categories. A notable exception is a rapidly developing literature building on a monopsony framework, first developed Joan Robinson (1933) and seminal work in the same vein by Manning (2003), which seeks to explore how various degrees of monopsony and firm wage-setting power affect the level of wages for different worker categories, and by extension inequality and wage income dispersion. The empirical evidence on the potential effects of firm wage setting power can dived into four categories, research that focuses on either *i)* quit- and recruiting responses to different wage levels; on *ii)* the link between wages and firm productivity; on *iii)* various forms of collusion and firm behavior aimed at suppressing employer mobility between firms, and *iv)* on the effects of concentration of a small number of employers. In sum, all these strands of the literature suggest non-negligible employer wage setting power, and where estimates of this power is most easily quantified in terms of effects on factual wage levels (the first two strands mentioned above) they suggest a mark-down of marginal revenue product of around 20-25% (for recent literature reviews, see Ashenfelter, Card, Farber, & Ransom, 2022; Berger, Herkenhoff, & Mongey, 2022; Lamadon, Mogstad, & Setzler, 2022; Sokolova & Sorensen, 2021).

In this paper, we build upon the last strand of the monopsony literature and analyze to what extent wage income inequality within Swedish local labor markets can be explained by varying degrees of employer concentration within local labor markets, both as measured across the urban hierarchy as related to cross-sectional differences in within- and between industry diversification, and over time. As we will further argue below, our paper hereby addresses an evident gap between two still largely separate literatures: On the one hand, the above-mentioned research efforts to address the root causes of macro level changes in dispersion in wages and income, now extended to include firm level factors related to firm productivity and employer concentration. On the other, the vast and still growing literature in regional science as concerns the causes and effects of agglomeration and explanations of the so-called urban wage premium (i.e., why larger cities pay more, as further discussed below). In this regard, our paper extends recent work both by Rinz (2022) which addresses employer concentration within US commuting zones and its effects on changes in both mean wage levels and inequality, as well as research on how varying degrees of local employer concentration can help explain the urban wage premium (Hirsch, Jahn, Manning, & Oberfichtner, 2022).

For comparative purposes, we start our analysis by first estimating the Swedish urban wage premium using common estimation methods in the literature, controlling for observable and unobservable individual worker characteristics, and testing to what extent the remaining variation in worker renumeration varies with local population size. By local industry and local labor market, we then use an AKM modelling framework to separate between worker- and firm-level fixed effects and analyze *i)* to what extent the average firm pay premium (enabled by higher firm productivity) varies with local labor market population size, and *ii)* to which degree employer concentration within local industries can help explain the variation in such firm pay premia. Finally, we address wage income inequality by applying our model (used in the previous steps of the analysis) to different income segments of the local labor market income distribution.

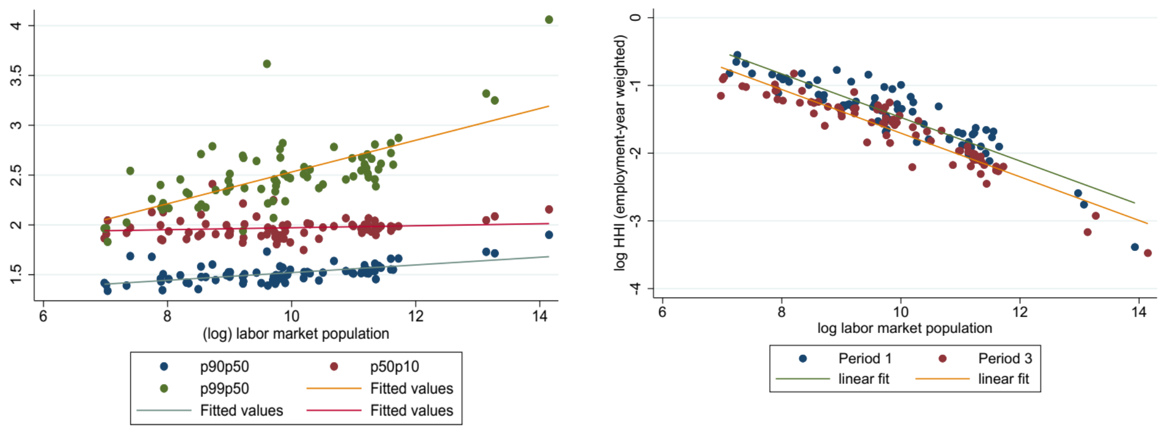
Our results suggest that, firstly, the urban wage premia (UWP) as measured by the firm-pay premia (firm fixed effects) is lower than basing the estimates on more standard estimators such as the Mincer equation. Second, the UWP variation across the urban population distribution is however the same; it is lower in smaller cities and labor markets but increases in a non-linear fashion with urban population size. It is also larger the higher the worker income, regardless of where within the urban population size distribution that we put our focus. In other words, we find higher levels of firm pay premia (and firm rent sharing) the larger the city or local labor market, and the higher the income of workers. Third, these higher levels of firm pay premia/rent-sharing are associated with lower labor market employer concentration, and vice versa, something which affects average wage income among firms accordingly. In fact, according to our estimates, reduced employer concentration explains all of the average UWP (firm pay premia) for all city size categories except for the three major metropolises. In these larger cities we have some remaining unexplained variation, but reduced employer concentration still accounts a sizeable share of the average UWP. Fourth, addressing wage income inequality by applying our model to different segments of the local labor market income distribution, we find that reduced average employer concentration in larger cities accounts for almost all variation in the (positive) link between city-size and wage inequality, except for the largest metropolises where it captures around 30-50 percent of variation depending on the income segment that we focus on.

Our paper is organized as follows, section 2 discusses our data and descriptive results as regards cross sectional differences in income inequality and labor market industry diversity and labor market concentration. Section 3 shows our chosen modelling framework and section discusses our results. Section 5 concludes.

2. Background and descriptive figures

Figure 1 below neatly illustrates our research problem. Starting with Panel A, we see that earnings inequality as measured by percentile ratios at the level of the local labor market increases with local population size. This inequality increase is related to top- and upper income levels, as exemplified by the 99/50 and 90/50 coefficients, whereas bottom level income inequality (the 50/10 ratio) remains constant across the population size distribution.

Figure 1. Earnings inequality (Panel A) and local labor market employer concentration (Panel B), as related to local population size.



Source: Mona database, Statistics Sweden

Resting on the assumption that wages correctly reflect the marginal productivity of workers, explanations of higher wages in larger cities – and by extension, inequality – usually focus on individual level productivity of workers in urban environments. The source of this higher individual productivity is most often related to three basic factors. Either *a)* to learning (sharing of knowledge), i.e., a situation in which human capital accumulation is faster in larger more population dense cities due to facilitated social interaction (Glaeser 1999; Glaeser & Maré 2001; Moretti 2004; Baum-Snow & Pavan 2012; De la Roca & Puga 2012); or to *b)* coordination effects, the "matching hypothesis", which suggests that cities create a context in which there is a better chance of bringing about a good match between workers and firms (Kim 1990; Wheeler 2006; Yankow 2006); or, finally, to *c)* sorting and self-selection, i.e. the notion that relatively higher worker productivity in larger cities is largely due to different types of innate abilities of workers living in and moving into these larger cities (see Combes et al. 2008, 2010; Matano & Naticchioni 2012).

Considerable effort has gone into disentangling these effects from one another, and even though there is still debate, a consensus is emerging that a large share of this urban wage premium can be ascribed to geographical sorting of individuals and differences in underlying worker ability (for overviews, see Rosenthal & Strange 2004; Puga 2010).

Given the increasing evidence that employer wage-setting power is non-negligible in many industries, the implicit underlying model in much of this UWP literature, which views firms essentially as price takers that cannot post wages below market rates without losing all workers, may however be faulty. If so, wages in any given industry may not only reflect workers' individual marginal productivity but also a mark-down from the marginal productivity of workers, a mark-down which in turn is positively corelated with employer concentration.

Returning to Figure 1, Panel B shows how such employer concentration, as measured by the Herfindahl-Hirschman index calculated using firm employment size and the number of firms within all local industries, on average drops by local labor market population size. Increasing inequality and decreasing employer concentration is therefore clearly negatively related to one another. Correlation is of course not causation, but to the extent that the Herfindahl-Hirschman index captures the number of potential employers for job seekers within these local labor markets (further discussed below), the displayed empirical patterns motivate our basic research question: To what extent do higher wages in larger labor markets, and the resulting higher labor market inequality, reflect more competition between employers within these labor markets?

3. Data and variable definitions

We use full population data from Statistics Sweden's Mona database, from which we have access to all individual level data as concerns educational attainment, employment status and place of work, as well as demographic information such as age and marital status. Necessary for the study at hand, our data also includes geo-coded firm- and establishment level employer information, including revenue, value added and industrial classification codes. The data stretches from 1996 to 2015, but since we need to be able to estimate both individual- and firm fixed effects, we choose instead to study three separate six-year time periods, 1996-2001, 2003-2008 and 2010-2015. By choosing shorter periods, the fixed effects may not only include more information, but it also allows us to construct a stacked panel which lets us track changes over time.

Our main unit of analysis are establishments nested within industries which in turn are nested within local labor markets, and for our purposes and the time-period that we analyze, Sweden can be divided into 75 local labor market regions which essentially correspond to commuter zones (Statistics Sweden, SCB). Based on these 75 larger geographical regions, we then define an industry specific local labor market as given by employers (establishments) belonging to the same 3-digit industry. From these industries we exclude the public- and financial sector which leaves us with 258 industries in total, including both manufacturing, services, and construction. Taking all 258 industries and the 75 labor market regions together would leave us with about 19 350 separate industry specific local labor markets. However, since far from all industries are represented in all the 75 regions, we end up with a total number of 11 792 labor markets in the final sample. These industry specific labor markets are distributed unevenly across regions, but as a rule the number of industries increases with local population size (see Figure 2, Panel A), going from the smallest regions with just 32 industries represented to the largest containing 249 separate industry categories. The total number of industry specific local labor market observations hereby amounts to 30 136 over the three time-periods.

Turning our measure of employer concentration, in our main approach we use the Herfindahl-Hirschman index (HHI) as estimated for the number of employees in establishments within each of our industry specific local labor market categories. For a given market, the HHI is commonly defined as the sum of the squared market shares for either firms or establishments within a market. In our case, the HHI is given by

$$HHI_{jkt} = \sum_{i=1}^{m} \left(\frac{Emp_{ijkt}}{\sum_{i=1}^{m} Emp_{ijkt}} \right)^2 \tag{1}$$

where Emp_{ijkt} represents the number of employees in establishment *i*, in industry *j*, and region *k* for time *t*, where *m* gives the number of firms in that market during *t.* Since we seek to model cross sectional relationships for three separate time periods, we opt for calculating each yearly HHI_{jkt} and average these yearly measures over each of our three separate time periods *p*. Hence, our period specific measure of HHI is given by $HHI_{jkp} = \frac{1}{T}\sum_{t=1}^{T} HHI_{jkt}$, where $T = 6$.[2]

[2] As a robustness check, we also calculate an alternative HHI measure based on the *average* number of employees for each firm as well as *average* number of firms over each separate time-period, given by $HHI_{jk\overline{p}} = \sum_{i=1}^{m} \left(\frac{\overline{Emp}_{ijk}}{\sum_{i=}^{m} \overline{Emp}_{ijk}} \right)^2$, where $\overline{Emp}_{ijk}$ gives the average number of employees in firm *i*, in industry *j*, and region *k* within the period *p*. This does however not change the outcome significantly.

4. Empirical strategy

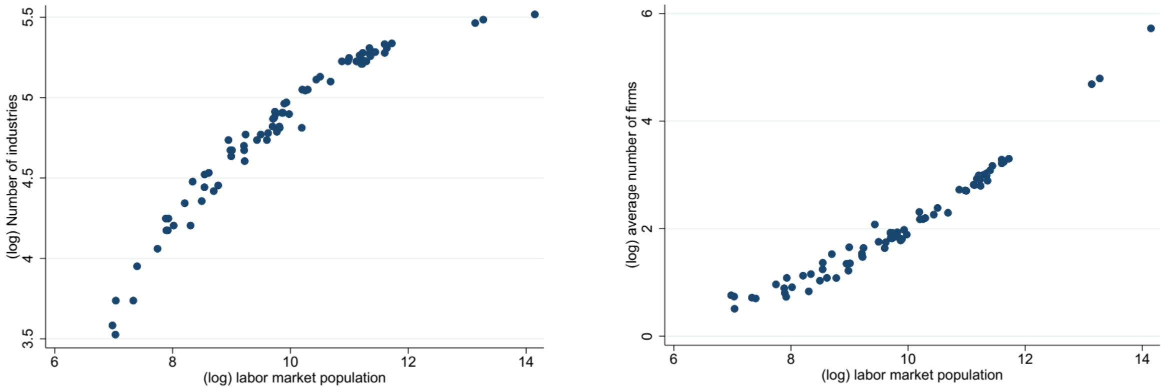
We start our analysis by estimating the general urban wage premium (UWP), across all workers and income categories, and the potential role of labor market concentration in explaining this premium. On this basis, we then move on to address levels of local wage income inequality in a second stage of the analysis.

Since both observed and unobserved individual worker heterogeneity on average varies across regions, and estimation of the UWP aims to capture regional level wage determinants that go beyond such variation, the standard procedure when estimating the UWP is to, firstly, run a Mincer equation that controls for both observed and unobserved worker characteristics while also including a region dummy variable for any income variation not captured by these controls. The estimated regional variation arrived at by way of including this dummy variable is then regressed on population size (or density), and the population size coefficient thus obtained equals the final UWP estimate (see e.g. Combes, Duranton, & Gobillon, 2008, 2010; De la Roca & Puga, 2017).

The traditional approach to UWP estimation does thereby not take firm level factors into account, the (potential) influence of which is either left in the error term or may result in biased regional dummy variable estimates (OVB). For example, it could be that high quality workers are more likely to be employed by high quality firms, firms which are also associated with paying higher shares of their rents to workers. If such high productivity firms are predominately located in more urban areas, we can suspect that the traditional UWP as estimated from the Mincer equation can be biased, and firm level factors risk being confused with individual level characteristics.

In our analysis, we aim to statistically model how local labor market employer concentration may affect average earnings in firms within local labor markets, and how the effect of such concentration varies with urban population size. Our modelling challenge in this regard is illustrated in Figure 2 below, which shows the extent to which both the number of industries (Panel A) and the average number of firms represented within these local industries (Panel B) varies with local population size (where average number of firms also shows large within-industry variation across the urban hierarchy, see Figure A1, Appendix 1). Any modelling approach aiming to capture the potential effect of local employer concentration thus needs to both account for within-industry competition among establishments in local labor markets, as well as controlling for industry fix effects across all industries, regardless of where these industries are situated in the local labor market population distribution.

Figure 2. The number of industries (Panel A) and average number of firms by industry (Panel B), across local labor market population size



Source: Mona database, Statistics Sweden

In our estimation strategy, we therefore choose to extend current urban wage premium estimation approaches by – firstly – using an AKM framework to partition wage income levels into individual worker fixed effects and firm fixed effects (the latter often referred to as the *firm pay premium*, see Abowd, Kramarz, & Margolis, 1999). Second, having controlled for all types of worker characteristics, we then use these firm fixed effects as dependent variable and regress firm pay premia on a categorical dummy variable signifying which industry and region that a firm belongs to (each firm's "industry-by-region" variable), hereby capturing the extent to which firm pay premia varies both by region and industry. Finally, to gauge to what extent firm pay premia embedded in our industry-by-region estimates varies by population size, and to address potential causes of this variation, we then regress our industry-by-region estimates on population size while subsequently adding regional level controls such as within industry employer concentration, our main focus in the analysis.

To compare our estimation strategy to more traditional UWP estimation approaches, however, we begin our modelling approach by estimating a Mincer model at the worker level of log earnings, i.e., a set of individual level-earnings regressions to control for unobserved characteristics at the individual level. Thus, for an individual *i* working in establishment *j(kl)*, which operates in industry *k* in the local labor market *l* at year *t*, we estimate the following specification:

$$ln\,Earnings_{ij(kl)t} = \alpha_i + \beta X_{ij(kl)t} + D_{kl}Market_{klt} + \epsilon_{ij(kl)t}, \quad (2)$$

where α_i and $X_{ij(kl)t}$ capture worker fixed effects and possible time varying observable characteristics, respectively, and $Market_{klt}$ is our industry-by-region dummy variable which takes the value of 1 if an individual works in a firm *j* operating in industry *k* located in labor market *l*. In the presence of individual fixed effects, the effect emanating from workers staying within the same establishment (*jkl*) during the whole period gets absorbed by α_i. Hence, the identification of D_{kl} (our industry-by-region dummy variable) comes from individuals that either (*i*) move between labor markets *l* (to a job for an establishment within the same or a different industry), or (ii) switches to a different establishment within the same labor market but in a different industry.

Due to the large number of industry-by-regions (*kl*-markets) it is however not possible to estimate our D_{kl} variables directly. Rather, we resort to absorbing the dummy variable estimates along with the individual fixed effects when estimating equation (2). The absorbed variable estimates (D_{kl}) are then recovered from the model post estimation. The results from regressing these industry-by-region estimates on local population size (i.e., our first UWP estimate, comparable to results arrived at using more traditional estimators) are shown in Table 1, panel A.

As discussed above, in case there are unobservable characteristics at the establishment level (e.g., firm specific pay premia), when using model (2) such unobserved heterogeneity is contained in the error term, $\epsilon_{ij(kl)t}$, unless appropriately controlled for. To address this ~~problem~~ aspect, we therefore choose to expand the Mincer equation by including the firm fixed effect directly into the model, resulting in the well-known AKM-model which stipulates that log earnings can be written as a linear function of fixed unobserved characteristics at both the worker- and firm level, together with an index of time varying covariates.

Thus, suppose that the error term in equation (2) includes the establishment fixed effect $\psi_{j(kl)}$ (i.e., $\epsilon_{ij(kl)t} = \psi_{j(kl)} + u_{ij(kl)t}$), which after substituting into (2) ~~becomes~~ writes as follows:

$$ln\,Earnings_{ij(kl)t} = \alpha_i + \psi_{j(kl)} + \beta X_{ij(kl)t} + D_{kl}Markets_{klt} + u_{ij(kl)t} \quad (3)$$

In model (3), since the establishments' industry affiliation is fixed by construction (and establishments rarely move across local labor markets) our industry-by-region variable $D_{kl}Markets_{klt}$ is to a large extent captured by the establishment fixed effects, $\psi_{j(kl)}$. In principle, we can estimate $\widehat{D_{kl}}$ by first backing out the firm fixed effect estimate $\widehat{\psi_{J(kl)}}$ from the AKM model, and then regress $\widehat{\psi_{J(kl)}}$ on our industry-by-region dummy variable ($Market_{klt}$), but as before in (2), this is not feasible due to the large number of kl-dummy variables. Luckily however, this modelling procedure is not necessary since we can arrive at the same estimate of $\widehat{D_{kl}}$ by simply taking the average of all estimated firm fixed effects within each industry-by-region category, i.e. averaging $\widehat{\psi_{J(kl)}}$ over all $j = 1, \dots, m_{kl}$ firms.[3] We repeat this procedure for each of the three period estimates of the AKM model, leaving us a stacked panel with at most three period observations for each local industry-by-labor market.

To then model our UWP and the potential effect stemming from labor market employer concentration, we use the following model:

$$\widehat{D_{klp}} = a + b \ln Pop_{lp} + cHHI_{klp} + d_{kp} + e_{kl} \tag{4}$$

where the dependent variable ($\widehat{D_{kl}}$) is the (estimated) average firm fixed effects at the local industry-by-region level (backed out from our estimates using model no. 3 above), and b is an elasticity corresponding to our preferred measure of UWP, representing the percentage change in average firm fixed effects from a percentage change in local labor market population. Following Hirsch et al. (2022), we then estimate the model with and without controls for the Herfindahl-Hirschman index (HHI) and examine the effect of labor market concentration on the UWP by analyzing how b thereby changes.

To conclude, our estimation of the urban wage premium is thus equivalent to the more standard approaches in the literature insofar that it is based on the variation in wage income that cannot be explained by *individual level* observed or unobserved factors. Our extension of these estimation approaches, however, allows for taking observable and unobservable firm level factors into account, and to further analyze how these firm pay premia vary across urban population size and within industries embedded in regions (our industry-by-region variable).[4]

[3] In the data treatment- and modelling do-file syntax that accompanies this paper (available upon request) we show that these two estimation methods yield the same results when applied to a smaller sample of the full population data.

[4] In this setting, all potential variation that uniquely stems from industry category are absorbed by the firm fixed effects because industry categorization is fixed at the firm level. The same holds for local labor markets since the local labor market is defined from the location of the establishment, and not from workers' residence. Thus, in contrast to the Mincer model, our controls for firm fixed effects does not leave us any variation on which to identify industry- or local labor market fixed effects by way of workers moving either between local labor markets but within the same industries, or between industries while remaining within a specific labor market. In the Mincer equation, however, such identification comes exclusively from these types of worker movements.

Causality identification and multicollinearity

There are two potential problems with our second step specification using (4) above. The first is that labor market concentration may be endogenous. Secondly, HHI at the local labor market level and local population size (*ln POP*) are to a certain degree ~~is~~ correlated, which makes estimation of each separate parameter a potential challenge.

As for the first problem, if there are shocks to the local industry which affect both earnings among firms and labor market concentration, HHI becomes endogenous and any attempt to causally interpret our findings becomes problematic. To remedy this problem, following Rinz (2022) and Azar, Marinescu, and Steinbaum (2022), we instrument our HHI_{jklp} by the predicted HHI from a weighted average of industry-by-region specific HHI across all other labor markets, using

$$HHI^{IV}_{jkl\overline{p}} = \sum_{r\neq j} HHI_{rklp} \frac{\overline{Emp}_{rklp}}{\sum_{r\neq j} Emp_{rklp}} \tag{5}$$

where – for time-period *p* – the summation is done over all industry-by-regions *r* except *j*. Our IV estimation strategy thereby represents a typical “leave-one-out” instrument.

To deal with the second potential problem related to multicollinearity, we examine an alternative model where we replace our local population variable *lnPop* with a set of dummy variables which capture where within the urban population structure an establishment resides. More specifically, we consider the following population dummy variables: (1) P_1 < 10 000, (2) 10 000 < P_2 < 100 000, (3) 100 000 < P_3 < 1 000 000, (4) P_4 > 1 000 000, which when substituted for *lnPop* in specification (4) gives us model (6).

$$\widehat{D_{klp}} = a + b_g P_{gp} + cHHI_{klp} + d_k + e_{klp} \tag{6}$$

In doing so we reduce the correlation between *lnPop* and HHI, from -0.51 to at most 0.36 for the population category with the highest correlation, P_1. [5] Bear in mind that in this alternative specification the interpretation of b_g does not strictly translate into the UWP, since it captures average differences in firm fixed effects across population size groups. However, by comparing estimates for different population categories we can infer how much higher the average wage premium is in more populated areas as compared to less populated areas. By comparing

[5] Correlations are weighted using the size of the industry-by-region local labor market. For the smaller size population groups , correlation with HHI is positive with $\mathrm{corr}(P_1, HHI) = 0.36$ and $\mathrm{corr}(P_2, HHI) = 0.23$. For larger population groups the correlation flips from positive to negative, with $\mathrm{corr}(P_3, HHI) = -0.14$ and $\mathrm{corr}(P_4, HHI) = -0.31$.

estimates with and without controls for HHI, this model also allows us to assess the role of HHI in shaping the UWP whilst lessening the potential problem of multicollinearity.

For (6), we also present the results for instrumenting HHI with the "leave-one-out" instrument, defined in (5). These estimates include industry fixed effects, and we thereby control for any differences in pay-premium at the industry level, differences which may be due to e.g., different national or international level of competition, capital intensity or industry specific human capital. All regressions are weighted by the size of the labor force in the kl-market.

Finally, to further probe the role of HHI, we extend the basic specification in (6) by incorporating the interaction between HHI and our population size dummy variables into the model (thereby allowing the UWP to be dependent of the level of HHI). Extending the specification in (6) we thus also consider the following model,

$$\widehat{D_{klp}} = a + b_g P_{gp} + cHHI_{klp} + \eta_g P_{gp} \cdot HHI_{klp} + d_{kp} + e_{klp}, \tag{7}$$

where η_g equals the change in $\widehat{D_{kl}}$ from a unit change in HHI for the population group g. If e.g., the pay premium is larger in more populated areas, we expect that less concentrated industries will have higher pay premiums compared to the more concentrated industries in the same region.

5. Results

In this section, we present results for the estimates of the urban wage premium and local labor market concentration at the level of industry-by-regions for the periods 1996-2001, 2003-2008 and 2010-2015. Table 1 shows the estimates for average earnings in Panel A (using model 2), and average firm fixed effects in Panel B from the first stage regressions using model 3. The first column in both panels shows the correlation between labor market population for the respective dependent variable, whereby the estimates represent an elasticity corresponding to either measure of the urban wage premium.

In column (2), we show the correlation between each of our two earnings measures and labor market concentration as given by the HHI. To the extent the UWP depends on the level of concentration in industries within the local labor market, we estimate the UWP in column (3) when controlling for the HHI. To deal with potential endogeneity of HHI, in column (4) we show estimates the same model but instrumenting for HHI using the "leave one out" instrument as defined in specification (5).

Starting with our UWP estimates in column (1), we see that using the earnings measure (Panel A) results in about three times as large an estimate as compared to using firm fixed pay premia in B (0.043 compared to 0.014). In other words, it is clear that a large share of the UWP as measured using more traditional approaches is related to firm level factors which influence wages. The elasticities imply that a 1 percent increase in local population is associated with 0.043 and 0.014 percent increase in earnings and the firm wage premium, respectively. Using Panel B estimates, and going from a population size of e.g., 10 000 to 100 000 and to 1 000 000, translates into a premium of 0.14 percent and 1.4 percent for each tenfold increase in the underlying size of the local labor market.[6]

Table 1. Results for the urban wage premium and labor market concentration

Panel A: Estimates for average earnings

	(1)	(2)	(3)	(4)
Dependent variable: average earnings (log)	OLS	OLS	OLS	IV HHI
Labor market population (log)	0.0430***		0.0430***	0.00817
	(0.00320)		(0.00207)	(0.00944)
HHI		-0.203***	0.000389	-0.447***
		(0.0126)	(0.0148)	(0.125)
Constant	12.11***	12.68***	12.11***	
	(0.0394)	(0.00445)	(0.0259)	
Observations	30,135	30,135	30,135	30,127
R-squared	0.878	0.838	0.878	-0.044

Panel B: Estimates using average firm fixed effects from first-step regressions

Dependent variable: average firm fixed effects	OLS	OLS	OLS	IV HHI
Labor market population (log)	0.0141***		0.0163***	0.00468*
	(0.00105)		(0.000843)	(0.00256)
HHI		-0.0486***	0.0282***	-0.121***
		(0.00471)	(0.00635)	(0.0305)
Constant	-0.173***	0.00968***	-0.206***	
	(0.0129)	(0.00162)	(0.0109)	
Observations	30,135	30,135	30,135	30,127
R-squared	0.645	0.611	0.647	-0.013

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[6] We have also estimated results in Panel B including time-variant firm level controls such as the average educational share of employees divided in 5 categories from basic to tertiary education; the share of female workers; and the share of workers from non-OECD countries (all control variables represent averages for the given period). These additional controls only marginally affect the magnitude of our estimates (tables available upon request).

Turning to industry specific local labor market concentration (in column 2) there is a negative correlation between both HHI and our individual earnings' measure in Panel A and HHI and the firm pay premia in Panel B. For earnings in Panel A, the correlation is substantially larger, -0.201 as compared to -0.049 for the firm pay premium. Thus, traversing the full range of the HHI from 0 to 1 (from low to high employer concentration) is associated with lower earnings, by approximately 4.9 percent in terms of the firm pay premium.

In column (3), when controlling for both population size as well as the level of labor market concentration, the UWP increases slightly with the coefficient on HHI turning positive. At the face of it, labor market concentration thus seems to be of limited importance for the UWP. As discussed in the previous section, there are however two concerns with the basic model specification used for these estimates. One is that HHI is potentially endogenous since local supply shocks may also be correlated with average earnings and our HHI estimates. The other problem is the degree to which local population size and labor market concentration are correlated with each other, potentially causing imprecise estimates due to high levels of multicollinearity. With a correlation between the two at about 0.5 this is likely a concern.

In Table 1 we address the first of these two concerns by instrumenting HHI with the "leave-one-out" instrument given in (5) above. As a result, we recover a negative coefficient on HHI, which more than doubles in size in both panels A and B. For the UWP on the other hand, instrumenting for HHI reduces the UWP to a point estimate of 0.008-0.005, statistically insignificant in Panel A and barely so in Panel B.

As for the second concern, multicollinearity is however likely a problem in these estimates. Moreover, we do not necessarily expect the UWP to be the same across the entire urban hierarchy and UWP estimates in previous Swedish studies (based on outcomes for internal migrants) have also been shown to vary in this regard (Ahlin, Andersson, & Thulin, 2014; Korpi & Clark, 2019). To address these issues, we therefore next examine model (6) where we replace our local population variable *lnPop* with four dummy variables capturing where within the urban population structure that an establishment resides.

The results are presented in Table 2, where column 1 shows the UWP (based on firm pay premia, our preferred measure) only including our population categories, column 2 shows the correlation with the Herfindahl-Hirschman Index (identical to column 2 in Table 1, Panel B), column 3 shows the results of adding the Herfindahl-Hirschman Index to our estimates using local population size categories, and column 4 shows the equivalent results using the instrumented HHI variable. In Column 3, we observe a similar pattern as we saw in Table 1 (Column 3) regarding estimates using the non-instrumented Herfindahl index while controlling for population size. Just as in Table 1, our HHI estimate turns positive when adding local population size (in this case our labor market population dummy variables). As previously in Table 1, in Column 4 our HHI estimate turns negative once we account for the potential endogeneity of the outcomes. We can also note that this is a somewhat smaller estimate for HHI as compared to when using our instrumented HHI in Table 1, at -0.066 log points as compared to -0.12 (Table 1, Column 4).

As for our regional dummy variables, when comparing columns 3 and 4 in Table 2, we also conclude that our pay premium estimates turn insignificant for population size categories 2 and 3 and is significantly reduced (by about half) for category 4, once we control for labor market concentration using our instrumented HHI variable (a finding very much in line with those found in Hirsch et. al., 2022, employing a similar approach using German data). It is therefore only among the largest labor markets (with a population exceeding 1 000 000 residents) where we find a positive and significant average pay premium which is not fully captured by our employer concentration variable, at 0.044 log points higher as compared to the reference category (the smallest labor market category)

Table 2. Urban wage premium from second-step regressions using local population categories.

VARIABLES	(1) OLS	(2) OLS	(3) OLS	(4) IV HHI
HHI		-0.0486***	0.0127**	-0.0663***
		(0.00471)	(0.00613)	(0.0209)
logRank = 2	0.0219***		0.0255***	0.00325
	(0.00445)		(0.00488)	(0.00743)
logRank = 3	0.0422***		0.0478***	0.0132
	(0.00448)		(0.00522)	(0.0101)
logRank = 4	0.0761***		0.0823***	0.0436***
	(0.00485)		(0.00617)	(0.0116)
Constant	-0.0436***		-0.0511***	
	(0.00433)		(0.00568)	
Observations	30,135	30,135	30,135	30,127
R-squared:	0.649	0.611	0.649	0.085

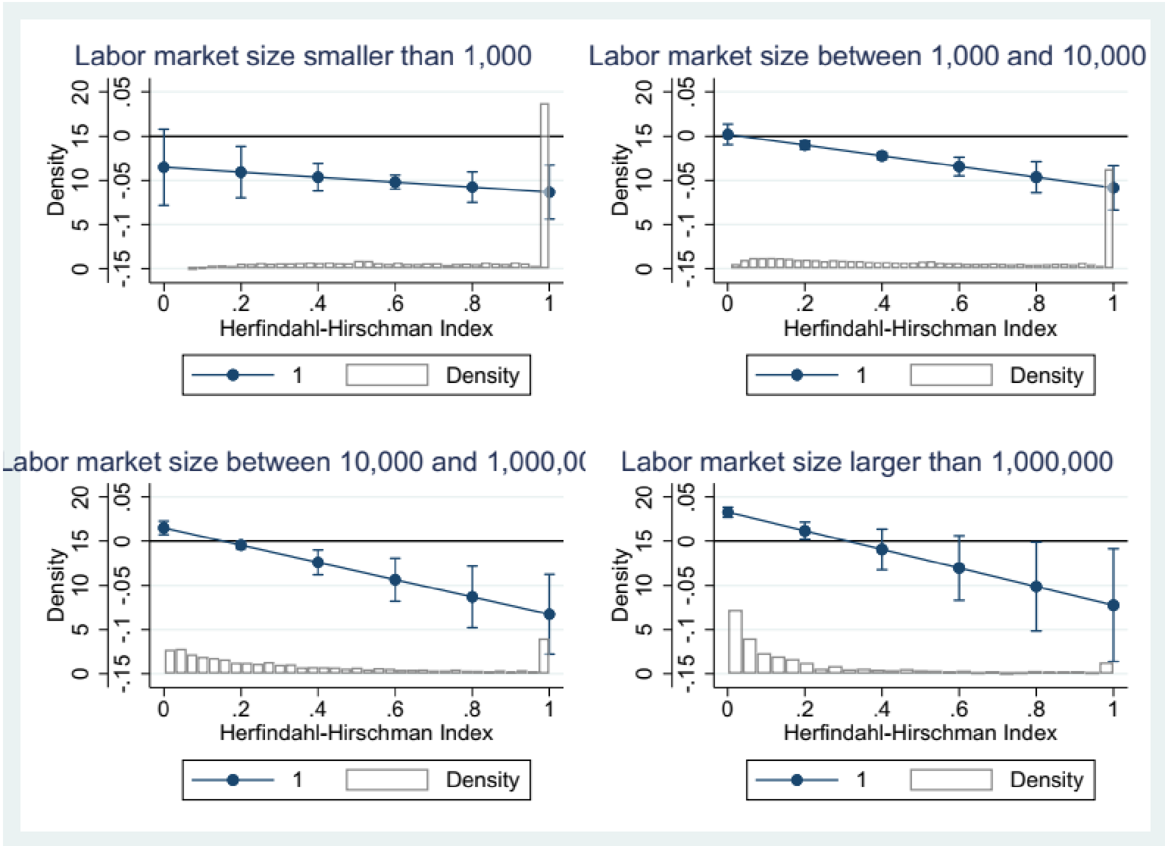
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Fallowing Hirsch et. al. (2022), our tentative conclusion thus far is therefore that the UWP (i.e., the UWP that goes beyond individual level productivity and is captured by firm-fixed effects) can to a large extent be explained by higher labor market concentration for smaller labor markets (population size dummy categories 1-3), and also significantly explains – reduces by about half – the UWP the coefficient estimate in population category 4 (the largest labor markets).[7]

Interaction effects

Our conclusion in this regard is further strengthened from estimating model (7), where we add variables for the interaction between our population categorical variables and our Herfindahl-Hirschman index. The model allows us to analyze to what extent our HHI estimates vary (within the different population categories) by the level of employer concentration within local industries.

Figure 3. Interaction effects: Herfindahl-Hirschman Index and the Urban Wage Premium using local population size categories.



Source: Mona database, Statistics Sweden

[7] To the extent that our dummy variable approach is successful in dealing with multicollinearity, we argue that the change in the sign of our HHI estimates when accounting for possible endogeneity of our outcomes, is first and foremost due to endogeneity rather than multicollinearity. With that said, multicollinearity is likely a factor affecting our results, and for this reason, we continue using the labor market groups to capture the UWP as our main approach going forward. To additionally probe this issue, we also deploy a spline regression setup, otherwise estimating the same models as in Table 1 (see Table XXX, appendix 1. NOTE: TO BE ADDED)

As illustrated in Figure 3 above, by plotting the results this way we readily see that the effect of labor market concentration on the UWP is negative all throughout, that is, the higher the value of the HH index, the lower the firm pay premium. We can also see that the positive UWP estimate in the largest population size category (4), stems from industries where local firm concentration is relatively low (i.e., for low values on the X-axis), and further, that the negative effect for the lower population size categories (1-3) is to a large extent driven by relatively more numerous industries with high HHI estimates. All these results strengthen our conclusion as regards the effect of employer concentration on the UWP, and that lower average UWP estimates estimated for smaller labor markets to a substantial degree is driven by higher employer concentration within these local labor markets.

Inequality and employer concentration

Next, to address the potential role of employer concentration in explaining local levels of wage income inequality, we estimate model no. 6 using four subsets of our sample, either corresponding to the tails of the distribution (below the 10^{th} and above the 90^{th} percentile), or to two broader income segments below and above the 50^{th} percentile (10<50, 50>90). Specifically, we test whether the role of employer concentration in explaining the urban wage premium, as detailed above in Tables 1 and 2, pertains only to certain segments of the income distribution.

For each subset of our sample, we estimate two versions of model no. 6. In the first we control only for our local population size categorical variables while our control for labor market employer concentration (the instrumented Herfindahl-Hirschman index) is added in the second. Focusing on results using the first of these two estimators, shown in Columns 1, 3, 5 and 7 in Table 3, we can readily see that while the UWP (as measured by the firm pay premium) pertains to all segments of the income distribution. It is however clearly larger the further up in the distribution that we put our focus, and the UWP is also larger within larger population size categories for all parts of the income distribution.

Turning to our IV estimates, in Columns 2, 4, 6 and 8, we find that the negative effect of our Herfindahl-Hirschman index (row 1) is significant all throughout. For all population size categories, this negative effect is also larger the further up the income distribution that we focus on, changing from -0.051 for our subset below the 10^{th} percentile, to -0.0755 for our subset above the 90^{th}. Noteworthy also, adding our control for employer concentration renders

Table 3. Urban wage premium from second-step regressions population categories

	<10		10>50		50>90		>90	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES:	OLS	IV_HHI	OLS	IV_HHI	OLS	IV_HHI	OLS	IV_HHI
H3_year		-0.0509**		-0.0658***		-0.0683***		-0.0746***
		(0.0221)		(0.0212)		(0.0211)		(0.0238)
logRank = 2	0.0130***	-0.00136	0.0197***	0.00115	0.0240***	0.00478	0.0306***	0.00957
	(0.00450)	(0.00781)	(0.00447)	(0.00754)	(0.00455)	(0.00745)	(0.00482)	(0.00818)
logRank = 3	0.0258***	0.00350	0.0367***	0.00790	0.0468***	0.0169*	0.0615***	0.0288**
	(0.00456)	(0.0106)	(0.00449)	(0.0103)	(0.00459)	(0.0102)	(0.00492)	(0.0113)
logRank = 4	0.0471***	0.0221*	0.0648***	0.0325***	0.0847***	0.0512***	0.115***	0.0785***
	(0.00495)	(0.0121)	(0.00486)	(0.0116)	(0.00499)	(0.0117)	(0.00536)	(0.0131)
Constant	-0.0733***		-0.0517***		-0.0324***		-0.0256***	
	(0.00438)		(0.00435)		(0.00442)		(0.00470)	
Observations:	30,135	30,127	30,026	30,018	29,508	29,500	30,135	30,127
R-squared:	0.640	0.005	0.648	0.039	0.627	0.111	0.598	0.184

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

most of our population size categorical variables either insignificant, or significantly reduces their coefficient size (see population category 4). When focusing on the 10>50 segment of the income distribution, adding our instrumented HHI variable to model 6 reduces the UWP estimate for population category 4 by about half, and by around 30 percent both for sample subset 50>90 and above the 90th percentile.

Thus, in line with our findings in Tables 1 and 2 comparing our UWP estimates with and without controlling for employer concentration, our tentative conclusion in that employer concentration explains most of the UWP, and that it is only for local population category 4, and category 3 above the 90th percentile, where we find a positive and significant UWP that is not fully accounted for by our employer concentration variable.

6. Concluding discussion

In this paper, we build upon recent progress within the literature on firm productivity, rent sharing and firm wage setting power and analyze to what extent the urban wage premium and wage income inequality within Swedish local labor markets can be explained by varying degrees of local labor market employer concentration, both as measured across the urban hierarchy as related to cross-sectional differences in within- and between industry diversification, and over time. As we argued by way of introduction, our paper hereby addresses an evident gap between two still largely separate literatures: On the one hand, the many research efforts to address the root causes of macro level changes in dispersion in wages and income, now extended to include firm level factors related to firm productivity and employer concentration. On the other, the vast and still growing literature in regional science as concerns the causes and effects of agglomeration and explanations of the so-called urban wage premium (i.e., why larger cities pay more).

We start our analysis by addressing the urban wage premium and its potential links to varying degrees of employer concentration in local labor markets. Instead of using a more traditional approach of estimating the urban wage premium (UWP) by way of a Mincer equation, explaining wage levels in terms of observable and unobservable individual level characteristics, we build upon an AKM-framework (Abowd et al., 1999) which also allows for estimating the contribution of firm fixed effects (so-called firm pay premia) when addressing the causes of different wage levels.

Thereby also controlling for individual level characteristics, we find that the contribution from firm level factors (firm productivity and rents) to average wage income increases with local labor market size (in the order of around 0.014 percent increase for every 1 percent increase in population size). Further, when we explore to what extent this average urban wage premium pertains to different segments of the local income distribution, we find that it is larger for workers with higher wages and that it increases non-linearly with labor market size (i.e., the UWP is larger for bigger metropolitan areas).

In terms of explaining the empirical findings, our results suggest that reduced employer concentration in larger cities captures (and explains) most of the UWP in labor markets with population sizes lower than the three major metropolitan areas, i.e., in cities with working age populations under 1000 000 inhabitants. In cities with local populations above that threshold, it explains around half of the UWP for income levels below the 50^{th} percentile of the local income distribution, and around 30 percent of the UWP above the 50^{th}.

The results are in line with recent studies which point to increasing dispersion in firm productivity and rent sharing as a potentially important factor when explaining changes in wage inequality over time (as suggested in e.g., Barth et.al., 2016). The results also strongly suggest that employer concentration is important for understanding regional income differences and the urban wage premium, a hitherto largely overlooked factor in this context.

References:

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

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics, 4*, 1043-1171.

Ahlin, L., Andersson, M., & Thulin, P. (2014). Market Thickness and the Early Labour Market Career of University Graduates: An Urban Advantage? *Spatial Economic Analysis, 9*(4), 396-419.

Ashenfelter, O., Card, D., Farber, H., & Ransom, M. R. (2022). Monopsony in the Labor Market: New Empirical Results and New Public Policies. *Journal of Human Resources, 57*(S), S1-S10. doi:10.3368/jhr.monopsony.special-issue-2022-introduction

Azar, J., Marinescu, I., & Steinbaum, M. (2022). Labor market concentration. *Journal of Human Resources, 57*(S), S167-S199.

Barth, E., Bryson, A., Davis, J. C., & Freeman, R. (2016). It's where you work: Increases in the dispersion of earnings across establishments and individuals in the United States. *Journal of Labor Economics, 34*(S2), S67-S97.

Berger, D., Herkenhoff, K., & Mongey, S. (2022). Labor market power. *American Economic Review, 112*(4), 1147-1193.

Card, D., Cardoso, A. R., Heining, J., & Kline, P. (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics, 36*(S1), S13-S70.

Cardoso, A. R. (1997). Workers or employers: who is shaping wage inequality? *Oxford Bulletin of Economics and Statistics, 59*(4), 523-547.

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics, 63*(2), 723-742.

Combes, P.-P., Duranton, G., & Gobillon, L. (2010). The identification of agglomeration economies. *Journal of economic geography, 11*(2), 253-266.

Davis, S. J., & Haltiwanger, J. (1991). *Wage dispersion between and within US manufacturing plants, 1963-1986*. Retrieved from

De la Roca, J., & Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies, 84*(1), 106-142.

Dunne, T., Foster, L., Haltiwanger, J., & Troske, K. R. (2004). Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment. *Journal of Labor Economics, 22*(2), 397-429.

Faggio, G., Salvanes, K. G., & Van Reenen, J. (2010). The evolution of inequality in productivity and wages: panel data evidence. *Industrial and Corporate Change, 19*(6), 1919-1951.

Hirsch, B., Jahn, E. J., Manning, A., & Oberfichtner, M. (2022). The urban wage premium in imperfect labor markets. *Journal of Human Resources, 57*(S), S111-S136.

Hsieh, C.-T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics, 124*(4), 1403-1448.

Katz, L. F. (1999). Changes in the wage structure and earnings inequality. *Handbook of labor economics, 3*, 1463-1555.

Korpi, M., & Clark, W. A. V. (2019). Migration and occupational careers: The static and dynamic urban wage premium by education and city size. *Papers in Regional Science, 98*(1), 555-574. doi:doi:10.1111/pirs.12328

Lamadon, T., Mogstad, M., & Setzler, B. (2022). Imperfect competition, compensating differentials, and rent sharing in the US labor market. *American Economic Review, 112*(1), 169-212.

Manning, A. (2003). *Monopsony in motion : imperfect competition in labor markets*. Princeton, N.J.: Princeton University Press.

Rinz, K. (2022). Labor market concentration, earnings, and inequality. *Journal of Human Resources, 57*(S), S251-S283.

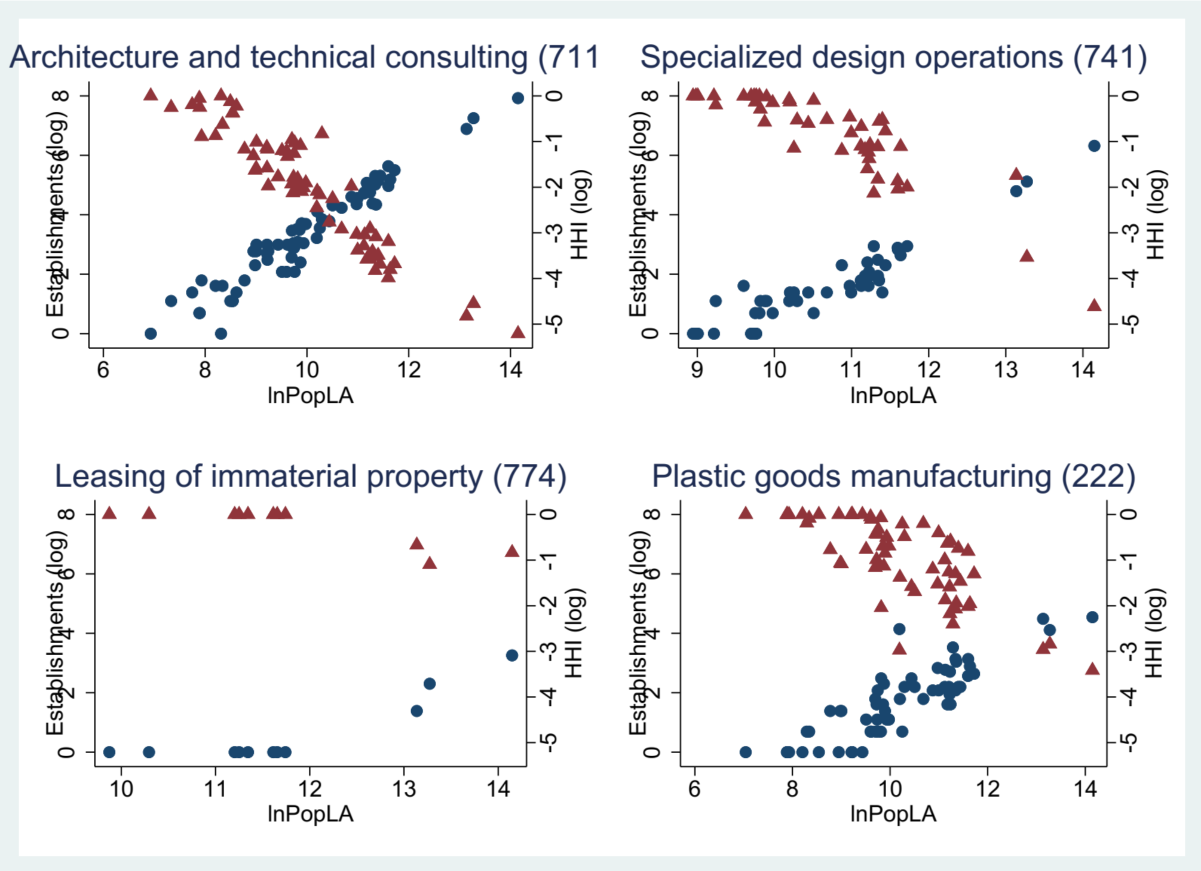
Robinson, J. (1933). *The Economics of Imperfect Competition*. London: MacMillan.

Skans, O. N., Edin, P.-A., & Holmlund, B. (2009). Wage dispersion between and within plants: Sweden 1985-2000. In *The structure of wages: an international comparison* (pp. 217-260): University of Chicago Press.

Slichter, S. H. (1950). Notes on the Structure of Wages. *The review of economics and statistics*, 80-91.
Sokolova, A., & Sorensen, T. (2021). Monopsony in labor markets: A meta-analysis. *ILR Review*, *74*(1), 27-55.
Syverson, C. (2011). What determines productivity? *Journal of economic literature*, *49*(2), 326-365.
Wright, G., Goldin, C., & Katz, L. F. (2009). The Race between Education and Technology. In: JSTOR.

Appendix:

Figure A1. Examples of between-industry variation in the number of local firms represented within a certain industry (y-axis, left), and the local Herfindahl-Hirschman Index (y-axis, right) as related to local population size (x-axis).



NOTE: Tringles (red) signify the log Herfindahl-Hirschman Index, and dots (blue) the log number of local establishment represented within the industry.