The Local Root of Wage Inequality

Hugo Lhuillier*
University of Chicago

Preliminary draft. Do not circulate.

Abstract

Wages vary substantially between and within cities. While wages are on average higher in larger cities, the real earnings of low-wage workers are lower. Using French matched employer-employee data, I document two novel facts that highlight the role of employers in shaping between- and within-city inequality jointly. First, high-paying jobs are concentrated in large cities whereas low-paying jobs are present throughout France. Second, the wage gains offered by large cities materialize over time as workers reallocate across employers. I propose a spatial framework that rationalizes these facts through two ingredients: heterogeneous employers and on-the-job search. Productive employers agglomerate in large cities to hire more workers. Fiercer competition for workers arises and pushes relatively productive employers to offer high wage. Higher average wage, faster growth, and greater within-city inequality follows. I estimate the model and find that employer sorting endogenously generates the two novel facts. The steeper ladder of large cities implies higher real lifetime earnings despite lower initial real wages. Finally, I quantify that the spatial concentration of workers boosts aggregate TFP and wages while increasing overall wage inequality.

*First version: August 2022. This version: April 2025. This article was born under the supervision of my advisors at Princeton University. I thank Gene Grossman, Richard Rogerson, Esteban Rossi-Hansberg and Gianluca Violante for their support and guidance in this project. I am indebted to Adrien Bilal and Gregor Jarosch for several insightful discussions. I am also grateful to Cécile Gaubert, John Grisby, Pablo Fajgelbaum, Eduardo Morales, Jonathan Payne, Charly Porcher and Steve Redding for valuable comments. This research benefited from financial support from the International Economics Section and the Simpson Center for Macroeconomics at Princeton University. Their support is gratefully acknowledged. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17—Centre d'accès sécurisé aux données—CASD). Special thanks to Marine Doux and Médianes for hosting the physical data in France.

[M]any urban dwellers suffer from extreme inequality. In a world with high and growing levels of urbanization, the future of inequality largely depends on what happens in cities [...].

United Nations, *World Social Report 2020*

Wages vary substantially across space. In urban hubs, average wages are high, and so are inequalities. There, high-income earners live alongside low-wage workers who face the high costs of living. To be concrete, Figure 1 plots the gross hourly wage distribution of each French commuting zone. The left panel covers nominal wages, and the right panel deflates wages by a citywide housing price index. In Paris, the average wage is 65% higher than in Lens, a mid-size city at the median of the wage distribution. Yet, considerable heterogeneity lies beneath these average gaps. Workers in the top 10% of Paris' distribution earn wages 112% higher than those at the top in Lens. This gap is 2% for workers in the bottom 10%. In fact, after accounting for housing prices, low-wage workers in Paris earn real wages 20% lower than their counterparts in Lens.

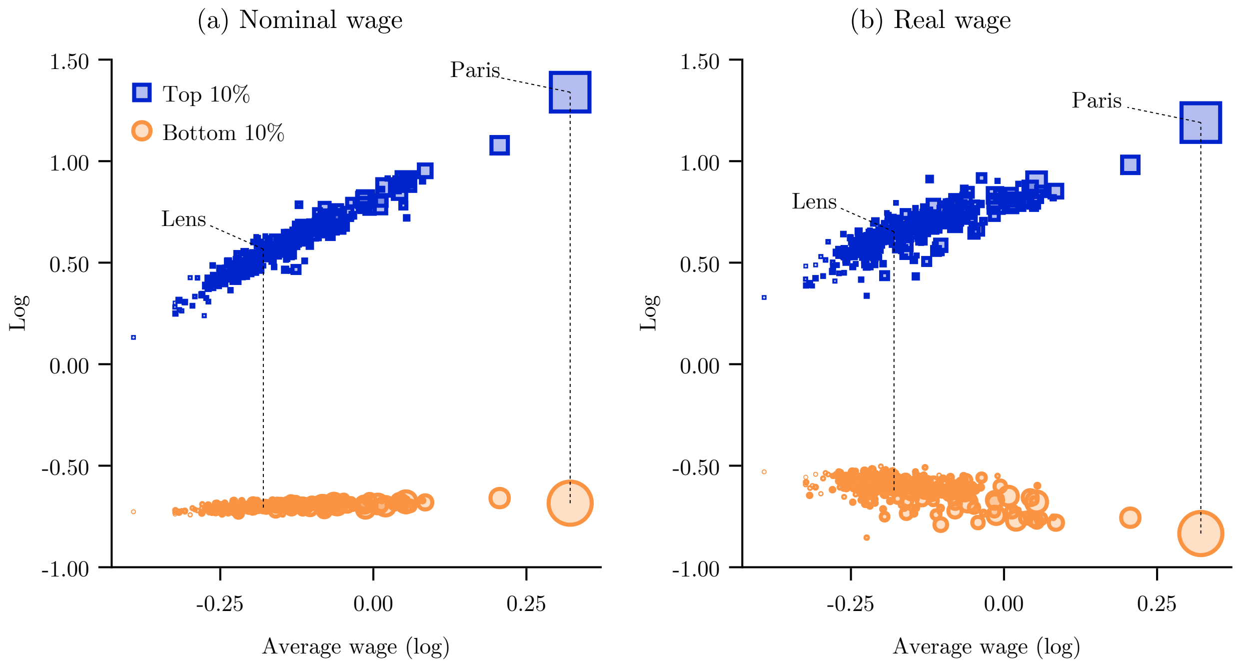
These large disparities have raised concerns among pundits and policy circles that modern metropolises, and their inner making, are at the root of national inequality (e.g., United Nations, 2020; Goldin, 2023). However, while within-city inequality explains the bulk of the aggregate wage dispersion, little is known as to its drivers. What explains the greater wage inequality of large cities? Why are workers accepting lower real wages there? What does within-city inequality reveals about the impact of cities on wages? And what are the consequences of agglomeration on wages and inequality?

This paper offers answers to these questions in four parts. First, I document that employers are instrumental in shaping local wages. Second, I provide a framework that generates spatial wage disparities through the sorting of employers across frictional local labor markets. Third, I estimate the model and show it rationalizes between- and within-city inequality. Fourth, I quantify that the spatial concentration of workers boosts aggregate wages and wage inequality.

Specifically, in the first part of the paper, I leverage French matched employer-employee data to document two novel facts about the importance of employers for local wages. First, I show that the concentration of high-paying jobs in a few large locations explains 31% of the between-city wage variance, and 33% of the variation in within-city inequality. To quantify the relative importance of employers, I estimate a mover design à la Abowd et al. (1999). I find that workers in cities twice larger earn employer wage premia 3.2% higher on average. However, these between-city differences in premia are concentrated amongst high-paying employers. Wage premia are -0.07% smaller in cities twice bigger when workers are in the bottom 10% of their local premia distribution. By contrast, they are 4.4% larger for workers in the top 10%. The spatial concentration of high-paying employers, together with the uniform presence of low-paying jobs throughout France, spurs higher and more dispersed premia in larger cities.

Second, I document that the wage gains offered by larger cities occur over time as workers

Figure 1: Gross hourly wage by commuting zone



Data source: French matched employer-employee (see Section 1.1). In both panels, each dot is a commuting zone. Panel (a) displays the average wage in the bottom 10% (orange circle) and top 10% (blue rectangle) the city wage distribution. Panel (b) displays the same statistics for real wages. Real wages are computed as nominal wages deflated by a citywide Cobb-Douglas price index with a housing expenditure share of 0.3. Nominal and real wages are normalized by their respective national average.

reallocate across employers. I estimate that starting wages are very similar across locations once netting out worker heterogeneity. For instance, the startup premia in Paris is only 0.5% higher than in Lens. However, the returns to job switching are higher in larger cities. I quantify that the wage gains upon a job switch are 0.3 percentage points higher in cities twice bigger (relative to a mean of 1.3%). This pattern holds after accounting for the sorting of heterogeneous workers across space and the impact of locations on learning. It also holds throughout the wage distribution and across occupations. As a result, substantial between-city differences in lifetime income arise. By age 55, I estimate that workers in Paris earn wages 13% higher than those in Lens.

Altogether, these two facts underscore the key role of employers in shaping local wages. They also provide new testable implications for frameworks that model spatial wage inequality. In the second part of the paper, I build such a theoretical framework that ties these two facts together through the sorting of employers across frictional labor markets.

Workers are ex-ante homogeneous. They choose where to reside to maximize their lifetime utility, which depends on the net present value of their expected lifetime income, housing prices, and local amenities. Workers continuously search for better-paying opportunities: they start their careers unemployed and progressively experience wage growth by climbing local job ladders. Employers differ in productivity and decide where to produce and hire workers. Search frictions constrain their effective size. As in Burdett and Mortensen (1998), wages are an effective hiring tool: offering higher wages allows employers to attract and retain more workers from the local competitors.

In equilibrium, productive employers agglomerate in large cities. Their location decision depends on three considerations: how many workers live there, how intense is the local competition as captured by the local wage offer distribution, and how expensive is commercial housing. Complementarity in production between productivity and size implies that productive employers have a stronger willingness to produce in larger locations. Sorting occurs despite the absence of TFP gaps because large cities allow employers to partially sidestep search frictions. Endogenous local competition dampens the size advantage of large cities: as productive employers agglomerate there, competition intensifies, and employers are required to offer higher wages to reach their target size. This offsetting competition channel implies that a range of mid-productivity employers are indifferent between producing in large cities with fierce competition or in small locations where it is milder when search frictions are not too strong.

The fiercer local competition in larger locations generate higher wages, faster wage growth, and greater local inequality. The fiercer competition indeed requires productive employers, for which a large share of their workforce comes from poaching competitors, to offer higher wages. By contrast, relatively unproductive employers, who hire most of their workforce from unemployment, are immune from the local competition. As a result, wages are on average higher in larger cities, but there are greater spatial variations in the right tail of the wage distributions than in the left. The job ladder of large locations steepens, job switching yields higher wage gains, and as consequence, workers initially accept lower real-paid jobs in anticipation of higher future real earnings.

In the third part of the paper, I estimate a quantitative version of the model and assess whether it can generate the spatial wage disparities observed in the data.

I extend the model in three significant ways. First, I allow cities to differ in Total Factor Productivity (TFP). Local TFPs capture exogenous fundamentals and reduced-form productivity spillovers as in standard quantitative spatial models (Redding and Rossi-Hansberg, 2017). Second, employers face flexible curvature in job creation cost functions that let them grow without raising wages too rapidly. They also face idiosyncratic entry costs to capture unobserved heterogeneity in location choice. Third, I add two additional reasons for why workers may accept lower real earnings in large cities: idiosyncratic preferences, and migration costs. The model is estimated on the employer fixed effects, which allows me to abstract from worker heterogeneity.

I provide a proof of identification for 16 of the 21 parameters by combining the strategy used in quantitative spatial models and wage-posting settings (Bontemps et al., 2000). The key part of the estimation consists of separately identifying the four channels shaping employer sorting: city size, search frictions, the dispersion in entry costs, and local TFPs. City sizes are readily observable. Search frictions are identified off workers flows in- and out- of employment and across employers. The dispersion in entry costs is estimated from the sensitivity of employers' location choice to local profit opportunities, where profits are computed from firm-level data on location, size, and wages, together with the structure of the model. Finally, TFPs are recovered as residuals to match each location's average wage net of the employer sorting predicted by the model.

I find that employer sorting drives wage disparities across space. My estimates of search frictions and entry costs dispersion align with those in the literature. Given these, productive employers have strong incentives to locate in large cities to maximize their size. This productivity selection suffices to explain the average wage gaps across commuting zones. As a result, local TFPs only account for 1.55% of the between-city wage variance. I estimate a size elasticity of TFP spillovers of 0.004, one order of magnitude lower than in the literature.

The estimation procedure explicitly targets the aggregate wage dispersion and city-specific average wages. It targets neither how wage inequality varies across space nor how steep are the local job ladders. I therefore use these two novel moments as over-identification tests to validate the empirical relevance of my framework.

The model quantitatively explains the differences in inequality across cities. As in the data, low-wage workers earn the same wage everywhere. By contrast, high-paying jobs are concentrated in large cities, and there are large spatial variation in the right tail of the wage distributions. I provide an exact decomposition of wage gaps across space into the direct effect of employer sorting and local competition. I find that the concentration of unproductive and productive employers in small and large places reduces employers' local market power and thus increases average wages. Further, I quantify that spatial differences in local competition entirely explains why large cities are more unequal. Absent local competition, productive employers in large cities would not offer substantially higher wages, and lower within-city inequality would follow.

Turning to the second over-identification exercise, I find that the steeper job ladder in large cities generates wage growth consistent with my estimated local returns to job switching. As a result, workers enjoy higher lifetime income. For instance, while unemployed workers in Paris earn real income 8% lower than in Lens, the net present value of their real lifetime earnings is 4% higher. Differences in lifetime earnings spillover to welfare disparities. If workers had access to the same job ladder no matter their birthplace, those born in Paris would be willing to give away 5% of their lifetime consumption not to be born in Lens. These welfare disparities rise to 10% once taking into account that workers born in Paris have a relatively easier access to better expected job opportunities. I therefore conclude that the steeper job ladder of large cities is a crucial component for why workers live there despite initial lower real earnings.

The two over-identification exercises validate the model's ability to generate between- and within-city inequality. Spatial wage disparities arise through employers sorting rather than spatial TFP gaps. However, despite the absence of productivity spillovers, agglomeration has local and aggregate impacts on wages and inequality, which I quantify in the fourth part of the paper.

I quantify the steady-state consequences of agglomeration on wages by computing counterfactual equilibria under random shocks to local amenities. I then trace out the response of wages at three levels of aggregation: the firm, the city, and the entire economy. In progress.

Related literature This paper relates to several strands of literature. The first is the literature that studies the drivers of local wage inequality (Glaeser et al., 2009; Baum-Snow and Pavan, 2013;

Eeckhout et al., 2014; Baum-Snow et al., 2018; Papageorgiou, 2022). This literature has primarily focused on the role played by worker heterogeneity. I emphasize that employers — and how workers reallocate across them — is a key driving force.

To separate the contribution of employer and worker heterogeneity, I estimate a two-way fixed effect model as in the original Abowd et al. (1999) and the long list of papers that followed (e.g. Card et al., 2013; Bonhomme et al., 2019; Song et al., 2019). Applying this method to local labor markets, Dauth et al. (2022) show that the high average wage of large location is partially explained by a better match between worker and firm, and Card et al. (2025) estimate the average wage impact of cities net of worker heterogeneity. While both papers focus on the differences in average wage across locations, I use this method to study within-city wage inequality.

I estimate that the reallocation of workers across employer spurs substantial wage growth, with larger gains in more unequal, bigger cities. This finding contributes to the literature on the dynamic impact of cities on wages (Baum-Snow and Pavan, 2012; Roca and Puga, 2017; Porcher et al., 2021; Eckert et al., 2022; Lhuillier, 2023). This literature has mostly investigated how cities shape learning. I find that the steeper job ladder in large cities — as measured by higher wages gains from job switching — explains 60% of the faster wage growth.

I microfound the existence of local job ladders by incorporating search frictions and on-the-job search into a quantitative spatial model, as in Schmutz and Sidibé (2019), Martellini (2020), and Heise and Porzio (2022). I contribute to these papers by allowing heterogeneous employers to sort across local labor markets, which I show is key to rationalize the steeper job ladder in large cities.[1] Most closely related to this paper, Lindenlaub et al. (2022) develop a framework where employers sort across frictional labor markets in which workers search on the job. Our papers complement each other: they study the impact sorting has on local labor shares whereas I analyze its impact for local wage inequality.

Finally, this paper connects to the literature on search frictions and monopsony power (Burdett and Mortensen, 1998; Bontemps et al., 2000; Engbom and Moser, 2021; Bilal and Lhuillier, 2021; Jarosch et al., 2021; Berger et al., 2022; Lamadon et al., 2022; Gouin-Bonenfant, 2022; Morchio and Moser, 2024). I contribute to it by showing that the sorting of employers across labor market shapes the rent-sharing process, reduces employers' market power, and increases aggregate wages.

The rest of this paper is organized as follows. Section 1 presents the two novel facts. Second 2 develops the theoretical framework that jointly accounts for them. Second 3 lays out the quantitative extensions of the model and its structural estimation. Second 4 concludes by assessing the quantitative relevance of the model.

[1]The role played by firm sorting on spatial inequality has also been documented by Combes et al. (2012), Behrens et al. (2014), Gaubert (2018), Bilal (2023), Oh (2023), and Franco (2024).

1 Two facts about spatial wage inequality

1.1 Data

I use employer tax records from France (*Déclaration Annuelle de Données Sociales*, DADS) between 2008 and 2019. This dataset comes in two formats. The first is a 4% representative panel that tracks the entire history of individuals in the labor market (*DADS panel*). The second is a repeated cross-section covering the universe of employed workers (*DADS salariés*). Both datasets provide information on workers' earnings, the number of hours worked, the establishment where they are employed and their occupation, the location where they work and live, along with demographic information.

This dataset has two advantages. First, its panel structure, combined with the information of workers' employers, allow to separate the relative contribution of workers and employers to wage dispersion. Second, its large sample provides sufficient statistical power to estimate precisely statistics at relatively granular geographical units.

I apply the same sample restrictions on both datasets. I focus my analysis on full-time employed workers between 25 and 55 years old. Workers employed in the public sector in France have their wage determined nationally by their tenure rather than based on their productivity or the local competitiveness of the labor market. I therefore keep in the sample workers employed in the private sector, and I exclude the education and health industries due to their large fraction of public servants. I abstract from the labor supply decision and use gross hourly wage as my measure of labor income. I deflate wages by the aggregate consumption price index and express them in 2018 euros. Appendix A.1 provides more details on the construction of the sample and Table A.1 some aggregate summary statistics. Between 2008 and 2019, aggregate wage inequality is constant in France.

I define a city as a commuting zone (CZ). A commuting zone is a statistical area defined by the French statistical agency (INSEE). It consists of a collection of contingent municipalities clustered together to reduce the commuting flows across them. Commuting zones are thus the natural geographical unit when thinking of local labor markets. There are 297 such areas in metropolitan France, which span the whole territory, with an average size of 31,823 employed workers. I use the terms commuting zones, cities, and locations interchangeably for the rest of the paper.

For computational feasibility and statistical accuracy, I sometimes group cities together according to their average wage. Specifically, I construct ten population-weighted deciles of average wages. The first decile contains 100 cities, Lens is in third group, and the tenth decile is composed of Paris and Saclay, the South-West suburb of Paris. Table A.2 provides summary statistics on each of the city cluster. This clustering accounts for 97% of the variation in average wage and wage variance across commuting zones.

1.2 Spatial wage inequality in France

Figure 1 displays the extent to which wage varies across cities in France. Panel (a) plots the average wage in the bottom 10% (orange circles) and top 10% (blue rectangles) of the city wage distribution against the unconditional average log wage.

Two patterns strike out. First, there are large average wage differences across CZs. The average wage in Paris is 65% higher than in Lens, the median city in the wage distribution. High-wage cities tend to be larger: workers in cities twice larger earn on average wages 8.3% higher. Decomposing the aggregate wage variance into a between-city and within-city component,

$$\mathbb{V}\text{ar}[\log w] = \underbrace{\mathbb{V}\text{ar}[\mathbb{E}(\log w)]}_{\text{Between-city}} + \underbrace{\mathbb{E}[\mathbb{V}\text{ar}(\log w)]}_{\text{Within-city}},$$

I find that spatial variation in average wages account for 10% of the total wage variance.

Second, within-city inequalities are far from homogeneous across locations. For instance, the standard deviation of log wages in Paris is 0.56, 58% higher than in Lens. For comparison, the standard deviation of log wages is 0.63 in the United States and 0.44 in France, countries often associated with high and low inequality.

Large spatial variations in the wages of high-paid workers drive the differences in within-city inequality. Workers in the top 10% of the wage distribution in Paris earn wages 87% higher than workers at the top in Lens. By contrast, low-wage workers earn similar wages everywhere.

Figure 1b repeats the same exercise with real wages. Real wages are computed as nominal wages deflated by a citywide Cobb-Douglas price index with a housing expenditure share of 0.2.[2]

Figure 1b reveals that low-wage workers earn lower real earnings in larger cities. The higher housing prices indeed more than offset the moderately higher nominal wages. For instance, workers in the bottom 10% of the wage distribution in Paris earn real wages 20% lower than their counterparts in Lens. Figure A.1 extends the analysis to all the deciles of the wage distribution; on average, workers in the bottom 40% of the city wage distribution earn lower real wages in larger places. A similar pattern holds when housing price are adjusted to take into account that workers in different part of the distribution sort across neighborhoods. It also holds amongst workers not born in the city, i.e., those who choose to live there.

Wages therefore vary substantially across space. The literature has primarily focused on the between-city component. However, this accounts for a relatively small share of the total wage variation, and is likely to underestimate the importance of cities on wages given the large spatial variations in within-city inequality. I now show that employers are instrumental in shaping both between- and within-city inequality.

[2] There does not exists publicly available price index at the commuting zone level in France. See Section 3.3 for details on the housing data.

1.3 *Fact #1: high-paying jobs are spatially concentrated*

I estimate the role of employers on local wage inequality via a two-way fixed effect model à la Abowd et al. (1999) —AKM henceforth. Specifically, I estimate

$$\log w_{it} = \alpha_i + \beta_{\ell(i,t)} x_{it} + \gamma_{j(i,t)} + \ \varepsilon_{it}, \tag{1}$$

where i is a worker, t is quarter-year, x_{it} denotes experience, and $\ell(i,t)$ and $j(i,t)$ are indices for the city group and job where i is employed at time t. A job is defined as an establishment $\times$ 4-digit occupation. The parameters α_i and γ_j are worker and job fixed effects. In addition, I allow cities to offer differential returns to workers' experience as captured by β_ℓ. I define workers' total effect as the sum of the static and dynamic components: $\psi_{it} \equiv \alpha_i + \beta_{\ell(i,t)} x_{it}$.

Workers who move between jobs identify separately worker and job fixed effects if worker mobility is conditionally random (Card et al., 2013). In this case, the job fixed effects capture wage premia —the additional wage a worker earns when working at a particular job. To limit well-known econometric difficulties linked to limited mobility bias, I follow Bonhomme et al. (2019) and group workers and jobs each in 100 equally populated groups based on their unconditional mean wage.[3] I then estimate equation (1) with OLS at the group level.[4]

The AKM model (1) is effectively a movers design. In particular, the relative level of the job fixed effects across cities are identified from workers switching locations. Accordingly, $\bar{\gamma}_\ell \equiv \mathbb{E}_\ell[\gamma_{j(i,t)}]$ is identical to standard estimates of the causal impact of cities on wages (as in e.g., Combes et al., 2008; Card et al., 2025). How much these explain of the between-city wage variance is given by the decomposition

$$\mathbb{V}\text{ar}[\mathbb{E}_\ell(\log w_{it})] = \underbrace{\mathbb{V}\text{ar}[\bar{\psi}_{\ell_{it}}] + \mathbb{C}\text{ov}[\bar{\psi}_{\ell_{it}}, \bar{\gamma}_{\ell_{it}}]}_{\text{Worker}} + \underbrace{\mathbb{V}\text{ar}[\bar{\gamma}_{\ell_{it}}] + \mathbb{C}\text{ov}[\bar{\psi}_{\ell_{it}}, \bar{\gamma}_{\ell_{it}}]}_{\text{Wage premia}},$$

where $\bar{\psi}_\ell \equiv \mathbb{E}_\ell[\psi_{it}]$ is the average worker contribution.

The AKM model (1) also allows to study the dispersion in wage premia within cities —and in that, the heterogeneous treatment of cities on wages. I quantify how the distribution of wage premia varies across space in two ways. First, I decompose the local wage variance into components due to worker heterogeneity, pay premia dispersion, and residuals:[5]

$$\mathbb{V}\text{ar}_\ell[\log w_{it}] = \underbrace{\mathbb{V}\text{ar}_\ell[\psi_{it}] + \mathbb{C}\text{ov}_\ell[\psi_{it}, \gamma_{j(i,t)}]}_{\text{Worker}} + \underbrace{\mathbb{V}\text{ar}_\ell[\gamma_{it}] + \mathbb{C}\text{ov}_\ell[\psi_{it}, \gamma_{j(i,t)}]}_{\text{Wage premia}} + \underbrace{\mathbb{V}\text{ar}_\ell[\varepsilon_{it}]}_{\text{Residuals}}. \tag{2}$$

Second, I project job fixed effects onto city size allowing for heterogeneous elasticities across the

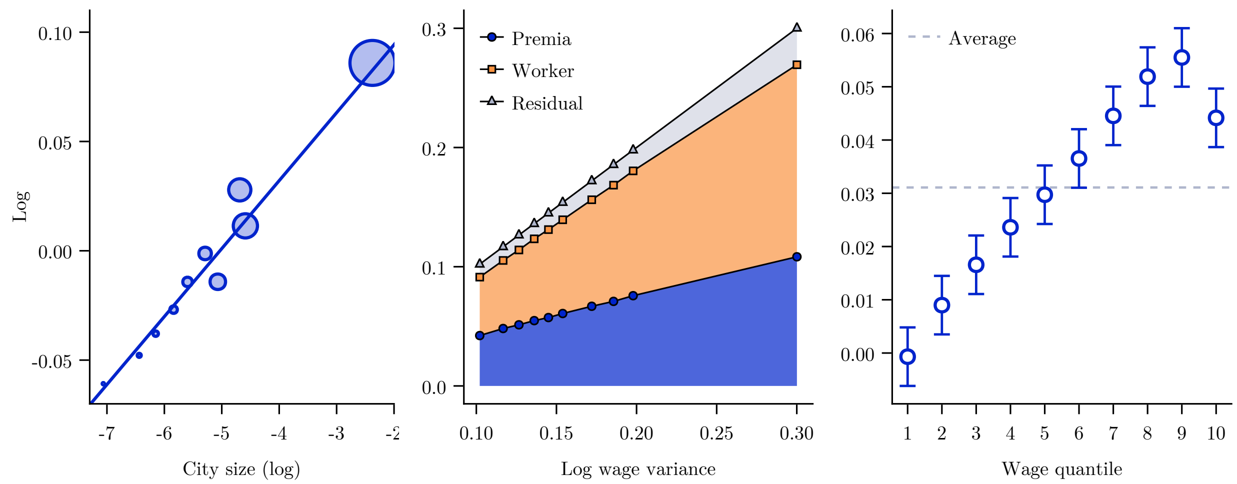
[3]The grouping for jobs is done within city groups to preserve the local nature of an establishment.

[4]The results are virtually identical when varying the number of groups between 10 and 200.

[5]In particular, projecting each term on the right-hand side onto the wage variance in a cross-city regression then quantifies the relative importance of worker, wage premia, and residual heterogeneity.

Figure 2: The drivers of spatial wage disparities

(a) Average wage premia (b) Within-city decomposition (c) Wage premia size elasticity



Panel (a) displays the variance decomposition (2). Each marker is a city group. The blue area represents the dispersion in wage premia, $\mathbb{V}\mathrm{ar}[\gamma_{it}] + \mathbb{C}\mathrm{ov}[\psi_{it}, \gamma_{j(i,t)}]$. The orange area represents worker heterogeneity, $\mathbb{V}\mathrm{ar}[\psi_{it}] + \mathbb{C}\mathrm{ov}[\psi_{it}, \gamma_{j(i,t)}]$. The grey area represents the residuals, $\mathbb{V}\mathrm{ar}[\varepsilon_{it}]$. Each measure is plotted against the location-specific wage variance, $\mathbb{V}\mathrm{ar}_\ell[\log w_{it}]$. Panel (b) presents the local distribution of job fixed effects. The orange circles and blue rectangles are the 1st and 9th deciles of the local job fixed effect distribution. Both are plotted against the location-specific wage variance. Job fixed effects are normalized by their national average.

fixed effect distributions:

$$\gamma_j = \mathrm{FE}_{q(j)} + \nu_{q(j)} \log m_{\ell(j)} + u_j. \tag{3}$$

In the above, $q(j)$ is the within-city decile in which job j belongs, and FE_q is a decile fixed effect. The elasticities ν_q thus measure how higher are wage premia in locations 1% larger amongst premia in the q^{th} decile of their local premia distribution. (3) is close to a quantile regression where quantiles are defined at the city level. While the job fixed effects are interpreted as causal estimates under the random mobility assumption, the parameters $\{\nu_q\}_{q=1}^{10}$ are viewed as summary statistics rather than causal elasticities.

Figure 2a plots the average wage premia by city group against the average city size. Wage premia are higher in bigger places: workers in cities twice larger earn a premia 3.1% higher on average. This elasticity is similar to Combes et al. (2008). In the aggregate, I find that wage premia accounts for a significant fraction —31%— of the between-city wage variance.

Figure 2b and Figure 2c turns to the novel, within-city analysis. Figure 2b presents the variance decomposition (2). Each marker represents a city group. The blue, orange, and grey area depicts the location-specific dispersion in job fixed effects, workers fixed effects, and residuals, projected against the city wage variance.

Figure 2b shows that wage premia are more dispersed in more unequal and larger cities. For instance, the dispersion in premia in Paris is 0.06 larger than in Lens, or 35% of the difference in wage variance (0.17). Generalizing to all cities, I find that the dispersion in job fixed effects explains 33% of the spatial variation in within-city wage inequality. Meanwhile, worker heterogeneity and the residuals explain 57% and 10%.

Figure 2c unpacks why are wage premia more dispersed in larger cities by plotting the city size elasticities estimated by (3).[6] Figure 2c reveals that high-paying jobs are concentrated in a few locations whereas low-paying jobs are present everywhere. Indeed, larger cities brings no wage gains for workers at the bottom of the pay premia distribution. By contrast, these gains are exacerbated for workers at the top. For instance, wage premia are about 5% higher in cities twice larger for workers in the top 20% of their local premia distribution.

Figure A.2b confirms non-parametrically the unequal allocation of jobs across space. Specifically, it displays where are located jobs in the bottom and top 10% of the aggregate premia distribution. Given that city groups are population-weighted, jobs are uniformly allocated across space if each city group contains about 10% of them —which holds for low-paying jobs. By contrast, 23% of jobs in the top 10% of the aggregate premia distribution are in Paris, whereas Lens and cities alike holds 6% of them.

Figure A.3 implements two robustness exercises. First, Figure A.3a separates the role of sorting —as measured by the correlation between the job and worker fixed effects— from the variance in job fixed effects. The within-city variance decomposition (2) indeed bundles the two together. I do so because, everything else equal, the covariance is an increasing function of job heterogeneity. When splitting the two, I find that the correlation between jobs and workers fixed effect is stable across commuting zones.[7]

Second, job fixed effects are estimated at the establishment $\times$ occupation. Given that I control for time-invariant unobserved worker heterogeneity and location-specific returns to experience, I indeed interpret pay differences between occupations for seemingly identical workers as pay premia. However, these may also capture time-varying unobserved skills or compensating differentials. Figure A.3b quantifies the relative importance of occupation, industry, and firm heterogeneity on local wage inequality. The industry component captures 6% of the between-city differences in inequality, suggesting that compensating differentials are not driving the result. The occupation and firm components are equally important: they both explain 47% of the spatial variations in wage variance. As such, pay premia still accounts for a sizeable portion of within-city wage inequality even when abstracting from occupation heterogeneity.

Figure 2 thus documents the importance of employers for spatial wage disparities. Large cities offer higher wages as high-paying jobs are concentrated there. Meanwhile, large cities do not benefit all workers equally as low-paying jobs are present throughout France. Greater within-city inequality follows. I now turn to show that the wage gains offered by large cities occur over time as workers reallocate across employers.

[6]Figure A.2a confirms these elasticites are not driven by outliers such as Paris by plotting non-parametrically the average premia in the bottom and top 10% of each city's premia distribution against that city size.

[7]There does not exist an exact decomposition that separates the role of sorting from that of worker and job heterogeneity. Instead, section A.2 derives a first-order approximation that separates the different channels. Figure A.3a implements this approximation.

1.4 Fact #2: the job ladder is steeper in unequal cities

To motivate the analysis in this section, Figure 3a presents cross-sectional evidence on how the reallocation of workers across jobs interact with their wage profile. Specifically, it depicts for two cities, Paris and Lens, the average wage as a function of the number of jobs an individual has worked at. The number of job held by worker i at time t is defined as the cumulative sum of job switches between age 25 and t:

$$\#\mathrm{J}_{it} = \sum_{\tau=\underline{t}_i}^{t} \mathrm{J2J}_{i\tau}. \tag{4}$$

A worker is said to switch jobs ($\mathrm{J2J}_{it} = 1$) if they transition between pairs of establishment $\times$ occupation within two consecutive quarters. In (4), $\underline{t}_i$ denotes the year in which worker i was 25.[8]

Figure 3a highlights that the wage gap between Paris and Lens grows as workers move along the job ladder. While this gap is 29% for workers who enter the labor force, it rises to 89% after 10 job switches.

Three distinct forces may shape the wage profiles shown in Figure 3a. First, heterogeneous workers may sort across locations, and may do so differently across their lifecycle. Second, workers which have occupied more jobs tend to be more experienced, and experience may be remunerated differentially throughout space. Third, switching jobs may bring higher gains in Paris.

I estimate the local impact of job switching on wages net of sorting and learning through the reduced form model

$$\log w_{it} = \mathrm{FE}_t + \mathrm{FE}_i + \alpha_{\ell(i,t)} + \beta_{\ell(i,t)} x_{it} + \gamma_{\ell(i,t)} \#\mathrm{J}_{it} + \varepsilon_{it}. \tag{5}$$

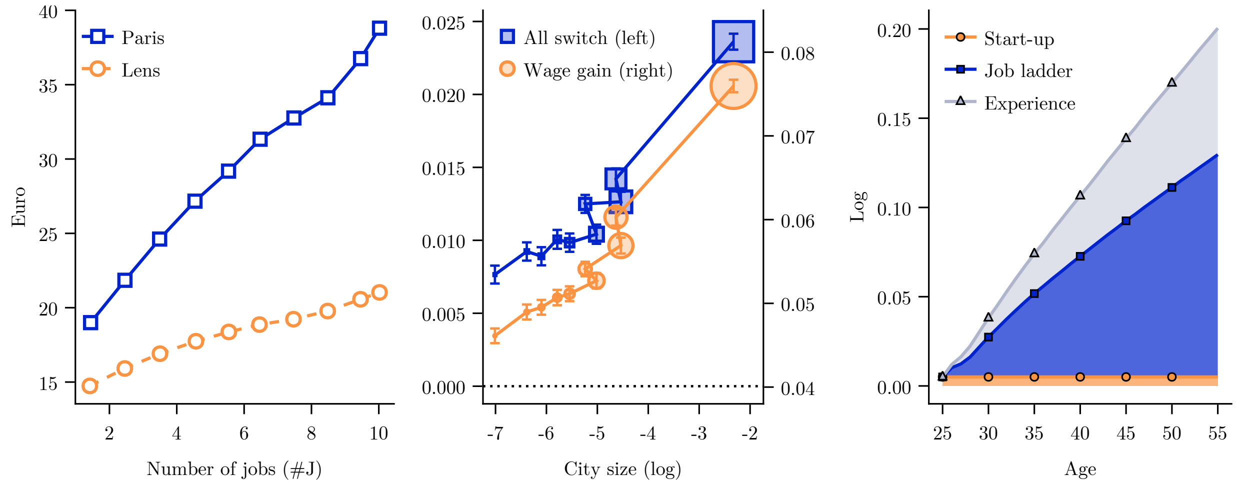
In (5), FE_i are worker fixed effects which control for the sorting of workers across space. The variable x_{it} is the experience of i at time t, and β_ℓ captures that particular cities may favor human capital accumulation. FE_t and α_ℓ are year and location fixed effects. Finally, $\#\mathrm{J}_{it}$ is the cumulative sum of job switches as defined by (4). To ensure that I am not misclassifying transitions through unemployment as job switches, I also estimate a version of (5) where I restrict $\#\mathrm{J}_{it}$ to job transitions associated with wage gains.

The main coefficients of interest are γ_ℓ: the local returns to job switching. Given the worker fixed effects and the city-specific returns to experience, γ_ℓ is identified from the wage growth of job switchers relative to the wage growth of job stayers. Identification then relies on random worker mobility across employers conditional on workers' unobserved characteristics and experience. The local returns to experience are in turn identified from the wage growth of job stayers. Finally, the location fixed effects α_ℓ measure the average wage in location ℓ, net of worker heterogeneity and

[8] For workers that I observe for the first time after age 25, I infer their prior number of switches based on their age. I estimate $\mathrm{J2J}_{it} = \alpha_t + \beta_{a_i} + u_{it}$, where a_i is the age of i. I then use the estimated age profile to infer workers' past number of switches: $\#\mathrm{J}_{it_i} = \sum_{\tau=\underline{t}_i}^{t_i} \mathrm{J2J}_{i\tau} = \sum_{a \leq a_{it_i}} \beta_a$, where t_i is the first year in which I observe i.

Figure 3: The local returns to job switching

(a) Hourly wage (b) Returns to job switching (c) Lifecycle wage decomposition



Left panel plots the average hourly wage by the cumulative number of job held as defined by (4). The orange circles are in Lens, the blue rectangles in Paris. Right panel plots the estimates $\{\gamma_\ell\}_{\ell=1}^{L}$ in (5) against cities' wage variance. The vertical bars represent 95th confidence intervals.

before any job switch or experience. It therefore measures the extra wage earned in location ℓ for seemingly identical workers upon entry in the labor market —which I call the startup premia.

The local returns to job switching are related to the pay premia estimated in Section 1.3. Indeed, both are identified from job switchers. However, whereas the job fixed effects estimate the wage differentials between any two employers, γ_ℓ measures the average wage change upon a job switch. It therefore quantifies the extent to which workers reallocate from low- to high-paying employers over time. If $\gamma_\ell > 0$, a job switch leads on average to a wage increase, consistent with the existence of a local job ladder. If $\gamma_{\ell'} > \gamma_\ell$, job switches yield larger wage gains in location ℓ' than ℓ, suggesting that the local job ladder is steeper. By contrast, if $\gamma_{\ell'} = \gamma_\ell$, any extra wage premia offered by location ℓ' must occur upon entry in the labor force, and is thus captured by the startup premia $\alpha_{\ell'}$.

Figure 3b plots the estimated returns to job switching against city size. The blue rectangles are the estimates considering all switches, whereas the orange circles restrict the attention to switches associated to wage gains. The startup premia and returns to experience are reported separately in Figure A.5.

I find that job switching fasten wage growth throughout France, but that the job ladder is steeper in larger places. On average, a job switch leads to a 1.3% wage increase. In Paris, these wage gains are 2.4%, whereas they are 0.9% in Lens. Generalizing to all cities, workers gain an additional 0.2 p.p. upon switching jobs in cities twice larger. The spatial differences in job ladder steepness are even steeper when focusing on job switches associated with wage gains (0.5 p.p.).

At the same time, I find that the startup premia are negligible. For instance, I estimate that entry-level wages in Paris exceed those in Lens by 0.5% after controlling for worker heterogeneity and the dynamic effects of cities. Finally, consistent with the prior literature, I find that the returns to experience are higher in larger cities.

Table A.4 in Appendix A.3 provides several robustness exercises. For expositional clarity, I replace the city-specific returns to job switches in (5) with an interaction between job switching and city size. Column (1) reports the baseline estimates from this modified specification for all job switches and column (2) for job switches associated with wage gains. Column (3) includes year-by-location fixed effects to control for location-specific wage trends. Human capital accumulation may not be fully captured by experience; then, the wage growth of movers may confound the gains from reallocation with learning. I control more flexibly for human capital accumulation in two ways. Column (4) includes occupation fixed effects, thereby identifying returns to job switching from moves between establishments within occupations. Column (5) introduces worker-specific returns to experience to account for heterogeneity in learning abilities. The estimates are very stable throughout.

Figure A.6 and A.7 conducts heterogeneity analysis. Specifically, Figure A.6 estimates the local returns to job switching by occupation. I find that the job ladder of large cities is steeper for every occupation, and particularly so for white-collar workers. Figure A.7 then proceeds with estimating the returns to job switching by within-city wage quartile. Here as well, I find that workers at every rank of the wage distribution experience higher wage gains upon job switching in larger cities.

The steeper job ladder of larger cities has sizeable consequences on lifetime income. According to equation (5), holding constant the distribution of worker fixed effects and observables across locations, the average wage of a worker with age a in location ℓ is[9]

$$\mathbb{E}_\ell[\log w_{it} \mid a] = \alpha_\ell + \beta_\ell \mathbb{E}[x_{it} \mid a] + \gamma_\ell \mathbb{E}\left[\#\mathrm{J}_{it} \mid a\right], \tag{6}$$

Furthermore, taking the average of (6) across the age distribution allows to quantify how much the higher wage premia of large cities stems from the startup premium versus the steeper ladder.[10]

Figure 3c uses (6) to compute the predicted wage difference between Paris and Lens along the lifecycle. There is virtually no wage discrepancy at entry in the labor market. The wage gap widens as workers reallocate across employers over time. By age 55, the steeper job ladder in Paris implies that workers earn wages 13% greater than in Lens. The higher returns to experience further boosts the wage gap by 7 p.p. Averaging across the lifecycle, I find that 92% of the between-city differences in average wage premia are caused by spatial variations in job ladder steepness.

[9]Figure A.4 depicts the average job switching rate by location (blue rectangles). Unconditional switching rates are higher in larger places. However, large cities attract young and skilled workers which are more inclined to switch jobs. The orange circles represent the local switching rate after controlling for worker unobserved heterogeneity and age. Then, job switching rates are constant across space.

[10]Specifically, $\mathbb{E}_\ell[\log w_{it}] = \alpha_\ell^{\mathrm{JL}} + \beta_\ell^{\mathrm{JL}} \mathbb{E}[x_{it}] + \gamma_\ell^{\mathrm{JL}} \mathbb{E}\left[\#\mathrm{J}_{it}\right] = \beta_\ell^{\mathrm{AKM}} \mathbb{E}[x_{it}] + \mathbb{E}_\ell[\gamma_{j(i,t)}^{\mathrm{AKM}}]$, where $\cdot^{\mathrm{AKM}}$ and $\cdot^{\mathrm{JL}}$ refer to the point estimates of (1) and (5) with a slight abuse of notation, and $\beta_\ell^{\mathrm{AKM}} = \beta_\ell^{\mathrm{JL}}$ by definition. Hence, $\mathbb{E}_\ell[\gamma_{j(i,t)}^{\mathrm{AKM}}] = \alpha_\ell^{\mathrm{JL}} + \gamma_\ell^{\mathrm{JL}} \mathbb{E}\left[\#\mathrm{J}_{it}\right]$.

1.5 Summary of testable implications

Section 1.3 and 1.4 document the importance of employers for spatial wage inequality through two novel facts. First, high-paying jobs are concentrated in large cities whereas low-paying jobs are present everywhere. Second, workers tend to start their career at these low-paying jobs, and large cities boost wages as workers reallocate across employers. Together, employers generate higher wages and greater within-city inequality in larger cities.

These two facts constitute testable implications against which we can benchmark existing theories of spatial wage disparities. First, TFP gaps would fail at generating within-city inequality. Second, to the extent that the facts control for unobserved worker heterogeneity and the effects of cities on learning, worker sorting cannot account for them. Third, *ex-post* productivity shocks or compensating differentials would not explain the steeper job ladder of larger locations.

In the next section, I build a model that jointly account for the two facts through the sorting of employers across frictional labor markets. I then use this novel framework to quantify the consequences of agglomeration on wages once accounting for between- and within-city wage inequality.

2 A spatial theory of wage premia

In this section, I develop a framework that combines search frictions as in Burdett and Mortensen (1998) with an otherwise standard quantitative spatial model (Redding and Rossi-Hansberg, 2017). I concentrate here on the model's core feature and characterize the equilibrium analytically. Section 3 presents quantitative extensions.

2.1 Setup

The economy is comprised of two types of agents. There is a unit mass of *ex-ante* homogeneous workers and a mass $M \leq 1$ of heterogeneous employers. Workers and employers meet within L locations. Time is continuous. There is no aggregate shock, and I focus on the steady state equilibrium.

Cities Cities are indexed by ℓ with $\ell \in \{1, 2, \dots, L\}$. They offer amenities A_ℓ to the workers that reside there. Cities differ in no other way *ex-ante*. I therefore order cities by their amenities: $A_1 < A_2 < \dots < A_L$. I later allow for TFP differentials. However, focusing initially on amenities highlights how employer sorting endogenously creates wage differences across cities.

Workers Workers are risk neutral, infinitely lived, and discount the future at rate ρ. They consume a freely traded good taken as the numéraire and local housing according to Cobb-Douglas preferences. Their lifetime utility is

$$\mathbb{E}_0 \left[\int_0^\infty \mathrm{e}^{-\rho t} \left(\frac{h_t}{\alpha} \right)^\alpha \left(\frac{c_t}{1-\alpha} \right)^{1-\alpha} dt \right],$$

where α is the expenditure share on housing. The expectation is taken over future labor income. Workers do not have access to a savings devise: if they reside in location ℓ and earn labor income y_t, they face the per-period budget constraint $p_\ell h_t + c_t \leq y_t$.

Unemployed workers earn unemployment insurance b, which can alternatively be interpreted as revenues from home production. They receive job offers at Poisson rate λ^u. When employed, workers supply inelastically one unit of labor, so their labor income equals their wage. They also receive job offers at Poisson rate λ^e and fall back into unemployment at rate δ. I assume for now that the contact rates are exogenous and constant across space, with $\delta < \lambda^e \leq \lambda^u$.[11]

Job offers are drawn from the local job offer distribution F_ℓ. That is, workers only search for jobs in the city where they live.[12] Workers are free to migrate across cities, but jobs are tied to a location. Accordingly, if a worker wants to switch locations, they must quit their job. These assumptions, while strong, are broadly consistent with the data: most job switches occur within cities, and workers who move between cities experience on average longer non-employment spells (Table A.5).

The allocation of workers across space is characterized by the triplet $\{m_\ell, e_\ell, u_\ell\}_{\ell=1}^L$, which denote the measures of total, employed, and unemployed workers in each location. Feasibility demands $\sum_{\ell=1}^L m_\ell = 1$.

Employers Employers are infinitely lived and discount the future at rate r. They produce the numéraire of the economy. They are *ex-ante* heterogeneous, indexed by their time-invariant productivity z. The aggregate distribution of productivity is Γ with support $[\underline{z}, \bar{z}]$, $\bar{z} \leq \infty$. I assume that Γ admits a finite and continuous density. The production — and revenues — generated by an employer with productivity z when they hire n workers is $R(z,n) = zn$. Given constant returns to scale, z is also the marginal product of labor (MPL) of a job at this employer. As such, I use employers and jobs interchangeably.[13] I assume that $\underline{z}$ is high enough relative to b so that any jobs is profitable in at least one city.

Employers hire workers in local frictional labor markets. The hiring process follows Burdett and Mortensen (1998). Employers have monopsony power over workers and are assumed to be atomistic. They post a single wage offer. They commit to a fixed and non-state-contingent wage that cannot be renegotiated throughout the employment spell. In particular, employers cannot make counteroffers when their workers receive alternative job opportunities. Rather, given their

[11] Figure A.4 shows that job switching rates are fairly constant across cities once controlling for worker heterogeneity.
[12] For a quantitative spatial model with wage posting and between-city search, see Heise and Porzio (2022).
[13] The framework does not model explicitly occupation. Rather, the productivity z captures the total productivity of the pair occupation $\times$ establishment.

target size, they optimally set their wage offer *ex ante* to maximize hiring.

Employers freely choose the city where they want to produce and hire workers. They pay unit housing cost r_ℓ to locate in city ℓ. The endogenous distribution of jobs across space is summarized by two objects: the mass of employers in each city, M_ℓ, and the local distribution of productivity, $d\Gamma_\ell$.[14] The allocation $\{M_\ell, d\Gamma_\ell\}_{\ell=1}^{L}$ is feasible if the number of jobs present throughout the economy is not greater than the total number of jobs available:

$$\sum_{\ell=1}^{L} M_\ell d\Gamma_\ell(z) \leq M d\Gamma(z). \tag{7}$$

Housing markets The residential and commercial housing markets are segmented. In each of them, absentee land owners supply the local housing stocks. The residential housing supply is given by $L_\ell = \bar{L} p_\ell^{\theta}$, where θ is the residential housing supply elasticity. Likewise, the commercial housing supply is given by $H_\ell = \bar{H} r_\ell^{\phi}$ for ϕ the commercial housing supply elasticity.

I characterize the steady state equilibrium in three steps. First, I derive the spatial allocation of workers and the local labor supply curves that result from it. I then solve for the local wage distributions given the spatial allocation of employers. Finally, I characterize the spatial allocation of employers.

2.2 Local labor supplies

The local labor supply curves follow from the job-switching behavior of workers together with their choice of location. Let U_ℓ denote the lifetime utility of an unemployed worker in location ℓ. This value satisfies the HJB equation

$$\rho U_\ell = \frac{A_\ell b}{P_\ell} + \lambda^u \int \max\{V_\ell(w) - U_\ell, 0\} dF_\ell(w), \tag{8}$$

where $P_\ell \equiv p_\ell^{\alpha}$ is the price index in location ℓ. The lifetime utility of unemployed workers in city ℓ consists of their contemporaneous real earnings adjusted for local amenities, and the expected value from future job opportunities. Unemployed workers choose their location to maximize their lifetime utility:

$$\bar{U} = \max_{\ell} U_\ell. \tag{9}$$

In equilibrium, unemployed workers are indifferent as to where to live.

By contrast, employed workers cannot move freely across locations. The lifetime utility of a

[14] Ultimately, the spatial allocation of employers is summarized by the measure of each job $M_\ell d\Gamma_\ell(z)$. Separating the two is useful to characterize sequentially the solution to the employer problem.

worker employed at wage w in location ℓ satisfies

$$\rho V_\ell(w) = \frac{A_\ell w}{P_\ell} + \lambda^e \int \max\{V_\ell(w') - V_\ell(w), 0\} dF_\ell(w') + \delta[U - V_\ell(w)], \tag{10}$$

which also accounts for the utility loss associated with falling back into unemployment.

Within a city, employed workers behave as income maximizers. Their utility is increasing with their wage; as a result, they climb the local job ladder by continuously accepting better-paying job offers. The flow of workers up the job ladder determines the number of employees at each wage. In particular, the distribution of wages amongst employed workers is related to the wage offer distribution according to[15]

$$G_\ell(w) = \frac{F_\ell(w)}{1 + k(1 - F_\ell(w))}. \tag{11}$$

The parameter $k \equiv \lambda^e/\delta$ summarizes the speed at which workers climb the job ladder relative to the rate at which they fall back to unemployment. Given the wage offer distribution F_ℓ and the employment distribution G_ℓ, the number of employed workers per wage offer w in location ℓ is

$$n_\ell(w) = \frac{1}{\theta_\ell} \frac{1 + k}{\left[1 + k(1 - F_\ell(w))\right]^2}, \tag{12}$$

where $\theta_\ell \equiv M_\ell/e_\ell$ is the labor market tightness in city ℓ. Equation (12) is the local labor supply curve. The supply curves slope upward: workers employed at high-wage jobs keep the same job for relatively longer since there are relatively fewer better-paying opportunities to switch to.

The labor supply curves differ across locations in two ways. First, the yield of a vacancy is lower in tighter labor markets. Second, the competition for workers — as represented by the job offer distribution — may be fiercer. In particular, in locations with a higher job offer distribution in the first order stochastic dominance sense, the labor supply curve shifts downward as workers transition more often to better paying opportunities. Search frictions, as summarized by k, determines the sensitivity of the supply curve to local competition. If frictions are high, workers transition infrequently across jobs, thus flattening the labor supply curve. In the limit of infinite frictions ($k \to 0$), the number of workers per job offer is entirely given by the market tightness.

Within a city, unemployed workers tradeoff a higher search efficiency against lower present earnings. They accept any wage offer greater than the city-specific reservation wage $\underline{w}_\ell$ given by

$$\underline{w}_\ell = b + (\lambda^u - \lambda^e) \int_{\underline{w}_\ell}^{\infty} \frac{1 - F_\ell(w)}{\rho + \delta + \lambda^e(1 - F_\ell(w))} dw. \tag{13}$$

In locations with a better job offer distribution, the option value of searching is higher, and so is the reservation wage. The gap between the search efficiency of unemployed and employed workers

[15]The derivation of the labor supply curve is similar to Burdett and Mortensen (1998) and detailed in Appendix B.1.

controls the sensitivity of the reservation wage to the job offer distribution. By contrast, the reservation wage is independent of local amenities and housing prices since these affect all workers equally.

Across space, unemployed workers choose where to live considering the future stream of job opportunities, housing prices and amenities. The lifetime utility of unemployed workers in location ℓ rewrites

$$U_\ell = \frac{A_\ell \mathcal{W}_\ell}{P_\ell}, \tag{14}$$

where $\mathcal{W}_\ell$ is the net present value of expected future income. When workers are infinitely patient, the net present value of expected future income equals expected income: $\mathcal{W}_\ell = \frac{1}{1+k^u}b + \left(1 - \frac{1}{1+k^u}\right)\mathbb{E}_\ell[w]$, where $\frac{1}{1+k^u}$ is the *ex-ante* probability to be unemployed. Equations such as (14) are standard in static spatial frameworks except that workers choose where to live based on their current, not future, earnings.

In cities with a better job offer distribution, expected future incomes are higher. In equilibrium, housing prices adjust to make unemployed workers indifferent across locations, and as a result, the real reservation wage must be lower for unemployed workers to be willing to live elsewhere. Said differently, workers on low rungs of the ladder accept jobs with relatively lower real earnings in anticipation of high future wage growth.

Proposition 1 summarizes the results on the bottoms of the local job ladders.

Proposition 1 (Reservation wage)**.**
Consider two locations, ℓ and ℓ', and ℓ has a higher wage offer distribution in the FOSD sense. Then the nominal reservation wage is higher in ℓ, $\underline{w}_\ell \geq \underline{w}_{\ell'}$, strictly if $\lambda^e < \lambda^u$. Meanwhile, the real reservation wage gross of amenities is strictly lower, $A_\ell \underline{w}_\ell / P_\ell < A_{\ell'} \underline{w}_{\ell'} / P_{\ell'}$.

The wage offer distributions are equilibrium objects determined by employers' optimal wage posting. I now solve for the wage offer distributions.

2.3 Local wage distributions

I focus on the employer problem in steady state assuming that their discount rate is low. Employers choose where to produce, how many workers to hire, and which wage to offer, to maximize their flow profits:

$$\pi(z) = \max_\ell \pi_\ell(z) = \max_\ell \left\{ \max_{w,n} \ R(z,n) - wn - r_\ell \quad \text{s.t.} \quad n \leq n_\ell(w) \right\}. \tag{15}$$

Employers' hiring decision is constrained by the local labor supply curves. Wages then become an effective hiring tool: a higher wage allows employers to poach and retain more workers from the competition. In addition, employers internalize that they can adjust their labor supply curve by changing location.

The solution to (15) is given by two joint fixed points. First, given a spatial allocation of employers, the wage posted by an employer depends on the local wage offer distribution, which itself is a function of the wage-setting strategies of other firms. Second, employers factor in local prices when choosing their production location, themselves a function of other employers' location choice. I thus solve (15) in two stages. First, I derive the optimal wages for a given spatial allocation of employers. I then characterize employers' location decision.

Within cities, the results from Burdett and Mortensen (1998) apply for any spatial allocation of employers, $\{M_\ell, \Gamma_\ell\}_{\ell=1}^{L}$. Employers are on their labor supply curve, $n = n_\ell(w)$. The wage offer distributions are continuous over the interval $[\underline{w}_\ell, \bar{w}_\ell]$.[16] The complementarity in the revenue function between employer productivity and their size implies that wages are strictly increasing in z within cities: more productive employers have a larger target size, and for that, offer higher wages. As a result, the rank of a job in the local wage offer distribution must correspond to the rank of the employer in the local employer distribution. That is, the wage offer distributions satisfy $F_\ell[w_\ell(z)] = \Gamma_\ell(z)$, where $w_\ell(z)$ is the optimal wage posted by z in city ℓ.

The wage distribution of city ℓ is determined by two functions. First, the number of workers employed at a job with productivity z is

$$n_\ell(z) \equiv n_\ell[w_\ell(z)] = \frac{1}{\theta_\ell} \frac{1+k}{[1+k(1-\Gamma_\ell(z))]^2}, \quad \forall z \in \operatorname{supp} \Gamma_\ell.$$

Employer size directly follows from the labor supply curves (12) and the rank-preserving condition $F_\ell[w_\ell(z)] = \Gamma_\ell(z)$. Second, the wage offered by an employer with productivity z is

$$w_\ell(z) = \underline{w}_\ell \left(\frac{n_\ell(\underline{z}_\ell)}{n_\ell(z)} \right) + \int_{\underline{z}_\ell}^{z} \zeta \left(\frac{n'_\ell(\zeta)}{n_\ell(z)} \right) \mathrm{d}\zeta, \quad \forall z \in \operatorname{supp} \Gamma_\ell, \tag{16}$$

where $\underline{z}_\ell \equiv \min \Gamma_\ell$ is the least productive employer in ℓ.

Optimal wages (16) result from the competition for workers along the local job ladder. At the bottom of the ladder, the least productive employers only hire workers from unemployment. The lowest wages thus reflects unemployed workers' outside option, $\underline{w}_\ell$. In the limit where $\lambda^e \to \lambda^u$, the least productive employers offer the unemployment insurance b and their wage is independent of the local competition. More productive employers offer higher wages to poach workers from their lower-paying competitors. This poaching game builds up throughout the job ladder, and the wage paid by an employer depends on the productivity of the other employers on the lower rungs. The higher the relative productivity of an employer, the more competition it faces.

The competition for workers turns spatial differences in employer productivity into spatial differences in wages. Proposition 2 its corollary characterize the shape of the local wage distributions

[16] Continuity holds independently of the local employer distribution. For instance, if some employer distribution has a hole, employers with productivity to the right of it have no incentives to offer wages strictly greater than the wage offered by employers to the left of the hole as this would yield the same size.

for a particular allocation of employers.

Proposition 2 (Spatial wage inequality)**.**
Consider two locations, ℓ and ℓ', such that the employer distribution in ℓ is higher in the FOSD sense. Then,

1. *The wage distribution in ℓ first order stochastically dominate that in ℓ', $G_\ell \succ_{FOSD} G_{\ell'}$. As a consequence, the wage gains upon a a job switch are larger, $\mathbb{E}_\ell[W/w \mid W > w] > \mathbb{E}_{\ell'}[W/w \mid W > w]$, and the real reservation wage is lower.*
2. *The top-to-bottom wage gap is larger in ℓ, $\bar{w}_\ell - \underline{w}_\ell > \bar{w}_{\ell'} - \underline{w}_{\ell'}$;*

When all employers are relatively more productive, the poaching competition intensifies. As a result, the wage distribution shifts up, and along with it the average wage. However, competition impacts employers differently depending on their relative productivity. Relatively unproductive employers, who primarily hire unemployed workers, are insulated from the local competition. Workers at the bottom of the wage distributions thus earn similar wages, no matter in which city they are and who hires them. By contrast, relatively productive employers face off all the local competition, and their workers enjoy disproportionally higher wages. Greater within-city wage inequality follows. Finally, workers experience faster wage growth as they reallocate from low- to high-paying jobs, and they accept lower real earnings when unemployed in anticipation of higher future real earnings.

The location decision of employers therefore spillovers onto wages. In cities where jobs are relatively productive, wages are relatively high not because they are offered by productive employers, but rather, because productive employers intensify the local competition. Wage disparities across and within cities arise without the need for TFP differentials.

Proposition 2 holds to the extent that local productivity distributions are ordered in term of first order stochastic dominance. I now show this is the case in equilibrium.

2.4 The spatial allocation of employers

The profits of an employer at its optimal wage offer and size is

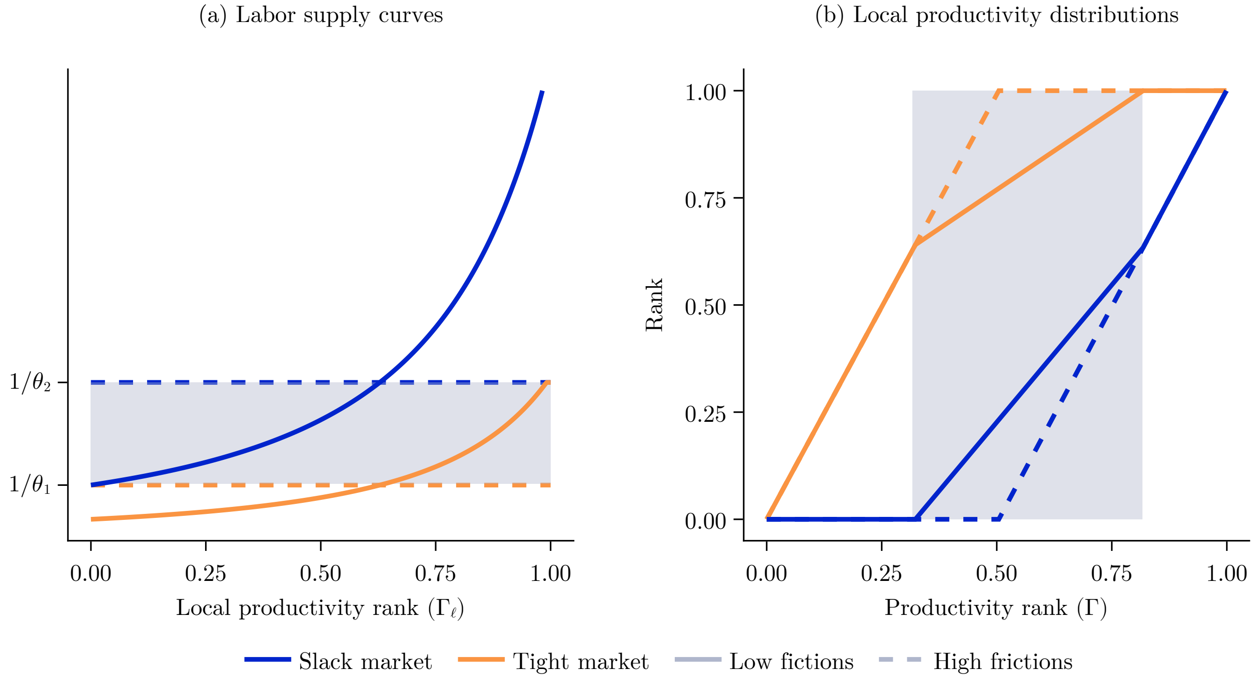
$$\pi_\ell(z) = [z - w(z, \Gamma_\ell)]\, n[\Gamma_\ell(z), \theta_\ell] - r_\ell.$$

where, with a slight abuse of notation, I have made explicit the dependence of the optimal wage and size on the local distribution of employers. Employers produce in the location that maximize their profits:

$$\pi_\ell(z) \geq \pi_{\ell'}(z) \text{ for all } \ell' \neq \ell \iff z \in \text{supp } \Gamma_\ell. \tag{17}$$

An equilibrium spatial allocation of employers is a tuple $\{M_\ell, \Gamma_\ell\}_{\ell=1}^{L}$ that satisfies feasibility (7) and profit maximization (17).

Figure 4: The spatial distribution of employers



The complementarity between productivity and size shapes the spatial allocation of employers. However, employer size is an equilibrium object. Labor market tightness determines the level of the supply curve. Local competition shapes employers' relative size. The profitability of each location thus depends itself on the spatial allocation of employers, and standard optimal transport techniques cannot be used to solve the assignment problem (17) (Chade and Eeckhout, 2020). Nevertheless, a sharp characterization of employers' location choice exists. I first show that productive employers concentrate in cities with slacker labor markets, and then establish that these are larger cities.

First, hold fix local labor market tightness. All else equal, productive employers have a stronger willingness to pay to produce in slacker labor markets to sidestep search frictions. In equilibrium, a tension arises between tightness and competition: when all productive employers agglomerate in the slackest labor market, the local competition intensifies, and most employers eventually hire only a few workers. The poaching competition thus introduces substitutability across locations. Weak search frictions increase this tradeoff by accelerating the reallocation of workers to high-paying employers. Proposition 3 formalizes this intuition.

Proposition 3 (Local productivity distributions)**.**
Locations with slacker labor market attract relatively more productive employers: $\theta_\ell < \theta_{\ell'}$ *if and only if* $\Gamma_\ell \succ \Gamma_{\ell'}$. *If search frictions are small relative to the difference in tightness,* $\sqrt{\frac{\theta_{\ell'}}{\theta_\ell}} < 1 + k$, *productivity distributions overlap,* $d\Gamma(supp\,\Gamma_\ell \cap supp\,\Gamma_{\ell'}) > 0$.

Figure 4(a) visually represents Proposition 3. The figure depicts the local labor supply curves,

n_ℓ, against employers' rank in the local productivity distribution. The blue and orange curves refer to locations with a slack and a tight labor market. The solid lines refer to an allocation with weak search frictions ($k > 0$) whereas the dashed lines depict an allocation with infinite frictions ($k \to 0$).

When search frictions are large, employers mostly hire from unemployment and the local labor supply curves are flat. Employers in slacker labor markets are uniformly larger. The complementarity between productivity and size then guarantees that if an unproductive employer find it profitable to produce there, then so must a productive employer. In addition, as in standard matching model, the uniformly greater size ensures that no (positive measure of) employers is indifferent between a slack and a tight labor market.[17]

When workers reallocate from low- to high-paying employers, the labor supply curves instead slope upwards and employers' relative productivity determines how many workers they can poach and retain from the competition. Sorting must still prevail. Within any city, relatively productive employers have a larger size than their unproductive competitors. Therefore, they still find it relatively more profitable to locate in slacker markets. However, the concentration of productive employers in slack labor markets intensify the local competition. Relatively unproductive employers struggle to hire workers and can now attain the same size in tighter labor markets with milder competition. Indifference follows for mid-productivity employers, and the sorting of employers across space is no longer perfect.

Either way, productive employers concentrate in slack markets to sidestep search frictions, and the local productivity distribution are ordered in the first order stochastic dominance.

In equilibrium, larger cities have slacker labor markets. Labor markets become tighter as the number of employers in the city rise, and so do housing prices. Both forces act as sources of congestion and prevent all employers to produce in the largest city. When the housing congestion is strong enough, large cities attract more employers but the labor market remains slacker.[18]

Proposition 4 (Existence)**.**
Suppose that $\lambda^e \approx \lambda^u$ and housing expenditures are large relative to unemployment insurance, $b\bar{H}^{1/\phi} \approx 0$. Then, in equilibrium, large cities attract more employers but have a slacker labor market: $m_\ell > m_{\ell'} \iff M_\ell > M_{\ell'} \iff \theta_\ell < \theta_{\ell'}$.

Figure 4(b) depicts the equilibrium spatial allocation of employers that correspond to the labor supply curves on the left panel. High housing prices in the large city prevent low productivity employers from producing there. When search frictions are large, the labor supply curve jump up from a tight to a slack labor market, and all employers above a productivity threshold produce in the larger location. When frictions are weaker, mid-productivity employers are indifferent between

[17] Whether employers prefer a slack or a tight labor market depends on the local housing prices, which are held fixed in Proposition 3. Accordingly, it can well be that all employers strictly prefer to locate in the tightest labor market. Then, Proposition 3 trivially holds as $\operatorname{supp}\Gamma_{\ell'} = \varnothing$ for all but one ℓ.

[18] Uniqueness is harder to establish as the underlying fixed point is non-linear and of infinite dimension. However, Proposition 4 holds in any equilibrium.

small and large places. Finally, productive employers locate in the largest city to maximize their size at the expense of higher wages and housing prices.

Figure A.8 confirms the prediction of Proposition 4 in the data. Labor market consistency implies that the tightness is the inverse of employers' average size. Figure A.8a shows that the average number of workers hired by employers is larger in bigger cities.

Propositions 1 to 4 thus show that the sorting of employers across space generates between- and within-city inequality once combined with frictional labor markets. Productive employers concentrate in large cities to sidestep search frictions. The local labor market competition intensifies and boosts wages. The higher productivity spillovers disproportionally to workers at the top of the job ladder without materializing for low-paid workers. As a result, larger cities offer higher wages, greater within-city inequality, and faster wage growth as workers reallocate from low- to high-paying employers.

I now estimate the model to assess whether this framework generates spatial wage inequality that quantitatively aligns with the data. I then use the estimated model to reassess the consequences of agglomeration on wages once between- and within-city inequality are taken into account.

3 Extended model and estimation

3.1 Quantitative model

Extensions I relax the strong assumptions imposed in Section 2. Across space, I allow cities to differ in terms of TFP T_ℓ. Local TFP depends on an exogenous component and reduced-form productivity spillovers, $T_\ell = \mathcal{T}_\ell m_\ell^\nu$, where ν is the spillover elasticity. Since unemployment insurance does not scale up with TFPs, higher TFP generates higher wages and wage inequality. It is therefore a quantitative question whether spatial wage disparities arise from employer sorting or TFP gaps.

On the employer side, I made two extensions. First, employers face idiosyncratic housing or entry costs, $\{\varepsilon_\ell\}_{\ell=1}^L$. Second, I let employers hire many workers without increasing wages too rapidly by posting several vacancies — or more generally, exerting endogenous hiring effort. Vacancies are costly and come at a convex cost

$$c(v) = \frac{v^{1+\gamma}}{1+\gamma},$$

where $\gamma > 0$ denotes the vacancy cost elasticity. Employer size reflects their position in the local job ladder together with their vacancy share:

$$n_\ell(w, v) = \frac{(1+k_\ell)e_\ell}{[1+k_\ell(1-F_\ell(w))]^2}\frac{v}{V_\ell}, \tag{18}$$

where V_ℓ is the aggregate number of vacancies posted in location ℓ,

$$V_\ell = M_\ell \int v_\ell(z) d\Gamma_\ell(z). \tag{19}$$

Search frictions are now city-specific and are determined by the matching function $\lambda_\ell^u = \mathcal{M}(u_\ell + \zeta e_\ell, V_\ell)/(u_\ell + \zeta e_\ell)$, where ζ is the relative search intensity of employed workers, $\lambda_\ell^e = \zeta \lambda_\ell^u$.

With these two new ingredients, an employer with productivity z and entry costs $\{\varepsilon_\ell\}_{\ell=1}^L$ solves

$$\pi(z, \{\varepsilon_\ell\}_{\ell=1}^L) = \max_{\ell,w,n,v} R(zT_\ell, n) - wn - c(v) - r_\ell - \varepsilon_\ell \quad \text{s.t.} \quad n \leq n_\ell(w, v), \tag{20}$$

where the total productivity of its jobs in location ℓ is zT_ℓ.

Finally, I introduce two additional reasons why workers may accept lower real earnings in large cities. First, I add migration costs. Workers permanently exit the labor force at Poisson rate χ, upon which they are replaced by their descendant. New entrants enter the labor force in the same location as their parent, and must pay a pecuniary flow migration cost κ to relocate elsewhere.[19] Migration costs imply that the welfare of unemployed workers are no longer equalized across locations.

Second, individuals have idiosyncratic preferences for each location. These preferences, which also ensure uniqueness of the equilibrium when they are sufficiently dispersed across workers, are drawn upon entry in the labor force and remain constant thereafter. The flow utility of a worker with preferences $\{\omega_{\ell'}\}_{\ell'=1}^L$ in location ℓ is

$$u_\ell(c, h, \{\omega_{\ell'}\}_{\ell'=1}^L) = A_\ell \omega_\ell \left(\frac{h}{\alpha}\right)^\alpha \left(\frac{c}{1-\alpha}\right)^{1-\alpha}.$$

A full description of the model, together with a definition of the equilibrium, is presented in Appendix C.1.

Parametric assumptions I impose the following parametric assumptions. The aggregate productivity distribution is Pareto with shape σ. Its scale is normalized to ensure a mean productivity of one. The idiosyncratic entry costs of employers are i.i.d. across space. They are drawn from Gumbel distribution with inverse dispersion ϑ. Workers' idiosyncratic preferences are i.i.d. Fréchet distributed with shape χ. Finally, the matching function is Cobb-Douglas with matching efficiency μ and elasticity ψ: $\mathcal{M}_\ell = \mu V_\ell^\psi (u_\ell + \zeta e_\ell)^{1-\psi}$.

3.2 Identification

Model inversion The estimation of the model is organized in two blocks. Its core is contained in the first block, which leverages the structure of the model to estimate the bulk of the parameters without the need for simulation. Proposition 5 proves identification of 16 of the 21 parameters.

[19] This migration cost can alternatively be interpreted as losses in flow utility.

Proposition 5 (Identification)**.**
Given aggregate data on flows in- and out- of employment, location-specific data on job switching rates, commercial and residential housing prices, and firm-level employment and wage data, the parameters α, b, δ, ζ, ξ, μ, ψ, M, γ, $\bar{L}$, θ, $\bar{H}$, and ϕ are identified. The dispersion in employers' idiosyncratic entry cost ϑ is identified conditional on small TFP differentials. Local amenities $\{A_\ell\}_{\ell=1}^{L}$ and the migration cost κ are identified given workers' discount factor ρ and idiosyncratic location preferences χ.

The proof is detailed in Appendix C.2. I present here the intuition behind the proof, focusing on the parameters that govern employer sorting. The location choice of employers is pinned down by four sets of parameters: the spatial allocation of workers, search frictions, the dispersion in entry cost, and TFPs. Three of those four parameters are identified by Proposition 5.

First, the spatial allocation of workers is observed and can be rationalized by the appropriate vector of amenities. Hence, it can be treated as a primitive for the purpose of identification.

Second, search frictions are identified from worker flows. The aggregate employment-to-unemployment rate is equal to δ. The labor force exit rate ξ is identified from the average career duration: $\xi = 1/\mathbb{E}[\text{career length}]$. The location-specific job switching rates J2J_ℓ identify the contact rates of employed workers λ_ℓ^e after accounting for rejected offers along the job ladder: $\text{J2J}_\ell = \delta((1+\delta/\lambda_\ell^e)\log(1+\lambda_\ell^e/\delta)-1)$. The aggregate unemployment-to-employment rate relative to the J2J rate pins down the search efficiency of employed workers ζ. Finally, the matching function is identified from the correlation between the local contact rates and labor market tightness, $\log \lambda_\ell^u = \log \mu + \psi \log \frac{V_\ell}{u_\ell + \zeta e_\ell}$, where the mass of vacancies V_ℓ is recovered from employer size and their position on the local job ladders.

Third, the dispersion in entry costs is identified from the correlation between local profit opportunities and employers' location choice. Indeed, given the Gumbel parametric assumption, the log-likelihood that an employer j with productivity z_j locates in ℓ is

$$\log \Omega_\ell(z_j) = H(z_j) + \vartheta \pi_\ell(z_j), \tag{21}$$

where $H(z) \equiv -\log \sum_{\ell'} \mathrm{e}^{\vartheta \pi_{\ell'}(z)}$.

To recover profits on the right-hand side of (21), I extend the insights of Bontemps et al. (2000) to a setting with local labor markets. Employer productivity is identified from wages net of markdowns. Specifically, the wage optimality condition behind (16) demands

$$z_j T_{\ell_j} \equiv \zeta_j = w_j + \left(\frac{1 + k_{\ell_j}(1 - F_j)}{2k_{\ell_j}}\right) \frac{\partial w_{\ell_j}(F_j)}{\partial F}, \tag{22}$$

where ℓ_j denote the production location of j, F_j their rank in the local wage offer distribution, and $w_\ell(F)$ is the inverse function of $F_\ell(w)$. ζ_j refers to the total productivity of j gross of the local TFP.

Vacancy costs are computed from employer size net of their position in the job ladder. Employers' rank, together with the search frictions they face, determines the number of workers they hire per vacancy through the labor supply curves. Their vacancy share v_j/V_{ℓ_j} are then recovered from their size using (18). The vacancy optimality condition $\log v_j/V_{\ell_j} = \frac{1}{1+\gamma}\log[(\zeta_j - w_j)n_j] - \log V_{\ell_j}$ identifies the vacancy cost elasticity γ. It also pins down the number of vacancies in each location, and therefore the vacancy cost paid by employer j: $c_j = v_j^{1+\gamma}/(1+\gamma)$.

Combining these steps, I obtain profits as revenues net of wages, vacancy costs, and housing costs: $\pi_{\ell_j}(\zeta_j) = (\zeta_j - w_j)n_j - c_j - r_{\ell_j}$. If TFPs are relatively similar across locations, the productivity of employers closely tracks their MPL. In that case, $\pi_\ell(\zeta_j) \approx \pi_\ell(z_j)$, and I can use (21) to directly estimate the dispersion in entry costs. In practice, TFP gaps may be large. Therefore, I instead target the conditional correlation from (21) in the second estimation block.[20]

Indirect inference The second step of the estimation calibrates 4 of the remaining 6 parameters by indirect inference. Given the parameters identified by Proposition 5, I simulate the model and minimize the distance between a vector of empirical statistics and the same statistics in the model. First, I estimate (21) in the model in the same way as in the data, and I set the entry costs dispersion to match the empirical conditional correlation. Second, I calibrate local TFPs to match average city wages net of the employer sorting predicted by the model. Given a vector of TFP, productivity spillovers are obtained from $\log T_\ell = \log \mathcal{T}_\ell + \nu \log m_\ell$. Third and last, I calibrate the productivity dispersion σ to match the aggregate wage variance.

3.3 Data

I briefly describe the data used for the estimation before turning to the estimation results. Further details are included in Appendix C.3.

I set a quarterly frequency and normalize all nominal variables by the aggregate average wage. To abstract from worker heterogeneity and focus on the causal impacts of cities on wages, I measure wages by the employed fixed effects estimated in Section 1.3.[21]

I solve the model for the 10 city groups defined in Section 1.1. There remains 297 local labor markets in the model, but each local labor market within a group is homogeneous. Increasing the number of city groups would increase measurement errors in the within-city employer fixed effect distributions.

I obtain residential housing prices from the *Carte des Loyers* (Rental Map), which estimates the average monthly rent per meter square at the municipal level. I do not have access to city-level

[20]The rest of the proof is standard. The Cobb-Douglas parameter is pinned down by the aggregate housing expenditure share. The unemployment insurance is given by the aggregate replacement rate. The housing supply parameters are obtained from the housing market clearing conditions. The migration cost is identified from the probabilities of workers leaving their hometowns relative to the chances of staying. Finally, amenities are recovered as a residual from the spatial allocation of workers.

[21]The conditional random mobility assumption behind the AKM specification is valid in random-search wage-posting models (Card et al., 2013). The log-linear specification is well-specified if workers' skills affect their wages multiplicatively and search frictions are constant across skills. When these assumptions fail, Bilal and Lhuillier (2021) shows that the AKM specification still fits well the data.

commercial rental rates. Instead, I define commercial housing prices as residential housing prices adjusted for the aggregate relative price of commercial to residential housing. I residualize residential housing prices by the mean worker fixed effect of each city to account for the sorting of workers across space.

The aggregate flows in and out of employment are obtained from the *Enquête emploi en continu* (Labor Force Survey). The local job switching rates are computed from the panel matched employer-employee data. Consistent with the model, I define a job switch as a transition between jobs within a location that leads to a wage increase. To account for the sorting of workers across space, I project worker-level job switching rates on worker and location fixed effects and use the latter for the estimation.

I read the housing expenditure share from the aggregate national accounts (INSEE, 2020) and set the replacement rate to 0.6 (OECD, 2025).

Finally, I externally calibrate the discount rate and the dispersion in workers' idiosyncratic location preferences. I set $\rho = 0.004$ to match an annual real interest rate of 5%. I follow the literature and set the taste shock dispersion to 2 (e.g., Fajgelbaum et al., 2019). These two parameters are relevant only for the estimation of amenities and migration costs.

3.4 Results

Table 1 presents the parameter estimates. Search frictions are relatively strong and constant across space. The aggregate contact rate for unemployed workers is 0.2. The job switching rate is substantially lower (3.7%). However, employed workers do not necessarily switch jobs when receiving a job offer. Accounting for rejected offers along the job ladder, I find that the search efficiency of employed workers is only 10% lower than that of the unemployed. I obtain a matching function elasticity of 0.15, towards the low end of the estimates found in the literature. However, this elasticity implies that the model fits well the between-city differences in contact rates with an R^2 of 0.89 (Figure C.10b).

The average labor supply elasticity in the model ($\mathbb{E}[\partial \log n_\ell(w)/\partial \log w]$) aligns with benchmark estimates. Unlike models that microfound upward-slopping labor supply curves with taste shocks, the labor supply elasticity in my framework is not a structural parameter. Instead, it depends jointly on search frictions and the vacancy cost elasticity. I estimate a vacancy cost elasticity of 2.05, well within the range found in the literature. Together with the estimated search frictions, the model produces an average labor supply elasticity of 5.92, close to the estimates of Lamadon et al. (2022) for the United States.

My estimate of the entry costs dispersion also aligns with those in the literature. I estimate $\vartheta = 1.86$, close to the conditional correlation of 1.26 implied by (21). Converting my estimate to a Fréchet elasticity, I obtain 1.46, between the 1.31 of Giroud and Rauh (2019) and the 2.63 of Fajgelbaum et al. (2018).

Lastly, I estimate that TFPs are very similar across locations once accounting for employer

Table 1: Parameter estimates

	Parameter	Target	Empirical moment	Simulated moment	Parameter estimate
		A. External calibration			
ρ	Discount factor	Annual interest rate 5%			0.004
χ	Taste shock dispersion	Fajgelbaum et al. (2019)			2.000
		B. Model inversion			
α	Cobb-Douglas preference	Housing exp. share	0.200		0.200
$\boldsymbol{A}$	Amenities	Unemployment distribution	Fig. C.9a		Fig. C.9a
b	Unemployment insurance	Replacement rate	0.600		0.638
κ	Migration cost	Migration rate	0.454		0.689
δ	Job destruction rate	EU rate	0.021		0.021
ζ	Rel. search efficiency	EE / UE rate	0.190		0.903
ξ	Exit rate	Average career length	120.0		0.008
μ	Matching efficiency	Avg. EE rate	0.037		0.285
ψ	Matching function elasticity	Correlation EE rate - vacancy	0.146		0.146
M	Mass employers	Average size	7.446		0.117
γ	Vacancy cost elasticity	Vacancy optimality (45)	2.024		2.024
$\bar{L}$	Residential housing supply level	Avg. residential price	0.488[a]		180.5
θ	Residential housing supply elasticity	Correlation prices - expenditures	0.094		9.666
$\bar{H}$	Commercial housing supply level	Avg. commercial price	1.214		0.162
ϕ	Commercial housing supply elasticity	Correlation prices - employer demand	0.209		4.790
		C. Indirect inference			
σ	Productivity dispersion	Wage variance	0.150	0.151	14.22
ϑ	Entry cost dispersion	Equation (21)	1.262	1.310	1.858
$\boldsymbol{T}$	TFP	Average wage	Fig. C.10a	Fig. C.10a	Fig. C.9b
ν	TFP spillovers	Correlation TFP - size	0.004	0.004	0.004

[a] Average residential price expressed in hundredths.

sorting (Figure C.9b). There are large disparities in city size, and search frictions are relatively strong. Productive employers therefore have strong incentives to concentrate in large cities to sidestep search frictions. For instance, the mean employment-weighted employer productivity in Paris is 23% larger than in Lens. By contrast, the average wage gap is 13%. The spatial sorting of employers thus suffices to rationalize the observed between-city wage gaps: the between-city TFP variance explains 0.59% of the total between-city productivity variance, and 1.55% of the between-city wage variance. Given the small dispersion in TFP, I estimate a spillover elasticity of 0.004, one order of magnitude smaller than typical estimates (e.g., Combes and Gobillon, 2015; Duranton and Puga, 2020).

The selection of employers across locations therefore explains between-city wage inequality. I now assess whether it also generates within-city inequality aligned with the facts documented in

Section 1.

4 Employer sorting, local competition, and spatial wage inequality

The estimation strategy proposed in Section 3.2 does not target neither how inequality varies across space nor the local wage returns to job switching. In this section, I use the two novel facts of Section 1 as over-identification exercises to assess the quantitative relevance of the model.

4.1 The impact of local competition on spatial wage inequality

Figure 5a displays the city size elasticities of the local wage deciles. The orange circles are the elasticities in the data as in Figure 2c. The blue rectangles are the same statistics in the model. The grey dashed line is the size elasticity of average wages, which is targeted in the estimation.

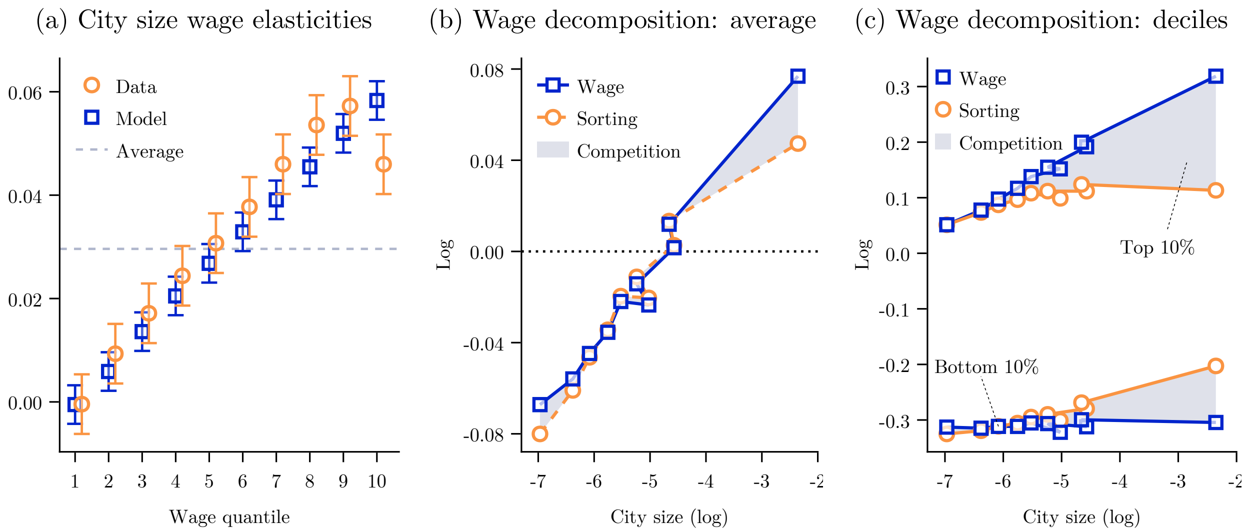
Greater inequality endogenously arises in larger cities as they disproportionally affects workers across the wage distribution. In both the data and the model, wages are on average higher in larger locations. However, there are substantial differences across the wage distributions. Workers at the bottom earn the same wage everywhere. For instance, in the model, workers in the bottom 10% earn wages 0.5% higher in Paris than in Lens, compared to 1.6% lower in the data. Meanwhile, workers at the top of the local wage distribution disproportionally gains from working in larger cities. For instance, the wage gap between Paris and Lens for workers in the top 10% of their local distribution is 20% in the model and 18% in the data. Overall, the model succeeds at generating greater spatial variations in the right tail of the local distributions.

Higher wages and greater inequality follows from the spatial concentration of productive employers and the fiercer competition that results. To decompose the relative importance of employer composition and local competition on wages, I define the markdown charged by employer z in location ℓ as $\mu_\ell(z) \equiv w_\ell(z)/zT_\ell$. Markdowns are the standard metrics to quantify labor market power. With this convention, a higher markdown $\mu_\ell(z)$ means less market power. I also define the (unweighted) average markdown charged by employer z as $\bar{\mu}(z) = \sum_\ell \mu_\ell(z)$. Then, for any wage partition $\mathcal{W}$ (e.g., wages in the bottom 10%), the average log wage in location ℓ amongst $\mathcal{W}$ can be decomposed into

$$\mathbb{E}_\ell[\log w \mid \mathcal{W}] = \mathbb{E}_\ell[\log z\bar{\mu}(z) \mid \mathcal{W}] + \mathbb{E}_\ell\left[\log \frac{\mu_\ell(z)}{\bar{\mu}(z)} \mid \mathcal{W}\right]. \tag{23}$$

The first term captures the direct effect of employer sorting on wages. As productive employers concentrate in large cities, the marginal products of labor (MPL) rise, and with that wages. However, monopsony power implies that wages and MPLs differ. In particular, productive employers tend to charge lower markdowns wherever they produce because their overall position in the job ladder shields them away from the competition (Gouin-Bonenfant, 2022). For instance, while the average markdown is 0.82, it is 4 percentage points lower (0.78) for employers in the top productivity decile.

Figure 5: The drivers of local wage inequality



The spatial sorting of employers therefore has a direct impact on both the local productivity and markdowns, and the composition channel reflects both dimensions.

The second term represents the impacts of local competition on wages. It is akin to a "within-firm" effect: by how much the local competition affects the markdowns charged by an employer above and beyond what they would usually price.

Figure 5b first applies this decomposition for average wages. The blue markers display the average wage in each city as in Figure 2a. The orange circles plot the composition effect. The difference between the two, summarized by the grey shaded areas, is the competition effect.

Local labor market competition boosts the wages of small and large cities. The direct impact of employer sorting accounts for 82% of the between-city wage variance, and all of the average wage variation amongst mid-size cities. Meanwhile, local competition boosts wages substantially at the tails of the city size distributions. In Paris, the spatial concentration of productive employers increases the average markdown by 2 percentage points. As a result, the average wage increases by 3%. Similarly, the agglomeration of unproductive employers in small cities also increases the relative competition they face, pushing the average markdown and wages by 1 p.p. and 1.3%. The spatial sorting of employers thus intensifies competition throughout France on average rather than redistributing away from small places.

Figure 5c then applies the decomposition (23) separately for wages in the bottom and top 10% of the local wage distribution. It reveals opposite impacts of competition for low- and high-wage workers.

Low-wage workers earn similar wages throughout space because the lack of competition faced by low-paying employers prevents spatial productivity differentials to materialize. For instance, workers in the bottom 10% of the wage distribution in Paris are hired by employers with a mean productivity 11% higher than workers at the bottom in Lens. However, employers in both places hire most of their workforce from unemployment. They therefore face similar competitive pressure, and the higher productivity in Paris does not spillover onto higher wages.

By contrast, the fiercer competition in large cities is what allows workers at the top of the distribution to earn high rents. While workers in the top 10% of the wage distribution in Paris are hired by the most productive employers, these employers generally charge substantially lower markdowns. As a result, the direct, compositional effect only generates wages in Paris 3% higher than in Lens. However, every employer in Paris is relatively more productive. This competitive pressure builds throughout the local job ladder as employers compete between each other, and forces employers in the top 10% of the wage distribution in Paris to offer markdowns 13 percentage points higher.

The differences in local competition across cities therefore flattens the bottoms of the wage distributions and sharpens the tops, generating the spatial differences in within-city inequality that we observe in the data.[22] Absent the fiercer local competition, larger cities would not be more unequal (Figure D.13b).

4.2 The consequences of the local job ladders on lifetime earnings

The second over-identification exercise assesses the ability of employer sorting in generating spatial differences in job ladder steepness. This is an important test as the reallocation of workers across employers is the core force driving within-city wage dispersion. Furthermore, the steepness of the local ladders determines the impact of cities on lifetime earnings.

Figure 6a plots the local returns to job switching, in the model on the y-axis and in the data on the x-axis. In the data, I use the estimates restricted to job switches with wage growth to ensure consistency with the model. In the model, the local returns are computed as the average wage growth conditional on a job switch:

$$\mathbb{E}_\ell\left[\frac{W'}{W} \mid \text{switch}\right] = \int \frac{1}{1-F_\ell(w)} \int_w \frac{w'}{w} dF_\ell(w') dG_\ell(w).$$

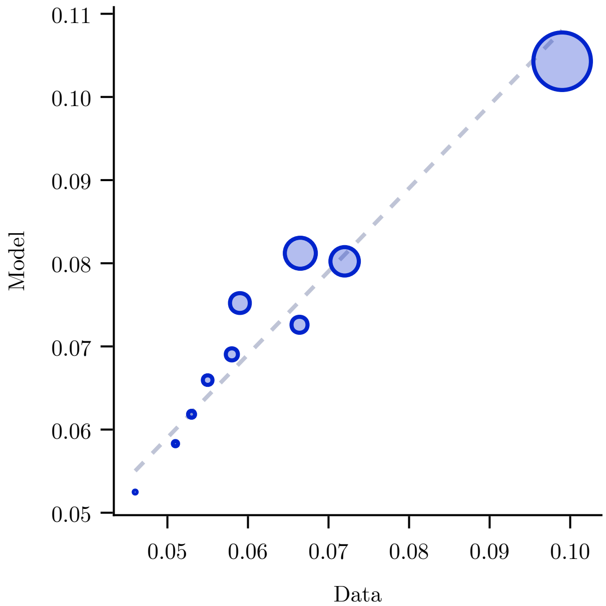
The aggregate returns to job switching in the model closely align with the empirical estimates: it leads to an average wage growth of 8%, slightly above the 7% estimated in the data. Across space, switching jobs generates faster wage growth in larger cities; the correlation between the model's estimates and those in the data is 94%. I therefore conclude that search frictions are a valid microfoundation for within-city wage inequality.

The steeper job ladder in larger cities generates higher lifetime real earnings for the workers there. The orange circles in Figure 6b shows the present real earnings of unemployed workers in each location. The blue rectangles are their net present value of real lifetime earnings. Unemployed workers earn lower real earnings in larger cities as unemployment insurance is constant throughout

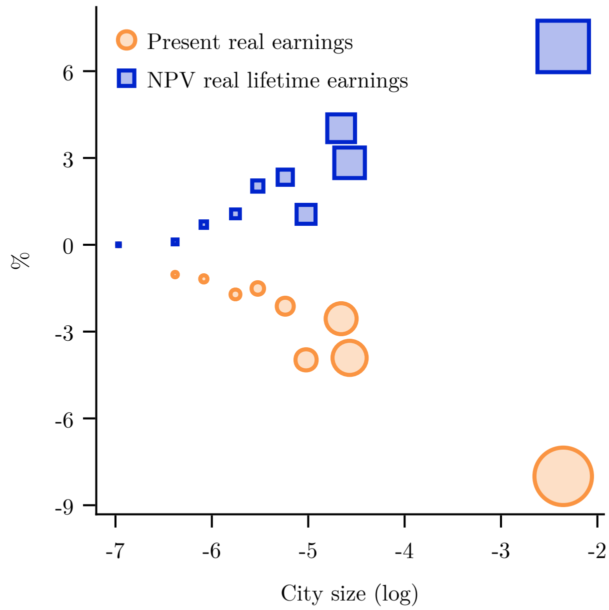
[22]Figure D.13a applies the wage decomposition (23) to all wage deciles. The top 30% of the wage distribution in larger cities benefit from the fiercer competition.

Figure 6: Welfare implications of the local job ladders

(a) Wage gains upon job switch



(b) Lifetime real earnings



space but housing prices are more expensive. However, real lifetime earnings are higher as workers have access to better future opportunities. I find that the net present value of unemployed workers' real lifetime earnings is 1.3% higher in cities twice larger.

Together, the two over-identification exercises validate the model's ability to generate between- and within-city inequality. Spatial wage disparities arise as productive employers agglomerate in large cities to partially sidestep search frictions. I now turn to quantifying what are the aggregate consequences of agglomeration on wages and wage inequality.

5 A new assessment of agglomeration benefits

References

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

Baum-Snow, N., Freedman, M., and Pavan, R. (2018). Why has urban inequality increased? *American Economic Journal: Applied Economics*, 10(4):1–42.

Baum-Snow, N. and Pavan, R. (2012). Understanding the city size wage gap. *The Review of economic studies*, 79(1):88–127.

Baum-Snow, N. and Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics*, 95(5):1535–1548.

Behrens, K., Duranton, G., and Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3):507–553.

Berger, D., Herkenhoff, K., and Mongey, S. (2022). Labor market power. *American Economic Review*, 112(4):1147–93.

Bilal, A. (2023). The geography of unemployment. *The Quarterly Journal of Economics*, 138(3):1507–1576.

Bilal, A. and Lhuillier, H. (2021). Outsourcing, inequality and aggregate output. Technical report, National Bureau of Economic Research.

Bonhomme, S., Lamadon, T., and Manresa, E. (2019). A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739.

Bontemps, C., Robin, J.-M., and Van den Berg, G. J. (2000). Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation. *International Economic Review*, 41(2):305–358.

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, pages 257–273.

Card, D., Heining, J., and Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, 128(3):967–1015.

Card, D., Rothstein, J., and Yi, M. (2025). Location, location, location. *American Economic Journal: Applied Economics*, 17(1):297–336.

Chade, H. and Eeckhout, J. (2020). Competing teams. *The Review of Economic Studies*, 87(3):1134–1173.

Combes, P.-P., Duranton, G., and Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of urban economics*, 63(2):723–742.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543–2594.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In *Handbook of regional and urban economics*, volume 5, pages 247–348. Elsevier.

Dauth, W., Findeisen, S., Moretti, E., and Suedekum, J. (2022). Matching in cities. *Journal of the European Economic Association*, 20(4):1478–1521.

Duranton, G. and Puga, D. (2020). The economics of urban density. *Journal of economic perspectives*, 34(3):3–26.

Eckert, F., Hejlesen, M., and Walsh, C. (2022). The return to big-city experience: Evidence from refugees in denmark. *Journal of Urban Economics*, page 103454.

Eeckhout, J., Pinheiro, R., and Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, 122(3):554–620.

Engbom, N. and Moser, C. (2021). Earnings inequality and the minimum wage: Evidence from brazil.

Fajgelbaum, P. D., Morales, E., Suárez Serrato, J. C., and Zidar, O. (2019). State taxes and spatial misallocation. *The Review of Economic Studies*, 86(1):333–376.

Fajgelbaum, P. D., Morales, E., Suárez Serrato, J. C., and Zidar, O. (2018). State taxes and spatial misallocation. *The Review of Economic Studies*, 86(1):333–376.

Franco, S. (2024). *Output market power and spatial misallocation*. The University of Chicago.

Gaubert, C. (2018). Firm sorting and agglomeration. *American Economic Review*, 108(11):3117–53.

Giroud, X. and Rauh, J. (2019). State taxation and the reallocation of business activity: Evidence from establishment-level data. *Journal of Political Economy*, 127(3):1262–1316.

Glaeser, E. L., Resseger, M., and Tobio, K. (2009). Inequality in cities. *Journal of Regional Science*, 49(4):617–646.

Goldin, I. (2023). To tackle inequality, start with cities. *Financial Times*. Accessed: 2025-02-27.

Gouin-Bonenfant, É. (2022). Productivity dispersion, between-firm competition, and the labor share. *Econometrica*, 90(6):2755–2793.

Heise, S. and Porzio, T. (2022). Labor misallocation across firms and regions.

INSEE (2020). Tableaux de l'economie française. Consommation des menages.

Jarosch, G., Nimczik, J. S., and Sorkin, I. (2021). Granular search, market structure, and wages.

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

Lhuillier, H. (2023). Should i stay or should i grow?

Lindenlaub, I., Oh, R., and Peters, M. (2022). Firm sorting across space.

Martellini, P. (2020). The city-size wage premium: Origins and aggregate implications.

Morchio, I. and Moser, C. (2024). The gender pay gap: Micro sources and macro consequences. Technical report, National Bureau of Economic Research.

OECD (2025). Benefits in unemployment, share of previous income.

Oh, R. (2023). Spatial sorting of workers and firms. Technical report, Technical report.

Papageorgiou, T. (2022). Occupational matching and cities. *American Economic Journal: Macroeconomics*, 14(3):82–132.

Porcher, C., Rubinton, H., and Santamaria, C. (2021). The role of establishment size in the city-size earnings premium.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9:21–58.

Roca, J. D. L. and Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, 84(1):106–142.

Schmutz, B. and Sidibé, M. (2019). Frictional labour mobility. *The Review of Economic Studies*, 86(4):1779–1826.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and Von Wachter, T. (2019). Firming up inequality. *The Quarterly journal of economics*, 134(1):1–50.

Topkis, D. M. (1998). *Supermodularity and complementarity.* Princeton university press.

United Nations (2020). World social report 2020: Inequality in a rapidly changing world.

A Two facts about spatial wage inequality

A.1 Data

Sample restrictions The employer tax records comes in two formats. The first format is a long-panel that tracks the labor market history of 4% of the French workforce. The second format is a repeated short-panel (two years) that provides information on the universe of jobs held by workers —where a job is defined a pair establishment × occupation. I apply the same restrictions on both datasets:

1. Exclude workers younger than 25 and older than 55 year old;
2. Exclude workers employed in the public sector;
3. Exclude the agriculture, education and health industries;
4. Keep only workers employed full-time;
5. Exclude workers that are non-employed for more than three years;
6. Exclude employment spell that lasts less than 30 days;
7. Exclude employment spell with no labor income or hours worked;

Construction of the panel Once a worker enters the long-panel, each of their employment spell are recorded. An employment spell is defined as a pair establishment × occupation. The dataset provides the start and end days of each employment spell. Workers can be observed multiple times within a given period if they work for multiple employers or if they switch employers. By contrast, a worker is observed only once per year if they work for a unique employer during that entire year.

I aggregate the data at the quarterly level. If workers hold multiple jobs within a quarter, I keep the job that provides the highest total labor income. I only keep the information (e.g., employer's ID, occupation, etc.) associated to that employment spell.

A.2 Accounting for sorting

The variance of log wages in any location reads

$$\mathbb{V}\mathrm{ar}_\ell[\log w_{it}] = \mathbb{V}\mathrm{ar}_\ell[\psi_{it}] + \mathbb{V}\mathrm{ar}_\ell[\gamma_{it}] + 2\rho_\ell(\psi_{it},\gamma_{it})\mathbb{S}\mathrm{d}_\ell[\psi_{it}]\mathbb{S}\mathrm{d}_\ell[\gamma_{it}] + \mathbb{V}\mathrm{ar}_\ell[\varepsilon_{it}],$$

where $\rho(x, y)$ is the correlation between (x, y). The third term can thus be large because the correlation between the employer and fixed effects is larger, or either are very dispersed. No exact decomposition exists to separate those three effects. Instead, consider a first order approximation of the wage variance around the point where spatial inequality is the same everywhere; i.e., $(\mathbb{V}\mathrm{ar}_\ell[\psi_{it}], \mathbb{V}\mathrm{ar}_\ell[\gamma_{it}], \rho_\ell(\psi_{it},\gamma_{it}), \mathbb{V}\mathrm{ar}_\ell[\varepsilon_{it}]) \approx (\mathbb{V}\mathrm{ar}[\psi_{it}], \mathbb{V}\mathrm{ar}[\gamma_{it}], \rho(\psi_{it},\gamma_{it}), \mathbb{V}\mathrm{ar}[\varepsilon_{it}])$. To a first order, we have

$$\begin{aligned}\mathbb{V}\mathrm{ar}_\ell[\log w_{it}] \approx \mathbb{V}\mathrm{ar}_\ell[\psi_{it}] + \rho(\psi_{it},\gamma_{it})(\mathbb{V}\mathrm{ar}_\ell[\psi_{it}] - \mathbb{V}\mathrm{ar}[\psi_{it}])\left(\frac{\mathbb{S}\mathrm{d}[\gamma_{it}]}{\mathbb{S}\mathrm{d}[\psi_{it}]}\right) +& \\ \mathbb{V}\mathrm{ar}_\ell[\gamma_{it}] + \rho(\psi_{it},\gamma_{it})\left(\mathbb{V}\mathrm{ar}_\ell[\gamma_{it}] - \mathbb{V}\mathrm{ar}[\gamma_{it}]\right)\left(\frac{\mathbb{S}\mathrm{d}[\psi_{it}]}{\mathbb{S}\mathrm{d}[\gamma_{it}]}\right) +& \\ 2\rho_\ell(\psi_{it},\gamma_{it})\mathbb{S}\mathrm{d}[\psi_{it}]\mathbb{S}\mathrm{d}[\gamma_{it}] + \mathbb{V}\mathrm{ar}_\ell[\varepsilon_{it}] + o_\ell,&\end{aligned} \tag{24}$$

where o_ℓ is the approximation error. The first line is the impact of worker heterogeneity on local wage inequality. The second line is the impact of employer heterogeneity on local wage inequality. Finally, the third line captures the impact of sorting and the residuals.

A.3 Tables

Table A.1: Worker-level aggregate summary statistics

Table A.2: City cluster-level summary statistics

Cluster	# CZs	Size	Avg. wage	St. dev.	P10	P90	Rent	Smallest CZ	Largest CZ
1	100	8,518	16.61	0.31	11.33	23.48	8.41	Le Blanc	La Roche-sur-Yon
2	53	15,924	17.34	0.34	11.35	25.38	9.58	Châteaudun	Vannes
3	41	21,214	18.04	0.35	11.54	26.73	9.62	Commercy	Metz
4	29	28,976	18.54	0.37	11.55	28.15	10.38	Tergnier	Tours
5	24	36,931	19.44	0.38	11.79	29.88	10.48	Houdan	Rouen
6	13	62,581	19.37	0.39	11.64	30.16	11.34	Wissembourg	Bordeaux
7	19	50,164	20.33	0.41	11.75	32.48	11.62	Ambert	Toulouse
8	8	101,150	21.43	0.43	11.92	34.61	13.96	Chinon	Roissy
9	8	91,755	22.39	0.45	12.00	36.28	14.54	Étampes	Lyon
10	2	924,781	27.63	0.52	12.53	46.09	22.25	Saclay	Paris

The columns are: the cluster ID, the number of commuting zones in the cluster, the average number of employed worker, the average wage, the standard deviation of log wages, the 10^{th} percentile of the wage distribution, the 90^{th} percentile of the wage distribution, the rent per meter square, the smallest commuting zone in the cluster, and the largest commuting zone in the cluster. All statistics are computed at the CZ level and then averaged at the cluster.

Table A.3: Between-city wage variance decomposition

Table A.4: The local returns to job switching

	(1)	(2)	(3)	(4)	(5)
# jobs	0.013	0.056	0.015	0.008	0.012
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
# jobs × ineq.	0.081	0.151	0.072	0.072	0.075
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
City × year F.E.			✓		
Occupation F.E.				✓	
Worker slope F.E.					✓
Switch	All	Wage gain	All	All	All
# Obs.	8,798,361	8,798,619	8,798,361	8,798,361	8,798,361
R^2	0.885	0.889	0.887	0.890	0.887

Dependant variable: log hourly wage. The indepedent variable # jobs is defined in (4), and ineq is the within-city wage variance. All regressions include: worker F.E., location F.E., experience, and experience interacted with local inequality. Standard errors are clustered at the worker level.

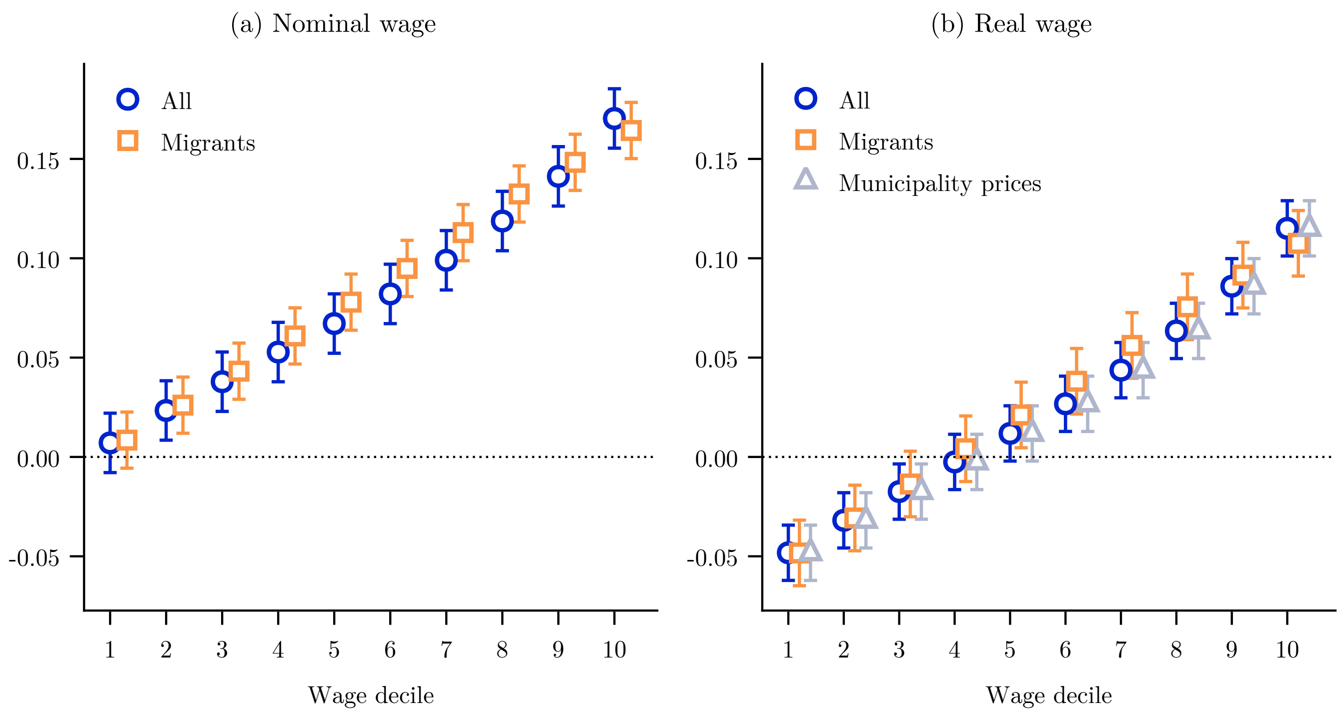
Table A.5: Job switching within and across commuting zones

	All flows	Flows with wage gain
Switching rate (%)	9.05	5.07
Within city (%)	93.6	94.2
Switchers' wage growth (%)	3.25	17.1
City stayers (%)	3.28	16.7
City movers (%)	2.71	23.7
Days between jobs	69.7	69.7
City stayers	64.6	64.6
City movers	144.5	144.5

Statistics computed at the quarterly frequency. Switching rate computed across jobs, where a job is an establishment × 4-digit occupation. Second column restricts the sample to job switches associated with positive wage growth. City stayers and movers defined based on their commuting zone of residence.

A.4 Figures

Figure A.1: City wage deciles and local inequality

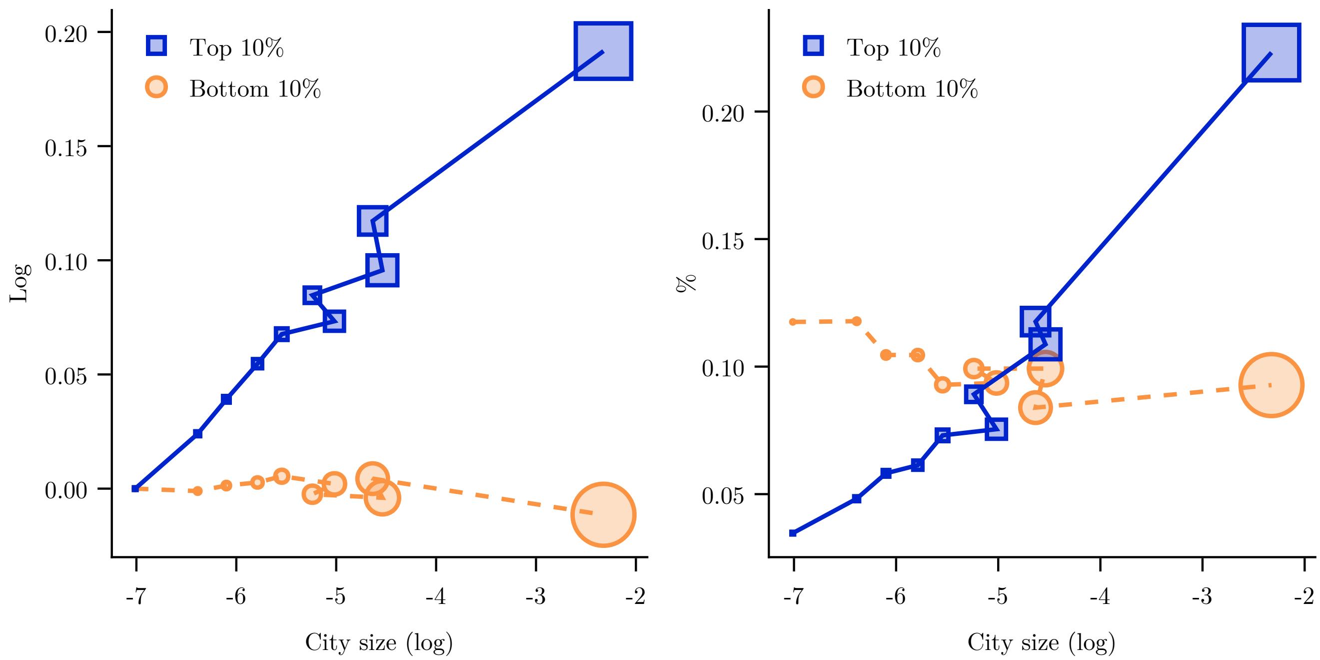


Both panels display the city size elasticity of wage deciles. Let $w_{q\ell}$ denotes the average wage in the qth decile of the wage distribution in location ℓ. Let m_ℓ denote location ℓ's size. The city elasticities are estimated by $\log w_{q\ell} = \alpha_q + \beta_q \log m_\ell + u_{q\ell}$. Panel (a) plots the $\hat{\beta}_q$ estimated on nominal wages. Panel (b) plots the $\hat{\beta}_q$ estimated on real wages. Real wages are computed as nominal wages deflated by a citywide Cobb-Douglas price index with a housing expenditure share of 0.2. The blue circles depict the estimates on the entire sample. The orange rectangles are estimated on workers living in different commuting zones than their birthplace. The gey triangles are estimated on wages deflated by a municipality housing price index. The municipality housing price index is computed based on the exposure of workers in a particular decile across municipalities. Specifically, $p^m_{\ell q} = \sum_{m \in \ell} \omega^q_{\ell m} p_m$, where $\omega^q_{\ell m}$ is the fraction of workers in decile q and CZ ℓ that lives in municipality m, and p_m is the average housing price in that municipality. The municipality price index is then $(p^m_{\ell q})^\alpha$. The vertical bars are 95th confidence intervals.

Figure A.2: Local distribution of employer fixed effects

(a) Average premia in bottom and top 10%

(b) Fraction of low- and high-paying jobs

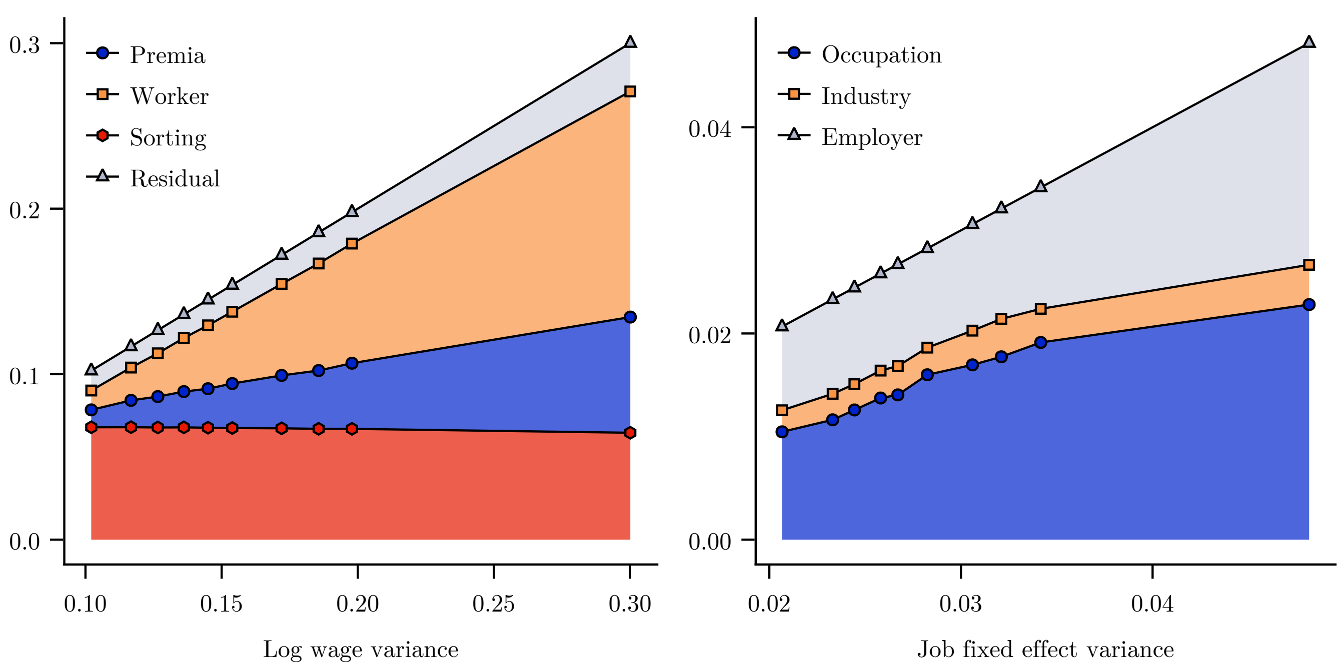


Panel (a) display the estimates of projections of city employer fixed effects deciles onto local wage inequality. Let ψ_ℓ^d denotes the dth employer fixed effects decile in location ℓ and V_ℓ location ℓ's variance of log wages. Each circle represents $\hat{\beta}^d$ estimated via $\psi_\ell^d = \alpha^d + \beta^d V_\ell + u_\ell^d$. The vertical bars are 95th confidence intervals. Panel (b) plots the city share of low- (orange circles) and high-paying (blue rectangles) jobs. Low and high-paying job defined as the 1st and 9th deciles of the national employer fixed effect distribution.

Figure A.3: Within-city wage variance decomposition — robustness

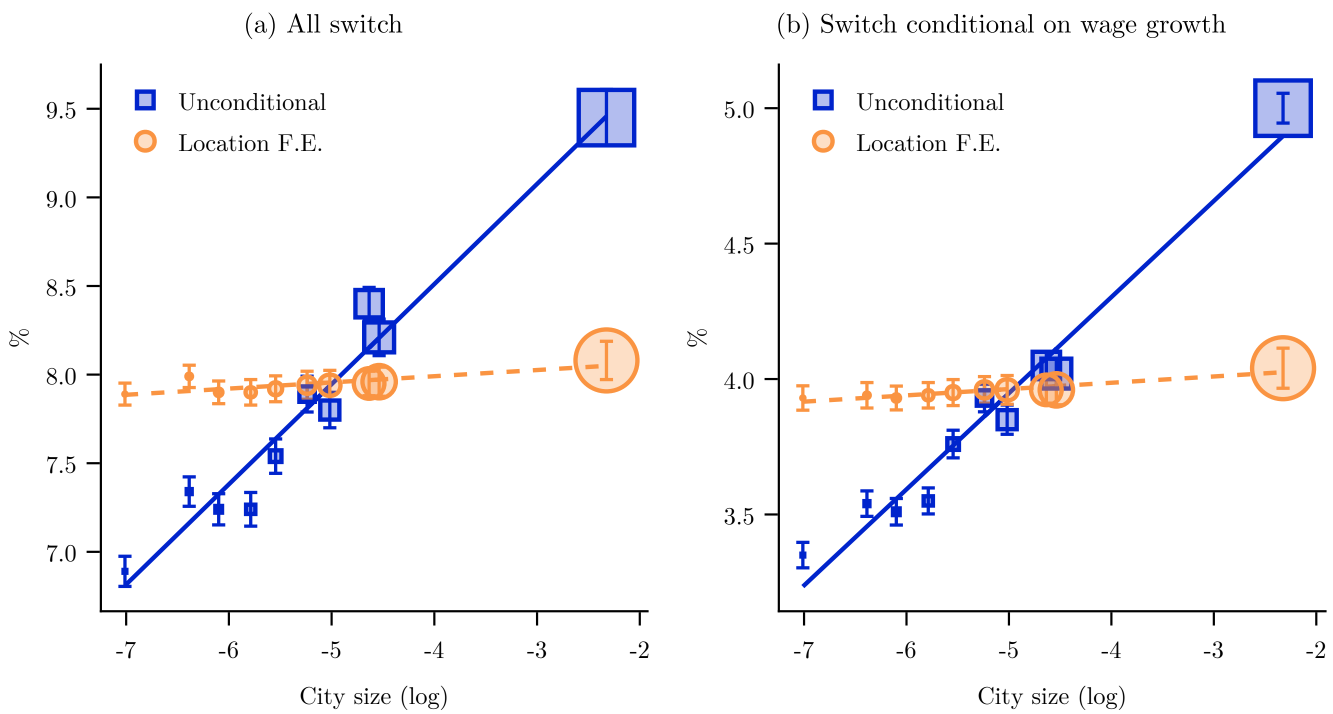
(a) Accounting for sorting

(b) Industry, occupation, and firm



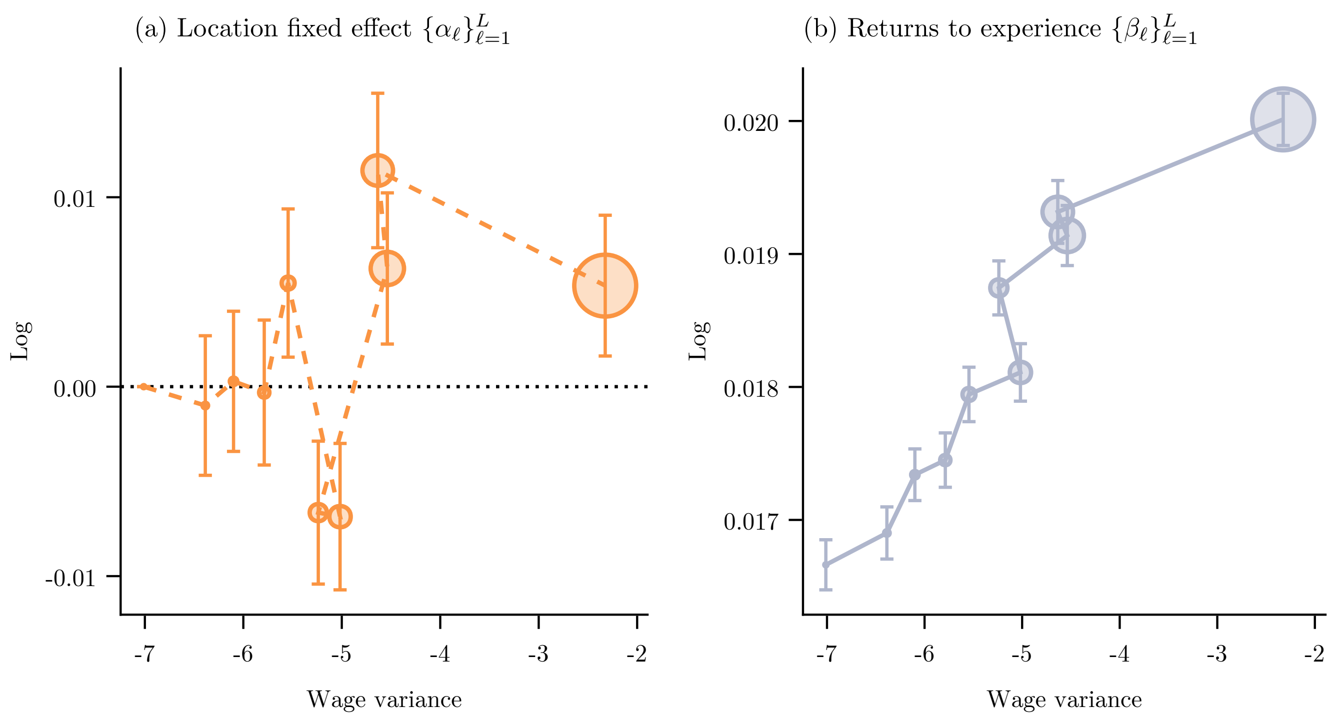
Panel (a) displays the variance decomposition (24). Each marker is a city group. The orange area represents the worker component (first line). The blue area represents the employer contribution (second line). The red area represents the sorting component (third line, first term). The grey area represents the residuals (third line, last two terms). Panel (b) decomposes the variance of the employer fixed effects into an occupation, industry, and establishment component by estimating $\gamma_j = \psi_{o(j)} + \zeta_{s(j)} + u_j$ for $o(j)$ and $s(j)$ the occupation and sector of job j. The blue area represents $\mathbb{V}\text{ar}[\psi_o] + \mathbb{C}\text{ov}[\psi_{o(j)}, \zeta_{s(j)}]$. The orange area represents $\mathbb{V}\text{ar}[\zeta_{s(j)}] + \mathbb{C}\text{ov}[\psi_{o(j)}, \zeta_{s(j)}]$ The grey area represents $\mathbb{V}\text{ar}[u_j]$.

Figure A.4: Local job switching rates



Left panel plots the average job switching rate by location. Job switch are defined as transition between pairs of establishment × occupation within two consecutive quarters. The blue rectangles are the unconditional job switch and the orange circles the average job switch net of worker heterogeneity as captured by the location fixed effects in (??). Right panel restricts the definition of a job switch to a transition between jobs associated with a wage gain.

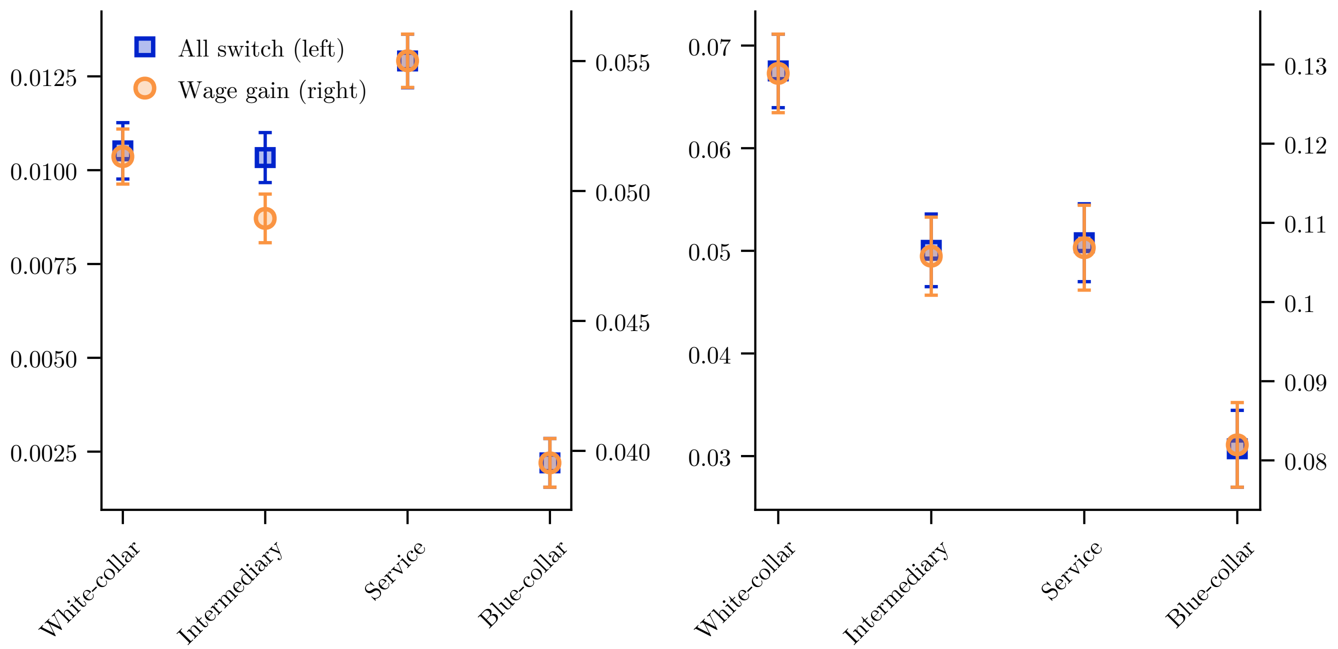
Figure A.5: Determinants of lifetime wages – other parameters



Left panel plots the city fixed effect $\{\alpha_\ell\}_{\ell=1}^L$ in (5). Right panel plots the returns to experience parameters $\{\beta_\ell\}_{\ell=1}^L$ in (5). In both panels, the parameters are plotted against the cities' wage variance. The vertical bars represent 95th confidence intervals.

Figure A.6: Local returns to job switching – heterogeneity by occupation

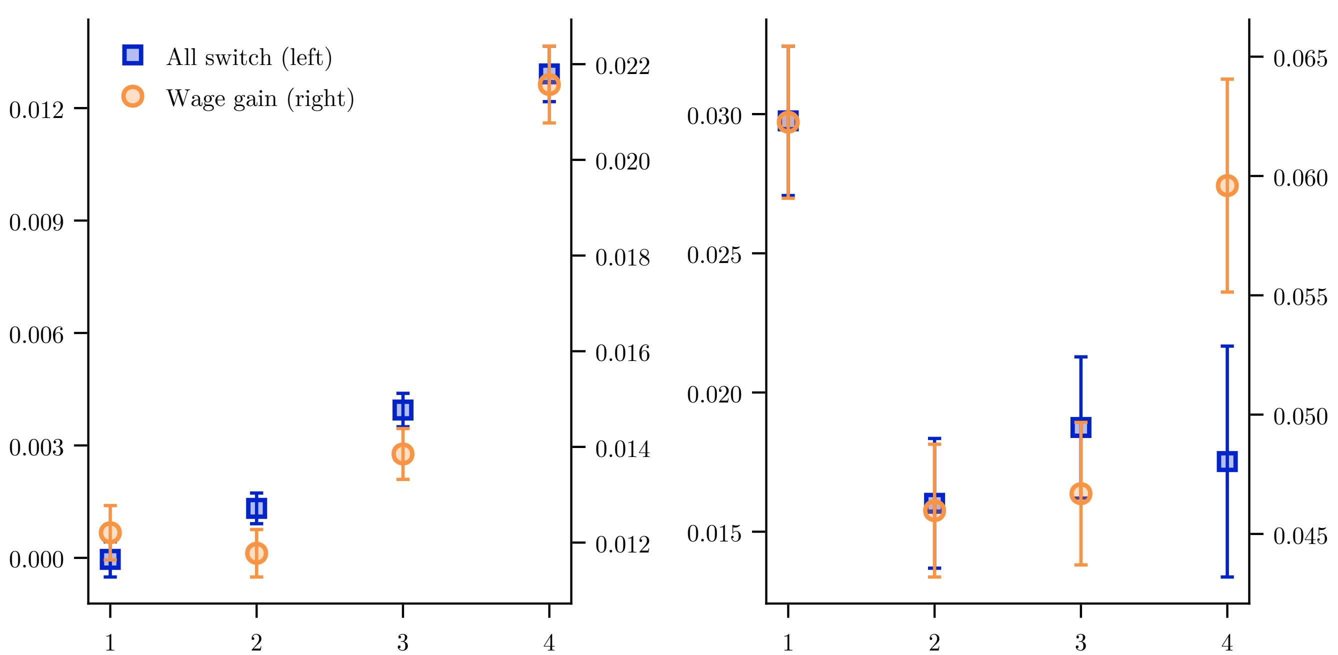
(a) Base return (b) Spatial heterogeneity



Both panels display point estimates of the returns to job switching by occupation. Specifically, I estimate $\log w_{it} = \text{FE}_t + \text{FE}_i + \text{FE}_{o(i,t)} + \alpha_{\ell(i,t)} + \beta^b x_{it} + \beta^s x_{it}\text{ineq}_{\ell(i,t)} + \gamma^b_{o(i,t)} z_{it} + \gamma^s_{o(i,t)} z_{it}\text{ineq}_{\ell(i,t)}$, where most variables are defined in (5), ineq_ℓ is ℓ wage variance, and $o(i,t)$ is the occupation of i at t. The left and right-panel display the point estimates of γ^b_o and γ^s_o. Standard errors are clustered at the city level.

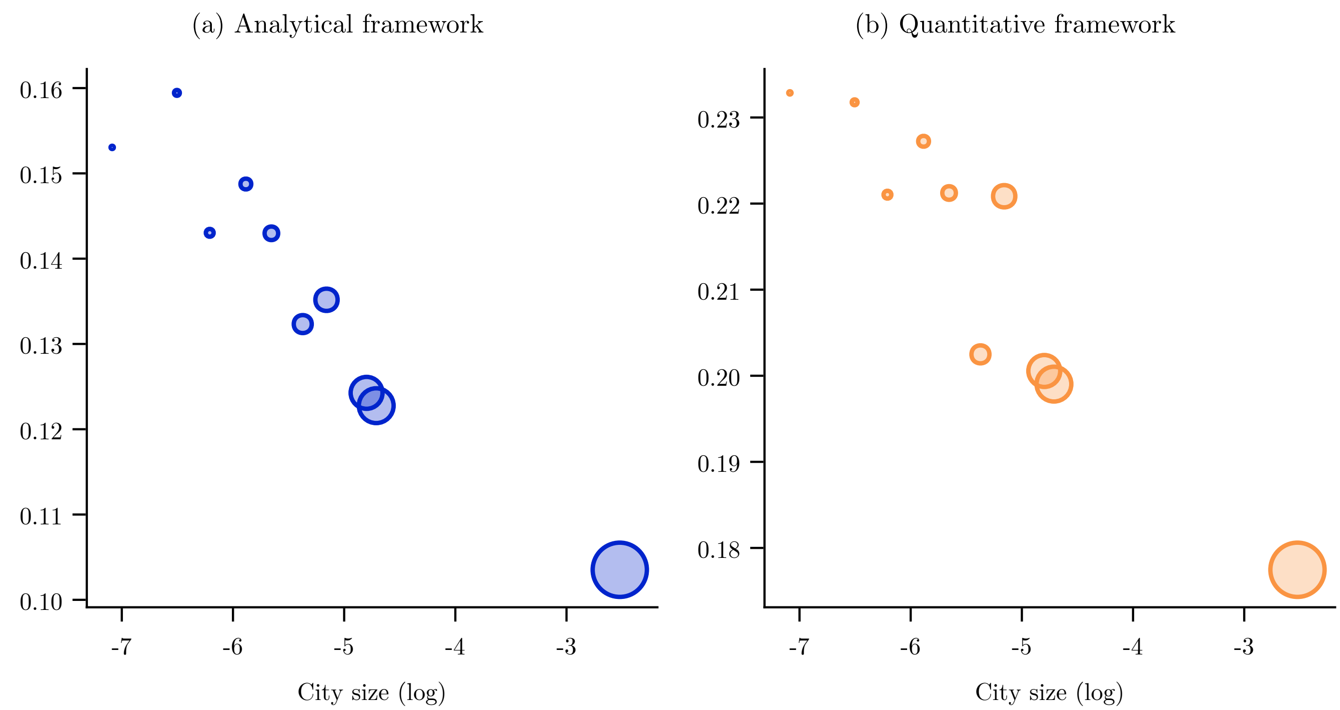
Figure A.7: Local returns to job switching – heterogeneity by wage quartile

(a) Base return (b) Spatial heterogeneity



Both panels display point estimates of the returns to job switching by wage quartile. Specifically, I estimate $\log w_{it} = \text{FE}_t + \text{FE}_i + \alpha_{\ell(i,t),q(i,t)} + \beta^b x_{it} + \beta^s x_{it}\text{ineq}_{\ell(i,t)} + \gamma^b_{q(i,t)} z_{it} + \gamma^s_{q(i,t)} z_{it}\text{ineq}_{\ell(i,t)}$, where most variables are defined in (5), ineq_ℓ is ℓ wage variance, and $q(i,t)$ is the within-city wage quartile of i at t. The left and right-panel display the point estimates of γ^b_q and γ^s_q. Standard errors are clustered at the city level.

Figure A.8: Labor market tightness by city



B A spatial theory of wage premia

B.1 Local supply curves

In location ℓ, the relative measure of unemployed and employed workers are such that the flow out of unemployment equates the flow into unemployment: $u_\ell/e_\ell = \delta/\lambda^u$. Consistency requires $u_\ell + e_\ell = m_\ell$, and therefore $u_\ell = \delta/(\lambda^u + \delta)m_\ell$ and $e_\ell = \lambda^u/(\lambda^u + \delta)m_\ell$.

The employment distribution, G_ℓ, is characterized by the within-city worker flows across jobs. In steady state, the flow of workers into the interval $[\underline{w}_\ell, w)$ has to be equal to the flow of workers out of the same interval, or $\lambda^u F_\ell(w)u_\ell = [\delta_\ell + \lambda^e \bar{F}_\ell(w)]e_\ell G_\ell(w)$. Solving for $G_\ell(w)$ and using $u_\ell/e_\ell = \delta/\lambda^u$ yields (11).

The labor supply curve is defined as the number of employed worker at wage w per wage offer, or $\lim_{\varepsilon\to 0^-}[G_\ell(w) - G_\ell(w-\varepsilon)]/[F_\ell(w) - F_\ell(w-\varepsilon)]e_\ell/v_\ell$. Taking the limit returns

$$n_\ell(w) = \frac{e_\ell}{v_\ell}\frac{1+k}{\left[1+k\bar{F}_\ell(w)\right]\left[1+k\bar{F}_\ell(w^-)\right]},$$

where $F_\ell(w^-) = \lim_{\varepsilon\to 0^-} F(w-\varepsilon)$. This is equivalent to equation (12) under a continuous wage offer distribution.

B.2 Reservation wage

Suppose for simplicity that F_ℓ admits a density. Differentiate the HJB of employed workers (10) with respect to w, integrate it back to w, and use $V_\ell(\underline{w}_\ell) = U$ to obtain

$$V_\ell(w) = U + \frac{A_\ell}{P_\ell} \int_{\underline{w}_\ell}^{w} \frac{1}{\beta + \delta + \lambda^e \bar{F}_\ell(w')} dw'.$$

Combine this expression into (8) and (10) to get

$$\beta U = \frac{A_\ell}{P_\ell} \left(b + \lambda^u \int_{\underline{w}_\ell} \frac{\bar{F}_\ell(w)}{\beta + \delta + \lambda^e \bar{F}_\ell(w)} dw \right), \tag{25}$$

$$\beta V_\ell(\underline{w}_\ell) = \frac{A_\ell}{P_\ell} \left(\underline{w}_\ell + \lambda^e \int_{\underline{w}_\ell} \frac{\bar{F}_\ell(w)}{\beta + \delta + \lambda^e \bar{F}_\ell(w)} dw \right) = \beta U. \tag{26}$$

Equating both equations and solving for $\underline{w}_\ell$ yield (13).

B.3 Proof of Lemma 1

Nominal Let $H_F(x)$ be the operator defined by

$$H_F(x) = \int_x \frac{\bar{F}(w)}{\beta + \delta + \lambda^e \bar{F}(w)} dw,$$

and $J_F(x)$ be defined by $J_F(x) = x - b - (\lambda^u - \lambda^e) H_F(x)$. The reservation wage solves $J_{F_\ell}(\underline{w}_\ell) = 0$. H satisfies two properties. First, $x \to H_F(x)$ is decreasing. Second, $F_1 \succ_{\text{FOSD}} F_2$ implies $H_{F_1}(x) > H_{F_2}(x)$ for any x. These properties carry through to J_F: $x \to J_F(x)$ is increasing and $F \to J_F(x)$ is weakly decreasing, strictly decreasing if $\lambda^u > \lambda^e$. Therefore, if $F_1 \succ_{\text{FOSD}} F_2$, $H_{F_1}(\underline{w}_2) \leq H_{F_2}(\underline{w}_2) = 0$, and it must be that $\underline{w}_1 \geq \underline{w}_2$, strictly if $\lambda^u > \lambda^e$.

Real From (26), the lifetime utility of a worker employed at the reservation wage is

$$\beta V_\ell(\underline{w}_\ell) = \frac{\underline{w}_\ell A_\ell}{P_\ell} \left(1 + \frac{\lambda^e H_F(\underline{w}_\ell)}{b + (\lambda^u - \lambda^e) H_F(\underline{w}_\ell)} \right),$$

where I have also used (13). We know that H_F is (strictly) increasing in F in the FOSD sense, and therefore so must $\lambda^e H_F / (b + [\lambda^u - \lambda^e] H_F)$. However, indifference (9) requires $V_\ell(\underline{w}_\ell) = \bar{U} = V_{\ell'}(\underline{w}_{\ell'})$. Therefore, $\underline{w}_\ell A_\ell / P_\ell$ is decreasing in F_ℓ in the FOSD sense.

B.4 Local wage offer distributions

I provide here the intuition for why there cannot be neither holes nor mass points in the wage offer distribution regardless of the shape of Γ_ℓ. For the full derivation, see Burdett and Mortensen (1998). To start, suppose there is a hole in F_ℓ between $\underline{w}$ and $\bar{w}$. We have $F_\ell(\underline{w}) \leq F_\ell(\bar{w})$, where the inequality is strict if there is a mass point at either $\underline{w}$ or $\bar{w}$. Therefore $n_\ell(\underline{w}) \leq n_\ell(\bar{w})$. However, by offering any wage in $(\underline{w}, \bar{w})$, an employer that used to post wage $\bar{w}$ would keep the same size while lowering its wage bill. This constitutes a profitable deviation, and therefore there cannot be holes in F_ℓ. Suppose now that there is a mass point at w. Take an employer with productivity z that offers wage w, and consider the deviation $w+\varepsilon$ for $\varepsilon > 0$ but small. For $\varepsilon \to 0^+$, we have $n(w+\varepsilon) > n(w)$ since there is a mass point at w, but $w+\varepsilon \to w$. Hence, the profit under wage offer w and $w+\varepsilon$ are respectively $(z-w)n(w) < (z-w-\varepsilon)n(w+\varepsilon)$, and offering wage $w+\varepsilon$ is a profitable deviation. This rules out any mass point in F_ℓ. Since neither arguments referred to Γ_ℓ, they hold for any Γ_ℓ.

I now show that $F_\ell[w_\ell(z)] = \Gamma_\ell(z)$. Since n_ℓ is strictly increasing in w, π_ℓ is continuously differentiable and strictly supermodular in (z, w). Directly applying Theorem 2.8.5. in Topkis (1998), it follows that w is strictly increasing in z. Given the ordering of wages, it must be that $F[w_\ell(z)] = \Gamma_\ell(z)$.

Finally, I derive the wage equation (16). Since there are no mass point in the wage offer distribution, we can take the first-order conditions of (17) with respect to w for almost every job productivity z,

$$\left(\frac{\partial R[z, n_\ell(w)]}{\partial n} - w\right)\frac{n'_\ell(w)}{n_\ell(w)} = 1.$$

Evaluating this equation at $w_\ell(z)$ and using the change of variable $n_\ell(z) = n_\ell[w_\ell(z)]$ yields

$$w'_\ell(z) = \frac{n'_\ell(z)}{n_\ell(z)}\left(z - w_\ell(z)\right), \tag{27}$$

after using $R(z, n) = z \cdot n$. Integrating this ODE with respect to w returns (16).

B.5 Proof of Proposition 2

Define $\eta_\ell(z) \equiv \theta_\ell n_\ell(z)$ as the tightness-independent size. Note that $\int \eta_\ell(z)\mathrm{d}\Gamma_\ell(z) = \Gamma_\ell(z)/(1 + k\bar{\Gamma}_\ell(z))$ is a proper cdf. Define $\eta^Q(q) \equiv \eta_\ell[\Gamma_\ell^{-1}(q)]$ as the size evaluated on rung q of the job ladder. Importantly, η^Q is identical across cities, which follows from search frictions being identical across space. Finally, let $\bar{w}_\ell$ denote the highest wage offered in city j. Throughout, I omit the city subscript whenever not necessary.

FOSD ordering (2.1) The wage paid by the employers at the q-th quantile of the wage distribution in city ℓ reads

$$\begin{aligned} w_\ell[\Gamma_\ell^{-1}(q)] &= \underline{w}_\ell\left(\frac{\eta^Q(0)}{\eta^Q(q)}\right) + \int_{\underline{z}_\ell}^{\Gamma_\ell^{-1}(q)} \zeta\left(\frac{\eta'_\ell(\zeta)}{\eta^Q(q)}\right)\mathrm{d}\zeta \\ &= \underline{w}_\ell\left(\frac{\eta^Q(0)}{\eta^Q(q)}\right) + \int_0^q \Gamma_\ell^{-1}(x)\left(\frac{\partial_x \eta^Q(x)}{\eta^Q(q)}\right)\mathrm{d}x, \end{aligned} \tag{28}$$

where the second equality follows from the change of variable $\Gamma_\ell(\zeta) \to x$. We guess and verify that $\Gamma_\ell \succ \Gamma_{\ell'}$ implies $F_\ell \succ F_{\ell'}$. So guess that the second ordering holds. From Proposition 1, we know that $\underline{w}_\ell \geq \underline{w}_{\ell'}$. Hence, for all $w \in [\underline{w}_{\ell'}, \underline{w}_\ell]$, we have $F_{\ell'}(w) \geq F_\ell(w) = 0$. From (28), $w_{\ell'}[\Gamma_\ell^{-1}(q)] < w_\ell[\Gamma_\ell^{-1}(q)]$ for all $q \in (0, 1]$. In particular, $\bar{w}_{\ell'} < \bar{w}_\ell$, so that $F_{\ell'}(w) = 1 > F_\ell(w)$ for all $w \in [\bar{w}_{\ell'}, \bar{w}_\ell)$. Finally, for any $w \in (\underline{w}_\ell, \bar{w}_{\ell'})$, (28) implies that $w_\ell[\Gamma_\ell^{-1}(q_\ell)] = w = w_{\ell'}[\Gamma_{\ell'}^{-1}(q_{\ell'})]$ if and only if $q_\ell > q_{\ell'}$. Hence, $F_{\ell'}(w) = F_{\ell'}\{w_{\ell'}[\Gamma_{\ell'}^{-1}(q_l)]\} = \Gamma_{\ell'}[\Gamma_{\ell'}^{-1}(q_{\ell'})] = q_{\ell'} > q_\ell = \Gamma_\ell[\Gamma_\ell^{-1}(q_\ell)] = F_\ell\{w_\ell[\Gamma_\ell^{-1}(q_\ell)]\} = F_\ell(w)$, where the second equality follows from the rank-preserving property of the job ladder. Concluding, we have

$F_{\ell'}(w) \geq F_\ell(w)$ for all w and $F_{\ell'}(w) > F_\ell(w)$ on the joint support of F_ℓ and $F_{\ell'}$, so that $F_\ell \succ F_{\ell'}$. From equation (11), it automatically follows that $G_\ell \succ G_{\ell'}$.

Inequality (2.2) Let $\Delta_{\ell\ell'} \equiv \bar{w}_{\ell'} - \underline{w}'_\ell - (\bar{w}_{\ell'} - \underline{w}_{\ell'})$ be the difference in top-to-bottom wage gap between city ℓ and ℓ'. Using (16), this difference writes

$$\Delta_{\ell\ell'} = (\underline{w}_\ell - \underline{w}_{\ell'})\left(1 - \frac{1}{(1+k)^2}\right) + \int_0^1 \left(\Gamma_\ell^{-1}(q) - \Gamma_{\ell'}^{-1}(q)\right)\left(\frac{\partial_x \eta^Q(q)}{\eta^Q(1)}\right) \mathrm{d}q.$$

If $\Gamma_\ell \succ \Gamma_{\ell'}$, this implies $\underline{w}_\ell \geq \underline{w}_{\ell'}$ and $\Gamma_\ell^{-1}(q) > \Gamma_{\ell'}^{-1}(q)$ for all $q \in (0,1)$, from which it follows that $\Delta_{\ell\ell'} > 0$. Note that this between-city difference in inequality is in level. Alternatively, the top-to-bottom wage gap in ratio in city ℓ is

$$\frac{\bar{w}_\ell}{\underline{w}_\ell} = \frac{1}{1+k^2} + \frac{1}{\underline{w}_\ell}\int_0^1 \Gamma_\ell^{-1}(q)\left(\frac{\partial_x \eta^Q(q)}{\eta^Q(1)}\right) \mathrm{d}x.$$

When $\lambda^e \to \lambda^u$, we have $\underline{w}_\ell \to b$, so that $\Gamma_\ell \succ \Gamma_{\ell'}$ also implies $\bar{w}_\ell/\underline{w}_\ell > \bar{w}_{\ell'}/\underline{w}_{\ell'}$ for $\lambda^e \approx \lambda^u$.

E2E gains (2.3) The ℓ are dropped as long as comparisons across cities are absent. The expected wage growth upon a job-to-job transition in a given city is

$$\Delta^{\text{EE}}(w) \equiv \mathbb{E}\left[\frac{w'}{w} \mid \text{switch from } w\right] = \frac{1}{\bar{F}(w)}\int_w \frac{w'}{w}\mathrm{d}F(w').$$

Consider this statistic for a worker employed on some rung q of the ladder,

$$\Delta^{\text{EE}}[F^{-1}(q)] = \frac{1}{1-q}\int_{F^{-1}(q)} \frac{w'}{F^{-1}(q)}\mathrm{d}F(w'). \tag{29}$$

For $q \to 0$, we know that $F^{-1}(q) \to \underline{w} \approx b$, where the second approximation comes from $\lambda^e \approx \lambda^u$. Hence, for $q \to 0$, we have

$$\Delta^{\text{EE}}[F^{-1}(q)] \to \int_b \frac{w'}{b}\mathrm{d}F(w')$$

If $\Gamma_\ell \succ G_{\ell'}$, then $F_\ell \succ F_{\ell'}$, and therefore $\Delta_\ell^{\text{EE}}(\underline{w}_\ell) < \Delta_{\ell'}^{\text{EE}}(\underline{w}_{\ell'})$. Since the inequality is strict and (29) is continuous in q, this holds for workers on low rungs of the ladder.

B.6 A competitive spatial matching model

Consider the simplest competitive spatial matching model. There are L cities that differ in size, m_ℓ, with $m_1 < m_2 < \cdots < m_L$. Cities are purposely homogeneous in TFP to study the implication for spatial wage inequality. There are heterogeneous firms indexed by their productivity z distributed according to Γ. Firms have a decreasing returns to scale technology: $R(z,n) = zn^\rho$.[23] Labor markets are competitive and segmented by locations. Employers pay housing cost r_ℓ to produce in location ℓ. The housing supply is $L_\ell = \bar{L}r_\ell^\chi$.

[23] Firm size is not well-defined absent DRS.

Employers solve

$$\pi(z) = \max_{\ell,n} \; zn^\rho - w_\ell n - r_\ell.$$

Their labor demand and optimal profits are

$$n_\ell(z) = \left(\frac{z\rho}{w}\right)^{\frac{1}{1-\rho}} \quad \text{and} \quad \pi(z) = \max_\ell \; (1-\rho)\rho^{\frac{\rho}{1-\rho}} \frac{z^{\frac{1}{1-\rho}}}{w^{\frac{\rho}{1-\rho}}} - r_\ell.$$

The spatial allocation of firms is described by $\{M_\ell, \Gamma_\ell\}_{\ell=1}^L$ for M_ℓ the measure of employers and Γ_ℓ the local productivity distribution.

The next proposition shows that wages are necessarily lower in larger cities.

Proposition B.1 (Wage distribution)**.**
In any equilibrium, wages are decreasing in city size: $m_\ell > m_{\ell'} \Rightarrow w_\ell < w_{\ell'}$*.*

Proof. Using the equilibrium condition for wages, the firm problem becomes

$$\max_\ell \; (1-\rho)\,\psi_\ell z^{\frac{1}{1-\rho}} - r_\ell,$$

where ψ_ℓ fully summarizes the profitability of ℓ:

$$\psi_\ell \equiv \left(\frac{m_\ell}{M_\ell \mathbb{E}_\ell\left[z^{\frac{1}{1-\rho}}\right]}\right)^\rho = \left(\frac{\rho}{w_\ell}\right)^{\frac{\rho}{1-\rho}}.$$

The constant ψ_ℓ is an equilibrium object that depends on the allocation of employers across space. However, given a distribution of $\{\psi_\ell\}_{\ell=1}^L$, the complementarity between (ψ_ℓ, z) implies that there is pure positive assortative matching between $\boldsymbol{\psi}$ and $\boldsymbol{z}$ — or pure negative assortative matching between $\boldsymbol{w}$ and $\boldsymbol{z}$ (Topkis, 1998).

We can now prove the proposition. From the expression for wages, wages are increasing in city size if

$$\frac{M_{\ell+1}}{M_\ell} \frac{\mathbb{E}_{\ell+1}\left[z^{\frac{1}{1-\rho}}\right]}{\mathbb{E}_\ell\left[z^{\frac{1}{1-\rho}}\right]} > \frac{m_{\ell+1}}{m_\ell} > 1.$$

That is, $M_\ell \mathbb{E}_\ell[z]$ must be increasing in m_ℓ. We also know that firms sort in the opposite direction of w_ℓ. Hence, $\mathbb{E}_\ell[z^{\frac{1}{1-\rho}}]$ is decreasing in m_ℓ. A necessary condition is therefore that M_ℓ is increasing in m_ℓ.

In any equilibrium with pure sorting, marginal firms must be indifferent between two locations. If w_ℓ is increasing in m_ℓ, this indifference is $\pi_\ell(\underline{z}_\ell) = \pi_{\ell+1}(\bar{z}_{\ell+1})$, or using the expression for profits and housing prices:

$$(1-\rho)\left[\psi_\ell - \psi_{\ell+1}\right] \underline{z}_\ell^{\frac{1}{1-\rho}} = \left(\frac{M_\ell}{\bar{L}}\right)^{1/\chi} - \left(\frac{M_{\ell+1}}{\bar{L}}\right)^{1/\chi} > 0.$$

Therefore, M_ℓ is decreasing in ℓ, and wages cannot be increasing in m_ℓ. □

B.7 Proof of Proposition 3

This section proves Proposition 3, and by doing so, uniquely characterizes the job distributions $\{\Gamma_\ell\}_{\ell=1}^L$ as a function of the vector of market tightness $\{\theta_\ell\}_{\ell=1}^L$. To do, I proceed in five parts. First, I show that the support of the distribution is convex in each city (Lemma B.2). I then prove that local job distributions are necessarily ranked in terms of FOSD (Lemma B.3). Lemma B.4 continues by deriving the pdf of the job allocation in each city, while Lemma B.5 proves the condition for the overlapping support in the job distribution. Finally, Algorithm 1 concludes by uniquely solving for the cutoffs of the job productivity in each city.

Lemma B.1.
The profit function $\pi_\ell(z)$ is continuously differentiable.

Proof. From (**??**), $\pi_\ell'(z)$ is differentiable. Differentiating with respect to z returns $\pi_\ell'(z) = n_\ell(z)$. Since F_ℓ has no mass point, $n_\ell(z)$ is continuous in z. □

Lemma B.2 (Convex support)**.**
The support of the job distribution in each city is an interval.

Proof. Suppose not. That is, for some city ℓ, there exists at least one hole in the support of Γ_ℓ. Wlog, suppose there is a unique hole, such that $\operatorname{supp}\Gamma_\ell = [\underline{z}_\ell, z] \cup [z+\varepsilon, \bar{z}_\ell]$ for some $\varepsilon > 0$. The profit maximization condition (17) then requires

1. $\pi_\ell(z') \geq \pi_{\ell'}(z')$ for all $z' \in [\underline{z}_\ell, z]$ and all cities ℓ',
2. $\pi_\ell(z') < \pi_{\ell^\star}(z')$ for all $z' \in (z, z+\varepsilon)$ and at least one city $\ell^\star$,
3. $\pi_\ell(z') \geq \pi_{\ell'}(z')$ for all $z' \in [z+\varepsilon, \bar{z}_\ell]$ and all cities ℓ'.

The first two conditions occur only if $\pi_\ell'(z') < \pi_{\ell^\star}'(z')$ for $z' \in B^+(z)$.[24] Using the Envelope condition, this rewrites $n_\ell(z') = n_\ell(z) = \pi_\ell'(z) < \pi_{\ell^\star}'(z') = n_{\ell^\star}(z')$, where the first equality follows from n_ℓ being constant outside of the support of Γ_ℓ. It follows that $\pi_\ell'(z') < \pi_{\ell^\star}'(z')$ for all $z' \in (z, z+\varepsilon)$, and therefore $\pi_\ell(z+\varepsilon) < \pi_{\ell^\star}(z+\varepsilon)$, a contradiction with the third condition. □

Lemma B.3 (First order stochastic dominance ordering)**.**
For two cities ℓ and ℓ' and a vector of tightness that is equilibrium compatible, $\theta_\ell < \theta_{\ell'}$ if and only if $\Gamma_\ell \succ \Gamma_{\ell'}$.[25]

Proof. I start by showing that if $\theta_\ell < \theta_{\ell'}$, then $\Gamma_\ell \succ \Gamma_{\ell'}$. For the sake of contradiction, suppose that $\bar{\Gamma}_{\ell'}(z) \geq \bar{\Gamma}_\ell(z)$ for some z. Then,

$$\pi_\ell'(z) = \frac{1}{\theta_\ell}\frac{1+k}{(1+k\bar{\Gamma}_\ell(z))^2} \geq \frac{1}{\theta_\ell}\frac{1+k}{(1+k\bar{\Gamma}_{\ell'}(z))^2} > \frac{1}{\theta_{\ell'}}\frac{1+k}{(1+k\bar{\Gamma}_{\ell'}(z))^2} = \pi_{\ell'}'(z). \tag{30}$$

If $\Gamma_\ell(z), \Gamma_{\ell'}(z) \in (0,1)$, Lemma B.2 implies that $\pi_\ell(z) = \pi_{\ell'}(z)$. However, from equation (30), $\pi_\ell'(z') > \pi_{\ell'}'(z')$ for $z' \in B^+(z)$, and therefore $\pi_\ell(z') > \pi_{\ell'}(z')$. Lemma B.2 then implies $z = \bar{z}_{\ell'}$ and $z < \bar{z}_\ell$, or $\bar{\Gamma}_{\ell'}(z) = 0 < \bar{\Gamma}_\ell(z)$, a contradiction.

[24] Throughout the appendix, $B(x)$ refers to an open ball around x, $B^+(x)$ an open ball to the right of x, and similarly $B^-(x)$ to an open ball to the left of x.

[25] An equilibrium compatible vector of market tightness is such that $\theta_\ell < \infty$ for all ℓ.

If $\Gamma_{\ell'}(z) = 0$ and $\Gamma_\ell(z) \in (0,1)$, then it must from Lemma B.2 that $\pi_{\ell'}(z) \leq \pi_\ell(z)$. Furthermore, the positive measure of jobs locating in city ℓ' requires the existence of $\underline{z}_{\ell'} \geq z$ such that $\pi_{\ell'}(\underline{z}_{\ell'}) \geq \pi_\ell(\underline{z}_{\ell'})$. If $\underline{z}_{\ell'} = z$, then it must be that $\pi_{\ell'}(\underline{z}_{\ell'}) = \pi_\ell(\underline{z}_\ell)$. However, (30) implies that $\pi'_{\ell'}(z') < \pi'_\ell(z')$ for all $z' \in B^+(\underline{z}_{\ell'})$, and therefore $\pi_{\ell'}(z') < \pi_\ell(z')$ for $z' \in B^+(\underline{z}_{\ell'})$, which contradicts Lemma B.2. Hence, it must be that $\underline{z}_{\ell'} > z$. But (30) implies $\pi'_{\ell'}(z') < \pi'_\ell(z')$ for all $z' \in (z, \underline{z}_{\ell'})$, and therefore $\pi_{\ell'}(\underline{z}_{\ell'}) < \pi_\ell(\underline{z}_{\ell'})$, a contradiction with the definition of $\underline{z}_{\ell'}$. It follows that $\underline{z}_{\ell'}$ does not exist, a contradiction.

Finally, suppose $\Gamma_\ell(z) = 1$, so that $z \geq \bar{z}_\ell$. If $\Gamma_{\ell'}(z) \in (0,1)$, then $\bar{z}_{\ell'} > \bar{z}_\ell$. Meanwhile, it must also be that $\pi_{\ell'}(\bar{z}_\ell) \leq \pi_\ell(\bar{z}_\ell)$. Finally, for all $z' \in [\bar{z}_\ell, \bar{z}_{\ell'})$, we have $\Gamma_{\ell'}(z') < \Gamma_\ell(z') = 1$, and therefore $\pi'_{\ell'}(z') < \pi'_\ell(z')$ from (30). However, this implies $\pi_{\ell'}(z') < \pi_\ell(z')$ for all $z' \in [\bar{z}_\ell, \bar{z}_{\ell'})$, which contradicts $\bar{z}_{\ell'} > \bar{z}_\ell$ from Lemma B.2. Hence, suppose that $\Gamma_{\ell'}(z) = 0$. However, in this case, a similar argument as in the previous paragraph shows that there does not exist a $\underline{z}_{\ell'}$, a contradiction. Hence, the only possibility is $\Gamma_{\ell'}(z) = 1$, a contradiction of $\Gamma_{\ell'} \succ \Gamma_\ell$.

The other direction automatically follows. Suppose not $\theta_\ell < \theta_{\ell'}$, such that $\theta_\ell \geq \theta_{\ell'}$. If $\theta_\ell > \theta_{\ell'}$, then we have already shown that $\Gamma_{\ell'} \succ \Gamma_\ell$, or not $\Gamma_\ell \succ \Gamma_{\ell'}$. If $\theta_{\ell'} = \theta_\ell$, it must be that $\Gamma_{\ell'}(z) = \Gamma_\ell(z)$ for all z. To see this, suppose for the sake of contradiction that $\bar{\Gamma}_{\ell'}(z) > \bar{\Gamma}_\ell(z)$ for some z. Note that, as in (30), this implies $\pi'_\ell(z) > \pi'_{\ell'}(z)$, and it is possible to re-use the same arguments as above to create a contradiction. By symmetry, $\bar{\Gamma}_\ell(z) > \bar{\Gamma}_{\ell'}(z)$ for some z also creates a contradiction. □

Lemma B.4 (Local job density)**.**
Define the functions $\mu^\ell : \mathbb{R}_+ \mapsto \mathbb{R}_+$ as the (employment-weighted) relative market tightness that employers z faces in city ℓ,

$$\mu^\ell(z) \equiv \frac{e_\ell \sqrt{\theta_\ell}}{\sum_{\ell' \in \mathcal{L}(z)} e_{\ell'} \sqrt{\theta_{\ell'}}} \quad \textit{for} \quad \mathcal{L}(z) \equiv \{\ell' \in \mathcal{J} : z \in [\underline{z}_{\ell'}, \bar{z}_{\ell'}]\}.$$

Then, if $z \in$ supp Γ_ℓ, the relative measure of z-jobs in ℓ is

$$\mathrm{d}\Gamma_\ell(z) = \left(\frac{M}{v_\ell}\right) \mu^\ell(z) \mathrm{d}\Gamma(z).$$

Before the proof, note that the function μ has three important properties. If all z-jobs locate in city ℓ, then $\mu^\ell(z) = 1$, so that $\mathrm{d}\Gamma_\ell(z) = M \mathrm{d}\Gamma(z)/v_\ell$. Second, $\mu^\ell(z)$ is decreasing in the number of overlapping cities, $\mathcal{L}(z)$. Finally, $\mu^\ell(z) \in [e_\ell\sqrt{\theta_\ell}/(\sum_{\ell' \in \{1,2,\ldots,L\}} e_{\ell'}\sqrt{\theta_{\ell'}}), 1]$, it is locally constant, and takes at most L values.

Proof. Take any $z \in$ supp Γ. If there exists a city ℓ so that $\pi_\ell(z) > \pi_{\ell'}(z)$ for all other ℓ', then all z employers locate in ℓ. The feasibility condition (7) then requires $v_\ell \mathrm{d}\Gamma_\ell(z) = M \mathrm{d}\Gamma(z)$. Otherwise, let $\mathcal{L}$ be the set of city for which $\pi_\ell(z) = \pi(z)$ for $\ell \in \mathcal{L}$. Unless $z \in \{\underline{z}_\ell, \bar{z}_\ell\}$ – a measure zero event, then this condition must also hold for a neighborhood around z. Hence, $n_\ell(z) = \pi'_\ell(z) = \pi'_{\ell'}(z) = n_{\ell'}(z)$ for any $(\ell, \ell') \in \mathcal{L}$. Differentiating $n_\ell(z) = n_{\ell'}(z)$ with respect to z implies

$$\frac{\mathrm{d}\Gamma_{\ell'}(z)}{\mathrm{d}\Gamma_\ell(z)} = \sqrt{\frac{\theta_\ell}{\theta_{\ell'}}},$$

for any $(\ell, \ell') \in \mathcal{L}$. But feasibility requires $\sum_{\ell' \in \mathcal{L}} v_{\ell'} \mathrm{d}\Gamma_{\ell'}(z) = M \mathrm{d}\Gamma(z)$, and therefore

$$\mathrm{d}\Gamma_\ell(z) = \frac{1}{\sqrt{\theta_\ell}} \frac{M \mathrm{d}\Gamma(z)}{\sum_{\ell' \in \mathcal{L}} e_{\ell'} \sqrt{\theta_{\ell'}}}.$$

Re-arranging and using the two functions $z \to \mu^\ell(z)$ and $z \to \mathcal{L}(z)$ yield the desired equation. □

An important aspect of Lemma B.4 is that, given $\mathcal{L}(z)$ and $\boldsymbol{\theta}$, the weighting function μ^ℓ is known. Hence, Lemma B.4 pins down the job distributions up to the thresholds $\{\underline{z}_\ell, \bar{z}_\ell\}_{\ell=1}^L$. Under perfect sorting, solving for these thresholds would be trivial. However, in this economy of matching with externality, perfect sorting does not always prevail.

Lemma B.5 (Overlapping supports)**.**
Take two cities ℓ and ℓ' so that $\theta_\ell < \theta_{\ell'}$. Then, supp $\Gamma_\ell \cap$ supp $\Gamma_{\ell'}$ has positive measure if and only if $\theta_{\ell'} < (1+k)^2\theta_\ell$.

Proof. I first prove $\Leftarrow$. Suppose $\theta_{\ell'} < (1+k)^2\theta_\ell$. For the sake of contradiction, suppose also that there is no overlap. Since $\theta_\ell < \theta_{\ell'}$, Lemma B.3 implies $\bar{z}_{\ell'} \leq \underline{z}_\ell$. Furthermore, $\pi_{\ell'}(\underline{z}_\ell) \leq \pi_\ell(\underline{z}_\ell)$ and $\pi_{\ell'}(z) > \pi_\ell(z)$ for $z \in B^-(\bar{z}_{\ell'})$. Hence, π_ℓ and $\pi_{\ell'}$ must cross at least once in $[\bar{z}_{\ell'}, \underline{z}_\ell]$ and π_ℓ crosses $\pi_{\ell'}$ from below; that is, there exists a $z^\star \in [\bar{z}_{\ell'}, \underline{z}_\ell]$ so that $\pi_\ell(z^\star) = \pi_{\ell'}(z^\star)$ and $\pi'_\ell(z) > \pi'_{\ell'}(z)$ for $z \in B^-(z^\star)$. By continuity, we therefore have $\pi'_\ell(z^\star) \geq \pi'_{\ell'}(z^\star)$. Furthermore, it must be that $\Gamma_{\ell'}(z^\star) = 1$. Combining these elements and using the Envelope theorem, this inequality writes

$$\frac{1}{\theta_\ell}\frac{1}{1+k} = n_\ell(\underline{z}_\ell) = \pi'_\ell(z^\star) \geq \pi'_{\ell'}(z^\star) = n_{\ell'}(\bar{z}_{\ell'}) = \frac{1+k}{\theta_{\ell'}},$$

where the first and last equality follows from n_ℓ and $n_{\ell'}$ being constant outside of their respective support. However, $\theta_{\ell'} < (1+k)^2\theta_\ell$, a contradiction.

We now prove the other direction. For this, suppose not $\theta_{\ell'} < (1+k)^2\theta_\ell$, or $\theta_{\ell'} \geq (1+k)^2\theta_\ell$. We want to show that this implies no overlap. Hence, for the sake of contradiction, suppose that there is overlap in the job distribution. Lemma B.2 then implies supp $\Gamma_\ell \cap$ supp $\Gamma_{\ell'} = [\underline{z}_\ell, \bar{z}_{\ell'}]$ with $\underline{z}_\ell < \bar{z}_{\ell'}$. Employers in $[\underline{z}_\ell, \bar{z}_{\ell'}]$ must be indifferent between the two cities, $\pi_\ell(z) = \pi_{\ell'}(z)$ for all $z \in [\underline{z}_\ell, \bar{z}_{\ell'}]$, and therefore $\pi'_\ell(z) = \pi'_{\ell'}(z)$. By continuity, this must hold at $\underline{z}_\ell$, or using the Envelope theorem,

$$\frac{1}{\theta_\ell}\frac{1}{1+k} = n_\ell(\underline{z}_\ell) = \pi'_\ell(\underline{z}_\ell) = \pi'_{\ell'}(\underline{z}_\ell) = n_{\ell'}(\underline{z}_\ell) = \frac{1}{\theta_{\ell'}}\frac{1}{(1+k\bar{\Gamma}_{\ell'}(\underline{z}_\ell))^2}.$$

However, $\theta_{\ell'} \geq (1+k)^2\theta_\ell$, and this equality cannot hold for any $\Gamma_{\ell'}(\underline{z}_\ell) \in (0,1)$, a contradiction. □

For a given $\boldsymbol{\theta}$, Lemma B.4 and B.5, together with the boundary conditions boundary conditions $\underline{z}_1 = \underline{z}$, $\bar{z}_L = \bar{z}$, $\Gamma_\ell(\underline{z}_\ell) = 0$ and $\Gamma_\ell(\bar{z}_\ell) = 1$, pins down uniquely the set of cutoffs $\{\underline{z}_\ell, \bar{z}_\ell\}_{\ell=1}^L$. These cutoffs are derived in Algorithm 1.

B.8 Proof of Proposition 4

The proof of Proposition 4 is broken up in several lemmas. Lemma B.6 proves that, in equilibrium, cities must necessarily have different market tightness. Lemma B.7 then derives the expression for the local prices

Order cities by $1/\theta_\ell$. Normalize $\underline{z}_0 = \underline{z}$.

for $\ell \in \{1, 2, \ldots, L\}$ **do**

Either $\ell = 1$ and $\underline{z}_\ell = \underline{z}$, or $\ell > 1$ and the lower bound has been found in the previous iteration (step 2 or 3).

1. **if** $\theta_{\ell-1} < \theta_\ell(1+k)^2$ **then** From Lemma B.5, $\underline{z}_\ell \in (\underline{z}_{\ell-1}, \bar{z}_{\ell-1})$. Furthermore, $\pi'_\ell(z) = \pi'_{\ell-1}(z)$ for $z \in (\underline{z}_\ell, \bar{z}_{\ell-1})$ implies

$$\frac{1}{\theta_\ell}\left(\frac{1+k}{\left(1+k\bar{\Gamma}_\ell(\bar{z}_{\ell-1})\right)^2}\right) = \frac{1+k}{\theta_{j-1}},$$

which solves for $\Gamma_\ell(\bar{z}_{\ell-1})$.

else From Lemma B.5 $\underline{z}_\ell = \bar{z}_{\ell-1}$ and $\Gamma_\ell(\bar{z}_{\ell-1}) = 0$.

2. **foreach** $\ell' \in \{\ell+1, \ldots, L\}$ **do**

if $\theta_\ell < \theta_{\ell'}(1+k)^2$ *and* $\theta_{\ell-1} > \theta_{\ell'}(1+k)^2$ **then**

From Lemma B.5 $\bar{z}_{\ell-1} < \underline{z}_{\ell'} < \bar{z}_\ell$. For $z \in (\underline{z}_{\ell'}, \bar{z}_\ell)$, indifference requires $\pi'_\ell(z) = \pi'_{\ell'}(z)$, or

$$\frac{1}{\theta_\ell}\left(\frac{1+k}{1+k\bar{\Gamma}_\ell(\underline{z}_{\ell'})^2}\right) = \frac{1}{\theta_{\ell'}}\left(\frac{1}{1+k}\right). \tag{31}$$

Furthermore, since $\underline{z}_{\ell'} > \underline{z}_{\ell-1}$, jobs in $(\max\{\underline{z}_{\ell'-1}, \bar{z}_{\ell-1}\}, \underline{z}_{\ell'})$ are located in cities $\{\ell, \ldots, \ell'-1\}$. Lemma B.4 then implies $\mathrm{d}\Gamma_\ell(z) = \mu^\ell_{\ell,\ell'-1} M \mathrm{d}\Gamma(z)/v_\ell$, so that

$$\Gamma_\ell(\underline{z}_{\ell'}) = \Gamma_\ell\left(\max\{\bar{z}_{\ell-1}, \underline{z}_{\ell'-1}\}\right) + \left(\frac{M}{v_\ell}\right)\mu^\ell_{\ell,\ell'-1}\left(\Gamma(\underline{z}_{\ell'}) - \Gamma\left(\max\{\bar{z}_{\ell-1}, \underline{z}_{\ell'-1}\}\right)\right).$$

Combined with equation (31), these two equations returns $\underline{z}_{\ell'}$.

end

end

3. The cdf must integrate to one. Letting $\ell^\star \equiv \max\{\ell, \sup\{\ell' : \theta_{\ell'} > (1+k)^2\theta_\ell\}\}$, this requires

$$1 = \Gamma_\ell(\max\{\bar{z}_{\ell-1}, \underline{z}_{\ell^\star}\}) + \left(\frac{M}{v_\ell}\right)\mu^\ell_{\ell,\ell^\star}\left(\Gamma(\bar{z}_\ell) - \Gamma\left(\max\{\bar{z}_{\ell-1}, \underline{z}_{\ell^\star}\}\right)\right),$$

which solves for $\bar{z}_\ell$.

end

Algorithm 1: Boundaries solver

that sustain the spatial job allocation. Finally, Lemma B.8 then orders cities based on their fundamentals.

Lemma B.6 (No symmetric equilibria)**.**

In equilibrium, no two cities can have the same labor market tightness.

Proof. Take two cities, ℓ and ℓ' with $e_\ell > e_{\ell'}$ (wlog). Suppose for the sake of contradiction that $\theta_\ell = \theta_{\ell'}$. Then, it must be that $v_{\ell'}/v_\ell = e_{\ell'}/e_\ell$, or $v_{\ell'} < v_\ell$. It follows that $r_{\ell'} < r_\ell$. At the same time, Lemma B.3 implies that $\Gamma_{\ell'} = \Gamma_\ell$, and therefore $\underline{z}_{\ell'} = \underline{z}_\ell = \underline{z}$. Furthermore, from equation (13), $\underline{w}_{\ell'} = \underline{w}_\ell \equiv \underline{w}$. Finally, employers must be indifferent between both locations. In particular, $\pi_{\ell'}(\underline{z}_{\ell'}) = (\underline{z} - \underline{w})n(\underline{z}) - r_{\ell'} = (\underline{z} - \underline{w})n(\underline{z}) - r_\ell = \pi_\ell(\underline{z}_\ell)$ must hold. However, this requires $r_{\ell'} = r_\ell$, a contradiction. □

Lemma B.7 (Local prices)**.**

Fix $\boldsymbol{\theta}$ and re-arrange cities so that θ_ℓ is decreasing in ℓ. Then, the spatial job allocation is sustained by a vector of local prices that satisfies the difference equation

$$\kappa_{\ell+1} - \kappa_\ell = \underline{z}_{\ell+1} n_{\ell+1}(\underline{z}_{\ell+1}) - \underline{z}_\ell n_\ell(\underline{z}_\ell) - \int_{\underline{z}_\ell}^{\underline{z}_{\ell+1}} n_\ell(\zeta)\mathrm{d}\zeta, \tag{32}$$

for κ_1 given. Furthermore, for two cities ℓ and ℓ', $\theta_{\ell'} < \theta_\ell \iff \kappa_\ell > \kappa_{\ell'}$.

Proof. For any city ℓ, take the city ℓ' such that $\theta_{\ell'} > \theta_\ell$ and there is no other third city l so that $\theta_{\ell'} > \theta_l > \theta_\ell$. If $z \in \operatorname{supp} \Gamma_\ell$ and $z \in \operatorname{supp} \Gamma_{\ell'}$, then it must be that $\pi_\ell(z) = \pi_{\ell'}(z)$ by profit maximization (17). By definition, $\underline{z}_\ell \in \operatorname{supp} \Gamma_\ell$. Hence, if $\underline{z}_\ell \in \operatorname{supp} \Gamma_{\ell'}$, then $\pi_\ell(\underline{z}_\ell) = \pi_{\ell'}(\underline{z}_\ell)$. For the sake of contradiction, suppose $\underline{z}_\ell \notin \operatorname{supp} \Gamma_{\ell'}$; that is, $\bar{z}_{\ell'} < \underline{z}_\ell$ from Lemma B.2. Feasibility (7) then requires that there exists at least one city l for which $\pi_{\ell'}(z) < \pi_l(z)$ for $z \in B^+(\bar{z}_{\ell'})$, and thus $1 = \Gamma_{\ell'}(z) > \Gamma_l(z) > \Gamma_\ell(z) = 0$. However, from Lemma B.3, this is possible in equilibrium only if city l satisfies $\theta_{\ell'} > \theta_l > \theta_\ell$, a contradiction. We therefore conclude that $\pi_\ell(\underline{z}_\ell) = \pi_{\ell'}(\underline{z}_\ell)$.

To prove the second statement, I start by showing that $\theta_\ell < \theta_{\ell'}$ implies $\kappa_\ell > \kappa_{\ell'}$. For that, take two cities ℓ and ℓ' that are adjacent in the $\boldsymbol{\theta}$-space with $\theta_\ell < \theta_{\ell'}$. Since n_ℓ is a strictly increasing function, equation (32) implies

$$\begin{aligned}
\kappa_\ell - \kappa_{\ell'} &\geq \underline{z}_\ell \left(n_\ell(\underline{z}_\ell) - n_{\ell'}(\underline{z}_\ell)\right) + \underline{z}_{\ell'} \left(n_{\ell'}(\underline{z}_\ell) - n_{\ell'}(\underline{z}_{\ell'})\right) \\
&= \underline{z}_\ell \left(n_\ell(\underline{z}_\ell) - n_{\ell'}(\underline{z}_\ell)\right) + \underline{z}_{\ell'} \left(n_{\ell'}(\underline{z}_\ell) - n_{\ell'}(\underline{z}_{\ell'})\right).
\end{aligned}$$

The second term is positive since $\underline{z}_\ell > \underline{z}_{\ell'}$. Regarding the first term, Lemma B.2 requires that $\pi'_{\ell'}(z) = n_{\ell'}(z) \leq n_\ell(z) = \pi'_\ell(z)$ for $z \in B^+(\underline{z}_\ell)$, which must also hold at $\underline{z}_\ell$. Hence $\kappa_\ell > \kappa_{\ell'}$. To prove the other direction, suppose that $\kappa_\ell > \kappa_{\ell'}$. In equilibrium, it must either be that $\Gamma_\ell \succ \Gamma_{\ell'}$ or $\Gamma_\ell \prec \Gamma_{\ell'}$. For the sake of contradiction, suppose the latter. From Lemma B.3, it must then be that $\theta_\ell > \theta_{\ell'}$. Using the same derivation as above, we then conclude that $\kappa_\ell < \kappa_\ell$, a contradiction. Hence, $\Gamma_\ell \succ \Gamma_{\ell'}$ and $\theta_\ell < \theta_{\ell'}$. □

Lemma B.8 (City ordering)**.**

Suppose that $\lambda^e \approx \lambda^u$ and $c \gg b$. Then, in equilibrium, $e_\ell > e_{\ell'} \iff v_\ell > v_{\ell'} \iff \theta_\ell < \theta_{\ell'}$.

Proof. From Lemma B.7, $\theta_\ell < \theta_{\ell'}$ if and only if $\kappa_\ell > \kappa_{\ell'}$. Using $\kappa_\ell = \underline{w}_\ell n_\ell(\underline{z}_\ell) + r_\ell$, it follows that $\theta_\ell < \theta_{\ell'}$

if and only if

$$\frac{b}{1+k}\left(\frac{1}{\theta_\ell}-\frac{1}{\theta_{\ell'}}\right)+c\left(v_\ell^\chi-v_{\ell'}^\chi\right)>0,$$

where $\lambda^e \approx \lambda^u$ implies $\underline{w}_\ell \approx b$ in both cities. If $c \gg b$, then $\kappa_\ell - \kappa_{\ell'} \approx c\left(v_\ell^\chi - v_{\ell'}^\chi\right) > 0 \iff v_\ell > v_{\ell'}$. Hence, $\theta_\ell < \theta_{\ell'}$ if and only if $v_\ell > v_{\ell'}$. Finally, for the sake of contradiction, suppose that $\theta_\ell < \theta_{\ell'}$ but $e_\ell < e_{\ell'}$. However, $\theta_\ell < \theta_{\ell'}$ holds only if $v_\ell > v_{\ell'}$. Together, $\theta_\ell = v_\ell/e_\ell > v_{\ell'}/e_{\ell'} = \theta_{\ell'}$, a contradiction. □

C Extended model and estimation

C.1 Quantitative model

Let κ denote the scale of the model. The model is scale invariant but this will be needed to map the model in the data since the empirical scale is unknown.

Workers We start with the worker problem. Their expected lifetime utility upon entry in the labor force is

$$U(\{\omega_\ell\}_{\ell=1}^L)=\mathbb{E}\left[\int \mathrm{e}^{-\beta}\omega_{\ell_t}A_{\ell_t}u(c_t,h_t)dt\right].$$

Workers are free to migrate as long as they are unemployed. Employed workers need to quit their jobs to migrate. Together with the facts that the economy is in steady state and that idiosyncratic preferences are time invariant, unemployed workers settle once in a location and then never choose to migrate again.

The problem of workers can therefore be broken down in two stages. In the later stage, workers are settled in location ℓ. They decide which jobs to accept and climb the local job ladder. The HJB equations describing the discounted lifetime utility of an unemployed and employed worker in location ℓ are

$$\beta U_\ell = \kappa b + \lambda_\ell^u \int \max\{V_\ell(w)-U_\ell,0\}dF_\ell(w),$$
$$\beta V_\ell(w) = w + \lambda_\ell^e \int \max\{V_\ell(w)-U_\ell,0\}dF_\ell(w) + \delta[U_\ell - V_\ell(w)].$$

These HJB equations are independent of workers' taste shocks, and therefore, need to be equalized across space. The job switching behaviors are identical to Section 2 and give rise to the same labor supply curves.

In the earlier stage, unemployed workers decide where to settle given the expected lifetime utility of unemployed workers in each location. This location choice solves:

$$U=\max_\ell \frac{\omega_\ell A_\ell U_\ell}{P_\ell}.$$

Given the Frechet shocks, the fraction of unemployed workers who decide to locate in city ℓ is

$$\frac{u_\ell}{\sum_{\ell'} u_{\ell'}} = \frac{[A_\ell U_\ell/P_\ell]^\chi}{\sum_{\ell'}[A_{\ell'}U_{\ell'}/P_{\ell'}]^\chi}, \tag{33}$$

where $\bar{u} \equiv \sum_\ell u_\ell$. The flows in and out of unemployment imply

$$e_\ell = k_\ell^u u_\ell \quad \text{and} \quad m_\ell = u_\ell + e_\ell = u_\ell \left(1 + k_\ell^u\right). \tag{34}$$

Employers Employers solve (20) where $c(v) = \kappa v^{1+\gamma}/(1+\gamma)$. The vacancy optimality condition demands

$$\left(zT_\ell - w_\ell(z)\right) n_\ell[w_\ell(z)] = \kappa v_\ell(z)^\gamma. \tag{35}$$

Potential profits in location ℓ after maximizing out vacancies are

$$\pi_\ell(zT_\ell) = \max_w \; c\kappa^{-\frac{1}{\gamma}} \left[\left(zT_\ell - w\right) n_\ell(w)\right]^{\frac{1}{c}} - r_\ell,$$

for $c \equiv \gamma/(1+\gamma)$ a constant. Given the spatial allocation of employers $\{\Gamma_\ell\}_\ell$, the solution to the employer's problem can be recasted as a system of two differential equations. For employers with $z \in \operatorname{supp} \Gamma_\ell$, the wage optimality condition is

$$\frac{n_\ell'[w_\ell(z)]}{n_\ell[w_\ell(z)]}\left(zT_\ell - w_\ell(z)\right) = \frac{2k_\ell \mathrm{d}F_\ell[w_\ell(z)]}{1 + k_\ell \bar{F}_\ell[w_\ell(z)]}\left(zT_\ell - w_\ell(z)\right) = 1.$$

Within a city, wages are increasing in productivity. Accordingly, letting $\Upsilon_\ell(z) \equiv F_\ell[w_\ell(z)]$ denote the rank of an employer in the local wage offer distribution, we obtain

$$\Upsilon_\ell(z) = \frac{M_\ell}{V_\ell} \int_{\underline{z}_\ell}^{z} v_\ell(x) \mathrm{d}\Gamma_\ell(x).$$

The number of workers hired by employer z in ℓ is then

$$n_\ell(z) = n_\ell[w_\ell(z), v_\ell(z)] = \frac{(1+k_\ell)e_\ell}{[1 + k_\ell(1 - \Upsilon_\ell(z))]^2} \frac{v_\ell(z)}{V_\ell}. \tag{36}$$

With endogenous vacancies, $\Upsilon_\ell(z) \neq \Gamma_\ell(z)$. Using $\Upsilon_\ell(z)' = dF_\ell[w_\ell(z)]w_\ell'(z)$, the wage optimality rewrites

$$w_\ell'(z) = \frac{2k_\ell \mathrm{d}\Upsilon_\ell(z)}{1 + k_\ell \bar{\Upsilon}_\ell(z)}[zT_\ell - w_\ell(z)]. \tag{37}$$

Given Υ_ℓ and the boundary condition $w_\ell(\underline{z}_\ell) = \underline{w}_\ell$, (37) is an ODE that yields $w_\ell(z)$. Meanwhile, the expression for $\Upsilon_\ell(z)$ implies $d\Upsilon_\ell(z) = M_\ell v_\ell(z) d\Gamma_\ell(z)/V_\ell$. Used in the vacancy optimality condition, we obtain a second differential equation:

$$\Upsilon_\ell'(z) = \frac{M_\ell}{V_\ell}\left[\left(\frac{zT_\ell - w_\ell(z)}{\kappa}\right) n_\ell(z)\right]^{\frac{1}{\gamma}} \Gamma_\ell'(z). \tag{38}$$

Given the condition $\Upsilon_\ell(\underline{z}_\ell) = 0$ and $w_\ell(z)$, (38) yields $\Upsilon_\ell(z)$.

The local productivity distributions are given by the local profit opportunities, π_ℓ, together with the employers' idiosyncratic costs. The inverse dispersion in the entry cost is ϑ/κ. Given their Gumbel distribution, the probability that a firm with productivity z produces in city ℓ is

$$\Omega_\ell(z) = \frac{e^{\vartheta \pi_\ell(zT_\ell)/\kappa}}{\sum_{\ell'} g_{\ell'} e^{\vartheta \pi_{\ell'}(zT_{\ell'})/\kappa}}. \tag{39}$$

The local productivity distribution in city ℓ is then

$$\Gamma_\ell(z) = \frac{M}{M_\ell}\int^z \Omega_\ell(x)\mathrm{d}\Gamma(x) \quad \Longleftrightarrow \quad \mathrm{d}\Gamma_\ell(z) = \frac{M}{M_\ell}\Omega_\ell(z)\mathrm{d}\Gamma(z), \tag{40}$$

and the mass of firm in city ℓ is

$$M_\ell = M\int \Omega_\ell(x)\mathrm{d}\Gamma(x). \tag{41}$$

The idiosyncratic entry costs ensure full support of the local productivity distribution, so that the boundary condition to (39) is $\Gamma_\ell(\underline{z}) = 0$.

Frictions The contact rates are given by the local matching function:

$$\lambda^u_\ell = \mu\left(\frac{V_\ell}{u_\ell + \zeta e_\ell}\right)^\xi. \tag{42}$$

Aggregate prices There are three sets of aggregate quantities. The reservation wages $\{\underline{w}_\ell\}_{\ell=1}^L$ are given by (13). Commercial housing prices clear the commercial housing market:

$$\frac{\alpha}{p_\ell}\left(u_\ell b\kappa + e_\ell \mathbb{E}_\ell[w]\right) = \bar{L}p_\ell^\theta. \tag{43}$$

The residential housing prices clear the residential housing market:

$$M_\ell = \bar{H}(r_\ell/\kappa)^\phi, \tag{44}$$

where the housing supply has been scaled by κ, $H_\ell = \bar{H}(r_\ell/\kappa)^\phi$.

Definition C.1 (Equilibrium)**.**
An equilibrium is a collection of size policy functions, $\boldsymbol{n}(z)$, wage policy functions, $\boldsymbol{w}(z)$, vacancy policy functions, $\boldsymbol{v}(z)$, employer-rank distributions, $\boldsymbol{\Upsilon}(z)$, productivity distribution, $\boldsymbol{\Gamma}(z)$, residential and commercial housing prices, $\boldsymbol{p}$ and $\boldsymbol{r}$, reservation wages, $\boldsymbol{w}$, contact rates, $\boldsymbol{\lambda^u}$ and $\boldsymbol{\lambda^e}$, vacancy and employer measures, $\boldsymbol{V}$ and $\boldsymbol{M}$, spatial distribution of workers, $\boldsymbol{u}$ and $\boldsymbol{e}$, so that:

1. *The wage, employer-rank and productivity distributions satisfy the differential equations* (37), (38) *and* (40)*, subject to the appropriate boundary conditions and the spatial allocation of employers $\boldsymbol{\Omega}(z)$ given by* (39)*;*
2. *The vacancy and size policy functions satisfy* (35) *and* (36)*;*
3. *The mass of vacancy and employers are given by* (19) *and* (41)*;*
4. *The mass of unemployed and employed workers are given by* (33) *and* (34)*;*
5. *The contact rates are given by* (42)*;*
6. *The reservation wages are given by* (13)*;*

7. The residential housing prices follows (43) *and* (44).

C.2 Proof of Proposition 5

Frictions The average UE rate, average EU rate, and location-specific average EE rate identifies the search frictions. First, the EU rate equates δ. Second, the location-specific average EE rates are

$$\mathrm{EE}_\ell = \lambda_\ell^e \int \bar{F}_\ell(w) dG_\ell(w) = \delta \left[\left(1 + \frac{1}{k_\ell}\right) \log\left(1 + k_\ell\right) - 1 \right].$$

The right-hand side is monotonically increasing in k_ℓ. Therefore, $k_\ell = h(\widehat{\mathrm{EE}}_\ell)$, for h the inverse of the above mapping. But $k_\ell = \zeta \lambda_\ell^u / \delta$, and $\zeta \lambda_\ell^u = \delta h(\widehat{\mathrm{EE}}_\ell)$. Third and last, the location-specific UE rate is λ_ℓ^u. Accordingly, the average UE rate is

$$\mathrm{UE} = \frac{\sum_\ell u_\ell \lambda_\ell^u}{\sum_\ell u_\ell} = \frac{1}{\zeta} \frac{1}{\sum_\ell \left(\frac{1}{\zeta \lambda_\ell^u}\right) \frac{e_\ell}{E}},$$

which follows from $u_\ell = \delta/(\delta + \lambda_\ell^u) m_\ell$ and $m_\ell = u_\ell + e_\ell = (\lambda_\ell^u + \delta) e_\ell / \delta$. In the denominator of the second term, e_ℓ / E refers to the employment share of location ℓ with $E = \sum_\ell e_\ell$. The denominator is therefore the employment-weighted mean of $1/\zeta \lambda_\ell^u$, which we know. The above expression thus identifies ζ. Given ζ, we separately identify $\{\lambda_\ell^u\}_\ell$.

Worker allocation The total mass of workers in ℓ reads $m_\ell = (\lambda_\ell^u + \delta) e_\ell / \lambda_\ell^u$. Consistency imposes $\sum_\ell m_\ell = 1$. Together:

$$\frac{1}{E} = \sum_\ell \left(\frac{\lambda_\ell^u + \delta}{\lambda_\ell^u}\right) \frac{e_\ell}{E}.$$

This expression identifies the aggregate mass of employed workers. We can then recover the location-specific measure of employed and unemployed workers via

$$e_\ell = \left(\frac{e_\ell}{E}\right) E \quad ; \quad u_\ell = \left(\frac{\delta}{\lambda_\ell^u}\right) e_\ell \quad ; \quad m_\ell = u_\ell + e_\ell.$$

We treat the spatial allocation of workers as given for now, and later show that any allocation can be rationalized by local amenities.

Unemployment insurance The UI rate is identified from the aggregate replacement rate. The location-specific replacement rate is $\mathrm{RE}_\ell = \kappa b / \mathbb{E}[w_\ell]$. The aggregate replacement rate is therefore

$$\mathrm{RE} = \frac{\sum_\ell \left(\frac{\kappa b}{\mathbb{E}[w_\ell]}\right) e_\ell}{\sum_\ell e_\ell}.$$

Inverting this expression identifies κb.

Employers Let j denote an employer in the sample. For each employer, we observe the wage it offers, w_j, its size, N_j, and its location, ℓ_j. From employers' location, we can compute the measure of firms in every location. Consistency indeed requires that the number of workers hired in each location equates the number of employed workers, $M_\ell \mathbb{E}_\ell[n(z)] = e_\ell$ holds. The average size of employers relative to the measure of employed workers thus identifies the number of employers: $m_\ell = e_\ell / \mathbb{E}_\ell[N(z)]$. The total measure of employers follows $M = \sum_\ell g_\ell m_\ell$.

Then, we can recover the productivity and vacancy of each employer from its size and wage:

1. Compute employers' rank in the local wage distribution, G_j.
2. Invert (11) to back out employers' rank in the wage offer distribution;

$$F_j = \frac{(1+k_\ell)G_j}{1+k_\ell G_j}.$$

3. Compute the number of workers per vacancy from employers' rank:

$$\eta_j = \frac{(1+k_\ell)e_\ell}{[1+k_\ell(1-F_j)]^2}.$$

4. Compute employers' vacancy share from their size: $v_j / V_{\ell_j} = n_j / \eta_j$.
5. Invert the wage optimality condition to identifies employers' productivity:

$$z_j T_{\ell_j} \equiv \zeta_j = w_j + \left(\frac{1+k_\ell(1-F_j)}{2k_\ell} \right) \left. \frac{\partial w}{\partial F} \right|_{w=w_j}.$$

In practice, I approximate $w \to F_\ell$ with a spline and use the approximation to compute $\partial w / \partial F$.

Given data on vacancy shares, the vacancy cost elasticity is identified the vacancy optimality condition. Equation (35) indeed implies

$$\gamma = \sqrt{\frac{\sum_\ell e_\ell \mathbb{V}\mathrm{ar}_\ell[\log\left((\zeta_j - w_j)\eta_j\right)]}{\sum_\ell e_\ell \mathbb{V}\mathrm{ar}_\ell[\log \frac{v_j}{V_\ell}]}}. \tag{45}$$

Given an estimate for γ, the measure of vacancy in each location is then recovered as a location-specific residual from the optimality condition:

$$(1+\gamma)\log V_\ell + \log \kappa = \mathbb{E}_\ell[\log(\zeta_j - w_j)\eta_j] - \gamma \mathbb{E}_\ell \left[\log \frac{v_j}{V_\ell} \right].$$

From (39), the log probability that an employer with total productivity ζ produces in ℓ is

$$\log \Omega_\ell^\zeta(\zeta) = \vartheta \pi_\ell(\zeta) - \log \sum_{\ell'} \mathrm{e}^{\vartheta \pi_{\ell'}(\zeta T_{\ell'}/T_\ell)} \approx \vartheta \pi_\ell(\zeta) + H(\zeta), \tag{46}$$

where $H(\zeta)$ is some location-independent function and the approximation is exact when there are no dispersion in TFP, $T_{\ell'} = T_\ell$ for all ℓ. According to (46), projecting employers' location probabilities onto the local profit opportunities while flexibly controlling for employers' total productivity identifies ϑ. The left-hand side is known from employers' productivity and their location choice. Given the parameters already identifies, realized profits can be computed

$$\pi_j = \pi_{\ell_j}(\zeta_j) = (\zeta_j - w_j)n_j - c(v_j) - r_{\ell_j}.$$

In practice, I instrument local profit opportunities with employers' wage to reduce measurement error in employers' total productivity. I take a discrete approximation of ζ and replace $H(\zeta)$ by employer-group fixed effect. To account for the measurement error introduced by this discretization, together with the fact that TFP differentials may not be small, I replicate the same estimation technique in the model and set ϑ to match the 2SLS coefficient obtained in the data.

Matching function The matching function (42) implies

$$\log \lambda_\ell^u = \log \mu + \psi \log \left(\frac{V_\ell}{u_\ell + \zeta e_\ell} \right) + u_\ell, \tag{47}$$

for ϵ_ℓ some measurement error. Estimating (47) by OLS given our estimates of $\{\lambda_\ell^u, V_\ell, u_\ell, e_\ell\}_{\ell=1}^L$ and ζ identifies the matching function parameters (μ, ψ).

Housing supply The residential housing market clearing condition reads $\bar{L}p_\ell^\theta = \alpha I_\ell / p_\ell$, or

$$\log p_\ell = \left(\frac{1}{1+\theta} \right) \log \left(\frac{\alpha}{\bar{L}} \right) + \left(\frac{1}{1+\theta} \right) \log I_\ell, \tag{48}$$

where $I_\ell = u_\ell \kappa b + e_\ell \mathbb{E}[w_\ell]$ are total expenditures in ℓ. Accordingly, given data on residential housing prices, estimating the above relationship by OLS identifies $\{\bar{L}, \theta\}$.

Turning to the commercial housing supply parameters, commercial housing market clearing demands $\bar{H}(r_\ell/\kappa)^\phi = m_\ell$, or

$$\log r_\ell = \log \kappa - \frac{1}{\phi} \log \bar{H} + \frac{1}{\phi} \log M_\ell. \tag{49}$$

Estimating (49) by OLS identifies ϕ and a bundle of $(\kappa, \bar{H})$.

Worker preferences We conclude with the identification of worker preferences, $(\alpha, \boldsymbol{A})$. The Cobb-Douglas parameter α equates the aggregate housing expenditure share.

Amenities are identified from the location choice of unemployed workers. Specifically, given the dispersion in preferences χ, the spatial allocation of unemployed workers $\boldsymbol{u}$, housing prices $\boldsymbol{P}$, and the lifetime utility of unemployed workers in each location $\boldsymbol{U}$, (33) can be inverted to obtain local amenities. The first two vetors are known. The latter can be computed given data on the wage offer distribution and the discount factor β. Indeed, following similar steps as in Section B.2, I obtain

$$\beta U_\ell = \kappa b + \lambda_\ell^u \int_{\underline{w}_\ell} \frac{1 - F_\ell(w)}{\beta + \delta + \lambda^e (1 - F_\ell(w))} dw.$$

This conclude the proof.

C.3 Data

To be written.

C.4 Figures

Figure C.9: City characteristics

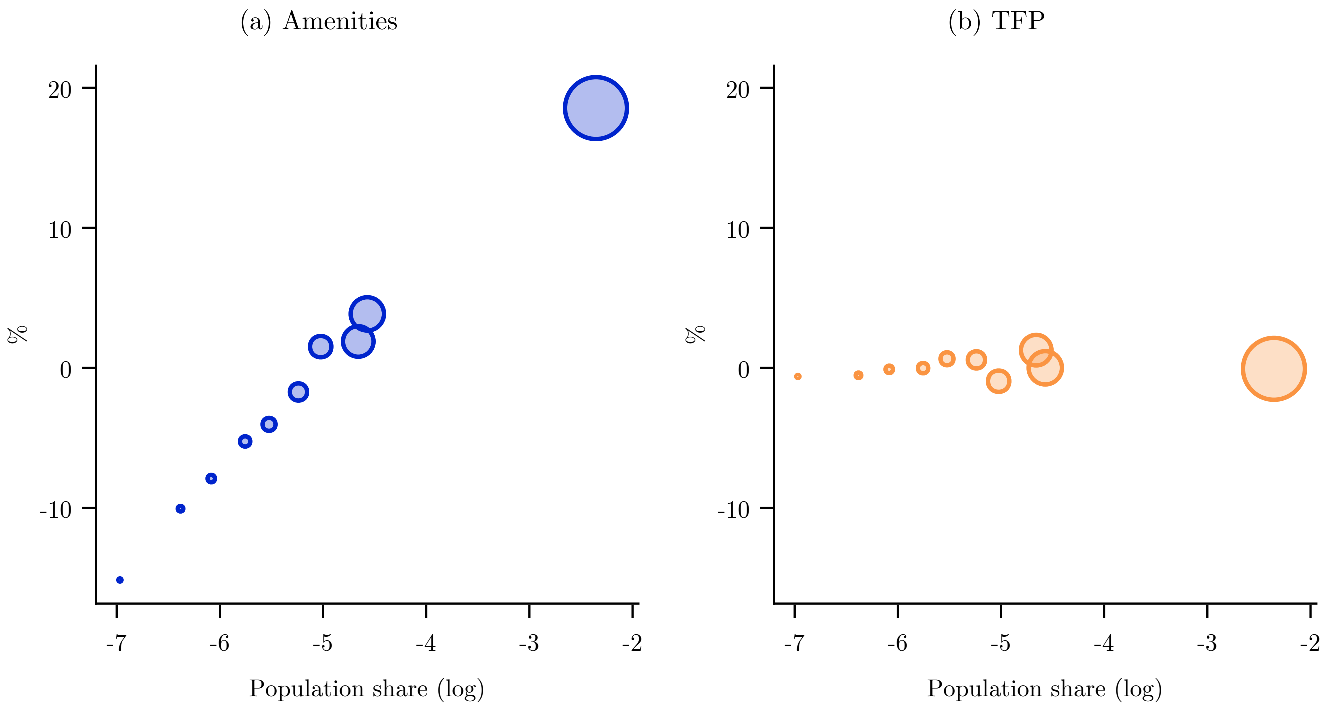


Figure C.10: Model fit

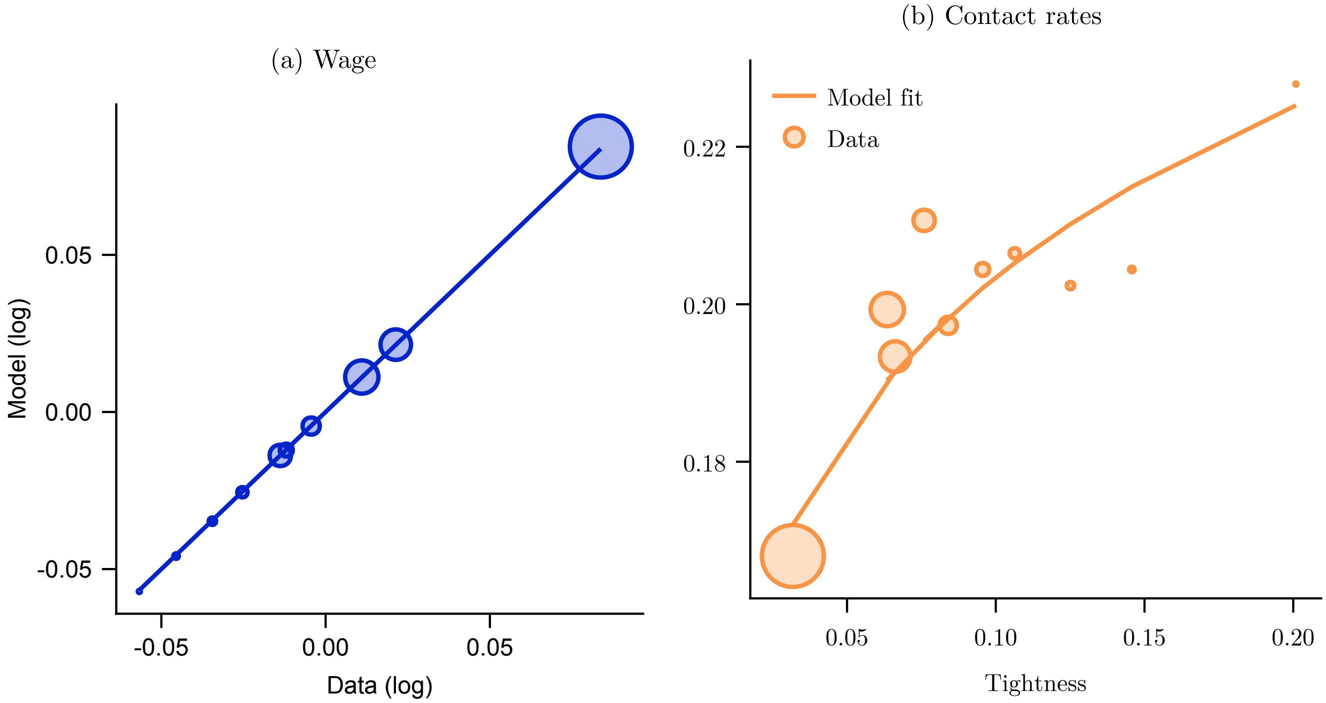
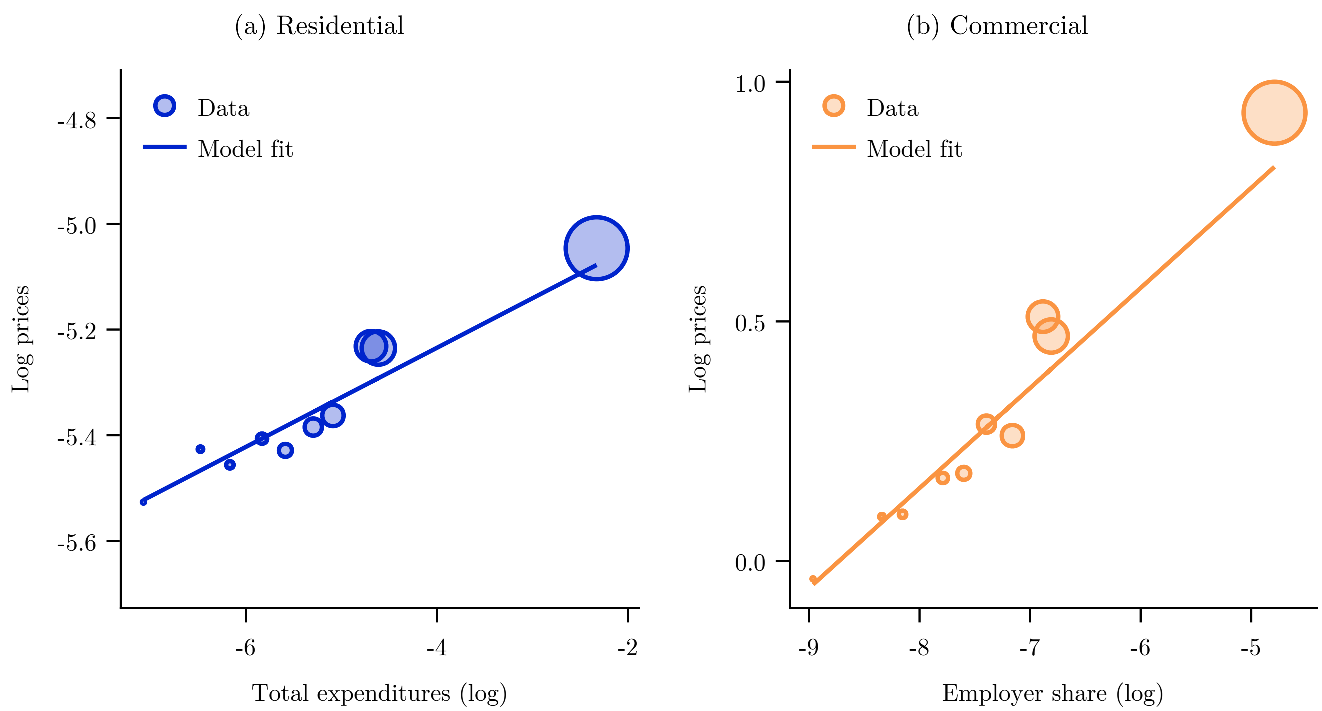


Figure C.11: Housing prices



D Quantitative exercises

D.1 Welfare decomposition

To be written.

D.2 Figures

Figure D.12: The drivers of local wage inequality

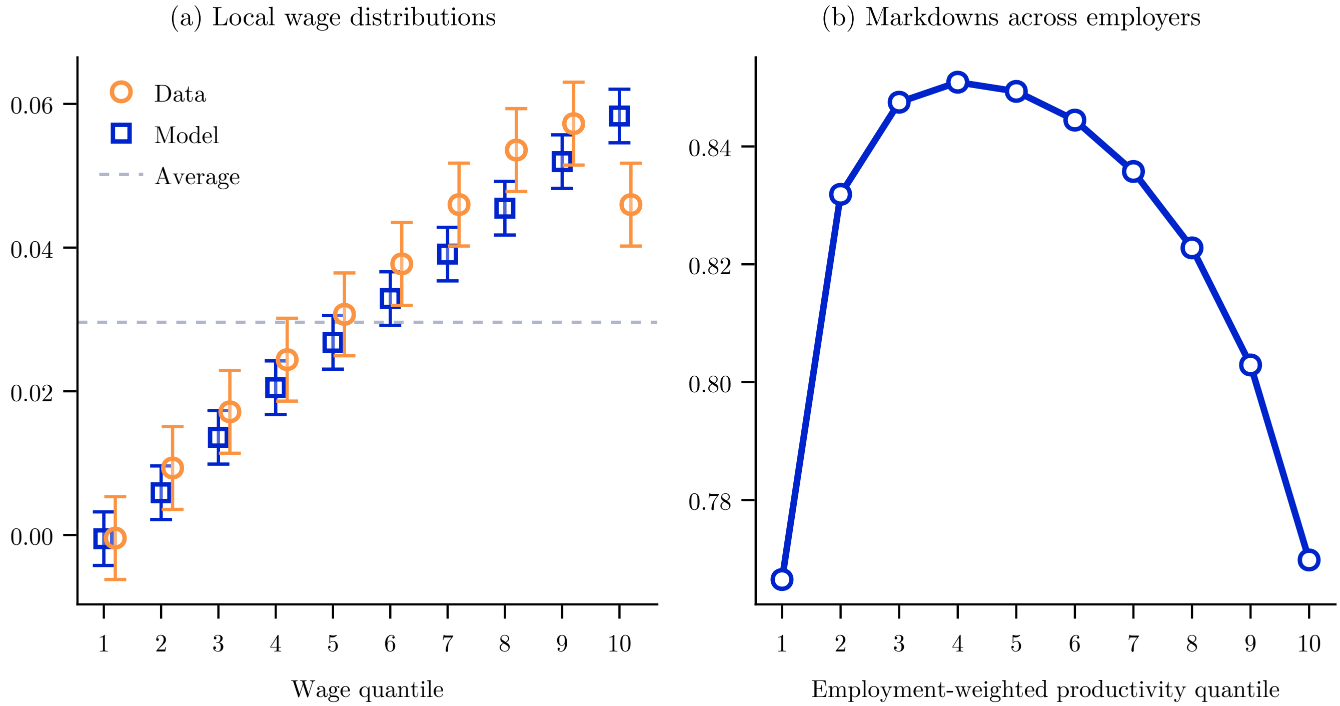


Figure D.13: The drivers of local wage inequality — robustness