The Local Root of Wage Inequality

Hugo Lhuillier*
University of Chicago

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 space. Second, the wage gains offered by large cities materialize over time as workers reallocate from low- to high-paying jobs. I propose a spatial framework that rationalizes these facts through two ingredients: heterogeneous employers and frictional local labor markets with on-the-job search. Productive employers agglomerate in large cities to hire more workers. Fiercer competition for workers arises. A higher average wage, faster growth, and greater within-city inequality follow. I estimate the model and quantify that local TFP gaps are minimal once I account for employers' incentives to sort by size. The steeper ladder of large cities implies higher lifetime real earnings for every local worker, including those with lower real wages.

*First version: August 2022. This version: June 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, Charly Porcher, Steve Redding, and Rob Shimer 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, wages are high, but so are inequalities. There, high-income earners live alongside low-wage workers who face 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. In particular, low-income earners in Paris —those in the bottom 10% of the wage distribution— earn only 2% more than low-income earners in Lens, and 20% *less* once accounting for housing prices.

Figure 1 highlights that, to understand how cities shape wages, it is important to study their inner making. However, while a vast body of research has studied wage gaps *across* space, little is known as to what causes within-city inequality. What drives spatial wage inequality? Why are wages higher but also more dispersed in larger cities? And why are some workers accepting lower real wages there?

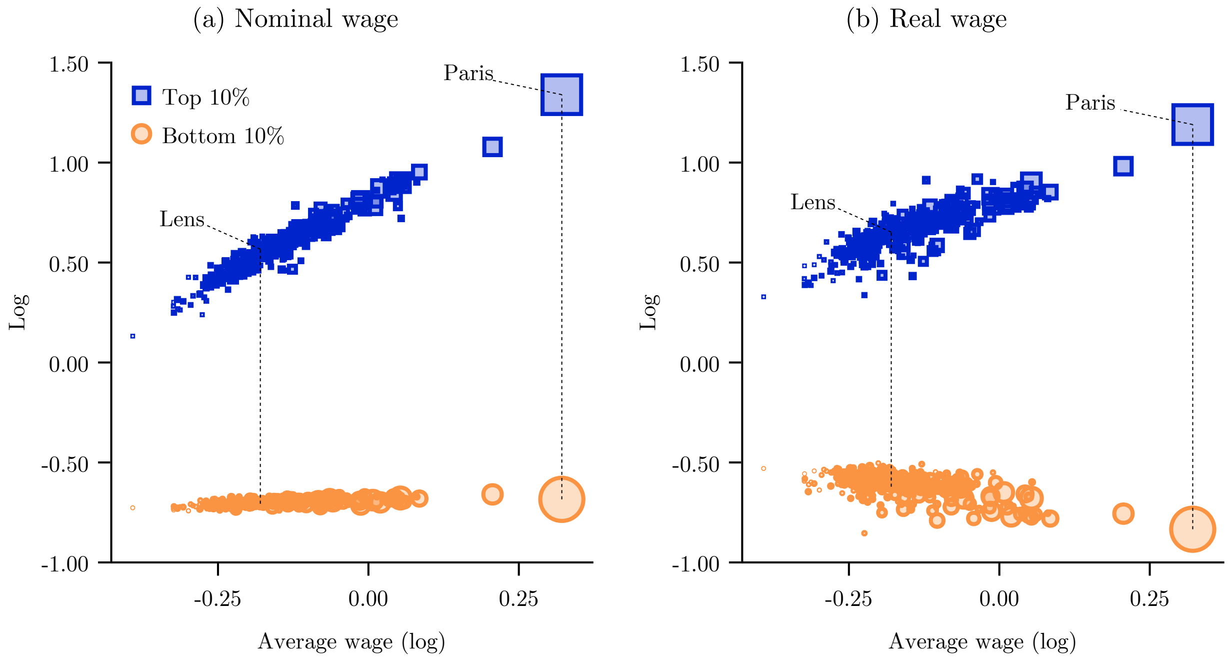
This paper offers answers to these questions in three parts. First, I document that employers shape between- and within-city wage inequality jointly. Second, I provide a framework that endogenously generates spatial wage disparities through the sorting of employers across frictional local labor markets. Third, I quantify that spatial TFP gaps are negligible once accounting for employer sorting, and that the net present value of lifetime real earnings is higher in larger cities —including for workers with low present real earnings.

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 high-paying jobs are concentrated in large cities while low-paying jobs are dispersed throughout France. To quantify the relative importance of jobs net of worker heterogeneity, I estimate a mover design à la Abowd et al. (1999) —AKM henceforth. Assuming conditional random mobility of workers across jobs, I estimate the wage premia offered by a job via that job's fixed effect.

I find that workers in commuting zones (CZ) twice larger earn an average wage premia 3.1% higher —or 33.6% of the between-city wage variance. These average gaps arise from the spatial concentration of high-paying jobs. For instance, while 17.9% of the jobs in Paris belong to the top 10% of the national job fixed effect distribution, this is only 6.9% of the jobs in Lens. By contrast, low-paying jobs —those in bottom 10% of the national fixed effect distribution— are equally present in both locations. As a result, larger locations are more unequal: the greater wage premia dispersion in larger cities accounts for 33.2% of the spatial differences in within-city wage variance.

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

Figure 1: Gross hourly wage by commuting zone



Data source: French matched employer-employee (see Section 1.1). 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 from low- to high-paying jobs. I estimate that starting wages are very similar across CZs once I control for unobserved worker heterogeneity. For instance, the startup premia in Paris is 0.4% lower than in Lens. However, the job ladder is steeper in the French capital. I estimate the local wage returns to job switching as the extra wage growth of job switchers relative to the wage growth of job stayers. I find that the wage gains upon a job switch are 0.3 percentage points (p.p.) higher in cities twice bigger (relative to a mean of 1.3%). This pattern holds within and across occupations, throughout the wage distribution, and after controlling for heterogeneous learning abilities across workers. As a result, wages diverge over workers' career, and the reallocation of workers across jobs explains 73% of the between-city wage variance net of worker heterogeneity.

These two facts underscore the key role of employers in shaping between- and within-city wage inequality. They also provide new testable implications for frameworks that study spatial wage disparities. Hicks neutral TFP gaps fail at generating local wage dispersion. Worker heterogeneity cannot explain wage premia. Finally, ex-post productivity or amenity shocks do not capture the dynamic impact of cities. In the second part of the paper, I therefore propose a new framework that generates spatial wage inequality via employer sorting 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 in the city where they live: 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 size. As in Burdett and Mortensen (1998), wages serve as a hiring tool: offering higher wages allows employers to attract and retain more workers from the local competitors.

I characterize the model analytically, which allows me to provide two new insights as to the drivers of spatial wage inequality.

First, productive employers agglomerate in large cities to partially sidestep search frictions. Employers' location decision depends on three considerations: how slack is the labor market, how intense is the local competition as captured by the 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 slacker markets. In equilibrium, markets are slacker in larger locations as housing prices price out unproductive employers. As productive employers agglomerate there, competition intensifies, and employers offer relatively higher wages to attain their target size.

Hiring frictions therefore generate a positive correlation between city size, employer productivity, and local wages without Total Factor Productivity (TFP) differentials. By contrast, when labor market are competitive, local TFPs are necessary to obtain a positive correlation between size and wages for otherwise productive employers agglomerate in locations with cheap inputs.

The second insight is that high-paying jobs are necessarily spatially concentrated, whereas low-paying jobs are necessarily present in every city. In the presence of monopsony power, the wage offered by employers depend on the competition they face rather than their own productivity. The fiercer competition in large cities requires productive employers, for which a large share of their workforce comes from poaching competitors, to offer higher wages. High-paying jobs are therefore concentrated in large cities as they exist only through the fierce local competition. By contrast, employers who hire most of their workforce from unemployment offer a wage close to unemployed's outside option. The wages at the bottom of the local ladders are independent of the local employer composition, and low-paying jobs exist in every city.

Altogether, the concentration of productive employers in large cities puts upward pressure on wages and local inequalities. The local job ladder steepens, job switching yields higher wage gains, and workers there initially accept lower real-paid jobs in anticipation of higher future real earnings.

Given these new insights, in the third part of the paper, I estimate a quantitative version of the model to re-assess the drivers of spatial wage disparities.

I extend the model in three ways. First, I allow cities to differ in Total Factor Productivity. High wages may then result from the agglomeration of productive employers or high TFP. 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 work at lower real earnings in large cities: idiosyncratic preferences, and migration costs. The model is estimated on the AKM job fixed effects, which allows me to abstract from worker heterogeneity.

I provide a proof of identification for 16 of the 20 parameters by combining the strategy used in quantitative spatial models (Redding and Rossi-Hansberg, 2017) 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 premia.

I find that the average gap in wage premia across CZs are entirely explained by employer sorting. 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, once interacted with the local competition for workers, suffices to explain why wage premia are on average higher in bigger places. Quantitatively, local TFPs account for 3.2% of the between-city wage premia variance.

The estimation procedure targets the aggregate wage premia variance and the local average premia. It targets neither how within-city inequality varies across space, nor how steep are the local job ladders. I therefore use the two novel facts documented in this paper as over-identification tests to validate the importance of employers for spatial wage disparities.

Testing first the model's prediction on within-city inequality, I find that the spatial concentration of productive employers does generate greater wage premia dispersion in bigger places. As in the data, low-paying jobs are dispersed throughout space, and as a result, 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 premia distributions.

Wage inequalities are greater in bigger locations as workers do not benefit equally from the fiercer local competition. Every employer in Paris is relatively more productive than in Lens. For instance, while workers in the bottom 10% of the wage distribution in Paris earn wages 3.2% higher than their counterparts in Lens, their employers' productivity is 11.9% greater. This productivity gap is not passed through onto workers as employers in both locations hire their workforce from unemployment. By contrast, workers at the top of the Paris' ladder extract more rents than any other workers because their employers compete locally with firms who are relatively more productivity. For instance, employers in the top 10% of the Paris' wage distribution charge markdown 19.3% higher than similar employers in other cities. Combined, I quantify that within-city inequality would be lower in larger cities were markdowns uniform across space.

Turning to the second over-identification exercise, I find that workers in larger cities enjoy higher lifetime real earnings, including those who earn lower real wages. The steeper job ladder in larger cities generates wage growth gains consistent with my estimated local returns to job switching. The faster growth spurs higher lifetime income. For instance, the net present value of new entrants' real lifetime earnings is 4.2% higher in Paris than in Lens —despite the fact that their real earnings are 8.1% lower. Generalizing to all locations, I quantify that new entrants in cities twice larger earn lifetime real earnings 1% higher.

The reallocation of workers across jobs, and the wage growth it begets, are thus key to the spatial distribution of economic activity. I conclude the paper with a counterfactual that simulates a decline in the average job switching rate by 1 p.p. through a national increase in search frictions.

A slower reallocation across employers depresses workers' expected lifetime income, and disproportionally so in large cities. Holding constant the spatial allocation of jobs, the extra lifetime earnings offered by Paris drops by 1.3 p.p., and as a result, the number of workers there shrinks by 12.7%. The consequences of stronger frictions on local productivity are *a priori* more ambiguous. On the one hand, unproductive employers can retain a greater share of their workers in large cities. On the other other hand, productive employers value relatively more the slackness of these places as it is costlier to poach workers from the competition. In net, the average productivity in Paris reduces from 18.9% to 14.7%, further depressing the earnings offered by large locations. In equilibrium, Paris' size shrinks by 22.2% as the expected lifetime earnings are 2.9 p.p lower than in the baseline equilibrium, highlighting the importance of considering the interactions between the local ladders and employers' spatial allocation.

Related literature This paper relates to several strands of literature. The first studies spatial wage inequality. A vast body of research, started by Marshall (1890), and then followed — amongst many others — by Glaeser and Maré (2001), Combes et al. (2008), Bacolod et al. (2009), Moretti (2011), and Moretti and Yi (2024), analyzes wage gaps *across* locations. Particularly related are Card et al. (2025) and Carry et al. (2025) who both document the importance of employers for between-city inequality. Motivated by the evidence that between-city inequality explains at most 20% of the aggregate wage variance, Glaeser et al. (2009), Baum-Snow and Pavan (2013), Eeckhout et al. (2014), Baum-Snow et al. (2018), and Papageorgiou (2022) analyze wage inequality *within* cities, primarily focusing on worker heterogeneity. I tie these two literature together by showing that the concentration of high-paying employers in large cities generate higher wages and greater local inequality.

Second, this paper contributes to the literature on the dynamic effects of cities on wages (Baum-Snow and Pavan, 2012; Roca and Puga, 2017; Porcher et al., 2023; Eckert et al., 2022; Lhuillier, 2023). Similar to Roca and Puga (2017), I estimate that starting wages are similar across locations once accounting for worker heterogeneity. However, whereas previous studies emphasize human capital accumulation, I document that the bulk of the faster wage growth in large cities occur from workers reallocating from low- to high-paying employers.

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 (2022), and

Heise and Porzio (2022). I contribute to these papers by allowing heterogeneous employers to sort across local labor markets, which I show is central to explain the steeper job ladder in larger cities. Lindenlaub et al. (2022) also develop a framework in which employers sort across frictional labor markets where workers search off- and on-the-job. Our papers complement each other: they study the impact sorting has on local labor shares whereas I analyze its impact on wage inequality.

Several other papers have emphasized the importance of firm sorting for spatial disparities, including Combes et al. (2012), Behrens et al. (2014), Gaubert (2018), Bilal (2023), Hong (2024), Franco (2024), and Kleinman (2024). In these frameworks, productive firms concentrate in large cities to benefit from high local productivity. By contrast, as in Oh (2023), I quantify that employer sorting is mainly driven by city size as it enables employers to mitigate search frictions.

Finally, this paper connects to the literature on search frictions and monopsony power (Burdett and Mortensen, 1998; Bontemps et al., 2000; Engbom and Moser, 2022; Bilal and Lhuillier, 2021; Jarosch et al., 2024; 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 markets shapes the rent-sharing process, reduces employers' market power in larger cities, and generates greater local inequality.

The rest of this paper is organized as follows. Section 1 presents the data and the two novel facts. Second 2 develops the novel theoretical framework. Second 3 lays out the quantitative extensions of the model and its structural estimation. Second 4 concludes by quantifying the consequences of employer sorting for spatial wage disparities.

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 allows to track workers throughout their careers. Second, its large sample size guarantees sufficient statistical power to precisely estimate statistics at relatively granular geographical units.

I apply the same sample restrictions to both datasets. I focus my analysis on full-time employed workers between 25 and 55 years old. Workers employed in the public sector have their wages 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 the 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.

I define a job as an establishment $\times$ 4-digit occupation, and a city as a commuting zone. 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, 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 group cities 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.1 provides summary statistics on each of the city cluster. This clustering accounts for 97% of the variation in average wage and wage variance across CZs.

1.2 Spatial wage inequality in France

Figure 1 displays the extent to which wages vary 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 wage.

Two patterns stand 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- and within-city component,

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

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 that are often associated with high and low inequality.[1]

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.

[1]Figure A.2 reproduces Figure 1 in the United States using data from the American Community Survey.

Figure 1b repeats the 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 0.3.[2]

Figure 1b shows 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 local wage distributions; on average, workers in the bottom 40% of their local wage distribution earn lower real wages in larger places. A similar pattern holds when housing prices are adjusted to take into account the heterogeneous exposure of workers across neighborhoods. It also holds amongst workers not born in the city, i.e., those who explicitly choose to live there.

Wages therefore vary substantially across space. Wages are on average higher in larger cities, but they are also more dispersed. I now show that employers are instrumental in shaping between- and within-city inequality jointly.

1.3 Fact #1: high-paying jobs are spatially concentrated whereas low-paying jobs are dispersed

I estimate how employers shape spatial wage inequality via a two-way fixed effect model à la Abowd et al. (1999) —AKM henceforth. Specifically, I estimate

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

where i is a worker, t is quarter-year, x_{it} denotes the number of years worked (i.e., experience), and $\ell(i,t)$ and $j(i,t)$ are indices for the city group and job where i is employed at time t. The worker fixed effects α_i^{AKM} control for time-invariant worker characteristics (e.g., education, gender) and other unobserved abilities. The parameters $\{\beta_\ell^{\text{AKM}}\}_{\ell=1}^{L}$ summarize the impact of experience on wages, allowing cities to offer differential returns to experience. I define the contribution of workers on wages as the sum of the static and dynamic components: $\psi_{it}^{\text{AKM}} \equiv \alpha_i^{\text{AKM}} + \beta_{\ell(i,t)}^{\text{AKM}} x_{it}$. Finally, γ_j^{AKM} are job fixed effects.

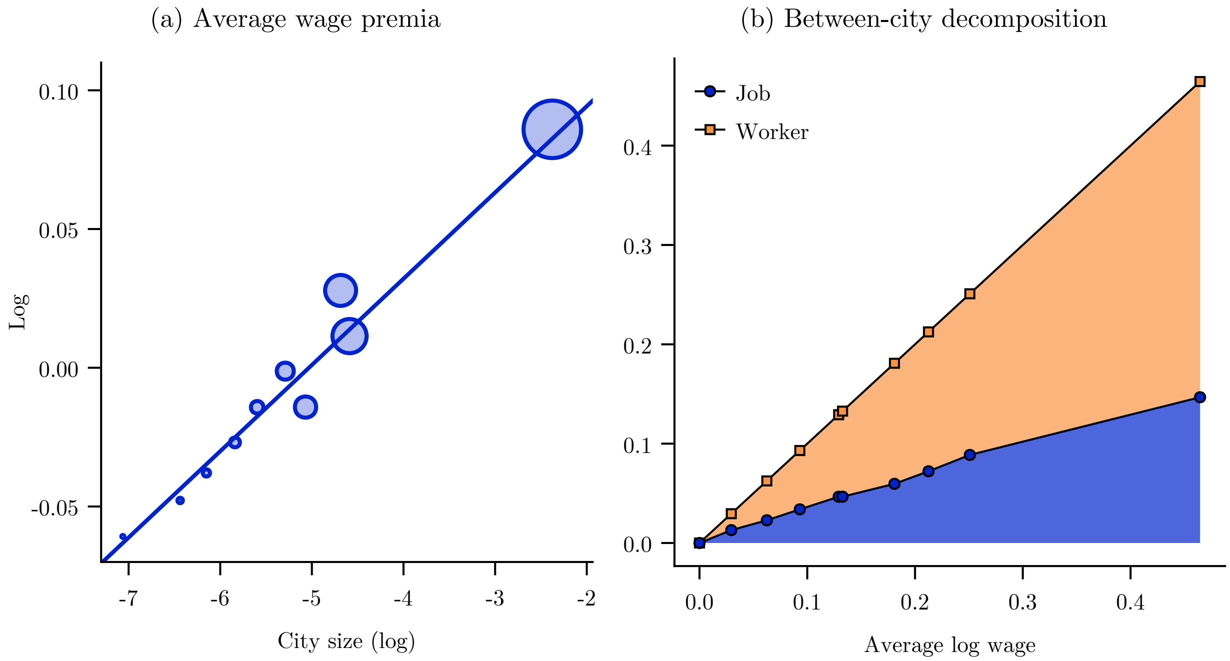
Workers who move between jobs identify the job fixed effects if the mobility of workers is random conditionally on these workers and their current jobs, $\mathbb{E}[\varepsilon_{it}^{\text{AKM}} \mid i, \{\ell(i,k), j(i,k), x_{ik}\}_k] = 0$ (Card et al., 2013). In this case, the fixed effects measure job wage premia —the additional wage a worker earns when working at a particular job.[3] Across space, the levels of the job fixed effects are identified from workers switching jobs across locations.

To limit well-known econometric difficulties linked to limited mobility bias, I follow Bonhomme et al. (2019) and group workers and jobs in 100 equally populated clusters based on their unconditional

[2] There does not exist a publicly available price index at the commuting zone level in France. See Section 3.2 for details on the housing data.

[3] The conditional random mobility assumption is trivially satisfied in the model I develop in Section 2. In addition, AKM models are usually estimated at the firm level. The identifying assumption behind (1) is thus relatively weaker as it allows for the mobility patterns to vary across establishments *and* occupations.

Figure 2: The role of employers in between-city wage inequality



Left panel displays the average job fixed effect per city group against the average city size in each group. Right panel displays the between-city wage decomposition (2) normalizing the average wage in the smallest location to zero. The blue area plots the contribution of jobs, $\mathbb{E}_\ell[\gamma_{j(i,t)}]$. The orange area plots the contribution of workers, $\mathbb{E}_\ell[\psi_{it}]$. In both panels, the fixed effects are obtained from (1).

mean wage.[4] I then estimate equation (1) by OLS at the group level.

I use the estimated job fixed effects to quantify the role of employers on between- and within-city wage inequality.

Starting with the between-city component, I find that larger locations offer higher wage premia on average. Figure 2a plots the average job fixed effect within each city group against the average city size in each group. I estimate that workers in cities twice as large earn an average premium 3.1% higher.

The higher premia in large cities account for a sizeable fraction of why wages are higher there. Using (1), I break down the average wage in each location into components due to worker and job heterogeneity:

$$\mathbb{E}_\ell[\log w_{it}] = \underbrace{\mathbb{E}_\ell[\psi_{it}^{\text{AKM}}]}_{\text{Worker}} + \underbrace{\mathbb{E}_\ell[\gamma_{j(i,t)}^{\text{AKM}}]}_{\text{Job}}. \tag{2}$$

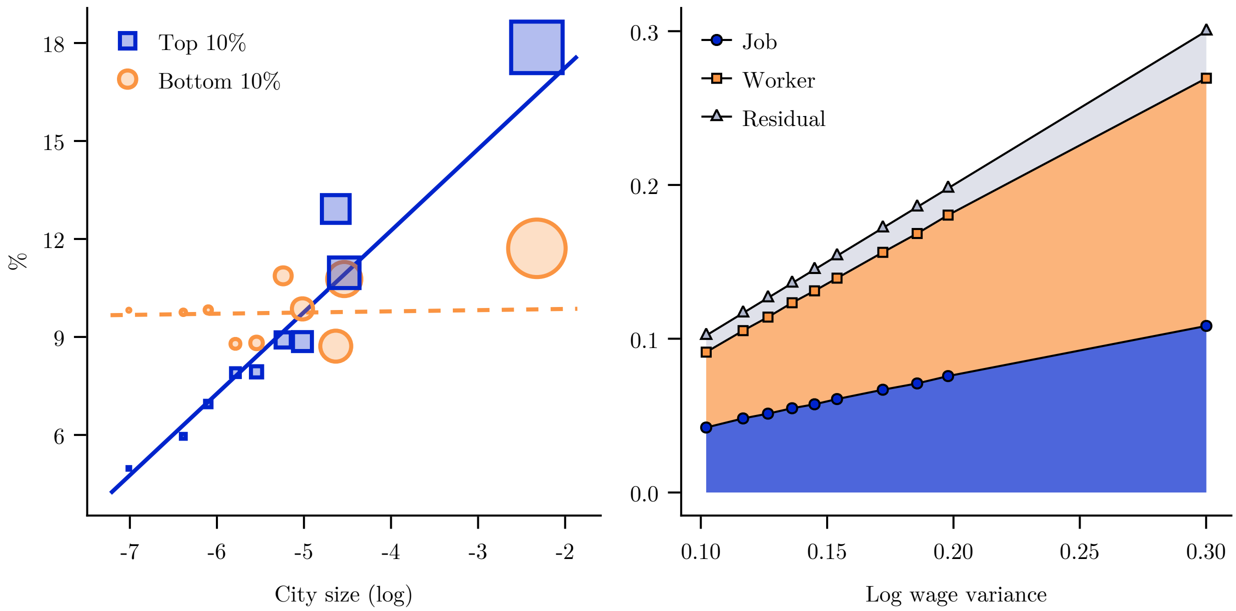
Figure 2b implements this decomposition. Each marker represents a city group. The blue and orange areas represent the average job and worker fixed effects in each city cluster, projected against the local average wage. Each component is expressed relative to the smallest location. I find that

[4] The grouping for jobs is done within city groups to preserve the local nature of an establishment. The results are virtually identical when varying the number of clusters between 10 and 200.

Figure 3: The role of employers in within-city wage inequality

(a) Job shares of low- and high-paying jobs

(b) Within-city decomposition



Left panel displays the employment share of low-paying (orange circles) and high-paying (blue rectangles) jobs in each city cluster. Low-paying (high-paying) jobs are defined as jobs with a fixed effect in the bottom 10% (top 10%) of the national job fixed effect distribution. Right panel displays the within-city wage decomposition (3). The blue area plots the contribution of jobs, $\mathbb{V}\text{ar}_\ell[\gamma_{it}] + \mathbb{C}\text{ov}_\ell[\psi_{it}, \gamma_{j(i,t)}]$. The orange area plots the contribution of workers, $\mathbb{V}\text{ar}_\ell[\psi_{it}] + \mathbb{C}\text{ov}_\ell[\psi_{it}, \gamma_{j(i,t)}]$. The grey area plots the residuals, $\mathbb{V}\text{ar}_\ell[\varepsilon_{it}]$. In both panels, the fixed effects are obtained from (1).

the gap in average wage premium across CZs explains 33.6% of the between-city wage variance.

The patterns and magnitudes documented in Figure 2 are similar to Combes et al. (2008) in France and Card et al. (2025) in the United States, and more generally, aligns with benchmark empirical estimates of the benefits of agglomeration for workers (e.g., Combes and Gobillon, 2015; Duranton and Puga, 2020). However, in contrast to standard estimates, the wage benefits offered by large cities are here explicitly tied to job premia. To the extent that there is more than one job per city, the spatial allocation of job premia then also matters for within-city inequality.

Figure 3a displays the spatial distribution of low- and high-paying jobs. Specifically, I identify in the data jobs whose fixed effect is in the bottom (low-paying) and top (high-paying) 10% of the unweighted national job fixed effect distribution. I then compute the fraction of jobs in each city that are either low- or high-paying. Figure A.4 extends the analysis to all the deciles of the job fixed effect distribution, and Figure A.3 repeats the exercise weighting the statistics by employment.[5]

The spatial allocation of jobs is strikingly different for low- and high-paying jobs. While low-paying jobs are uniformly allocated across space, high-paying jobs are disproportionally present in

[5]Job shares are the relevant statistics to study the spatial allocation of jobs. However, wage statistics rely on jobs' employment shares. The two statistics are virtually identical for low-paying jobs. High-paying jobs hire relatively more (fewer) workers in larger (smaller) cities, thus amplifying their importance for between- and within-city wage inequality.

the largest cities. For instance, 17.9% of jobs in Paris are high-paying, whereas this is 6.9% of them in Lens.

The spatial concentration of high-paying jobs rationalizes why large cities offer higher wage premia on average.[6] At the same time, it also implies greater within-city inequality as low-paying jobs are present throughout France. Using (1), I decompose the within-city wage variance into components due to worker heterogeneity, pay premia dispersion, and the residuals:

$$\begin{aligned}\mathbb{V}\mathrm{ar}_\ell[\log w_{it}] = \underbrace{\mathbb{V}\mathrm{ar}_\ell[\psi_{it}^{\mathrm{AKM}}] + \mathbb{C}\mathrm{ov}_\ell[\psi_{it}^{\mathrm{AKM}}, \gamma_{j(i,t)}^{\mathrm{AKM}}]}_{\text{Worker}} + \\ \underbrace{\mathbb{V}\mathrm{ar}_\ell[\gamma_{it}^{\mathrm{AKM}}] + \mathbb{C}\mathrm{ov}_\ell[\psi_{it}^{\mathrm{AKM}}, \gamma_{j(i,t)}^{\mathrm{AKM}}]}_{\text{Jobs}} + \underbrace{\mathbb{V}\mathrm{ar}_\ell[\varepsilon_{it}]}_{\text{Residuals}}.\end{aligned} \tag{3}$$

Figure 3b displays the variance decomposition. 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.[7]

Figure 3b shows that the spatial distribution of jobs is key to understand why larger cities are more unequal. For instance, the dispersion in pay premia in Paris is 0.06 larger than in Lens, or 35% of the difference in wage variance. Generalizing to all cities, the job fixed effects explain 33% of the spatial variation in within-city wage inequality. Meanwhile, worker heterogeneity and the residuals account for 57% and 10%.

Figure 2 and Figure 3 thus document the importance of jobs in shaping between- and within-city inequality. Appendix A.4 assesses the robustness of these findings through two exercises.

First, I quantify separately the role of establishments, occupations, and industries. The baseline specification (1) is implemented at the job (establishment $\times$ occupation) level. 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. I separate the role of occupations and industries from that of establishments by projecting the job fixed effects onto occupation, industry and establishment fixed effects:

$$\psi_j^{\mathrm{AKM}} = \delta_{o(j)}^{\mathrm{AKM}} + \phi_{\iota(j)}^{\mathrm{AKM}} + \psi_{f(j)}^{\mathrm{AKM}} + \nu_j. \tag{4}$$

In the above, $o(j)$, $\iota(j)$, and $f(j)$ refer to the occupation, industry and establishment of job j. I estimate (4) by weighted OLS where the weights are the employment shares of each job.

Figure A.6 uses (4) to decompose the relative importance of establishment, industry, and

[6] Figure A.5 plots counterfactual average job fixed effects when only the bottom x^{th} percentile of the national fixed effect distribution are included. Figure A.5 shows that the concentration of the top 20% highest paying job in the largest cities account for 73.2% of the between-city dispersion in average wage premium.

[7] The level of the worker and job fixed effects are not separately identified, which is why the between-city decomposition in Figure 2 is normalized to the smallest location. By contrast, the level of the variances in (3) is identified.

occupation on between- and within-city wage inequality. Industries explain very little of it. If anything, the industries present in the largest locations pay on average slightly lower wages after accounting for occupation and establishment heterogeneity. The small role played by industries suggest that compensating differentials are not the core force driving Figure 2 and Figure 3.

Occupations do matter for the spatial distribution of job wage premia, but relatively less than establishments. High-paying occupations are relatively more present in larger cities, which explains 44.2% of the between-city wage premia variance. At the same time, some low-paying occupations are present throughout France, and occupations account for 26.1% of the spatial differences in within-city wage premia dispersion. The establishments fixed effects account for the remainder: 58.1% of the between-city dispersion and 73.9% of the within-city dispersion in job fixed effects. As such, employers still explain a sizeable portion of spatial wage inequality even if wage gaps across occupations fully capture productivity differentials.

The second robustness exercise shows that the sorting of high-wage workers into high-wage jobs does not explain why larger cities are more unequal. The within-city variance decomposition (3) bundles the job fixed effect variance with the covariance between the worker and job fixed effects. I do so because, everything else equal, the covariance is an increasing function of job heterogeneity. In Section A.2, I provide a first-order approximation of the within-city wage variance that isolates the impact of sorting as measured by the *correlation* between worker and job fixed effects. Figure A.7 implements this approximation. I find the correlation between the job and worker fixed effects to be stable across CZs. As such, while this sorting component explains a non-trivial amount of the aggregate wage variance, it does not rationalize the greater wage inequality in larger cities.

Workers in large cities thus earn higher wages because they have access to higher-paying jobs. However, large cities do not benefit all workers equally as low-paying jobs are present throughout France. Greater within-city inequality follows. I now show that the wage gains offered by large cities occur over time as workers reallocate from low- to high-paying jobs.

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

To motivate the analysis, Figure 4a presents cross-sectional evidence on how the reallocation of workers across jobs interact with their wage profile. 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 jobs held by worker i at time t is defined as the cumulative sum of job switches between age 25 and t:

$$\#\mathrm{J}_{it} = \sum_{\tau=\underline{t}_i}^{t} \mathrm{J2J}_{i\tau}. \tag{5}$$

A worker is said to switch jobs ($\text{J2J}_{it} = 1$) if they transition between pairs of establishments $\times$ occupations within 90 days. In (5), $\underline{t}_i$ denotes the year in which worker i was 25.[8]

Figure 4a documents a rising wage gap between Paris and Lens as workers move along the job ladder. While this gap is 28.9% for workers who enter the labor force, it rises to 88.7% after 9 job switches.

Three distinct forces may shape the wage profiles shown in Figure 4a. 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, the job ladder may be steeper in Paris.

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

$$\log w_{it} = \alpha_i^{\text{JL}} + \beta_{\ell(i,t)}^{\text{JL}} x_{it} + \delta_{\ell(i,t)}^{\text{JL}} + \lambda_{\ell(i,t)}^{\text{JL}} \#\text{J}_{it} + \varepsilon_{it}^{\text{JL}}. \tag{6}$$

As in (1), α_i^{JL} 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^{JL} captures that some cities may favor human capital accumulation. The parameters $\delta_{\ell(i,t)}^{\text{JL}}$ are location fixed effects. Finally, $\#\text{J}_{it}$ is the cumulative sum of job switches as defined in (5). To ensure that I am not misclassifying transitions through unemployment as job switches, I also estimate a version of (6) where I restrict $\#\text{J}_{it}$ to job transitions associated with wage gains.

There are two coefficients of interest in (6). First, the local returns to job switching as estimated by λ_ℓ^{JL}. Given the worker fixed effects and the city-specific returns to experience, λ_ℓ^{JL} 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 jobs conditional on workers' unobserved characteristics and experience.[9] If $\lambda_\ell^{\text{JL}} > 0$, job switching spurs faster wage growth. If $\lambda_{\ell'}^{\text{JL}} > \lambda_\ell^{\text{JL}}$, this extra wage growth is greater in ℓ' than ℓ, suggesting a steeper local job ladder.[10]

The second parameters of interest are the location fixed effects δ_ℓ^{JL}. These fixed effects measure the average wage in location ℓ, net of worker heterogeneity and before any job switch or experience. They therefore quantify the extra wage earned in location ℓ for seemingly identical workers upon entry in the labor market —the local startup premium.

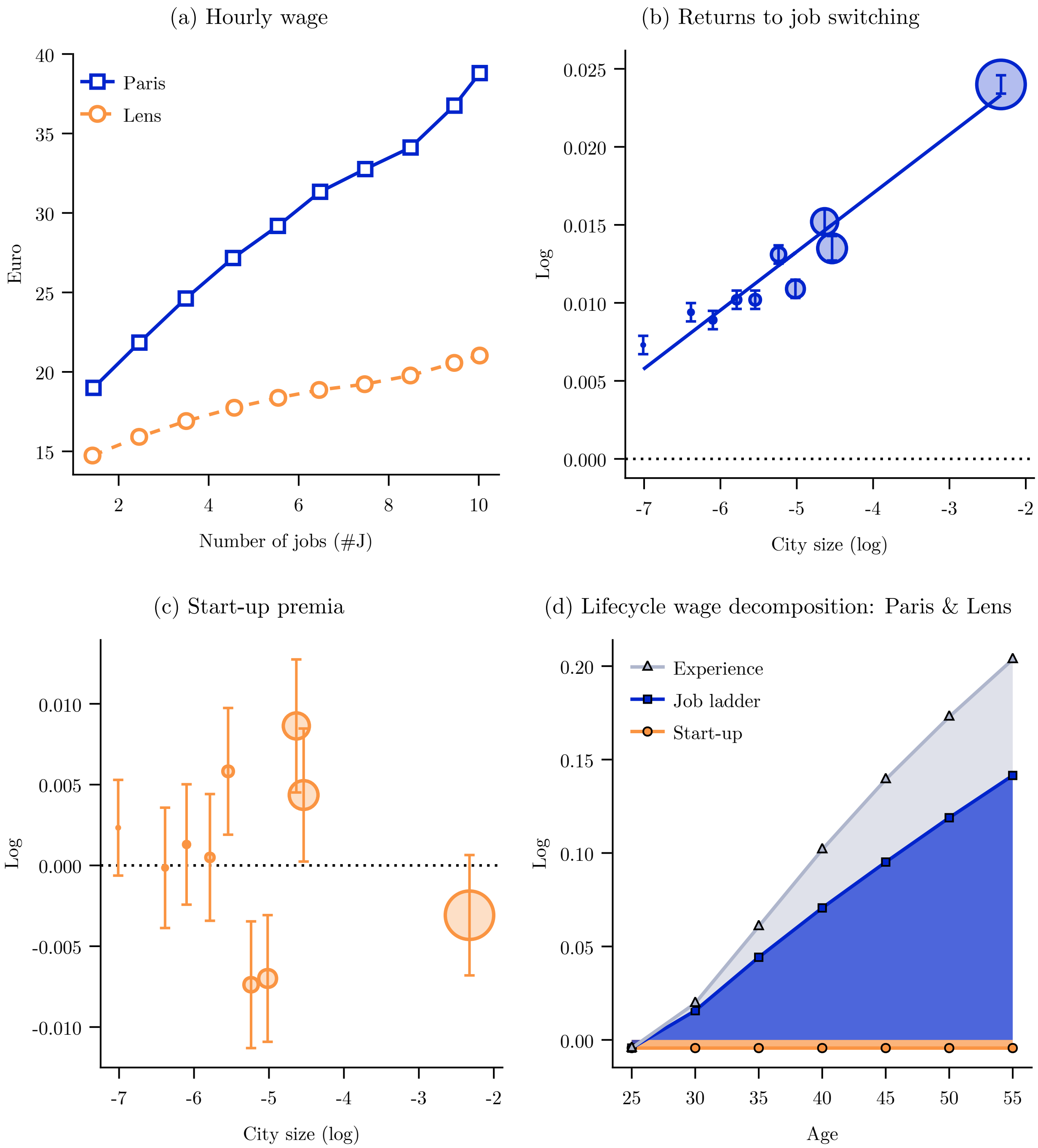
Together, these two sets of parameters provide a complete picture of how jobs shape wages throughout space and time. Workers may earn high wages in large cities thanks to a high start-up

[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 $\text{J2J}_{it} = \alpha_t + \beta_{a_i} + u_{it}$, where a_i is the age of i. I then use the estimated age profile to infer workers' past number of switches: $\#\text{J}_{it_i} = \sum_{\tau=\underline{t}_i}^{t_i} \text{J2J}_{i\tau} = \sum_{a \leq a_{it_i}} \beta_a$, where t_i is the first year in which I observe i.

[9]The identification assumption is thus stricter than for the AKM model (1). However, (6) does not suffer from the limited mobility bias, which allows for further heterogeneity analysis. The two approach are therefore complementarity. Figure A.11 re-estimate (6) replacing the left-hand side variable with the job fixed effects of Section 1.3. The between-city differences in returns to job switching are very similar when estimated on wages or job fixed effects.

[10]The local returns to experience are in turn identified from the wage growth of job stayers. Consistent with the prior literature, I find that they are higher in larger cities (Figure A.8).

Figure 4: The local returns to job switching



Panel (a) plots the average hourly wage by number of jobs held, in blue for Paris and in orange for Lens. The number of jobs held are defined in (5). Panel (b) plots the point estimates for the local return to job switching, $\{\lambda_\ell^{JL}\}_{\ell=1}^{L}$. Panel (c) plots the location fixed effects, $\{\delta_\ell^{JL}\}_{\ell=1}^{L}$, normalized to zero. Both sets of estimates are plotted against the average city size in each city cluster. The vertical bars in (c) and (d) represent 95th confidence intervals with standard errors clustered at the individual level. Panel (d) decomposes the lifecycle wage profile between Paris and Lens according to (7). The orange area represents the difference in location fixed effects, $\Delta_\ell \delta_\ell^{\mathrm{JL}}$. The blue area represents the contribution of job switching, $\Delta_\ell \lambda_\ell^{\mathrm{JL}} \mathbb{E}_\ell[\#J_{it} \mid a]$. The grey area represents the contribution of experience, $\Delta_\ell \beta_\ell^{\mathrm{JL}} \mathbb{E}_\ell[x_{it} \mid a]$. The point estimates used in (b), (c), and (d) are obtained form (6).

premium. Alternatively, they may access those high-wage opportunities over time as they reallocate across jobs.

Figure 4b presents the estimated returns to job switching by city cluster. Figure 4c displays the local startup premia. Both estimates are plotted against the average city size in each city cluster.

I find that job switching fastens wage growth throughout France, and even more so in larger places. On average, a job switch leads to an extra 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 0.3 p.p. upon switching jobs in cities twice larger. The spatial differences in job ladder steepness are even greater when focusing on job switches associated with wage gains (0.5 p.p.) as shown in Figure A.11b.

By contrast, I estimate a negligible startup premium in large cities. For instance, entry-level wages in Paris are 0.4% lower than in Lens after controlling for worker heterogeneity and the dynamic effects of cities. Doubling the size of a city in the cross-section is associated with a 0.1% decrease in the startup premium, with a p-value of 0.52.

Table A.2 in Appendix A.3 provides several robustness exercises. For expositional clarity, I replace the non-parametric local returns to job switches in (6) 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. Column (4) and (5) address the concerns that human capital accumulation may not be fully captured by experience; then, the wage growth of movers may confound the gains from reallocation with learning. Specifically, column (4) includes occupation fixed effects, thereby identifying returns to job switching from moves between establishments within occupations. Column (5) introduces worker-specific experience slope to account for heterogeneity in learning abilities. In all cases, job switching brings faster wage growth in larger cities, and the estimates are very stable throughout.

Do the gains from reallocation occur only to a subset of workers? Figure A.9 and A.10 address this question by conducting heterogeneity analysis. Figure A.9 estimates the local returns to job switching by occupation. I find that the job ladder of large cities is steeper for every occupation, including for blue-collar workers. Figure A.10 reports estimates by wage quartile. Here as well, I find that workers at every rank of the wage distribution experience larger wage gains when they switch job in bigger cities. The steeper job ladder of larger cities has sizable consequences on lifetime earnings. According to equation (6), the lifecycle wage profile in location ℓ net of worker heterogeneity is

$$\mathbb{E}_\ell[\log w_{it} - \alpha_i^{\mathrm{JL}} \mid a] = \delta_\ell^{\mathrm{JL}} + \beta_\ell^{\mathrm{JL}} \mathbb{E}_\ell[x_{it} \mid a] + \lambda_\ell^{\mathrm{JL}} \mathbb{E}_\ell\left[\#\mathrm{J}_{it} \mid a\right], \tag{7}$$

where a is a particular age.

Figure 4d uses (7) to decompose the difference in lifecycle wage profile between Paris and Lens.

Given the small estimates of the startup premium, 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 14.1% greater than in Lens. The higher returns to experience further boosts the wage gap by 6.2 p.p.

Figure A.12 generalizes this decomposition to all locations. Over time, the steeper ladder in larger cities explains 69.8% of the between-city wage growth variance (Figure A.12b).[11] The remainder is due to experience. Averaging across the lifecycle, the reallocation of workers to better-paying jobs explain the majority (73%) of the between-city wage variance net of worker heterogeneity (Figure A.12a). By contrast, the startup premia accounts for none.

1.5 Summary of testable implications

Sections 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 careers at these low-paying jobs, and large cities boost wages as workers reallocate across employers. Together, employers generate higher wages and greater 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 wage dispersion. Second, to the extent that (1) and (6) 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 thus build a new theory of spatial wage disparities that jointly account for the two facts through the sorting of employers across frictional labor markets.

2 A spatial theory of wage premia

The framework 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 extends the model quantitatively.

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

[11] As made clear by (7), the job ladder of large cities may boost wages by generating greater wage gains upon a switch ($\lambda_\ell^{\mathrm{JL}}$) as well as more frequent switches ($\mathbb{E}_\ell[\# J_{it}]$). Figure A.12b separates both channels. I find that the greater gains upon a switch drive most (88.2%) of the impact of local ladders on wage growth.

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: $A_1 < A_2 < \cdots < A_L$. I later allow for TFP differentials. 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 device: 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 revenue from home production. They receive job offers at Poisson rate λ^u. When employed, workers supply inelastically one unit of labor, so their 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 $\delta < \lambda^e \leq \lambda^u$.[12]

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.[13] 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 them experience longer non-employment spells (Table A.3).

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 λ. 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

[12] Figure C.17 shows that job switching rates are fairly constant across cities once controlling for worker heterogeneity.
[13] For a quantitative spatial model with wage posting and between-city search, see Heise and Porzio (2022).

employers and jobs interchangeably.[14] I assume that $\underline{z}$ is high enough relative to b so that all jobs are 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 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, $d\Gamma_\ell$.[15] 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:

$$\sum_{\ell=1}^{L} M_\ell d\Gamma_\ell(z) \leq M d\Gamma(z). \tag{8}$$

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 in three steps. First, I derive the spatial allocation of workers and the local labor supply curves. I then solve for the local wage distributions given the spatial allocation of employers. Finally, I characterize the spatial allocation of employers.

2.2 Job search

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

$$\rho U_\ell = \frac{A_\ell b}{P_\ell} + \lambda^u \int \max\{V_\ell(w) - U_\ell, 0\} dF_\ell(w), \tag{9}$$

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

[14] The framework does not model explicitly occupation. Rather, the productivity z captures the total productivity of the pair occupation $\times$ establishment.

[15] 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.

utility:

$$\bar{U} = \max_{\ell} U_\ell. \tag{10}$$

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 worker employed at wage w in location ℓ satisfies

$$\rho V_\ell(w) = \frac{A_\ell w}{P_\ell} + \lambda^e \int \max\{V_\ell(w') - V_\ell(w), 0\} dF_\ell(w') + \delta[U - V_\ell(w)], \tag{11}$$

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 ladder by continuously accepting better-paying job offers. The flows of workers up the job ladder determine 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[16]

$$G_\ell(w) = \frac{F_\ell(w)}{1 + k(1 - F_\ell(w))}. \tag{12}$$

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

$$n_\ell(w) = \frac{1}{\theta_\ell} \frac{1 + k}{\left[1 + k(1 - F_\ell(w))\right]^2}, \tag{13}$$

where $\theta_\ell \equiv M_\ell/e_\ell$ is the labor market tightness in city ℓ. Equation (13) 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.

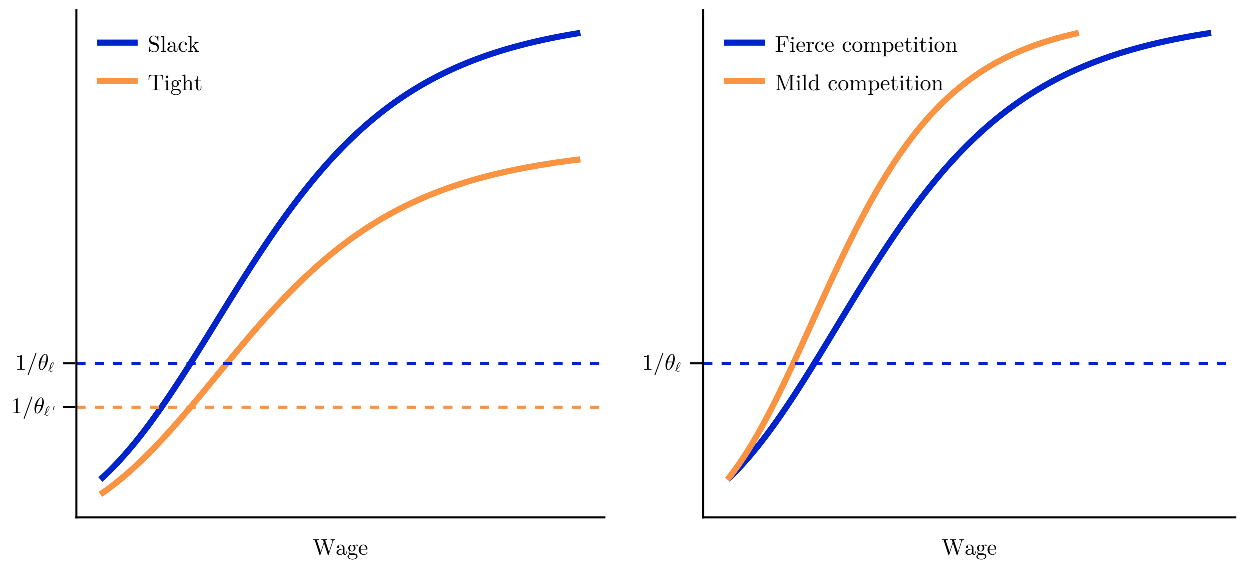
The labor supply curves differ across locations in two ways as shown in Figure 5. First, the yield of a vacancy is lower in tighter labor markets (Figure 5a). Second, the competition for workers —as represented by the job offer distribution— may be fiercer. In particular, in locations with a higher job offer distribution in the first order stochastic dominance (FOSD) sense, the labor supply curve shifts downward as workers transition more often to better paying opportunities (Figure 5b). Search frictions, as summarized by k, determine 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, as depicted by the dashed lines in Figure 5.

Within a city, unemployed workers tradeoff a higher search efficiency against lower present

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

Figure 5: Local labor supply curves

(a) The impact of market tightness (b) The impact of market competition



earnings. They accept any wage offer greater than the city-specific reservation wage $\underline{w}_\ell$ given by

$$\underline{w}_\ell = b + (\lambda^u - \lambda^e) \int_{\underline{w}_\ell}^{\infty} \frac{1 - F_\ell(w)}{\rho + \delta + \lambda^e(1 - F_\ell(w))} dw. \tag{14}$$

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 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

$$U_\ell = \frac{A_\ell \mathcal{W}_\ell}{P_\ell}, \tag{15}$$

where $\mathcal{W}_\ell$ is the net present value of expected future income. When workers are infinitely patient, this 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 (15) are standard in static spatial frameworks à la Rosen (1979) and Roback (1982) 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 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, $\frac{A_\ell \underline{w}_\ell}{P_\ell} < \frac{A_{\ell'} \underline{w}_{\ell'}}{P_{\ell'}}$.

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's problem in steady state assuming a low discount rate ($\lambda \to 0$). Employers choose where to produce, how many workers to hire, and which wage to offer, to maximize their flow profits:

$$\pi(z) = \max_\ell \pi_\ell(z) = \max_\ell \left\{ \max_{w,n} \; R(z,n) - wn - r_\ell \quad \text{s.t.} \quad n \leq n_\ell(w) \right\}. \tag{16}$$

Employers' hiring decisions are 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 (16) consists of 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 employers. Second, employers factor in local prices when choosing their production location, themselves a function of other employers' location choice. I thus solve (16) 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, $\{M_\ell, d\Gamma_\ell\}_{\ell=1}^L$. Employers are on their labor supply curve, $n = n_\ell(w)$. The wage offer distributions are continuous over the interval $[\underline{w}_\ell, \bar{w}_\ell]$.[17] 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 that

[17] 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.

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 $z \in \operatorname{supp}\Gamma_\ell$ is

$$n_\ell(z) \equiv n_\ell[w_\ell(z)] = \frac{1}{\theta_\ell} \frac{1+k}{[1+k(1-\Gamma_\ell(z))]^2}. \tag{17}$$

Employer size directly follows from the labor supply curves (13) and the rank-preserving condition $F_\ell[w_\ell(z)] = \Gamma_\ell(z)$. Second, the wage offered by an employer with productivity $z \in \operatorname{supp}\Gamma_\ell$ is

$$w_\ell(z) = \underbrace{\underline{w}_\ell}_{\substack{\text{Outside}\\ \text{option}}} \times \underbrace{\left(\frac{n_\ell(\underline{z}_\ell)}{n_\ell(z)}\right)}_{\substack{\text{Emp. share}\\ \text{from unemp.}}} + \underbrace{\left(1 - \frac{n_\ell(\underline{z}_\ell)}{n_\ell(z)}\right)}_{\substack{\text{Emp. share}\\ \text{from poaching}}} \times \underbrace{\mathbb{E}^{n'_\ell}[Z \mid Z \leq z]}_{\substack{\text{Competitors'}\\ \text{avg. productivity}}}, \tag{18}$$

where $\underline{z}_\ell \equiv \min\Gamma_\ell$ is the least productive employer in ℓ, and the expectation is taken with respect to the measure $dn_\ell(z)$.

Employers' monopsony power implies that the wage they offer depends on the competition they face. How fierce this competition is depends on two factors: how many workers an employer poaches from their competitors, and how willing these competitors are to retain their workers. The location where employers produce shapes the second dimension, but the effects are heterogeneous across employers depending on their relative productivity.

The least productive employers in a city indeed only hire workers from unemployment as they cannot compete with the other firms. The wage they offer thus reflects unemployed workers' outside option, $\underline{w}_\ell$. In the limit where $\lambda^e \to \lambda^u$, this outside option coincides with the unemployment insurance, b. Workers at the bottom of the local ladders thus earn similar wages, no matter in which city they are and who hires them. Said differently, low-paying jobs are present in every location.

More productive employers offer higher wages to poach workers from their lower-paying competitors. This poaching game builds up throughout the job ladder: the higher the relative productivity of an employer, the more competition it faces, and therefore the more the wage it offers depends on the local productivity distribution. When all employers are relatively more productive, the poaching competition intensifies. The wage distribution shifts up, and along with it the average wage. However, these wage gains are concentrated at the top of the ladder where employers face off all the local competition.

Combined, the local concentration of productive employers steepens the job ladder. Workers experience faster wage growth when switching job as they have access to a relatively greater number of high-paying jobs. As a result, they are willing to accept lower real earnings when unemployed in anticipation of higher future real earnings.

Proposition 2 formally characterizes local wages as a function of the productivity distributions.

Proposition 2 (Spatial wage inequality)**.**
Consider two locations, ℓ and ℓ', such that the employer productivity distribution in ℓ is higher in the FOSD sense, $\Gamma_\ell \succ_{FOSD} \Gamma_{\ell'}$. 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 job switch are larger, $\mathbb{E}_\ell[\frac{W}{w} \mid W > w, w] > \mathbb{E}_{\ell'}[\frac{W}{w} \mid W > w, w]$, and the real reservation wage is lower.*
2. *The top-to-bottom wage gap is larger in ℓ, $\bar{w}_\ell - \underline{w}_\ell > \bar{w}_{\ell'} - \underline{w}_{\ell'}$, and $\frac{\bar{w}_\ell}{\underline{w}_\ell} > \frac{\bar{w}_{\ell'}}{\underline{w}_{\ell'}}$ if $\lambda^u - \lambda^e \geq 0$ is not too large;*

Proposition 2 replicates the two facts that called for a new theory of spatial wage inequality. It 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 is

$$\pi_\ell(z) = [z - w(z, \Gamma_\ell)]\, n[\Gamma_\ell(z), \theta_\ell] - r_\ell. \tag{19}$$

where, with a slight abuse of notation, I have made explicit the dependence of the optimal wage and size on the spatial allocation of jobs. Employers produce in the location(s) that maximize their profits:

$$\pi_\ell(z) \geq \pi_{\ell'}(z) \text{ for all } \ell' \neq \ell \iff z \in \text{supp } \Gamma_\ell. \tag{20}$$

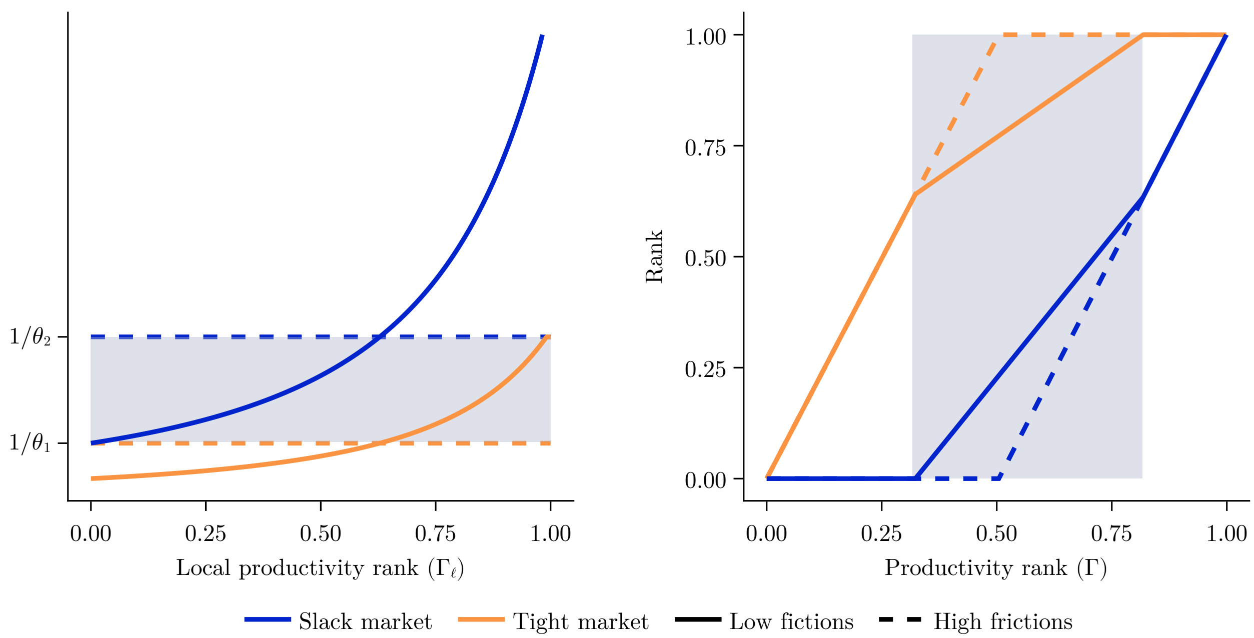
An equilibrium spatial allocation of employers is a tuple $\{M_\ell, d\Gamma_\ell\}_{\ell=1}^L$ that satisfies feasibility (8) and profit maximization (20).

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 (20) (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 the 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 few workers as they slide down the job ladder. The poaching competition thus introduces substitutability across locations. Weak search frictions increase this tradeoff by accelerating

Figure 6: The spatial allocation of employers

(a) Labor supply curves (b) Local productivity distributions



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,* $d\Gamma(\mathit{supp}\,\Gamma_\ell \cap \mathit{supp}\,\Gamma_{\ell'}) > 0$*.*

Figure 6(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. 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.[18]

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

[18]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.

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 distributions are ordered in the first order stochastic dominance sense.

In equilibrium, larger cities have slacker labor markets. Large cities are relatively attractive, and a greater number of employers locate there. As a result, commercial housing prices rise, and not every employer produces in the largest city. In particular, this congestion force implies that the labor market in larger cities remains slacker despite having more jobs.[19] Proposition 4 states this result formally.[20]

Proposition 4 (Labor market slackness)**.**
Suppose that $\lambda^e \approx \lambda^u$. 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 6(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 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.

Propositions 1 to 4 thus show that the sorting of employers across space generates between- and within-city inequality. Productive employers concentrate in large, slack cities to sidestep search frictions. The local labor market competition intensifies. 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.

Before confronting the quantitative prediction of the model with the data, it is worth emphasizing the unique feature of this model: spatial wage inequality arises without local TFP gaps. Search frictions are crucial to arrive at this conclusion.[21] To make this point, I write down in Appendix B.8

[19] Figure C.18 confirms the prediction of Proposition 4 in the data. Labor market consistency implies that the tightness is the inverse of employers' average size, $M_\ell \mathbb{E}_\ell[n_\ell(w)] = e_\ell$. Figure C.18a shows that the average number of workers per employer is larger in bigger cities.

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

[21] This prediction would generally hold in any framework with local upward slopping labor supply curves. Search

a general model of employer sorting with *competitive* labor markets. In this environment, firms sort across space to access high revenue TFP and cheap inputs. If TFPs are homogeneous across locations, productive employers concentrate in large cities only if wages are relatively lower. Said differently, high local TFP in large cities is necessary to generate higher wages —even in the presence of firm sorting.[22] Search frictions break this necessity result because employers sort to access slacker markets, and they are willing to pay higher wages for that.

Propositions 1 to 4 do not contain quantitative predictions; I therefore turn to estimating the model to assess how well it aligns with the data, and through this, re-assess the drivers of spatial wage disparities.

3 Extended model and estimation

Extensions I relax the strong assumptions imposed until now. Across space, I allow cities to differ in terms of TFP, T_ℓ. Local TFP complements employer's productivity: the MPL of a worker hired by employer z in location ℓ is zT_ℓ. I assume that local TFPs are exogenous.[23] 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:

$$n_\ell(w, v) = \frac{(1+k_\ell)e_\ell}{[1+k_\ell(1-F_\ell(w))]^2}\frac{v}{V_\ell}, \tag{21}$$

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

$$V_\ell = M_\ell \int v_\ell(z) d\Gamma_\ell(z). \tag{22}$$

Search frictions are now city-specific and are determined by the matching function $\lambda_\ell^u = \mathcal{M}(u_\ell +$

frictions provide a microfoundation for the supply curves, and they generate reallocation from low- to high-paying employers over time as in the data.

[22] These local TFPs could be either exogenous or endogenous (e.g., knowledge spillovers, market access, etc.). Regardless, they need to be (i) local and (ii) affects the marginal product of labor.

[23] This assumption turns out to be without loss of generality since that I estimate homogeneous local TFPs.

$\zeta e_\ell, V_\ell)/(u_\ell + \zeta e_\ell)$, where ζ is the relative search intensity of employed workers, $\lambda^e_\ell = \zeta \lambda^u_\ell$.

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

$$\pi(z, \{\varepsilon_\ell\}_{\ell=1}^L) = \max_{\ell,w,n,v} \; R(zT_\ell, n) - wn - c(v) - r_\ell - \varepsilon_\ell \quad \text{s.t.} \quad n \leq n_\ell(w, v). \tag{23}$$

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 ξ, upon which they are replaced by their descendant.[24] 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.[25] 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: $\mathcal{M}_\ell = \mu V_\ell^\psi (u_\ell + \zeta e_\ell)^{1-\psi}$.

3.1 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 20 parameters.

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

[24] Entry and exist of workers is required to have well-defined migration flows.

[25] This migration cost can alternatively be interpreted as losses in flow utility. I assume away heterogeneity in migration costs across locations because the city clusters do not contain geographic information (e.g., distance).

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 costs, and TFPs. Three of these 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. Given the Gumbel parametric assumption, the log-likelihood that an employer j with productivity z_j locates in ℓ is

$$\log\Omega_\ell(z_j) = H(z_j) + \vartheta\pi_\ell(z_j), \tag{24}$$

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

To recover profits on the right-hand side of (24), 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 (18) demands

$$z_j T_{\ell_j} \equiv \zeta_j = w_j + \left(\frac{1+k_{\ell_j}(1-F_j)}{2k_{\ell_j}}\right)\frac{\partial w_{\ell_j}(F_j)}{\partial F}, \tag{25}$$

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 (21). 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 (24) to directly estimate the dispersion in entry costs. In practice, TFP gaps may be large. I address this concern in two ways. First, I include location fixed effects in (24). Second, instead of reading the entry cost dispersion from (24), I target the conditional correlation in the second estimation block allowing for arbitrary TFPs.[26]

Indirect inference The second step of the estimation calibrates jointly 3 of the remaining 5 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 calibrate local TFPs as residuals to match the average wage of each city net of the employer sorting predicted by the model. Second, I calibrate the productivity dispersion σ to match the aggregate wage variance. Third, I estimate (24) in the model in the same way as in the data, and I set the entry cost dispersion to match the empirical conditional correlation.

3.2 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 nominal variables by the aggregate average wage. To abstract from worker heterogeneity, I measure wages by the employer fixed effects estimated in Section 1.3.[27]

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

I obtain residential housing prices from the *Carte des Loyers* (Rental Map). I do not have access to city-level 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.

[26] 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.

[27] 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.

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 (Figure C.17b).

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 5 as in Caliendo et al. (2019). As shown in Proposition 5, these two parameters are relevant only for the estimation of amenities and migration costs.

3.3 Results

Table 1 presents the parameter estimates. As in Section 3.1, I focus the discussion around the parameters that shape employers' location choice and wage strategy.

Search frictions are relatively strong. The aggregate contact rate for unemployed workers is 0.2, and the job destruction rate gross of the exit rate is 0.03. The job switching rate is substantially lower than the UE rate at 3.8%. However, employed workers do not necessarily switch jobs when receiving a job offer. Accounting for rejected offers, I find that the search efficiency of employed workers is only 6.3% lower than that of the unemployed. Employers are able to partially sidestep these frictions by posting more vacancies as I estimate a relatively low vacancy cost elasticity ($\gamma = 2.1$). These parameters are all aligned with benchmark estimates in the literature.

Search frictions are constant across space. I estimate that the job switching rates are homogeneous across cities after accounting for worker sorting (Figure C.17). Meanwhile, Figure C.18b shows that larger cities have a slacker labor market, consistent with Proposition 4. These two patterns are reconciled with the absence of congestion from vacancies in the matching functions ($\psi \approx 0$).[28]

My estimate of the entry cost dispersion lines up with the recent papers that study firms' location choice. I estimate $\vartheta = 1.7$, close to the conditional correlation given by (24). Converting my estimate to a Fréchet elasticity, I obtain 1.8, between the 1.3 of Giroud and Rauh (2019) and the 2.6 of Fajgelbaum et al. (2018).

Given these parameters, I estimate that TFPs are homogeneous across locations. There are large variations in city size, and search frictions are binding. Productive employers therefore have strong incentives to concentrate in large cities to maximize their size. Figure 7a shows that the mean employment-weighted employer productivity in Paris is 18.9% larger than in Lens. Meanwhile,

[28]This estimate departs from standard values in search-and-matching model that often set $\psi = 0.5$ (Petrongolo and Pissarides, 2001). At the same time, several papers in the spatial literature argue that large cities facilitate the matching between workers and firms (e.g., Moretti, 2012; Moretti and Yi, 2024). In this setting, this thicker market externality takes place if $\psi < 0$. My matching function estimate lies between these two literatures.

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	Caliendo et al. (2019)			5.000
		B. Model inversion			
α	Cobb-Douglas preference	Housing exp. share	0.200		0.200
$\boldsymbol{A}$	Amenities	Employment distribution	·		·
b	Unemployment insurance	Replacement rate	0.600		0.681
κ	Migration cost	Migration rate	0.454		0.220
δ	Job destruction rate	EU rate	0.021		0.021
ζ	Rel. search efficiency	EE / UE rate	0.193		0.937
ξ	Exit rate	Average career length	120.0		0.008
μ	Matching efficiency	Avg. EE rate	0.038		0.197
ψ	Matching function elasticity	Correlation EE rate - tightness	0.000		0.000
M	Mass employers	Average size	7.193		0.121
γ	Vacancy cost elasticity	Vacancy optimality (51)	2.115		2.115
$\bar{L}$	Residential housing supply level	Avg. residential price	0.478[a]		183.8
θ	Residential housing supply elasticity	Correlation prices - expenditures	0.095		9.546
$\bar{H}$	Commercial housing supply level	Avg. commercial price	1.196		0.165
ϕ	Commercial housing supply elasticity	Correlation prices - employer demand	0.209		4.788
		C. Indirect inference			
σ	Productivity dispersion	Wage variance	0.180	0.179	9.48
ϑ	Entry cost dispersion	Equation (24)	1.617	1.592	1.709
$\boldsymbol{T}$	TFP	Average wage	Fig. 7a	Fig. 7a	Fig. 7a

[a] Average residential price expressed in hundredths.

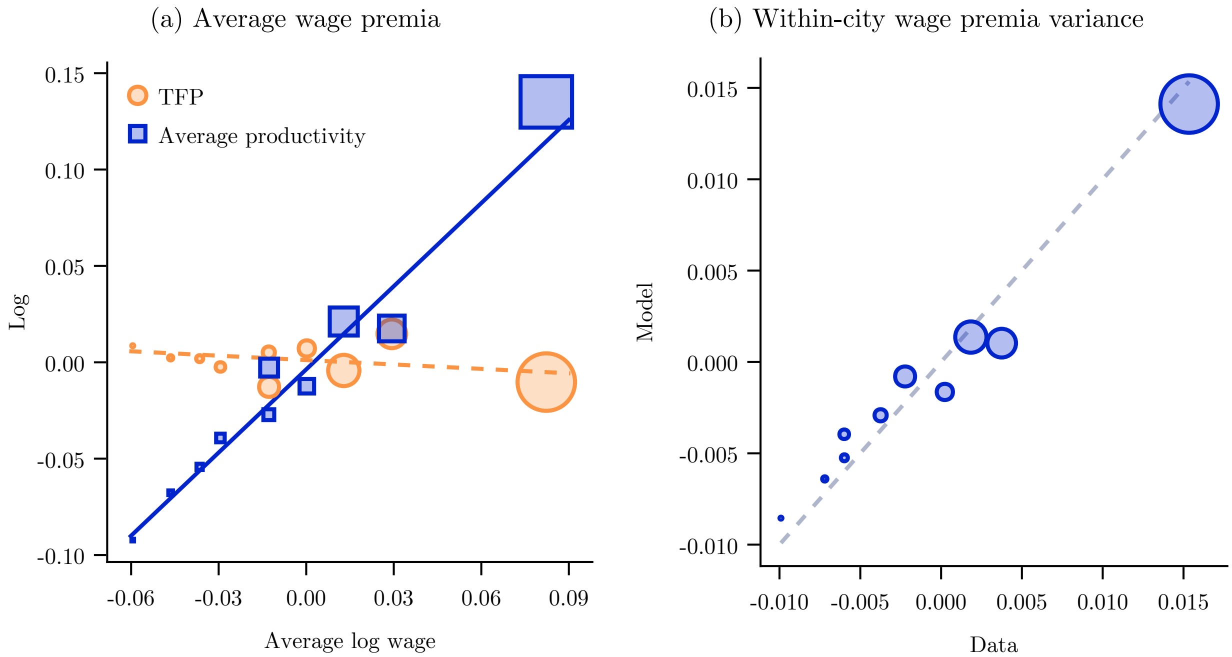
the gap in average wage premia is 11.9%. Once taking into account employers' market power, the sorting of employers across locations thus suffices to explain why the average wage premia is greater in larger places.[29] The between-city TFP variance accounts for 1.5% of the between-city MPL variance, and 3.2% of the between-city wage variance.[30]

The spatial agglomeration of productive employers lies at the root of the between-city wage premia differences. I now assess whether it also generates the facts that motivated this framework.

[29]The difference between employers' productivity and their wage is determined by the labor supply elasticity they face. In this framework, this elasticity is not a structural parameter. Instead, it depends jointly on the search frictions, the vacancy cost elasticity, and the spatial allocation of employers. Altogether, the model produces an average labor supply elasticity ($\mathbb{E}[\partial \log n_\ell(w)/\partial \log w]$) of 5.5 and an average markdown ($\mathbb{E}[w_\ell(z)/(zT_\ell)]$) of 0.74, which align closely with the estimates of Lamadon et al. (2022) for the United States.

[30]The TFPs measured here capture labor productivity since the model is estimated on wage data. It may thus be that large cities amplify firm's productivity, but this is not passed-through onto wages.

Figure 7: Spatial wage inequality in the data and the model



4 Employer sorting and within-city inequality

The estimation strategy proposed in Section 3.1 does not target how inequality or the local returns to job switching vary across space. In addition, the model is substantially over-identified: it only has the aggregate productivity distribution to explain the wage premia distribution of every location. In this section, I thus use the two novel facts of Section 1 as over-identification exercises.

4.1 The consequences of local competition on wages

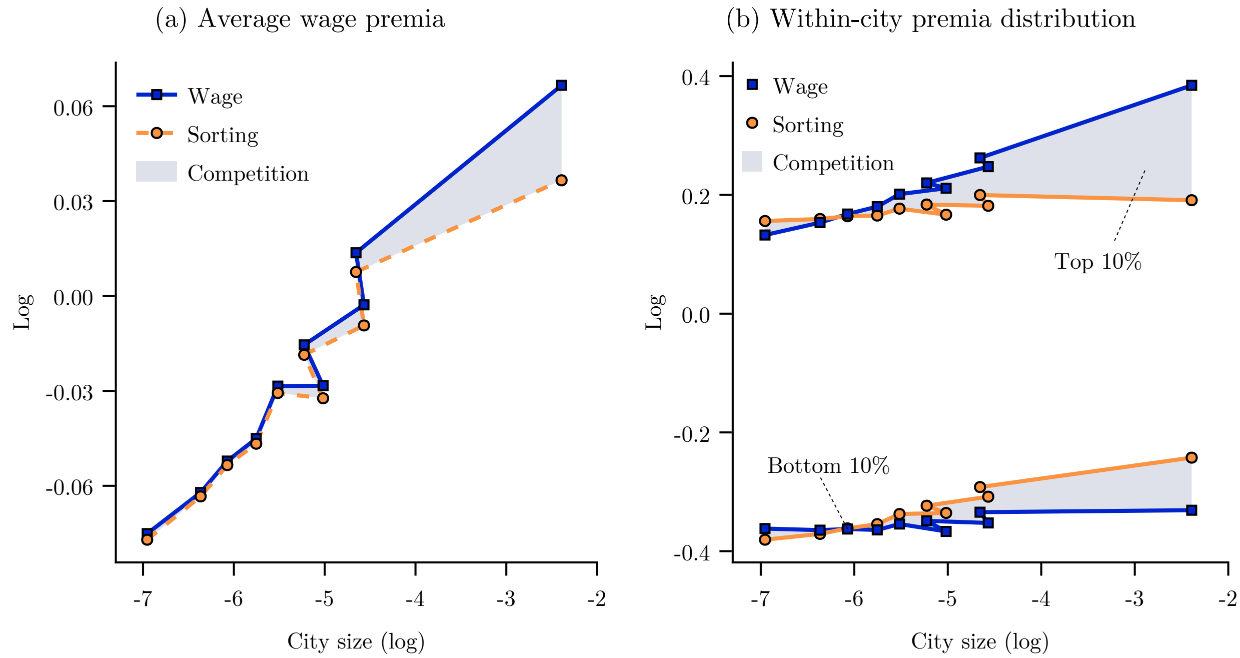
I first start by asking whether larger cities are more unequal in the model. Figure 7b displays the within-city wage premia dispersion, in the data on the x-axis and in the model on the y-axis. The grey dashed line is the 45 degree line.

I find that the concentration of productive employers in large cities does generate greater within-city inequality. The differences in local inequality aligns well with the data.

To understand how different parts of the wage distributions vary across space, I project each city's wage deciles on the size of the city. Figure D.19 plots these decile-specific city size elasticity, in orange in the data and in blue in the model.

As in the data, greater inequality arises in larger cities because high-paying jobs are spatially concentrated whereas low-paying jobs are spatially dispersed. Workers at the bottom of the local wage distribution indeed earn the same wage everywhere. For instance, workers in the bottom 10% in Paris earn wages 3.2% higher than workers in the bottom 10% in Lens (compared to 1.7% lower in the data). By contrast, workers at the top of the local ladder disproportionally gain from working

Figure 8: The impact of employer composition and local competition on wages



in larger cities: the wage gap between Paris and Lens for workers in the top 10% of their local distribution is 21.7% (compared to 15.3% in the data).

Given the homogeneity of TFP across locations, the spatial differences in wages ultimately follow from the allocation of employers. However, cities are not a simple clustering of jobs: they shape wages by determining the competition faced by employers.

To decompose the relative importance of employer sorting and local competition, 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 means less market power. I also define the (unweighted) average markdown charged by employer z as $\bar{\mu}(z) = L^{-1}\sum_\ell \mu_\ell(z)$. Then, for any wage partition $\mathcal{W}$, the average log wage in location ℓ can be decomposed into

$$\mathbb{E}_\ell[\log w \mid \mathcal{W}] = \mathbb{E}_\ell[\log z\bar{\mu}(z) \mid \mathcal{W}] + \mathbb{E}_\ell\left[\log \frac{\mu_\ell(z)}{\bar{\mu}(z)} \mid \mathcal{W}\right]. \tag{26}$$

The first term captures the direct effect of employer sorting on wages. As productive employers concentrate in large cities, the MPL rise. 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 0.74, it is 11.5 p.p. lower for employers in the top productivity decile (Figure D.21a). The firm term thus summarizes how the employer composition affects wages assuming employers charge constant markdowns across space.[31]

The second term represents the impact of cities on wages through local competition. 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 price on average.

Figure 8a begins by applying this decomposition to average wages. The blue markers display the average wage premia by location. The orange circles plot the composition effect. The difference between the two, summarized by the grey shaded areas, is the competition effect.

Local competition boosts substantially wages in the largest cities. The spatial concentration of productive employers in Paris indeed intensifies the local competition for workers. As a result, the average markdown is 1.5 p.p. higher than what it would given the local pool of employers, and wages rise by 3%. At the same time, the spatial concentration of productive employers does not redistribute competition away from the smaller cities. The spatial sorting of employers indeed implies that they compete with relatively more similar firms, which (weakly) decreases their market power. Altogether, local competition explains 21.1% of the between-city wage premia variance.

Figure 8b then applies the decomposition (26) separately for wages in the bottom and top 10% of the local wage distribution. Figure 8b reveals that not every employer is equally affected by the fiercer competition of large cities.

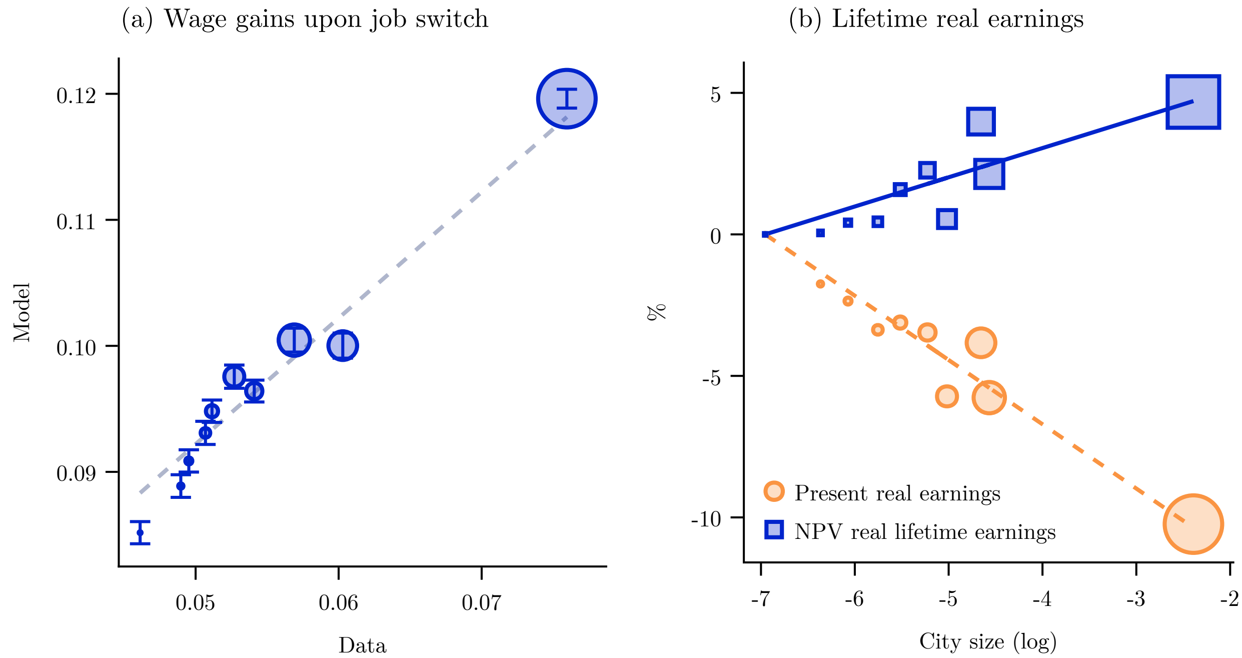
On the one hand, low-paying jobs exist throughout space because the lack of bargaining power from unemployed workers prevents spatial productivity differentials from materializing. For instance, employers in the bottom 10% of the wage distribution in Paris are 7.6% more productive than employers at the bottom in Lens. If they were to set a uniform markdown, they would offer wages 11.9% higher. However, in both places, these employers hire most of their workforce from unemployment. They therefore face similar competitive pressure, offer a wage close to unemployed workers' reservation wages, and the higher productivity in Paris does not spillover onto higher wages.

On the other hand, the highest-paying jobs are spatially concentrated because they rely on the intense competition unique to large cities. While employers in the top 10% of the wage distribution in Paris are substantially more productive than employers at the top in Lens (33.1%), they also tend to charge substantially lower markdowns all else equal. Combined, the wage gap between Paris and Lens at the top of wage distribution would be 2.7% were markdowns uniform. However, every employer in Paris is relatively more productive. This competitive pressure builds throughout the local ladder as employers poach each other. As a result, employers in the top 10% of the wage distribution in Paris offer markdowns 10.1 p.p. higher than they would typically charge.

This heterogeneous competition effect across employers is unique to this framework. It arises from the interaction between employer sorting and local frictional labor markets. Crucially, it is this mechanism that lets the model replicate the first fact documented in Section 1.3. Figure D.20 indeed

[31]The uniform markdown $\bar{\mu}(z)$ is not an equilibrium object. In particular, it may correspond to the markdown employers would use if there were a single city. In Figure D.21b, I solve the model with homogeneous locations ($\boldsymbol{A} = 1, \boldsymbol{T} = 1$), and depict the function $z \to \bar{\mu}(z)$ together with the equilibrium markdown in this counterfactual model. In practice, the two distributions are very similar.

Figure 9: The consequences of local ladders on wage growth and lifetime earnings



shows that, were markdowns uniform across locations, the agglomeration of productive employers in large cities would generate *lower* inequality.

4.2 How local ladders shape earnings

Proposition 2 shows that the concentration of high-paying jobs in large cities implies a steeper local job ladder. In the second over-identification exercise, I assess what are the consequences of the local ladders for wage growth, lifetime earnings, and the spatial distribution of economic activity.

Figure 9a 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 associated 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 model overpredicts the aggregate gains of job switching. On average, job switchers experience an average wage increase of 9.9% in the model; this is 5.6% in the data. However, the between-city differences in the returns to job switching aligns well with the data. Workers who switch jobs in Paris enjoy wage increase 2.6 p.p. higher than in Lens in the model, compared to 2.9 p.p. in the data. Throughout space, the correlation between the model and the empirical estimates is 96.1%.

The steeper ladder in larger cities generates faster growth, and therefore higher lifetime real earnings, for workers there. The orange circles in Figure 9b shows the present real earnings of new entrants in each location. The blue rectangles are the net present value (NPV) of their expected real lifetime earnings.

New entrants earn lower real earnings in larger cities. They start as unemployed, and while housing is more expensive in larger cities, unemployment insurance is constant throughout space. However, real lifetime earnings are higher as workers have access to better future opportunities. For instance, while the current real earnings of new entrants in Paris is 8.1% lower than in Lens, their real lifetime earnings are 4.2% higher. Generalizing to all cities, I find that the NPV of real lifetime earnings for new entrants is 1% higher in cities twice larger.[32]

The reallocation of workers across jobs, and the wage growth it begets, are thus key to the spatial distribution of economic activity. I conclude this paper with a counterfactual that quantifies their importance. Specifically, I simulate a decrease in the job switching rate via an increase in search frictions. I set the counterfactual frictions so that the job switching rate is one percentage point lower —a trend observed in the United States and other countries during the late 1990s and early 2000s (e.g., Engbom, 2019).[33]

I solve for three counterfactual economies. In the first, which I refer to as the worker partial equilibrium, I compute the new spatial allocation of workers holding constant the wage offer distributions. In the second, I let employers re-optimize their decisions holding constant the spatial distribution of workers —the employer partial equilibrium. Finally, the third counterfactual is the general equilibrium. In all cases, I compare the economies in steady state.

Figure 10 traces the response of lifetime earnings and average employer productivity in Paris and Lens to the decline in the job switching rate.[34] Appendix D.1 generalizes the analysis to all the French cities.

I first detail the partial equilibrium results, summarized by the dashed lines. When frictions are stronger, workers stay employed relatively longer at low-paying jobs. Holding constant the wage offer distributions, lifetime earnings shrink, and disproportionally so in larger cities. For instance, the gap in new entrants' lifetime earnings between Paris and Lens drops by 2.2 p.p.. As a result, large locations become less attractive. The effects are sizable: in partial equilibrium, a one percentage point decrease in the job switching rate leads to 12.7% decrease in the number of workers in Paris, and a 5.2% increase in Lens' size (Figure D.23b).

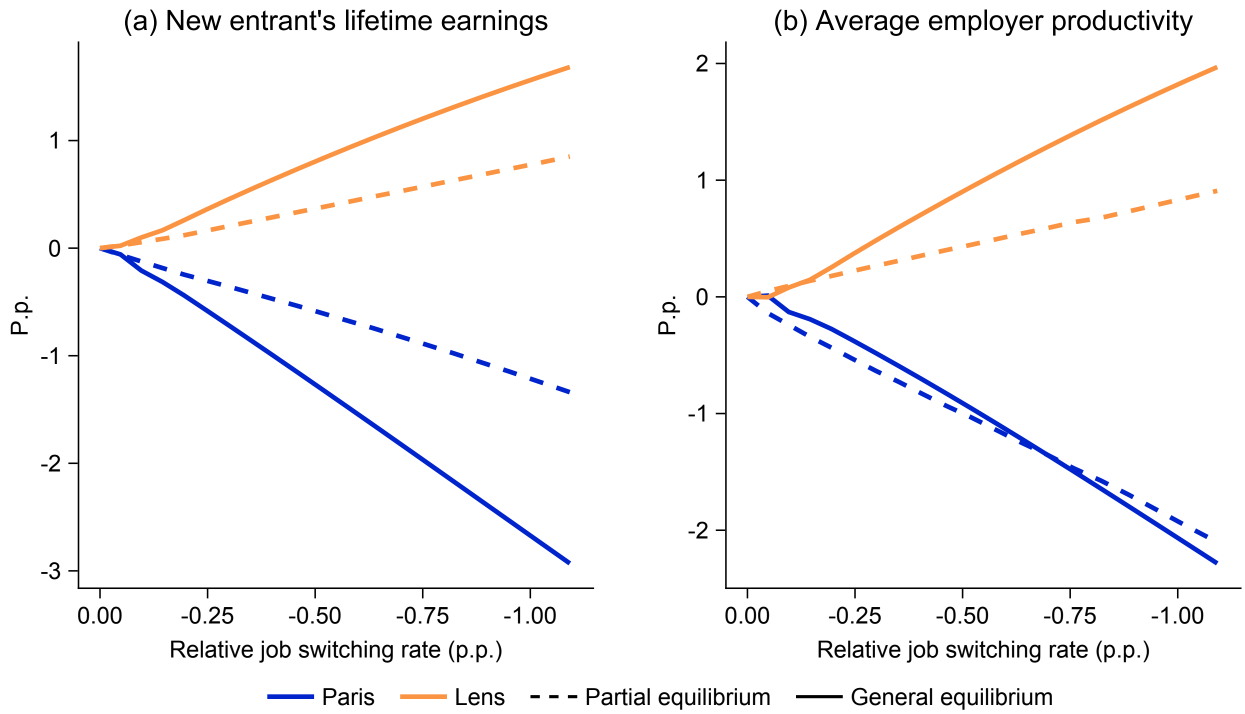
The partial equilibrium consequences on local productivity are *a priori* more ambiguous. On the one hand, large cities are more attractive for unproductive employers, who can now retain a greater fraction of their workers. On the other hand, productive employers value relatively more the slackness of large cities since it is relatively harder for them to poach workers from their competitors. Figure 10(b) shows that the first force dominate in partial equilibrium. Holding constant city sizes,

[32]There are also differences in lifetime earnings across workers within cities. However, when the discount rate is relatively low, these differences are small. As a result, the between-city differences in average lifetime earnings are similar to that of new entrants, with a city-size gradient of 0.011.

[33]The change in search friction is uniform across locations due to the absence of congestion in the matching function.

[34]Stronger frictions also affect aggregate lifetime earnings and productivity. Figure 10 expresses the two variables relative to their respective national average to net out the aggregate effect.

Figure 10: Search friction and the spatial distribution of activity



an increase in search frictions reduces the between-city differences in employer average productivity. The effects are also large: the gap in average employer productivity between Paris and Lens drops from 18.9% to 15.9%.[35]

In general equilibrium, the response of workers interact with that of employers. Productive employers find it more profitable to produce in small cities —which are now relatively larger— thereby amplifying the partial equilibrium increase in productivity. In the major cities, the decrease in productivity further depresses lifetime earnings. These general equilibrium interactions amplify significantly the partial equilibrium effects. The average productivity gap between Paris and Lens reduces to 14.7%. The extra lifetime earnings offered by Paris to new entrants are 2.9 p.p. lower than in the baseline economy —twice the size of the partial equilibrium response. The number of workers in Paris drops by 22.2%, and those in Lens rises by 10.1%.

Two key takeaways thus arise from Figure 10. First, local job ladders are instrumental for the spatial distribution of economic activity and wage disparities. Second, the general equilibrium interplay between employers' spatial allocation and local labor markets frictions are of first order importance.

[35] At the same time, *more* employers locate in large cities as higher frictions increase the profitability of these locations for both unproductive and productive employers (Figure D.24a).

5 Conclusion

In this paper, I document two novel facts about the importance of employers for spatial wage inequality. First, while high-paying jobs are concentrated in large cities, low-paying jobs are present throughout France. Second, workers access high wages in large cities by reallocating over time from low- to high-paying jobs. I argue that standard spatial models fail to rationalize these facts. Instead, I propose a new framework with two core features: spatial sorting of employers and frictional local labor markets. I show, theoretically and quantitatively, that the model replicates the novel empirical patterns. Hiring frictions imply that productive employers concentrate in large cities to maximize their size; higher wages and greater within-city inequality follow despite the absence of local TFP differentials. Workers in large cities earn higher lifetime earnings despite having lower real earnings upon entry as they have access to better future opportunities. When workers reallocate at a slower pace across jobs, the comparative advantage of large cities declines for workers and employers, and spatial disparities dampen.

This framework opens the door to new questions. What are the consequences of minimum wage reforms on local wage inequality? Should location-specific minimum wages be favored over a uniform policy? What does the rise of superstar firms, as in Song et al. (2019) and Rossi-Hansberg et al. (2021), imply for local labor market competition and spatial wage disparities? And does our understanding of place-based policies change when employers compete for workers along local job ladders? These questions, which can be answered within this new framework, are all fruitful avenues for future research.

References

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

Bacolod, M., Blum, B. S., and Strange, W. C. (2009). Skills in the city. *Journal of Urban Economics*, 65(2):136–153.

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.

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.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica*, 87(3):741–835.

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.

Carry, P., Kleinman, B., and Nimier-David, E. (2025). Location effects or sorting? Evidence from firm relocation.

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.

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. (2019). Firm and worker dynamics in an aging labor market.

Engbom, N. and Moser, C. (2022). Earnings inequality and the minimum wage: Evidence from brazil. *American Economic Review*, 112(12):3803–3847.

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.

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. and Maré, D. C. (2001). Cities and skills. *Journal of labor economics*, 19(2):316–342.

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

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.

Hong, G. (2024). Two-sided sorting of workers and firms: Implications for spatial inequality and welfare.

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

Jarosch, G., Nimczik, J. S., and Sorkin, I. (2024). Granular search, market structure, and wages. *Review of Economic Studies*, 91(6):3569–3607.

Kleinman, B. (2024). Wage inequality and the spatial expansion of firms.

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.

Marshall, A. (1890). *The Principles of Economics*. McMaster University Archive for the History of Economic Thought.

Martellini, P. (2022). Local labor markets and aggregate productivity.

Morchio, I. and Moser, C. (2024). The gender pay gap: Micro sources and macro consequences.

Moretti, E. (2011). Local labor markets. In *Handbook of labor economics*, volume 4, pages 1237–1313. Elsevier.

Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

Moretti, E. and Yi, M. (2024). Size matters: Matching externalities and the advantages of large labor markets.

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

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

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

Petrongolo, B. and Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic literature*, 39(2):390–431.

Porcher, C., Rubinton, H., and Santamaría, C. (2023). The role of establishment size in the city-size earnings premium. *Journal of Urban Economics*, 136:103556.

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

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, 90(6):1257–1278.

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

Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, pages 74–104.

Rossi-Hansberg, E., Sarte, P.-D., and Trachter, N. (2021). Diverging trends in national and local concentration. *NBER Macroeconomics Annual*, 35(1):115–150.

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.

A Two facts about spatial wage inequality

A.1 Data

A.1.1 France

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 $\times$ 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 $\times$ 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.1.2 United States

Wages I obtain wage date from the American Community Survey. I use sample restrictions similar to the French data:

1. Exclude workers younger than 25 and older than 55 year old;
2. Exclude workers employed in the public sector;
3. Exclude the agriculture industry;
4. Keep only workers employed full-time ($\geq$ 30 hours per week and 48 weeks per year);
5. Exclude workers with no labor income;

I define hourly wages as yearly labor income over the number weeks worked last year and the usual amount of hours worked per week. I truncate the left-tail of the hourly wage distribution by US

State at the 2.5% to reduce measurement error in hours worked.[36]

Housing prices I obtain rent data from the American Community Survey. To ensure consistency, I compute local rents on the same sample as local wages for those who report positive rents.

Employment I obtain total private non-farm employment by US State from the Bureau of Labor Statistics.

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 large, or because they 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

$$\begin{aligned}
\mathbb{V}\mathrm{ar}_\ell[\log w_{it}] = \mathbb{V}\mathrm{ar}_\ell[\psi_{it}] &+ \rho(\psi_{it}, \gamma_{it})(\mathbb{V}\mathrm{ar}_\ell[\psi_{it}] - \mathbb{V}\mathrm{ar}[\psi_{it}]) \left(\frac{\mathbb{S}\mathrm{d}[\gamma_{it}]}{\mathbb{S}\mathrm{d}[\psi_{it}]}\right) + \\
&\mathbb{V}\mathrm{ar}_\ell[\gamma_{it}] + \rho(\psi_{it}, \gamma_{it}) \left(\mathbb{V}\mathrm{ar}_\ell[\gamma_{it}] - \mathbb{V}\mathrm{ar}[\gamma_{it}]\right) \left(\frac{\mathbb{S}\mathrm{d}[\psi_{it}]}{\mathbb{S}\mathrm{d}[\gamma_{it}]}\right) + \\
&2\rho_\ell(\psi_{it}, \gamma_{it})\mathbb{S}\mathrm{d}[\psi_{it}]\mathbb{S}\mathrm{d}[\gamma_{it}] + \mathbb{V}\mathrm{ar}_\ell[\varepsilon_{it}] + o_\ell,
\end{aligned} \tag{27}$$

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.

[36]Labor incomes are already right-truncated in the ACS.

A.3 Tables

Table A.1: 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, 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.2: The local returns to job switching

	(1)	(2)	(3)	(4)	(5)
# jobs	0.013	0.054	0.016	0.009	0.012
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
# jobs × log size	0.375	0.742	0.318	0.333	0.346
	(0.008)	(0.012)	(0.008)	(0.008)	(0.008)
City × year F.E.			✓		
Occupation F.E.				✓	
Worker slope F.E.					✓
Switch	All	Wage gain	All	All	All
# Obs.	8,798,361	8,798,361	8,798,361	8,798,361	8,798,361
R^2	0.886	0.889	0.887	0.890	0.887

Dependant variable: log hourly wage. The indepedent variable # jobs is defined in (5), and size is the average number of employed workers per city in each city cluster. All regressions include: worker F.E., location F.E., experience, and experience interacted with log city size. Standard errors are clustered at the worker level. The coefficients on the second line are scaled to represent the marginal effect of doubling city size.

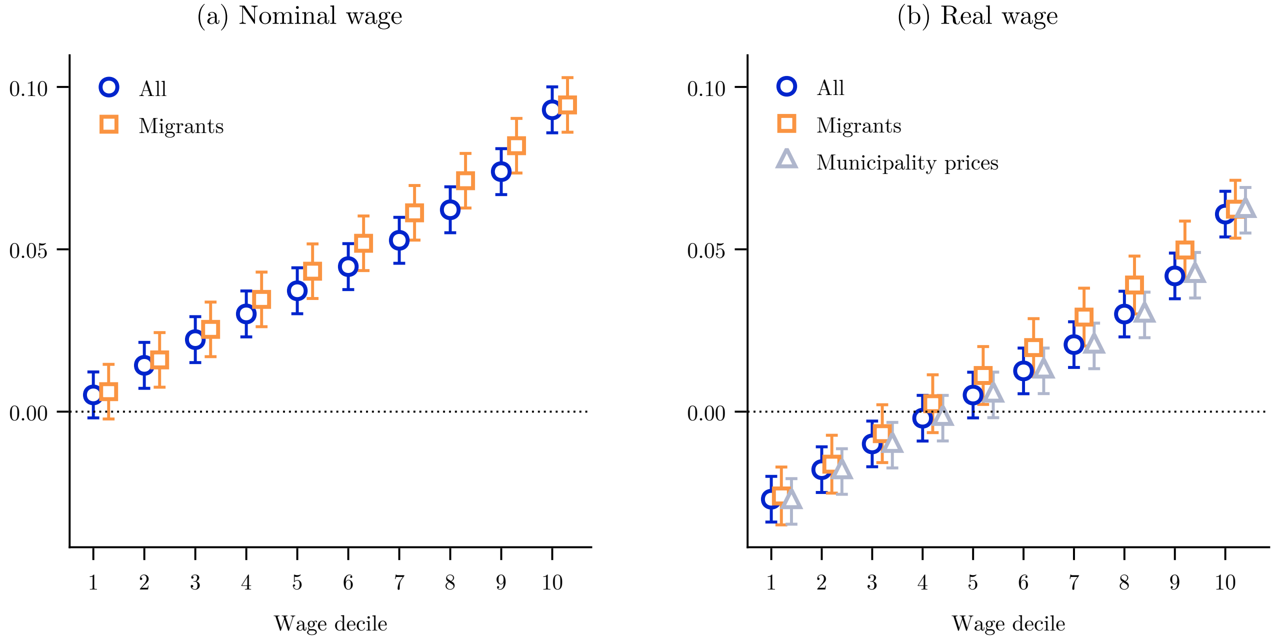
Table A.3: Job switching within and across commuting zones

	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 $\times$ 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.

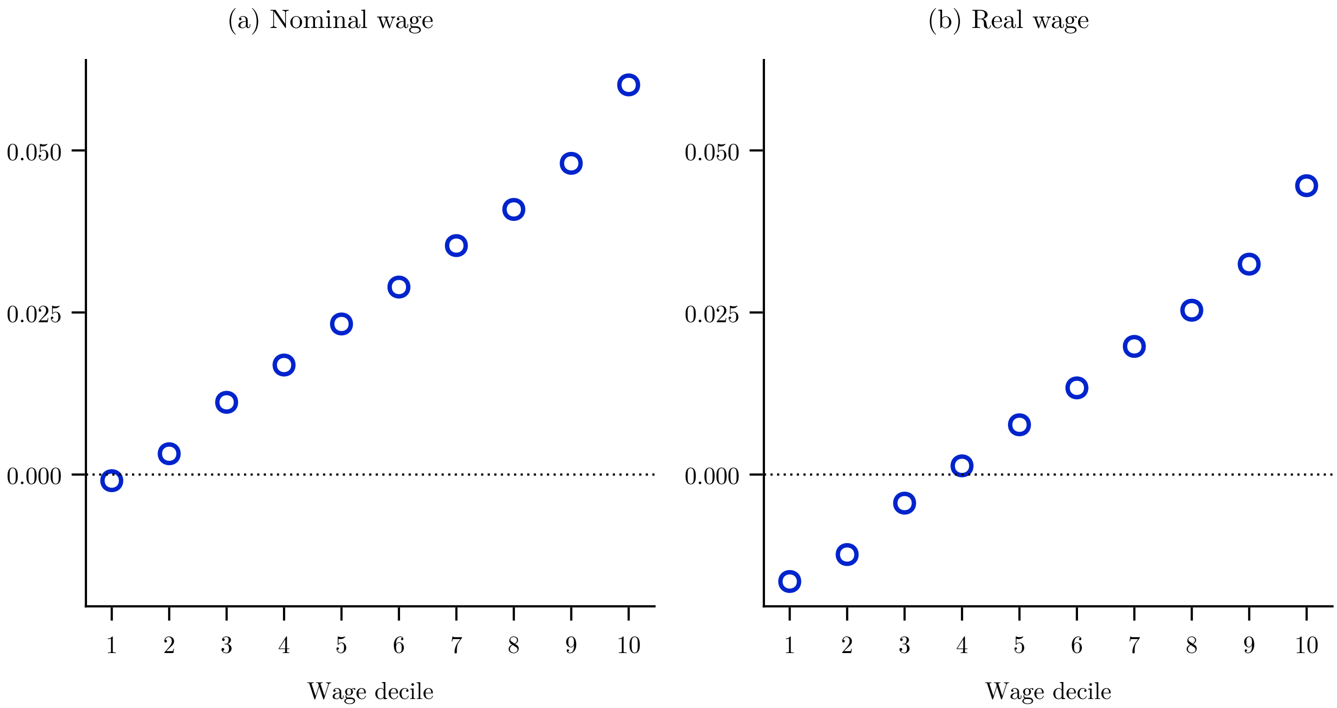
A.4 Figures

Figure A.1: Wage distribution by commuting zone



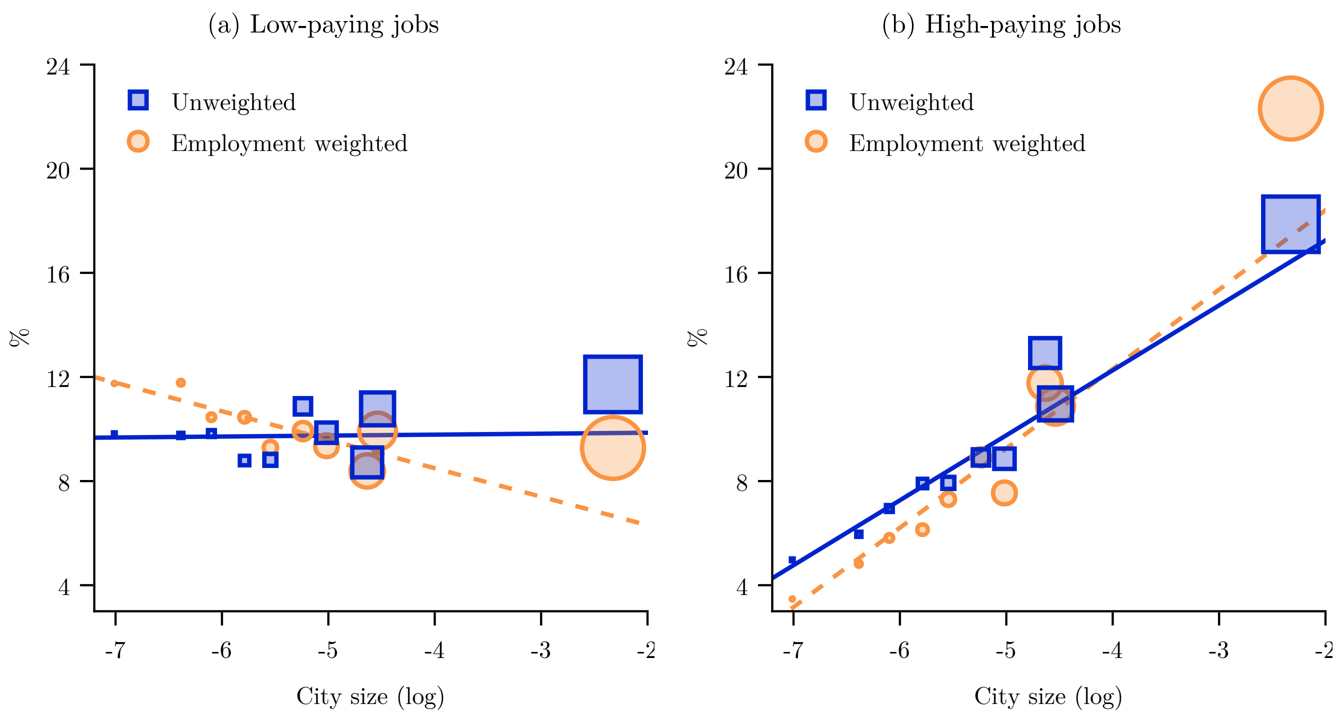
Both panels display the city size elasticity of wage deciles. Let $w_{q\ell}$ denotes the average wage in the qth decile of the wage distribution in location ℓ. Let m_ℓ denote location ℓ's size. The city elasticities are estimated by $\log w_{q\ell} = \alpha_q + \beta_q \log m_\ell + u_{q\ell}$. Panel (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 0.3. 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: Wage distributions by US state



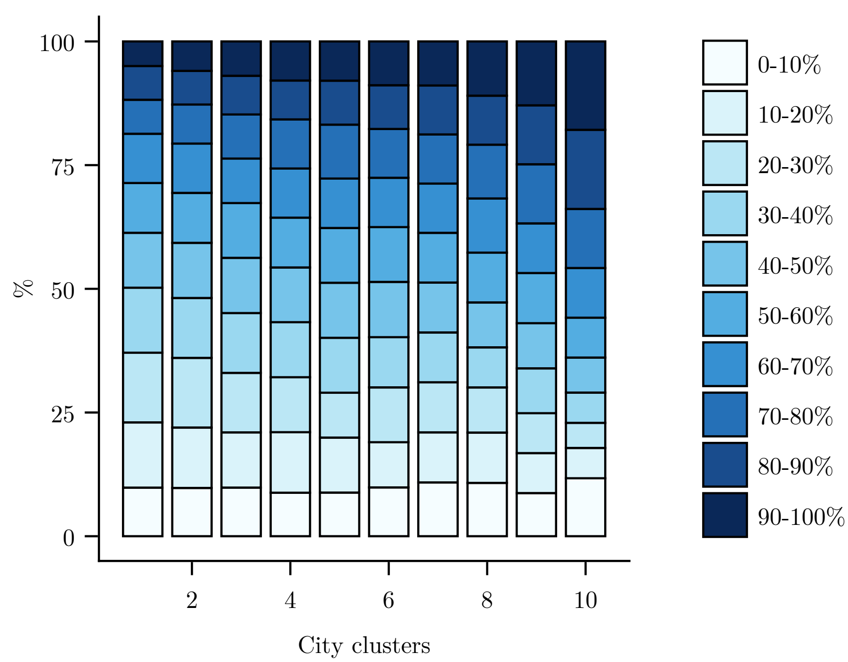
Data source: American Community Survey (wages and rent) and the Bureau of Labor Statistics (employment by state) –see Section A.1.2. Both panels display the state size elasticity of wage deciles. Let $w_{q\ell}$ denotes the average wage in the qth decile of the wage distribution in state ℓ. Let m_ℓ denote state ℓ's size. The state 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 statewide Cobb-Douglas price index with a housing expenditure share of 0.35

Figure A.3: Local share of low- and high-paying jobs



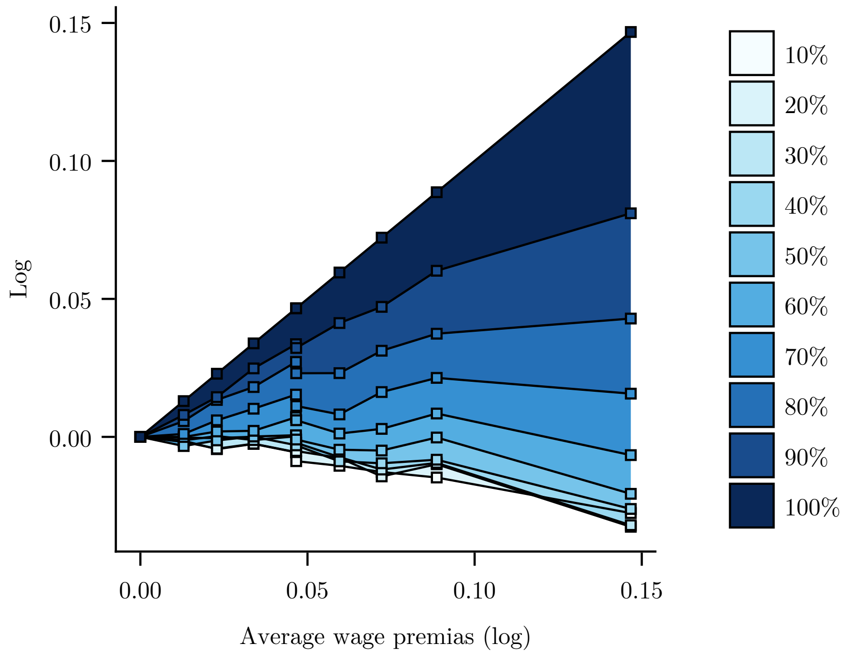
Left-panel depicts the job shares (blue rectangles) and employment shares (orange circles) of low-paying jobs in each city. Right-panel reproduces the exercice for high-paying jobs. Low- and high-paying jobs are defined as jobs belonging in the bottom and top 10% of the national job fixed effect distribution, unweighted and employment-weighted for the job and employment shares respectively. The job fixed effects are obtained from (1).

Figure A.4: Spatial distribution of job wage premia



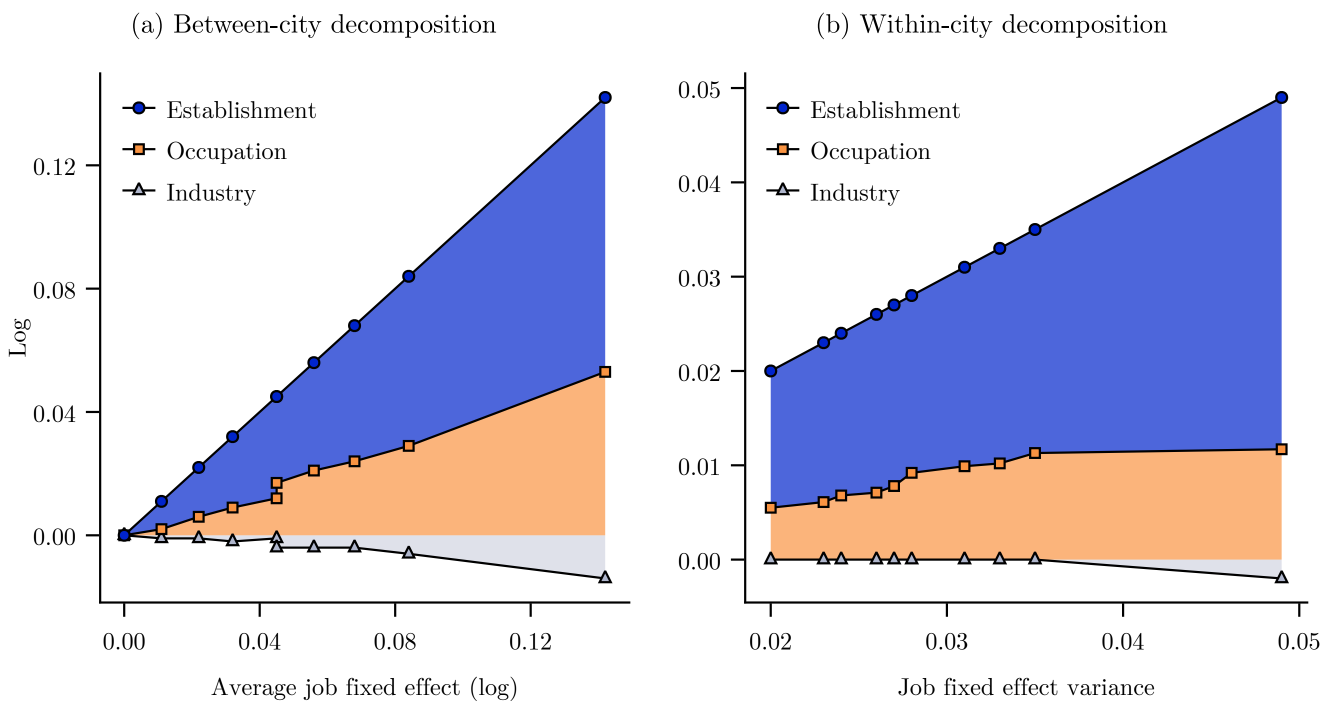
Panel displays the local job shares of jobs in the x^{th} decile of the aggregate job fixed effect distribution. The job fixed effects are obtained from (1).

Figure A.5: Counterfactual average wage premia



Panel displays the counterfactual average job fixed in each city cluster when only the bottom $x\%$ of the aggregate job fixed effect distribution are included. The average job fixed effect is normalized to zero in the smallest location for all counterfactuals. The job fixed effects are obtained from (1).

Figure A.6: The role of occupation, industry, and establishment on spatial wage inequality



Panel (a) displays the decomposition $\mathbb{E}_\ell[\psi_j] = \mathbb{E}_\ell[\phi_{\iota(j)}] + \mathbb{E}_\ell[\delta_{o(j)}] + \mathbb{E}_\ell[\psi_f(j)]$. The grey area represents $\mathbb{E}_\ell[\phi_{\iota(j)}]$, the orange area $\mathbb{E}_\ell[\delta_{o(j)}]$, and the blue area $\mathbb{E}_\ell[\psi_f(j)]$. Panel (b) plots the decomposition of the within-city job fixed effect variance. The grey area plots the dispersion in industry FE, $\mathbb{V}\text{ar}_\ell[\phi_{\iota(j)}] + \mathbb{C}\text{ov}[\delta_{o(j)}, \phi_{\iota(j)}] + \mathbb{C}\text{ov}[\phi_{\iota(j)}, \psi_f(j)]$. The orange area plots the dispersion in occupation FE, $\mathbb{V}\text{ar}_\ell[\delta_{o(j)}] + \mathbb{C}\text{ov}[\delta_{o(j)}, \phi_{\iota(j)}] + \mathbb{C}\text{ov}[\delta_{o(j)}, \psi_f(j)]$. The blue area plots the remainder, $\mathbb{V}\text{ar}_\ell[\psi_{f(j)}] + \mathbb{V}\text{ar}_\ell[\nu_j] + \mathbb{C}\text{ov}[\psi_{f(j)}, \phi_{\iota(j)}] + \mathbb{C}\text{ov}[\delta_{o(j)}, \psi_f(j)]$. The fixed effects are obtained from (1) and (4).

Figure A.7: The role of sorting for within-city wage inequality

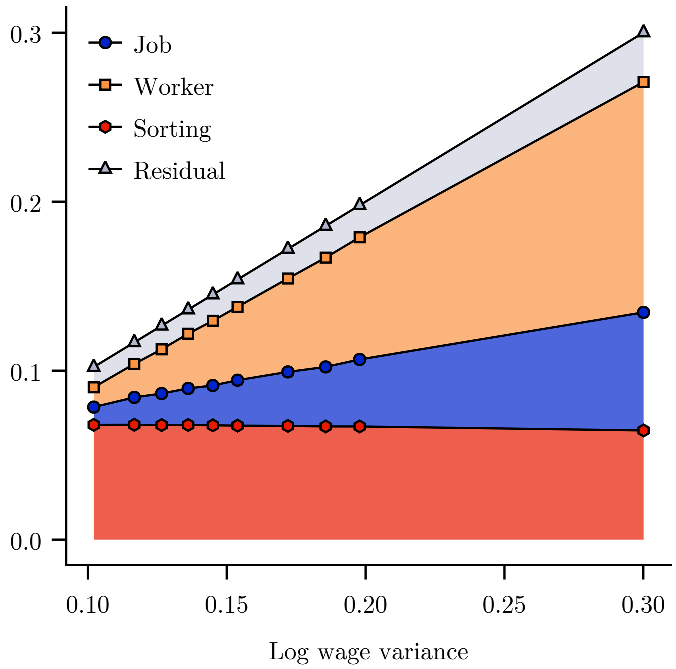


Figure displays the first-order approximation of the variance decomposition (27). Each marker is a city group. The orange area represents the worker component, $\mathbb{V}\mathrm{ar}_\ell[\psi_{it}] + \rho(\psi_{it}, \gamma_{it})(\mathbb{V}\mathrm{ar}_\ell[\psi_{it}] - \mathbb{V}\mathrm{ar}[\psi_{it}]) \left(\frac{\mathbb{S}\mathrm{d}[\gamma_{it}]}{\mathbb{S}\mathrm{d}[\psi_{it}]}\right)$. The blue area represents the employer contribution, $\mathbb{V}\mathrm{ar}_\ell[\gamma_{it}] + \rho(\psi_{it}, \gamma_{it}) (\mathbb{V}\mathrm{ar}_\ell[\gamma_{it}] - \mathbb{V}\mathrm{ar}[\gamma_{it}]) \left(\frac{\mathbb{S}\mathrm{d}[\psi_{it}]}{\mathbb{S}\mathrm{d}[\gamma_{it}]}\right)$. The red area represents the sorting component, $\rho_\ell(\psi_{it}, \gamma_{it})\mathbb{S}\mathrm{d}[\psi_{it}]\mathbb{S}\mathrm{d}[\gamma_{it}]$. The grey area represents the approximation residuals. The job fixed effects are obtained from (1).

Figure A.8: The local returns to experience

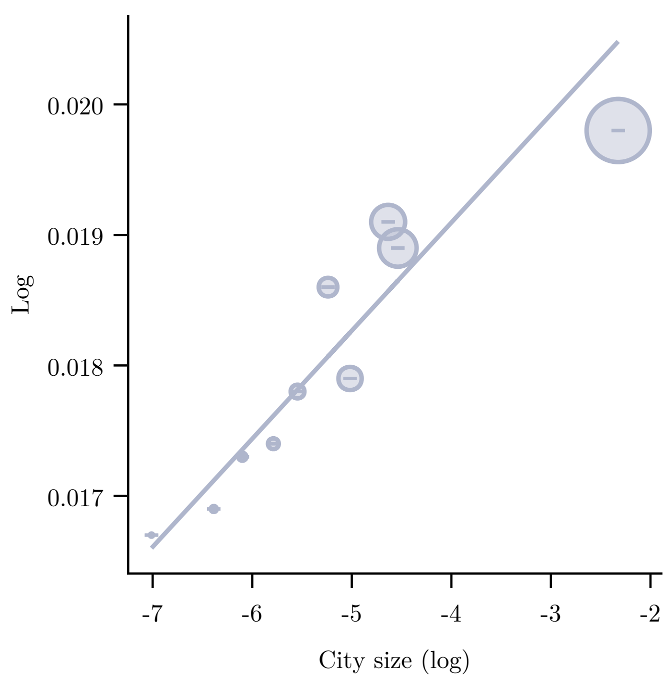
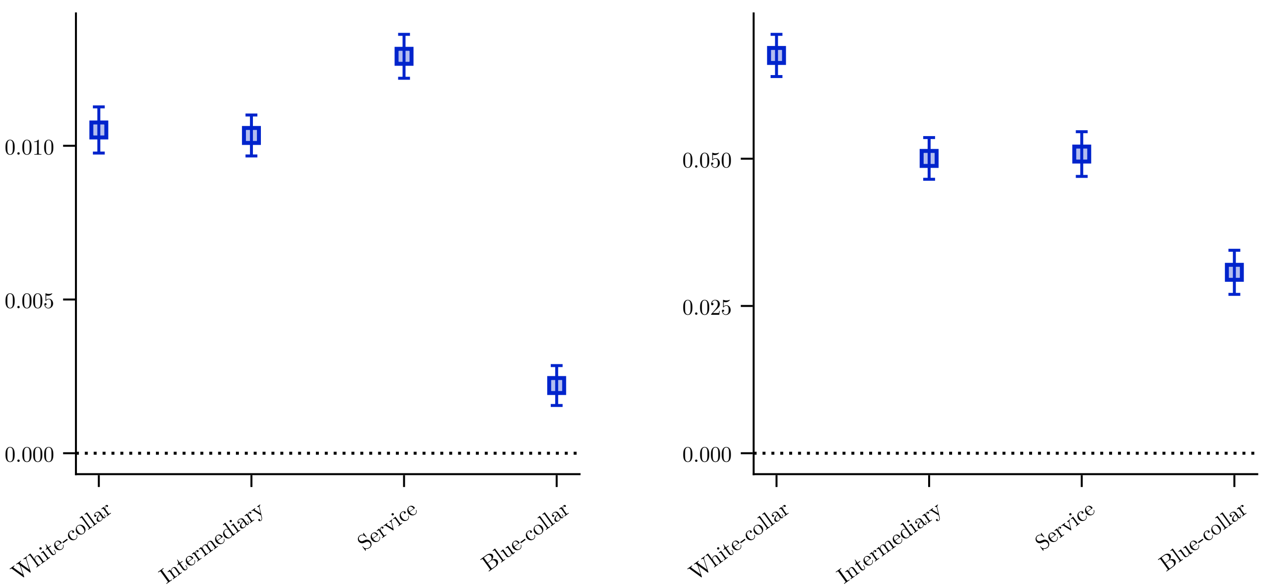


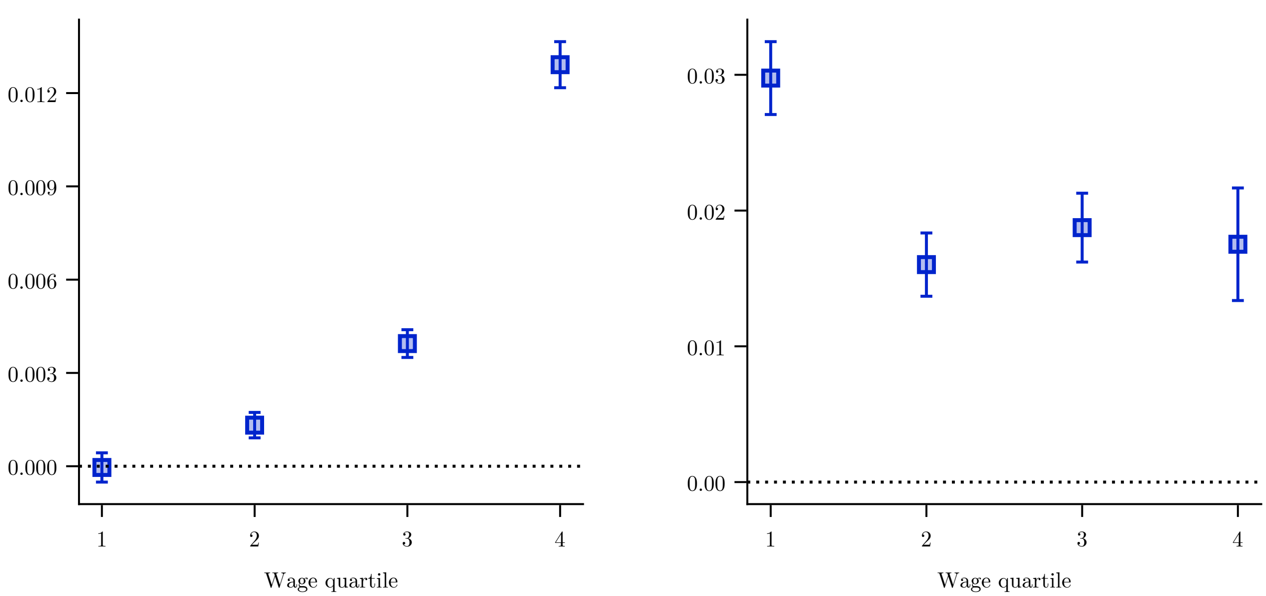
Figure plots the returns to experience parameters $\{\beta_\ell\}_{\ell=1}^L$ in (6). The vertical bars represent 95th confidence intervals. Standard errors are clustered at the individual level.

Figure A.9: The local returns to job switching by occupation



Both panels display point estimates of the returns to job switching by occupation. Specifically, I estimate $\log w_{it} = \text{FE}_t + \text{FE}_i + \text{FE}_{o(i,t)} + \alpha_{\ell(i,t)} + \beta^b x_{it} + \beta^s x_{it} \log \text{size}_{\ell(i,t)} + \gamma^b_{o(i,t)} J_{it} + \gamma^s_{o(i,t)} J_{it} \log \text{size}_{\ell(i,t)}$, where most variables are defined in (6), size_ℓ is the average city size in cluster ℓ, 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 individual level.

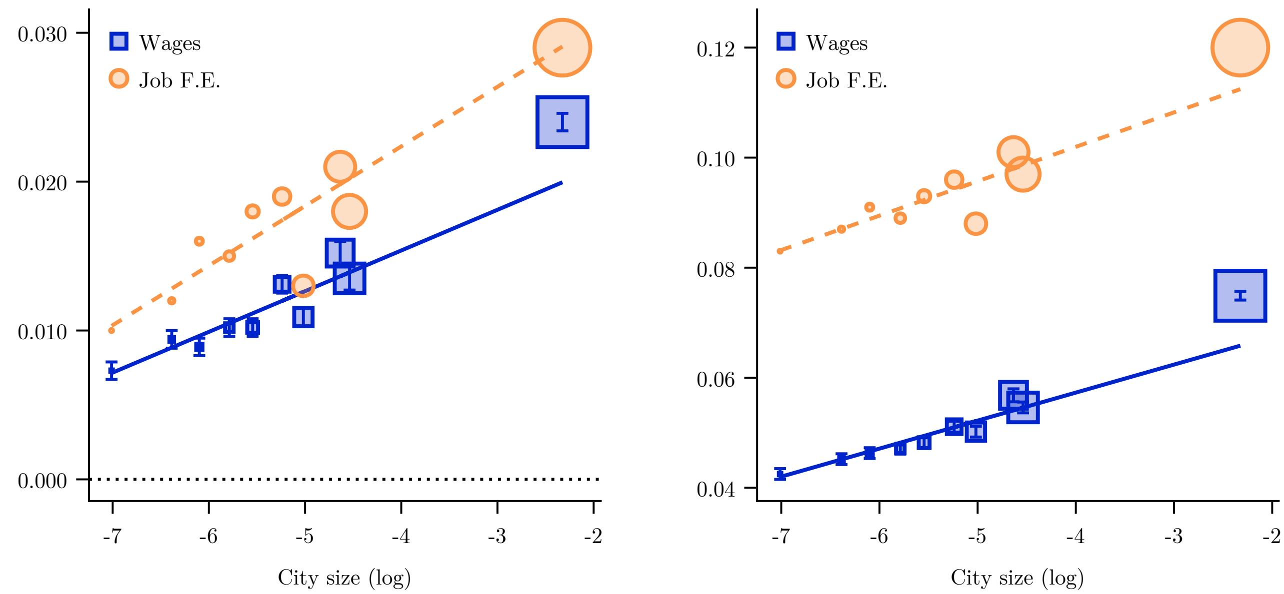
Figure A.10: The local returns to job switching by wage quartile



Both panels display point estimates of the returns to job switching by wage quartile. Specifically, I estimate $\log w_{it} = \text{FE}_t + \text{FE}_i + \alpha_{\ell(i,t),q(i,t)} + \beta^b x_{it} + \beta^s x_{it} \text{size}_{\ell(i,t)} + \gamma^b_{q(i,t)} z_{it} + \gamma^s_{q(i,t)} z_{it} \text{size}_{\ell(i,t)}$, where most variables are defined in (6), size_ℓ is the average city size in cluster ℓ, 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 individual level.

Figure A.11: Estimating the local returns to job switching with AKM fixed effects

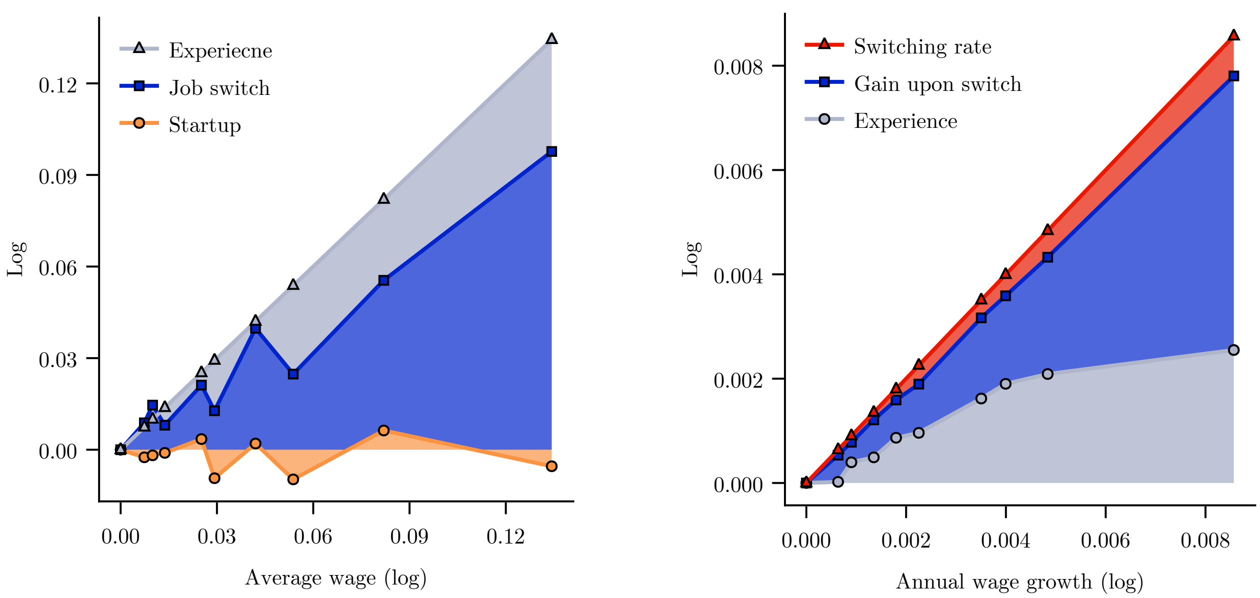
(a) All switch (b) Switch with gains



Blue rectangles display the estimates from (6), on the left using all switches and on the right using switches associated with wage growth. Orange circles display the average change in AKM fixed effect for switchers, on the left using all switches, and on the right using switches for which the change in fixed effect is positive. Job fixed effects estimated from (1).

Figure A.12: The consequences of startup premia, job ladders, and experience on local wages

(a) Average wage (b) Average wage growth



Left panel decomposes the between-city average wage gap net of worker heterogeneity, $\mathbb{E}_\ell[\log w_{it} - \alpha_i^{\mathrm{JL}}]$. The orange area depicts $\delta_\ell^{\mathrm{JL}}$. The blue area represents $\lambda_\ell^{\mathrm{JL}}\mathbb{E}_\ell[J_{it}]$. The grey area shows $\beta_\ell^{\mathrm{JL}}\mathbb{E}_\ell[x_{it}]$. Right panel decomposes the between-city average wage growth gap, $\mathbb{E}_\ell[\Delta_t \log w_{it}] = \beta_\ell^{\mathrm{JL}}\mathbb{E}_\ell[\Delta_t x_{it}] + \mathbb{E}[\lambda_\ell^{\mathrm{JL}}]\mathbb{E}_\ell[\mathrm{J2J}_{it}] + (\lambda_\ell^{\mathrm{JL}} - \mathbb{E}[\lambda_\ell^{\mathrm{JL}}])\mathbb{E}_\ell[\mathrm{J2J}_{it}]$. The grey area depicts $\beta_\ell^{\mathrm{JL}}\mathbb{E}_\ell[\Delta_t x_{it}]$. The blue area represents $(\lambda_\ell^{\mathrm{JL}} - \mathbb{E}[\lambda_\ell^{\mathrm{JL}}])\mathbb{E}_\ell[\mathrm{J2J}_{it}]$. The red area shows $\mathbb{E}[\lambda_\ell^{\mathrm{JL}}]\mathbb{E}_\ell[\mathrm{J2J}_{it}]$. In both decompositions, all variables are expressed relative to the smallest location. Point estimates obtained from (6).

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 (12).

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

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

where $F_\ell(w^-) = \lim_{\varepsilon\to 0^-} F(w-\varepsilon)$. This is equivalent to equation (13) 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 (11) with respect to w, integrate it back to w, and use $V_\ell(\underline{w}_\ell) = U$ to obtain

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

Combine this expression into (9) and (11) to get

$$\rho U = \frac{A_\ell}{P_\ell}\left(b + \lambda^u\int_{\underline{w}_\ell}\frac{\bar{F}_\ell(w)}{\rho+\delta+\lambda^e\bar{F}_\ell(w)}dw\right), \tag{28}$$

$$\rho V_\ell(\underline{w}_\ell) = \frac{A_\ell}{P_\ell}\left(\underline{w}_\ell + \lambda^e\int_{\underline{w}_\ell}\frac{\bar{F}_\ell(w)}{\rho+\delta+\lambda^e\bar{F}_\ell(w)}dw\right) = \rho U. \tag{29}$$

Equating both equations and solving for $\underline{w}_\ell$ yield (14).

B.3 Proof of Lemma 1

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

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

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 imply that 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 (29), the lifetime utility of a worker employed at the reservation wage is

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

where I have also used (14). 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 (10) 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 derive here the wage equation (18). This section exactly follows Burdett and Mortensen (1998). I first show that, for any spatial allocation of employers $\{\Gamma_\ell\}_\ell$, there cannot be holes or mass points in the local wage offer distributions.

To start, suppose there is a hole in F_ℓ between $\underline{x} \geq \underline{w}_\ell$ and $\bar{x} \leq \bar{w}_\ell$. We have $F_\ell(\underline{x}) \leq F_\ell(\bar{x})$, where the inequality is strict if there is a mass point at either $\underline{x}$ or $\bar{x}$. Therefore $n_\ell(\underline{x}) \leq n_\ell(\bar{w})$. However, by offering any wage in $(\underline{x}, \bar{x})$, an employer that used to post wage $\bar{x}$ 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 $w \in [\underline{w}_\ell, \bar{w}_\ell]$. 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 $n_\ell(w + \varepsilon) > n_\ell(w)$ 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 $(z - w)n_\ell(w) < (z - w - \varepsilon)n_\ell(w + \varepsilon)$, and offering wage $w + \varepsilon$ is a profitable deviation. This rules out mass point in F_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 (18). Since there are no mass point in the wage offer distribution, we can take the first-order conditions of (20) with respect to w for any $z \in \text{supp}\ \Gamma_\ell$,

$$(z - w)\,\frac{n_\ell'(w)}{n_\ell(w)} = 1.$$

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

$$w'_\ell(z) = \frac{n'_\ell(z)}{n_\ell(z)}\left(z - w_\ell(z)\right). \tag{30}$$

Integrating this ODE with respect to w and using the boundary condition $w_\ell(\underline{z}_\ell) = \underline{w}_\ell$ returrns

$$w_\ell(z) = \underline{w}_\ell\left(\frac{n_\ell(\underline{z}_\ell)}{n_\ell(z)}\right) + \int_{\underline{z}_\ell}^{z} \zeta\left(\frac{n'_\ell(\zeta)}{n_\ell(z)}\right) \mathrm{d}\zeta, \quad \forall z \in \operatorname{supp}\Gamma_\ell. \tag{31}$$

Since $\int_{\underline{z}_\ell}^{z} n'_\ell(t)dt = n_\ell(z) - n_\ell(\underline{z}_\ell)$, we have

$$w_\ell(z) = \underline{w}_\ell\left(\frac{n_\ell(\underline{z}_\ell)}{n_\ell(z)}\right) + \left(1 - \frac{n_\ell(\underline{z}_\ell)}{n_\ell(z)}\right)\int_{\underline{z}_\ell}^{z} t\left(\frac{n'_\ell(t)}{\int_{\underline{z}_\ell}^{z} n'_\ell(x)dx}\right) dt,$$

which corresponds to (18).

B.5 Proof of Proposition 2

FOSD ordering (2.1) We first show that $\Gamma_\ell \succ_{\text{FOSD}} \Gamma_{\ell'}$ implies $F_\ell \succ_{\text{FOSD}} F_{\ell'}$. Let w_ℓ^q denote the q-th quantile of the wage offer distribution in ℓ, $q = F_\ell(w_\ell^q)$. Likewise, let z_ℓ^q denote the firm that offers the q-th quantile of the wage offer distribution in ℓ, $w_\ell(z_\ell^q) = w_\ell^q$. Combining the two and using the rank preserving property of F_ℓ returns $q = F_\ell[w_\ell(z_\ell^q)] = \Gamma_\ell(z_\ell^q)$, or $z_\ell^q = \Gamma_\ell^{-1}(q)$. Using the wage expression (31), the q-th quantile of the wage offer distribution is

$$w_\ell^q = w_\ell[\Gamma_\ell^{-1}(q)] = \underline{w}_\ell\left(\frac{1+k(1-q)}{1+k}\right)^2 + \int_0^q \Gamma_\ell^{-1}(u)\frac{2k[1+k(1-q)]^2}{[1+k(1-u)]^3}du, \tag{32}$$

where the above expression uses the fact that employment size (net of market tightness) is constant across space given a rank on the ladder. $F_\ell \succ_{\text{FOSD}} F_{\ell'}$ follows from (32) using a guess-and-verify. Guess that $F_\ell \succ_{\text{FOSD}} F_{\ell'}$. Proposition 1 implies $\underline{w}_\ell \geq \underline{w}_{\ell'}$. Meanwhile, $\Gamma_\ell^{-1}(u) \geq \Gamma_{\ell'}^{-1}(u)$ for all u, and strictly for some. Hence, $w_\ell^q \geq w_{\ell'}^q$ for all q, and strictly for some. Finally, since G_ℓ is a monotonically increasing function on F_ℓ, $G_\ell \succ G_{\ell'}$.

Inequality (2.2) We first prove that $\bar{w}_\ell/\underline{w}_\ell$ is increasing in Γ_ℓ when $\lambda^u - \lambda^e \geq 0$ is small. When $\lambda^u = \lambda^e$, we have $\underline{w}_\ell = b \perp \ell$, and $w_\ell^q/\underline{w}_\ell$ increasing in Γ_ℓ for all q follows from Proposition 2.1. In addition, $\underline{w}_\ell$ is continuous in $\lambda^u - \lambda^e$ and w_ℓ^q is independent of λ^u. Hence, $w_\ell^q/\underline{w}_\ell$ is increasing in Γ_ℓ for $\lambda^u - \lambda^e$ not too large. This holds for $\bar{w}_\ell = w_\ell^1$.

Second, we show that $w_\ell^q - \underline{w}_\ell$ is increasing in Γ_ℓ. Let $\Delta_{\ell\ell'}^q \equiv w_{\ell'}^q - \underline{w}'_\ell - (w_{\ell'}^q - \underline{w}_{\ell'})$ be the difference in quantile q-to-bottom wage gap between city ℓ and ℓ'. Using (32), this difference writes

$$\Delta_{\ell\ell'}^q = (\underline{w}_\ell - \underline{w}_{\ell'})\left(\frac{1+k(1-q)}{1+k}\right)^2 + \int_0^q \left(\Gamma_\ell^{-1}(u) - \Gamma_{\ell'}^{-1}(u)\right)\frac{2k[1+k(1-q)]^2}{[1+k(1-u)]^3}du$$

If $\Gamma_\ell \succ \Gamma_{\ell'}$, this implies $\underline{w}_\ell \geq \underline{w}_{\ell'}$ and $\Gamma_\ell^{-1}(u) > \Gamma_{\ell'}^{-1}(u)$ for all $q \in (0,1)$, from which it follows that $\Delta_{\ell\ell'}^q > 0$ for all q.

B.6 Proof of Proposition 3

This section solves for the local productivity distributions $\{\Gamma_\ell\}_{\ell=1}^L$ given local market tightness $\{\theta_\ell\}_{\ell=1}^L$. I focus my attention to cities in which there are both workers and firms, i.e. $0 < \theta_\ell < \infty$. The proof has 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 density of the productivity distributions. Lemma B.5 proves the condition for the overlapping support.

Lemma B.1 (Envelope theorem)**.**
The profit function $\pi_\ell(z)$ is continuously differentiable with $\pi_\ell'(z) = n_\ell(z)$

Proof. Differentiability follows from the Envelope theorem. Then, $\pi_\ell'(z) = n_\ell(z)$. Since F_ℓ has no mass point (Section B.4), $n_\ell(z)$ is continuous, and $\pi_\ell \in \mathcal{C}^1$. □

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

Proof. Suppose not. That is, for some city ℓ, there exists at least one hole in 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 (20) 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 imply (i) $\pi_\ell(z) = \pi_{\ell^\star}(z)$ and (ii) $\pi_\ell'(z) < \pi_{\ell^\star}'(z)$. Since $z \to n_\ell(z)$ is constant outside the support of Γ_ℓ and strictly increasing in it (see (17)), Lemma B.1 and (ii) implies $n_\ell(z') = \pi_\ell'(z') < \pi_{\ell^\star}(z') = n_{\ell^\star}(z')$ for all $z' \in (z, z+\varepsilon)$. Together with (i), it then must be that $\pi_\ell(z+\varepsilon) < \pi_{\ell^\star}(z+\varepsilon)$, a contradiction. □

Lemma B.3 (First order stochastic dominance ordering)**.**
For two cities ℓ and ℓ', $\theta_\ell < \theta_{\ell'}$ if and only if $\Gamma_\ell \succ \Gamma_{\ell'}$.

Proof. I start by showing that $\theta_\ell < \theta_{\ell'} \Rightarrow \Gamma_\ell \succ \Gamma_{\ell'}$. For the sake of contradiction, suppose that there exists a set $\mathcal{Z} \subseteq [\underline{z}, \bar{z}]$ of positive measure satisfying the three following conditions:

1. $\min \mathcal{Z} \geq \min(\underline{z}_\ell, \underline{z}_{\ell'})$;
2. $\max \mathcal{Z} \leq \max(\bar{z}_\ell, \bar{z}_{\ell'})$;
3. $\bar{\Gamma}_{\ell'}(z) \geq \bar{\Gamma}_\ell(z)$ for all $z \in \mathcal{Z}$.

The first two conditions are needed since, trivially, $\Gamma_\ell(z) = \Gamma_{\ell'}(z) = 0$ for all $z < \min(\underline{z}_\ell, \underline{z}_{\ell'})$ and $\bar{\Gamma}_\ell(z) = \bar{\Gamma}_{\ell'}(z) = 0$ for all $z > \max(\bar{z}_\ell, \bar{z}_{\ell'})$. The third condition implies $\max \mathcal{Z} \leq \bar{z}_{\ell'}$.

Lemma B.1 and (17) implies

$$\pi'_\ell(z) = \frac{1}{\theta_\ell}\frac{1+k}{[1+k\bar{\Gamma}_\ell(z)]^2} \geq \frac{1}{\theta_\ell}\frac{1+k}{[1+k\bar{\Gamma}_{\ell'}(z)]^2} > \frac{1}{\theta_{\ell'}}\frac{1+k}{[1+k\bar{\Gamma}_{\ell'}(z)]^2} = \pi'_{\ell'}(z), \tag{33}$$

for all $z \in \mathcal{Z}$.

Suppose first that $\min \mathcal{Z} \geq \underline{z}_{\ell'}$, such that $\mathcal{Z} \subseteq \operatorname{supp} \Gamma_{\ell'}$. In that case, profit maximization (20) demands $\pi_{\ell'}(z) \geq \pi_\ell(z)$ for all $z \in \mathcal{Z}$. If $\pi_{\ell'}(z) = \pi_\ell(z)$ for any $z \in \mathcal{Z}$, then (33) implies $\pi_{\ell'}(z') < \pi_\ell(z')$ for $z' \in B^+(z) \subset \mathcal{Z}$, a contradiction.[37] Hence, it must be that $\pi_{\ell'}(z) > \pi_\ell(z)$ for all $z \in \mathcal{Z}$, and $\min \mathcal{Z} \geq \bar{z}_\ell$. The inequality (33) thus simplifies to

$$\pi'_\ell(z) = \frac{1+k}{\theta_\ell} > \frac{1}{\theta_{\ell'}}\frac{1+k}{[1+k\bar{\Gamma}_{\ell'}(z)]^2} = \pi'_{\ell'}(z),$$

for all $z \in [\bar{z}_\ell, \max \mathcal{Z}]$. Meanwhile, profit maximization (20) requires $\pi_{\ell'}(\bar{z}_\ell) \leq \pi_\ell(\bar{z}_\ell)$. Together, these two inequalities imply $\pi_{\ell'}(z) < \pi_\ell(z)$ for all $z \in \mathcal{Z}$, which contradicts $\mathcal{Z} \subseteq \operatorname{supp} \Gamma_{\ell'}$ by profit maximization.

Suppose then that $\min \mathcal{Z} < \underline{z}_{\ell'}$, and therefore, $\min \mathcal{Z} \geq \underline{z}_\ell$ from the first condition on the definition of $\mathcal{Z}$. Then, for all $z \in [\min \mathcal{Z}, \underline{z}_{\ell'}]$, (33) becomes

$$\pi'_\ell(z) = \frac{1}{\theta_\ell}\frac{1+k}{[1+k\bar{\Gamma}_\ell(z)]^2} > \frac{1}{\theta_{\ell'}}\frac{1}{1+k\bar{\Gamma}_{\ell'}(z)} = \pi'_{\ell'}(z). \tag{34}$$

In addition, profit maximization (20) demands $\pi_\ell(z) \geq \pi_{\ell'}(z)$ for all $z \in [\min \mathcal{Z}, \underline{z}_{\ell'}]$. Together, these inequalities imply $\pi_\ell(\underline{z}_{\ell'}) > \pi_{\ell'}(\underline{z}_{\ell'})$, a contradiction with profit maximization.

$\theta_\ell < \theta_{\ell'} \Leftarrow \Gamma_\ell \succ \Gamma_{\ell'}$ automatically follows. Suppose not $\theta_\ell < \theta_{\ell'}$, i.e., $\theta_\ell \geq \theta_{\ell'}$. I show that $\Gamma_\ell \not\succ \Gamma_{\ell'}$. If $\theta_\ell > \theta_{\ell'}$, I have already shown that $\Gamma_{\ell'} \succ \Gamma_\ell$. If $\theta_{\ell'} = \theta_\ell$, it must be that $\Gamma_{\ell'} = \Gamma_\ell$ almost everywhere for other similar contradictions as above can be constructed. □

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,

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

Given $\{\theta_\ell\}_{\ell=1}^L$ and a vector of cutoffs $\{\underline{z}_\ell, \bar{z}_\ell\}_{\ell=1}^L$, the productivity distribution in ℓ is unique and

[37] 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.

given by

$$\mathrm{d}\Gamma_\ell(z) = 1\{z \in [\underline{z}_\ell, \bar{z}_\ell]\}\left(\frac{M}{M_\ell}\right)\mu^\ell(z)\mathrm{d}\Gamma(z).$$

Proof. Consider the productivity set $\mathcal{Z}^{\mathcal{L}} = \{z : \mathcal{L}(z) = \mathcal{L}\}$, for $\mathcal{L}$ some element of the power set of L. $\mathcal{Z}^{\mathcal{L}}$ satisfies two properties. First, $\mathcal{Z}^{\mathcal{L}}$ is convex for all $\mathcal{L} \in \mathcal{P}(L)$.[38] Second, $\cap_{\mathcal{L}\in\mathcal{P}(L)}\mathcal{Z}^{\mathcal{L}} = \text{supp}\ \Gamma$.

Take first $\mathcal{L} \in \{\{1\}, \{2\}, \ldots, \{L\}\}$. Let ℓ denote the unique location in $\mathcal{L}$. The feasibility condition (8) requires $M_\ell \mathrm{d}\Gamma_\ell(z) = M\mathrm{d}\Gamma(z)$ for all $z \in \mathcal{Z}^{\mathcal{L}}$.

Take then $\mathcal{L} \in \mathcal{P}(L)\backslash\{\{1\}, \{2\}, \ldots, \{L\}\}$. Then, for all $z \in \mathcal{Z}^{\mathcal{L}}$, profit maximization (20) demands $\pi_\ell(z) = \pi_{\ell'}(z)$ for all $(\ell, \ell') \in \mathcal{L}$. The convexity of $\mathcal{Z}^{\mathcal{L}}$ then implies $\pi'_\ell(z) = n_\ell(z) = n_{\ell'}(z) = \pi_{\ell'}(z)$ for all $(\ell, \ell') \in \mathcal{L}$ and all $z \in \mathcal{Z}^{\mathcal{L}}$. Differentiating that equality one more time with respect to z yields

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

for any $(\ell, \ell') \in \mathcal{L}$ and all $z \in \mathcal{Z}^{\mathcal{L}}$. But feasibility (8) requires $\sum_{\ell'\in\mathcal{L}} M_{\ell'}\mathrm{d}\Gamma_{\ell'}(z) = M\mathrm{d}\Gamma(z)$, and therefore

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

for all $\ell \in \mathcal{L}$ and all $z \in \mathcal{Z}^{\mathcal{L}}$. □

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, profit maximization (20) requires $\pi_{\ell'}(\underline{z}_\ell) \leq \pi_\ell(\underline{z}_\ell)$ 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 $z^\star \in [\bar{z}_{\ell'}, \underline{z}_\ell]$ 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 (Lemma B.1), 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 Lemma B.1, we have

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

[38] Fix $\mathcal{L} \in \mathcal{P}(L)$. Take $(z, z') \in \mathcal{Z}^{\mathcal{L}}$. We have $\lambda z + (1-\lambda)z' \in \mathcal{Z}^{\mathcal{L}} \iff \mathcal{L}(\lambda z + (1-\lambda)z') = \mathcal{L}(z) = \mathcal{L}(z') = \mathcal{L}$. So take $\ell \in \mathcal{L}$. We know $(z, z') \in [\underline{z}_\ell, \bar{z}_\ell]$, and therefore $\lambda z + (1-\lambda)z' \in [\underline{z}_\ell, \bar{z}_\ell]$ for all $\lambda \in (0, 1)$.

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 $\text{supp}\,\Gamma_\ell \cap \text{supp}\,\Gamma_{\ell'} = [\underline{z}_\ell, \bar{z}_{\ell'}]$ 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 Lemma B.1, this must hold at $\underline{z}_\ell$, and therefore

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

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. $\square$

B.7 Proof of Proposition 4

The proof of Proposition 4 is broken up in three lemmas. Lemma B.6 derives the expression for the commercial housing prices that sustain the spatial job allocation. Lemma B.7 then shows that tighter market must have higher prices and more employers. Lemma B.8 then concludes by proving that larger cities have a tighter market. Lemma B.7 and Lemma B.8 are proved under $\lambda^u = \lambda^e$. Since these lemmas are strict inequalities and the model is continuous in ζ, they also hold for $\lambda^u \approx \lambda^e$.

Lemma B.6 (Housing 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 housing prices $\{r_\ell\}_{\ell=1}^L$ that satisfies the difference equation

$$r_{\ell+1} = r_\ell + \left(\underline{z}_{\ell+1} - \underline{w}_{\ell+1}\right) n_{\ell+1}(\underline{z}_{\ell+1}) - \left(\underline{z}_\ell - \underline{w}_\ell\right) n_\ell(\underline{z}_\ell) - \int_{\underline{z}_\ell}^{\underline{z}_{\ell+1}} n_\ell(\zeta)\mathrm{d}\zeta, \tag{35}$$

subject to the boundary condition $r_1 = \left(\frac{M-\sum_{\ell>1} M_\ell}{\bar{H}}\right)^{1/\phi}$.

Proof. I first prove (35). Fix 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 the locations are inversely ordered by θ_ℓ, such ℓ' exists for all $\ell \in \{2, \ldots, L\}$. Lemma B.2, Lemma B.3, and feasibility (8) together imply that $\underline{z}_\ell \in \text{supp}\,\Gamma_{\ell'}$. Profit maximization (20) then requires $\pi_\ell(\underline{z}_\ell) = \pi_{\ell'}(\underline{z}_\ell)$. Re-arranging: $r_\ell = (\underline{z}_\ell - \underline{w}_\ell)n_\ell(\underline{z}_\ell) - \pi_{\ell'}(\underline{z}_\ell)$. Meanwhile, the fundamental theorem of calculus implies $\pi_{\ell'}(\underline{z}_\ell) = \pi_{\ell'}(\underline{z}_{\ell'}) + \int_{\underline{z}_{\ell'}}^{\underline{z}_\ell} \pi'_{\ell'}(\zeta)d\zeta = \pi_{\ell'}(\underline{z}_{\ell'}) + \int_{\underline{z}_{\ell'}}^{\underline{z}_\ell} n_{\ell'}(\zeta)d\zeta$ where the second equality follows from Lemma B.1. Using $\pi_{\ell'}(\underline{z}_{\ell'}) = (\underline{z}_{\ell'} - \underline{w}_{\ell'})n_{\ell'}(\underline{z}_{\ell'}) - r_{\ell'}$ and combining with the original expression yields (35). Housing market clearing in each location yields $r_\ell = \left(\frac{M_\ell}{\bar{H}}\right)^{1/\phi}$. Feasibility requires $\sum_\ell M_\ell = M$. The two expressions together generate the boundary condition r_1. $\square$

Lemma B.7 (Mass of employers)**.**
Suppose that $\lambda^u = \lambda^e$*. Then,* $\theta_\ell < \theta_{\ell'} \iff r_\ell > r_{\ell'} \iff M_\ell > M_{\ell'}$*.*

Proof. The second equivalence directly follows from $r_\ell = (M_\ell/\bar{H})^{1/\phi}$, so we only need to focus on the first. I first prove $\theta_\ell < \theta_{\ell'} \Rightarrow r_\ell > r_{\ell'}$. Take two cities ℓ and ℓ' that are adjacent in the $\boldsymbol{\theta}$-space with $\theta_\ell < \theta_{\ell'}$. Lemma B.3 implies $\underline{z}_\ell > \underline{z}_{\ell'}$. Since $n_{\ell'}$ is strictly increasing, equation (35) implies

$$r_\ell - r_{\ell'} > (\underline{z}_\ell - b)\, n_\ell(\underline{z}_\ell) - (\underline{z}_{\ell'} - b)\, n_{\ell'}(\underline{z}_{\ell'}) > 0,$$

where I have also used $\underline{w}_\ell = b$ when $\zeta = 1$, and the second inequality follows from $n_\ell(\underline{z}_\ell) > n_{\ell'}(\underline{z}_{\ell'})$ and $\min_\ell \underline{z}_\ell > b$ by assumption.

I now prove $\theta_\ell < \theta_{\ell'} \Leftarrow r_\ell > r_{\ell'}$. Suppose not $\theta_\ell < \theta_{\ell'}$, i.e., $\theta_\ell \geq \theta_{\ell'}$. If $\theta_\ell > \theta_{\ell'}$, the previous argument yields $r_\ell < r_{\ell'}$. If $\theta_\ell = \theta_{\ell'}$, then Lemma B.3 implies $\Gamma_\ell = \Gamma_{\ell'}$, and equation (35) yields $r_\ell = r_{\ell'}$. Together, we thus have $r_\ell \leq r_{\ell'}$, i.e, not $r_\ell > r_{\ell'}$. □

Lemma B.8 (City ordering)**.**
Suppose that $\lambda^u = \lambda^e$*. Then,* $e_\ell > e_{\ell'} \iff \theta_\ell < \theta_{\ell'}$*.*

Proof. I first show $e_\ell > e_{\ell'} \Rightarrow \theta_\ell < \theta_{\ell'}$. For the sake of contradiction, suppose that $\theta_\ell \geq \theta_{\ell'}$. Since $e_\ell > e_{\ell'}$, it must be that $M_\ell > M_{\ell'}$, a contradiction with Lemma B.7.

I now turn to $e_\ell > e_{\ell'} \Leftarrow \theta_\ell < \theta_{\ell'}$. Suppose not $e_\ell > e_{\ell'}$, i.e. $e_\ell \leq e_{\ell'}$. If $e_\ell < e_{\ell'}$, the previous argument implies $\theta_\ell > \theta_{\ell'}$, i.e., not $\theta_\ell < \theta_{\ell'}$. So suppose $e_\ell = e_{\ell'}$. Then, $\theta_\ell/\theta_{\ell'} = M_\ell > M_{\ell'} > 1 \iff M_\ell > M_{\ell'}$, which also contradicts Lemma B.7. □

B.8 A competitive spatial matching model

Consider a general spatial matching model with competitive labor markets. There is a measure M of heterogeneous firms indexed by their productivity z distributed according to Γ. Firms use a single input, labor. They face a decreasing returns to scale technology: $R(z, n) = zn^\rho$. Decreasing returns could arise due to span-of-control costs or love-for-variety across differentiated goods.[39]

There are L cities. Cities differ in their size, $\{m_\ell\}_{\ell=1}^L$, and their revenue TFP, $\{T_\ell\}_{\ell=1}^L$. City size can either be exogenous or endogenous (e.g., free entry, preference shocks, migration costs, etc.). Likewise, local TFPs can either be exogenous or endogenous (e.g., productivity spillovers, market access, etc.). Both local characteristics are treated as given by firms.

Labor markets are competitive and segmented by locations. The law of one price holds within locations, and the local wage is denoted w_ℓ. Employers pay housing cost r_ℓ to produce in location ℓ. The housing supply is $L_\ell = \bar{L} r_\ell^\chi$.

[39] Decreasing returns to scale are required for the firm boundary to be well-defined in the absence of search frictions. The model of Section 2 could be extended to allow for DRS without much consequences (see Bilal and Lhuillier, 2021).

Employers solve

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

Their labor demand and optimal profits are

$$n_\ell(z) = \left(\frac{T_\ell z \rho}{w_\ell}\right)^{\frac{1}{1-\rho}} \quad \text{and} \quad \pi(z) = \max_\ell \; \kappa_1 \psi_\ell z^{\frac{1}{1-\rho}} - r_\ell,$$

where $\kappa_1 \equiv (1-\rho)\rho^{\frac{\rho}{1-\rho}}$ is a parametric constant, and ψ_ℓ summarizes the profitability of ℓ:

$$\psi_\ell \equiv \left(\frac{T_\ell}{w_\ell^\rho}\right)^{\frac{1}{1-\rho}}.$$

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. As in the model of Section 2, the spatial allocation of firms is determined by the profit condition (20) and the feasibility condition (8). Finally, labor market clearing in every location demands

$$m_\ell = M_\ell \int n_\ell(z) d\Gamma_\ell(z) = M_\ell \left(\frac{T_\ell \rho}{w_\ell}\right)^{\frac{1}{1-\rho}} \mathbb{E}\left[z^{\frac{1}{1-\rho}}\right]. \tag{36}$$

The allocation of firms across space is determined by the complementarity between firms' productivity and local profitability, ψ_ℓ. More productive firms have a higher willingness to pay to access high local profitability. The profitability of a location depends in turn on its local revenue TFP, and on the price of the inputs. Absent revenue TFP, productive firms thus sort across space to access cheap labor. As a result, there is a negative correlation between the average firm productivity and local wages. In equilibrium, large cities attract more employers, which are relatively more productive, because they offer cheap labor.

The next proposition states this result formally.

Proposition B.1 (Spatial wage inequality in competitive matching models)**.**
Suppose that revenue TFPs are homogeneous, $T_\ell = 1$. In any equilibrium,

- *Employers are more productive in larger cities, and they are more of them: $m_\ell > m_{\ell'}$ implies $M_\ell > M_{\ell'}$ and $\mathbb{E}_\ell[z] > \mathbb{E}_{\ell'}[z]$;*
- *Wages are lower in larger cities: $m_\ell > m_{\ell'}$ implies $w_\ell < w_{\ell'}$.*

Proof. Suppose that TFPs are homogeneous, $T_\ell = 1$ for all ℓ. Thus, order cities by city size without loss of generality, $m_1 < m_2 < \cdots < M_L$.

Using the equilibrium condition for wages (36), local profitability rewrites

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

where the constant of proportionally is a function of ρ.

The constant ψ_ℓ is thus 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).

With these initial remarks, we start by proving the second part of the proposition. For the sake of contradiction, suppose that wages are increasing in city size, $w_1 < w_2 < \cdots < w_L$. From the expression for wages, it must then be that

$$\frac{M_{\ell+1}}{M_\ell} \frac{\mathbb{E}_{\ell+1} \left[z^{\frac{1}{1-\rho}} \right]}{\mathbb{E}_\ell \left[z^{\frac{1}{1-\rho}} \right]} > \frac{m_{\ell+1}}{m_\ell} > 1. \tag{37}$$

That is, $M_\ell \mathbb{E}_\ell[z]$ must be increasing in m_ℓ. We also know from Topkis (1998) 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 for (37) is therefore that M_ℓ is increasing in m_ℓ.

In any equilibrium with pure sorting, marginal firms must be indifferent between two locations. Since w_ℓ is increasing in m_ℓ and firms sort in the opposite direction of w_ℓ, this indifference is $\pi_\ell(\underline{z}_\ell) = \pi_{\ell+1}(\bar{z}_{\ell+1})$. Using the expression for profits and housing prices, the indifference condition reads

$$\kappa_2 \left[\psi_\ell - \psi_{\ell+1} \right] \underline{z}_\ell^{\frac{1}{1-\rho}} = \left(\frac{M_\ell}{\bar{L}} \right)^{1/\chi} - \left(\frac{M_{\ell+1}}{\bar{L}} \right)^{1/\chi} > 0,$$

for $\kappa_2 > 0$ a parametric constant. Therefore, M_ℓ is decreasing in ℓ, and (37) does not hold –a contradiction.

The first part of the proposition automatically follows from Topkis (1998). $\square$

C Extended model and estimation

C.1 Quantitative model

Let K 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. The expected lifetime utility of a worker born in l upon entry in the labor force is

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

where $\beta \equiv \rho + \xi$ is workers' effective discount rate, and $\kappa_{l\ell}$ is the migration cost from l to ℓ. Workers are free to migrate as long as they are unemployed. Employed workers need to quit their jobs to migrate.[40] Together with the facts that the economy is in steady state and that idiosyncratic preferences are time invariant, workers settle in a location upon their entry in the labor force and 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

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

These HJB equations are independent of workers' taste shocks, and therefore, need not 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, workers decide where to settle given the expected lifetime utility of unemployed workers in each location:

$$U_l(\boldsymbol{\omega}) = \max_{\ell} \ \frac{\kappa_{l\ell}\omega_\ell A_\ell U_\ell}{P_\ell}.$$

This location problem is very tractable due to three features. First, since the economy is in steady state, the measure of people who live in ℓ upon entry is the same as the stationary measure of people in ℓ. Second, the birth process implies that the measure of workers born in ℓ is the same as the measure of people living in ℓ. Third, the Fréchet shocks smooth the discreteness of the location choice. With these remarks in mind, let $m_{l\ell}$ denote the measure of workers born in l and who choose to start their career in ℓ. The spatial allocation of new entrants is given by

$$\frac{m_{l\ell}}{m_\ell} = \frac{[\kappa_{l\ell}A_\ell U_\ell/P_\ell]^\chi}{\sum_{\ell'}[\kappa_{l\ell'}A_{\ell'}U_{\ell'}/P_{\ell'}]^\chi}. \tag{38}$$

Then, the total measure of people living in ℓ is $m_\ell = \sum_l m_{l\ell}$.

[40] I assume a homogeneous migration cost, $\kappa_{l\ell} = \kappa 1\{\ell \neq l\}$. However, since the model is solved at the city-group level, these migration costs have to be adjusted to reflect that there is more than one location per group. See Section C.3 for more details.

Search The derivations of the local labor supply curves and reservation wages follow Section B.1 and Section B.2 with the addition of exit shocks. The reservation wage thus reads

$$\underline{w}_\ell = Kb + (\lambda^u_\ell - \lambda^e_\ell) \int_{\underline{w}_\ell} \frac{\bar{F}_\ell(w)}{\rho + \xi + \delta + \lambda^e_\ell \bar{F}_\ell(w)} dw. \tag{39}$$

The ins- and outs- of employment deterime the local measure of employed workers

$$u_\ell = \left(\frac{\delta + \xi}{\lambda^u_\ell + \delta + \xi} \right) m_\ell. \tag{40}$$

Likewise, the flows up the job ladder determine the distribution of wages amongst employed workers,

$$G_\ell(w) = \frac{(\delta + \xi) F_\ell(w)}{\delta + \xi + \lambda^u_\ell \bar{F}_\ell(w)}.$$

These three expressions make clear that it is the joint separation rate, $\bar{\delta} = \delta + \xi$, that matters for the flows of workers across employment and jobs. I therefore define $k_\ell = \lambda^e_\ell / \bar{\delta}$ with a slight abuse of notation.

Employers Employers solve (23) where $c(v) = Kv^{1+\gamma}/(1+\gamma)$. The vacancy optimality condition demands

$$(zT_\ell - w_\ell(z))\, n_\ell[w_\ell(z)] = K v_\ell(z)^\gamma. \tag{41}$$

Potential profits in location ℓ after maximizing out vacancies are

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

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 $z \in \text{supp } \Gamma_\ell$, the wage optimality condition is

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

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

$$n_\ell(z) = n_\ell[w_\ell(z), v_\ell(z)] = \frac{(1+k_\ell)e_\ell}{[1+k_\ell(1-\Upsilon_\ell(z))]^2}\frac{v_\ell(z)}{V_\ell}. \tag{42}$$

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

$$w'_\ell(z) = \frac{2k_\ell \mathrm{d}\Upsilon_\ell(z)}{1+k_\ell\bar{\Upsilon}_\ell(z)}[zT_\ell - w_\ell(z)]. \tag{43}$$

Given Υ_ℓ and the boundary condition $w_\ell(\underline{z}_\ell) = \underline{w}_\ell$, (43) 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:

$$\Upsilon'_\ell(z) = \frac{M_\ell}{V_\ell}\left[\left(\frac{zT_\ell - w_\ell(z)}{K}\right)n_\ell(z)\right]^{\frac{1}{\gamma}}\Gamma'_\ell(z). \tag{44}$$

Given the condition $\Upsilon_\ell(\underline{z}_\ell) = 0$ and $w_\ell(z)$, (44) 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/K. Given their Gumbel distribution, the probability that a firm with productivity z produces in city ℓ is

$$\Omega_\ell(z) = \frac{\mathrm{e}^{\vartheta\pi_\ell(zT_\ell)/K}}{\sum_{\ell'} g_{\ell'}\mathrm{e}^{\vartheta\pi_{\ell'}(zT_{\ell'})/K}}. \tag{45}$$

The local productivity distribution in city ℓ is then

$$\Gamma_\ell(z) = \frac{M}{M_\ell}\int^z \Omega_\ell(x)\mathrm{d}\Gamma(x) \quad \Longleftrightarrow \quad \mathrm{d}\Gamma_\ell(z) = \frac{M}{M_\ell}\Omega_\ell(z)\mathrm{d}\Gamma(z), \tag{46}$$

and the mass of firm in city ℓ is

$$M_\ell = M\int \Omega_\ell(x)\mathrm{d}\Gamma(x). \tag{47}$$

The idiosyncratic entry costs ensure full support of the local productivity distribution, so that the boundary condition to (45) is $\Gamma_\ell(\underline{z}) = 0$.

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

$$\lambda^u_\ell = \mu\left(\frac{V_\ell}{u_\ell + \zeta e_\ell}\right)^\xi. \tag{48}$$

Aggregate prices There are three sets of aggregate quantities. The reservation wages $\{\underline{w}_\ell\}_{\ell=1}^L$ are given by (14). Commercial housing prices clear the commercial housing market:

$$\frac{\alpha}{p_\ell}\left(u_\ell bK + e_\ell \mathbb{E}_\ell[w]\right) = \bar{L}p_\ell^\theta. \tag{49}$$

The residential housing prices clear the residential housing market:

$$M_\ell = \bar{H}(r_\ell/K)^\phi, \tag{50}$$

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

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* (43)*,* (44) *and* (46)*, subject to the appropriate boundary conditions and the spatial allocation of employers* $\boldsymbol{\Omega}(z)$ *given by* (45)*;*
2. *The vacancy and size policy functions satisfy* (41) *and* (42)*;*
3. *The mass of vacancy and employers are given by* (22) *and* (47)*;*
4. *The mass of unemployed and employed workers are given by* (40) *and* $e_\ell = m_\ell - u_\ell$*;*
5. *The contact rates are given by* (48)*;*
6. *The reservation wages are given by* (39)*;*
7. *The residential housing prices follows* (49) *and* (50)*.*

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 $u_\ell = \delta/(\delta+\lambda^u_\ell)m_\ell$ and $m_\ell = u_\ell + e_\ell = (\lambda^u_\ell+\delta)e_\ell/\delta$. 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 $1/\zeta\lambda^u_\ell$, which we know. The above expression thus identifies ζ. Given ζ, we separately identify $\{\lambda^u_\ell\}_\ell$.

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

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

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

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

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 $\text{RE}_\ell = Kb/\mathbb{E}[w_\ell]$. The aggregate replacement rate is therefore

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

Inverting this expression identifies Kb.

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, $M_\ell\mathbb{E}_\ell[n_j] = e_\ell$ holds. The average size of employers relative to the measure of employed workers thus identifies the number of employers: $M_\ell = e_\ell/\mathbb{E}_\ell[n_j]$. The total measure of employers follows $M=\sum_\ell m_\ell$.

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 (12) 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:

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

In practice, I approximate $w \to F_\ell$ with a spline and use the approximation to compute $\partial w/\partial F$. Figure C.13 shows the quality of the spline approximation for the cities of Lens and Paris. Figure C.14 presents the distribution of wages and estimated MPLs for Paris and Lens. The MPL distributions are more dispersed with a thicker right tail than the wage distributions.

Given data on vacancy shares, the vacancy cost elasticity is identified the vacancy optimality condition. Equation (41) indeed implies

$$\gamma = \sqrt{\frac{\sum_\ell e_\ell \mathbb{V}\text{ar}_\ell[\log\left((\zeta_j - w_j)\eta_j\right)]}{\sum_\ell e_\ell \mathbb{V}\text{ar}_\ell[\log \frac{v_j}{V_\ell}]}}. \tag{51}$$

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

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

From (45), the log probability that an employer with total productivity ζ produces in ℓ is

$$\log \Omega_\ell^\zeta(\zeta) = \vartheta\pi_\ell(\zeta) - \log \sum_{\ell'} \mathrm{e}^{\vartheta\pi_{\ell'}(\zeta T_{\ell'}/T_\ell)} \approx \vartheta\pi_\ell(\zeta) + H(\zeta), \tag{52}$$

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 (52), 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 estimate (52) with three modifications. First, I instrument local profit opportunities with employers' wage to reduce measurement errors. The instrument is relevant since, within a city, wages, productivity, and profits are all monotonic functions of each other. The instrument satisfies the exclusion restriction because wages are raw data. Second, I cluster employers in groups of MPL within cities and replace $H(\zeta)$ with a group fixed effect. I use five groups per city. Third, to account for possible dispersion in TFP –or other unobserved local characteristics that affect employers' location choice– I introduce location fixed effects. Combining these three modifications, I estimate

$$\log \Omega_{q\ell} = \mathrm{FE}_\ell + \mathrm{FE}_q + \beta \pi_{q\ell} + u_{\ell q}, \tag{53}$$

where q is a group, $\Omega_{q\ell}$ is the share of firms q in ℓ, and $\pi_{q\ell}$ is the average profits of employers in group q in location ℓ. I estimate (53) with 2SLS, instrumenting $\pi_{q\ell}$ by the average wage offered by employer in group q in location ℓ. Figure C.15 shows the first-stage, reduced-form, and second-stage of the IV procedure in the data. The markers are the data, and the lines the linear fit.

As explained in the main text, to account for the possibility that the TFP differentials are not entirely soaked up by the location fixed effects, as well as the measurement error introduced by the discretization, I replicate (53) in the model. I then set ϑ to match the 2SLS coefficient obtained in the data.

Matching function The matching function (48) implies

$$\log \lambda_\ell^u = \log \mu + \psi \log \left(\frac{V_\ell}{u_\ell + \zeta e_\ell} \right) + u_\ell, \tag{54}$$

for ϵ_ℓ some measurement error. Estimating (54) 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

$$\log p_\ell = \left(\frac{1}{1+\theta} \right) \log \left(\frac{\alpha}{\bar{L}} \right) + \left(\frac{1}{1+\theta} \right) \log I_\ell, \tag{55}$$

where $I_\ell = u_\ell K b + e_\ell \mathbb{E}[w_\ell]$ 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 $\bar{H}(r_\ell/K)^\phi = m_\ell$, or

$$\log r_\ell = \log K - \frac{1}{\phi} \log \bar{H} + \frac{1}{\phi} \log M_\ell. \tag{56}$$

Estimating (56) by OLS identifies ϕ and a bundle of $(K, \bar{H})$. Figure C.16 shows the fit of the model with the data. Although simple, the model covers 88% and 93% of the empirical variation in residential and commercial housing prices.

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

The bilateral flows of workers born in l and working in ℓ relative to the number of stayers identify the migration costs given the dispersion in preferences:

$$\frac{m_{l\ell}}{m_{ll}}\frac{m_{\ell l}}{m_{\ell\ell}} = \left(\frac{\kappa_{l\ell}}{\kappa_{ll}}\frac{\kappa_{\ell l}}{\kappa_{\ell\ell}}\right)^{\chi}.$$

Once assumed that migration costs are uniform across space and scaled to reflect the presence of multiple locations within city groups (see Section C.3), the above equation simplifies to

$$\frac{m_{l\ell}}{m_{ll}}\frac{m_{\ell l}}{m_{\ell\ell}} = g_l g_\ell \left(\frac{\kappa^{2\chi}}{[1+(g_l-1)\kappa^{\chi}][1+(g_\ell-1)\kappa^{\chi}]}\right), \quad \forall \ell \neq l,$$

where g_ℓ is the number of locations in city group ℓ.

Finally, amenities are identified from the location choice of workers. Equation (38), together with $m_\ell = \sum_l m_{l\ell}$, imply

$$\nu_\ell = m_\ell \left(\sum_l \frac{m_l \kappa_{l\ell}^{\chi}}{\sum_{\ell'} \kappa_{l\ell'}^{\chi}\nu_{\ell'}}\right)^{-1},$$

where $\nu_\ell \equiv (A_\ell U_\ell / P_\ell)^{\chi}$. Given $(\kappa, \chi, \boldsymbol{m})$, the above equation identifies $\boldsymbol{\nu}$ up to a normalization. We normalize mean amenities to unity. Given $(\chi, \boldsymbol{\nu})$, housing prices $\boldsymbol{P}$, and the net present value of unemployed's lifetime earnings, $\boldsymbol{U}$, we can invert $\boldsymbol{\nu}$ to obtain amenities. We compute unemployed's lifetime earnings from

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

where all the elements on the right-hand side have already been identified. This conclude the proof.

C.3 Data

Wages I measure wages through the job fixed effects estimated in (1) on the panel-version of the matched employer-employee dataset (see Section 1.1 and A.1). To be consistent with that model and the facts documented in Section 1, I define the size of a job as the average number of workers per job within each job cluster.[41]

Job flows Worker flows across jobs are also obtained from the panel-version of the matched employer-employee dataset. To be consistent with the AKM model (1), I define switches as the job-group level. That is, I set $\text{J2J}_{it} = 1$ if worker i (i) switches job cluster within 90 days, (ii)

[41] The fact that the clusters are employment weighted within cities mean that the total number of workers per cluster is constant. However, the average number of workers per job in each cluster can still vary.

that job switch takes place within the city of origin, and (iii) the switch is associated with a wage increase.[42]

Young, educated workers are known to transition more frequently across jobs. They are also known to be disproportionally present in large cities. I therefore residualize job flows to account for this worker heterogeneity. Specifically, I estimate

$$\mathrm{J2J}_{it} = \alpha_i + \beta a_{it} + \gamma_{\ell(i,t)} + u_{it},$$

where α_i is a worker fixed effects, a_{it} is a control for age, and γ_ℓ is a location fixed effect. I then define the local average job switching rate as the location fixed effect.

Housing prices I measure residential housing prices via rents. I obtain data on rents from the "Carte des Loyers" (Rental Map), which I aggregate at the commuting zone level.

I do not have data on commercial housing prices. Rather, I suppose that the relative housing prices are the same in the commercial and residential markets. I adjust the scale of the commercial housing prices in two ways. First, I adjust them by the mean difference in housing prices between both markets. Second, I adjust commercial housing prices by the average housing size in the commercial market since employers do not face an intensive margin. Both statistics are obtained from the Valeur Foncières.

I also adjust residential housing prices to capture the presence of worker heterogeneity in the data. To do so, I extend the model to allow for skill heterogeneity. Suppose workers differ in their skill, s, whose distribution in location ℓ is Φ_ℓ. Employers cannot target their hiring to specific skills, and skills shift wages proportionally: a worker with skill s employed at firm z in location ℓ earns $sw_\ell(z)$.

This extended model as two key properties. First, the AKM specification is well-specified, and the worker fixed effects in (1) estimate s. Second, only the residential housing block of the model is affected by worker heterogeneity. Specifically, residential housing expenditures in ℓ are now $p_\ell L^d_\ell = \alpha \mathbb{E}_\ell[s](u_\ell b + e_\ell \mathbb{E}_\ell[w])$. The market clearing condition (55) becomes

$$\log p_\ell = \frac{1}{1+\theta} \log \frac{\alpha}{\bar{L}} + \frac{1}{1+\theta} \log(u_\ell b + e_\ell \mathbb{E}_\ell[w]) + \frac{1}{1+\theta} \log \mathbb{E}_\ell[s]. \tag{57}$$

Accordingly, we can estimate θ and $\bar{L}$ while controlling for $\mathbb{E}_\ell[s]$ using the average worker fixed effects in each location. Since this extension eventually only matters for residential housing prices, I do not use elsewhere and revert to the original model without skill heterogeneity.s

Migration flows I obtain migration flows from the panel-version of the matched employer-employee dataset. Specifically, I know workers' birthplace and their current workplace, which allows me to compute $m_{l\ell}$ for every pair of city clusters (l, ℓ).

[42] Alternatively, I could focus on switches associated with an increase in the job fixed effect. The two statistics are closely aligned.

Migration costs have to be adjusted to reflect the fact that the model is solved at the city-cluster level. Migration costs indeed capture the cost to move away from your birthplace. However, absent any adjustments, workers born in clusters with many locations would face a lower migration costs as they could move freely across locations within a cluster. The adjustment can be derived by writing down two models, one at the city level in which each city within a cluster is homogeneous, and another model at the city-cluster level. There exists a unique adjustment to the migration costs in the city-cluster model that render both models identical:

$$\kappa_{l\ell} = \begin{cases} \left(\dfrac{1+(g_l-1)\kappa^{\chi}}{g_l}\right)^{\frac{1}{\chi}} & \text{if } l = \ell, \\ \kappa & \text{if } l \neq \ell, \end{cases}$$

where g_l is the number of locations in cluster l.

C.4 Figures

Figure C.13: Wage offer distributions

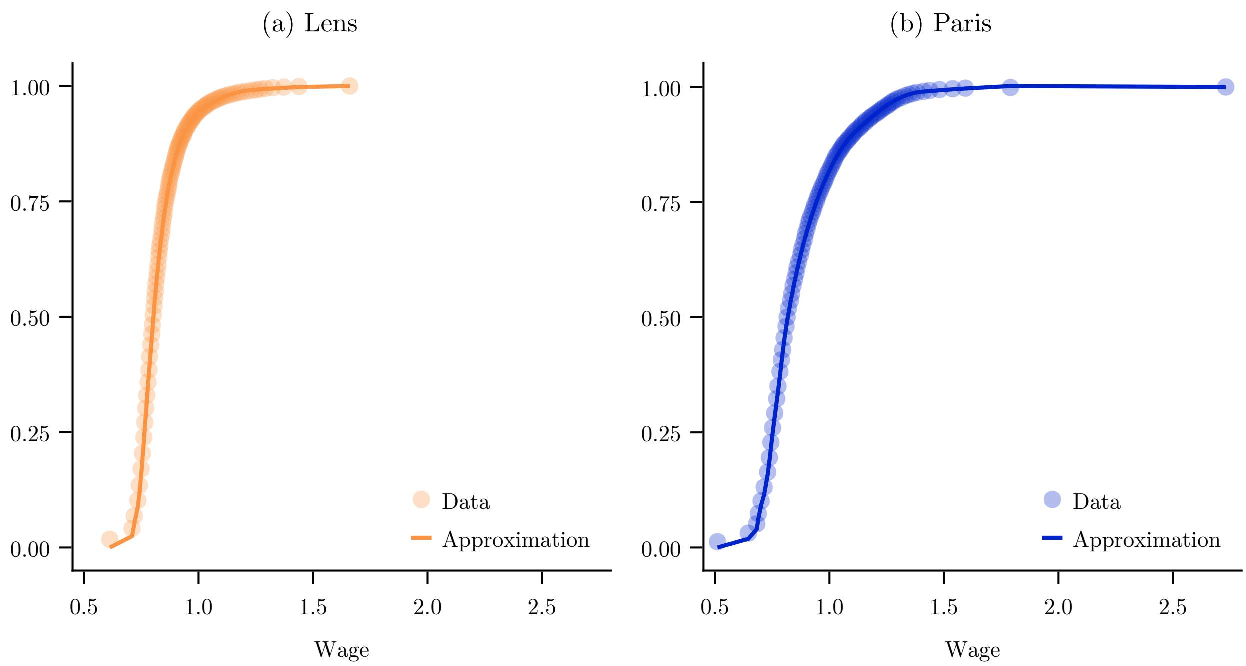


Figure C.14: Wage and marginal products of labor distributions

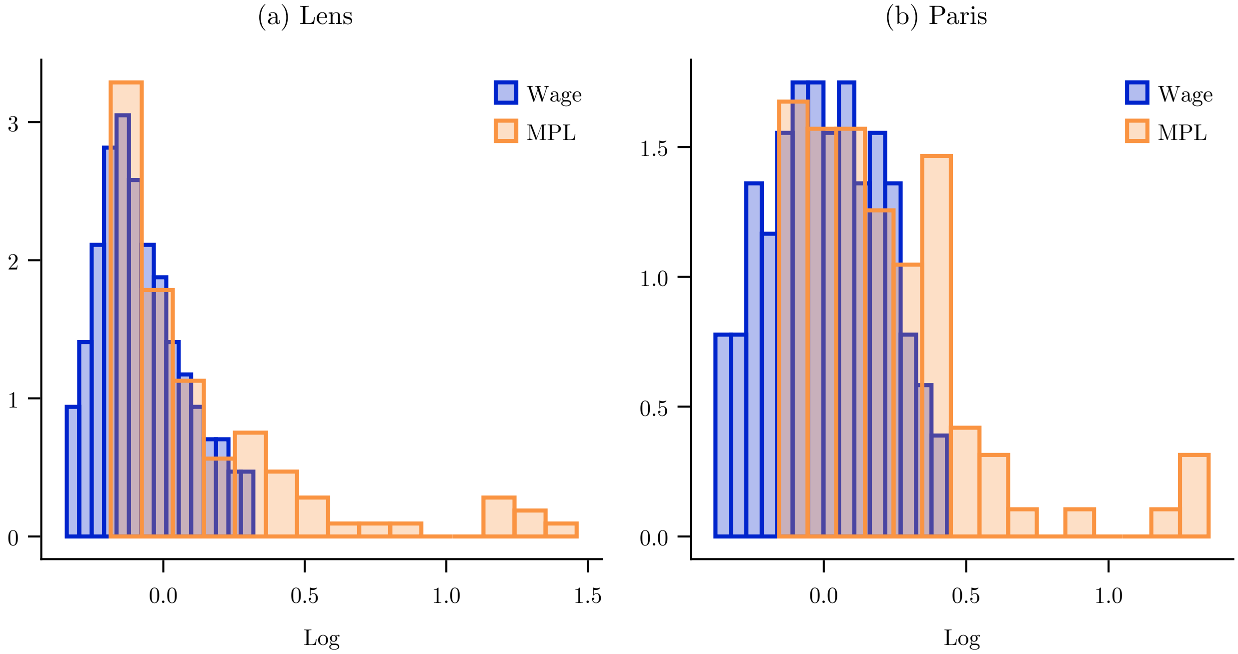


Figure C.15: Entry cost dispersion estimation

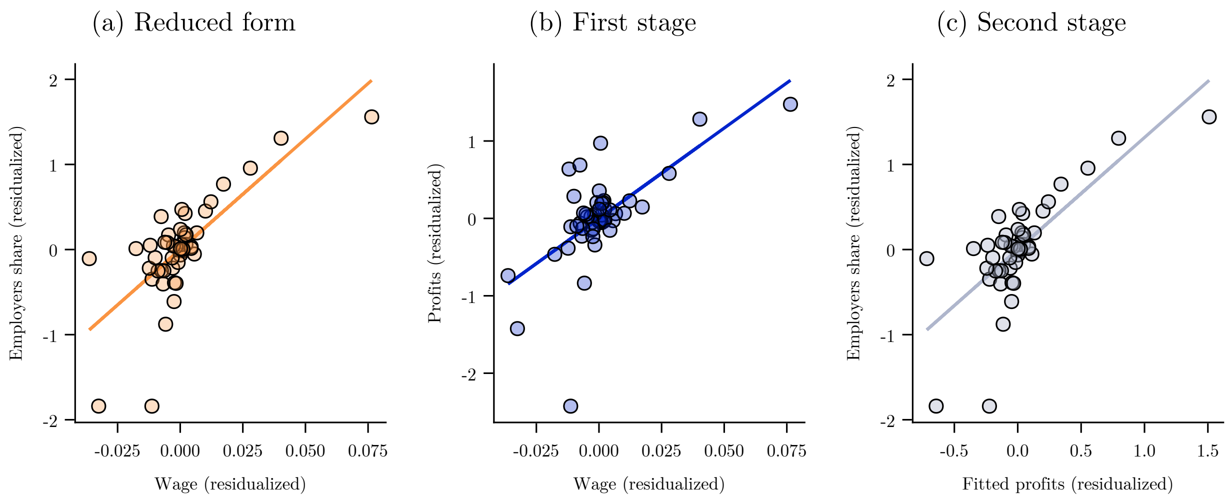


Figure C.16: Housing prices

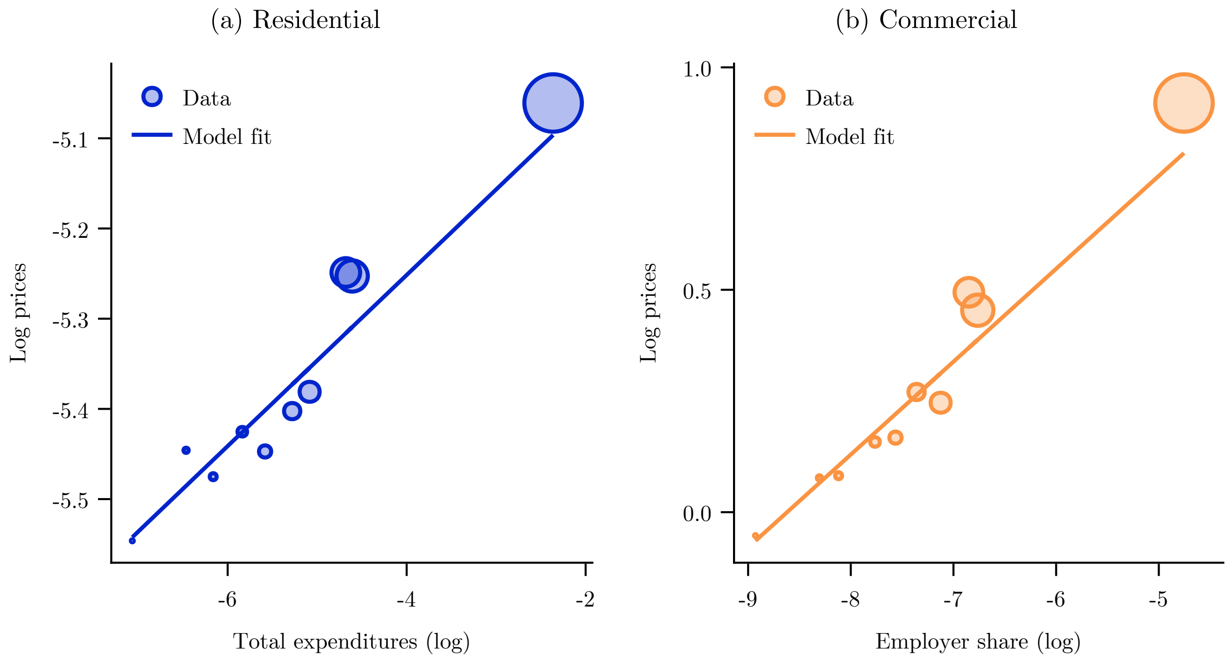


Figure C.17: Local job switching rates

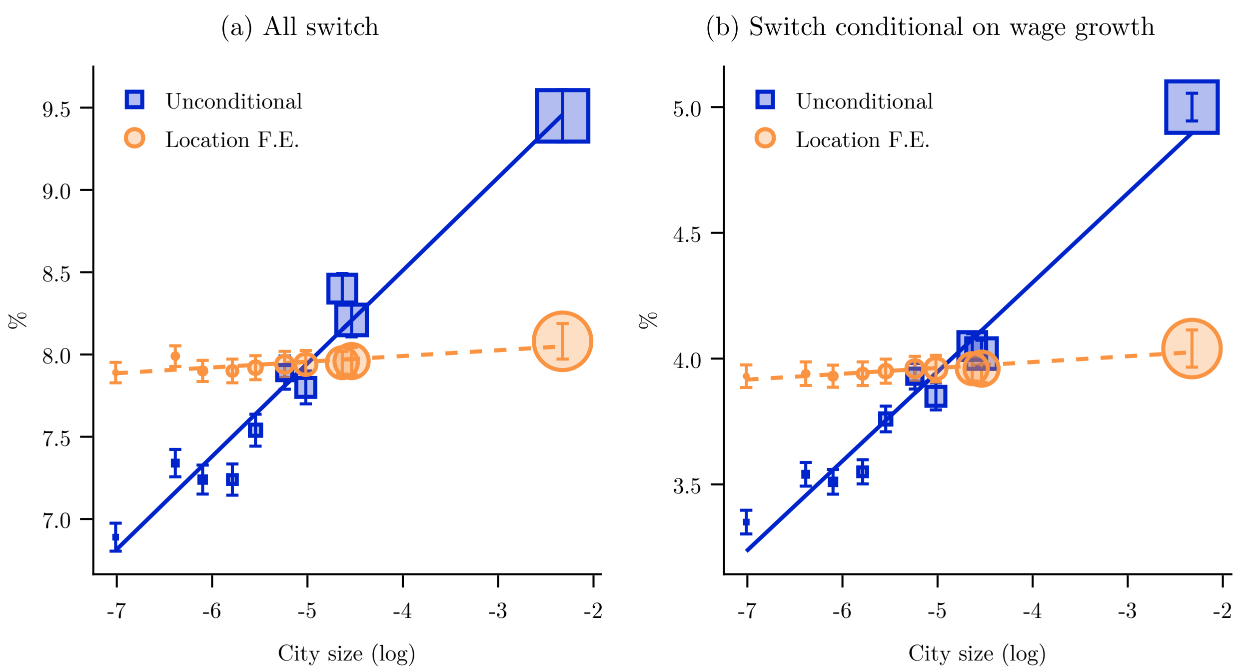
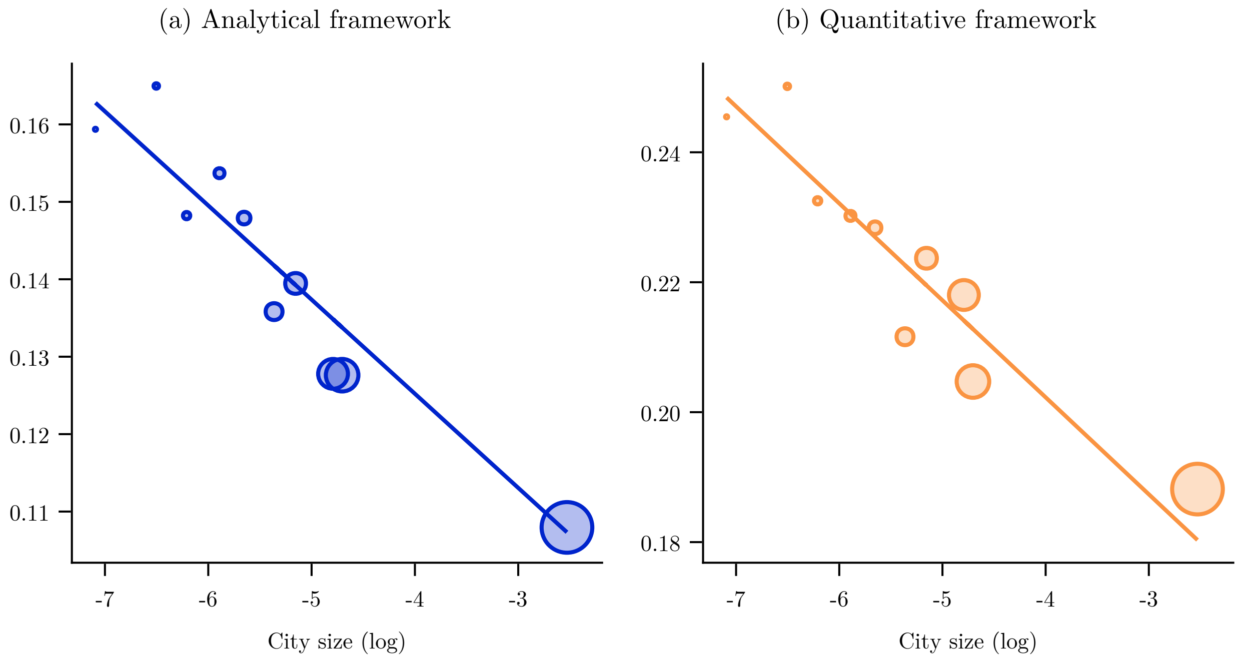


Figure C.18: Labor market tightness by city

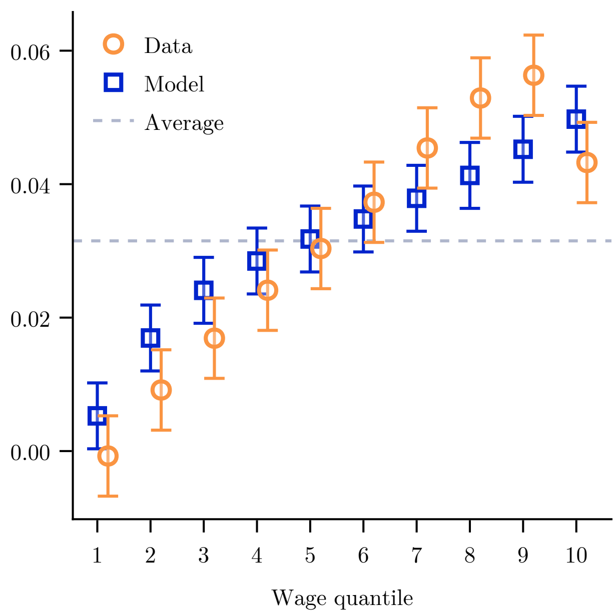


Left panel shows the labor market tigtheness as defined in the theoretical model: $\theta_\ell = e_\ell/M_\ell$. The right panel shows the labor market tightness in the quantitative model: $\theta_\ell = (u_\ell + \zeta e_\ell)/V_\ell$. Aggregate vacancies are measured in the data by comining employers' size with their rank in the job ladder (see Section C.2).

D Quantitative exercises

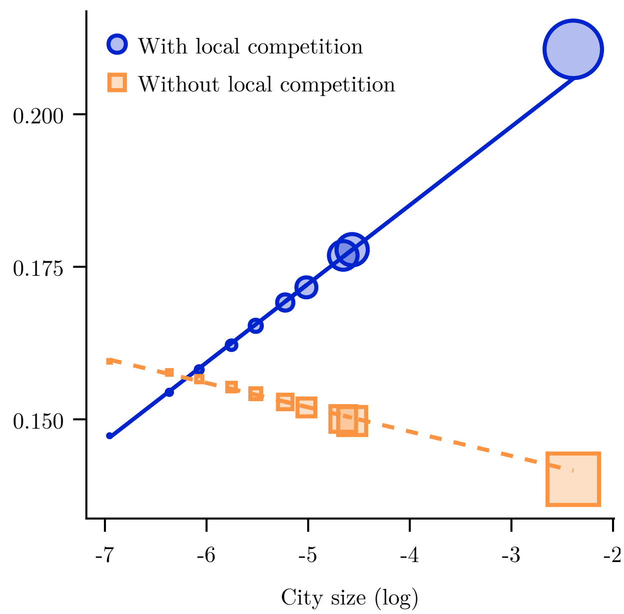
D.1 Figures

Figure D.19: Local wage distributions



Let $w_{\ell q}$ be the q-th quantile of city ℓ's wage distribution, and m_ℓ the city size. Estimate $\log w_{\ell q} = \alpha_q + \beta_q \log m_\ell + u_{\ell q}$ by OLS. The figure plots the estimated $\{\beta_q\}_q$, in orange in the data and in blue in the model.

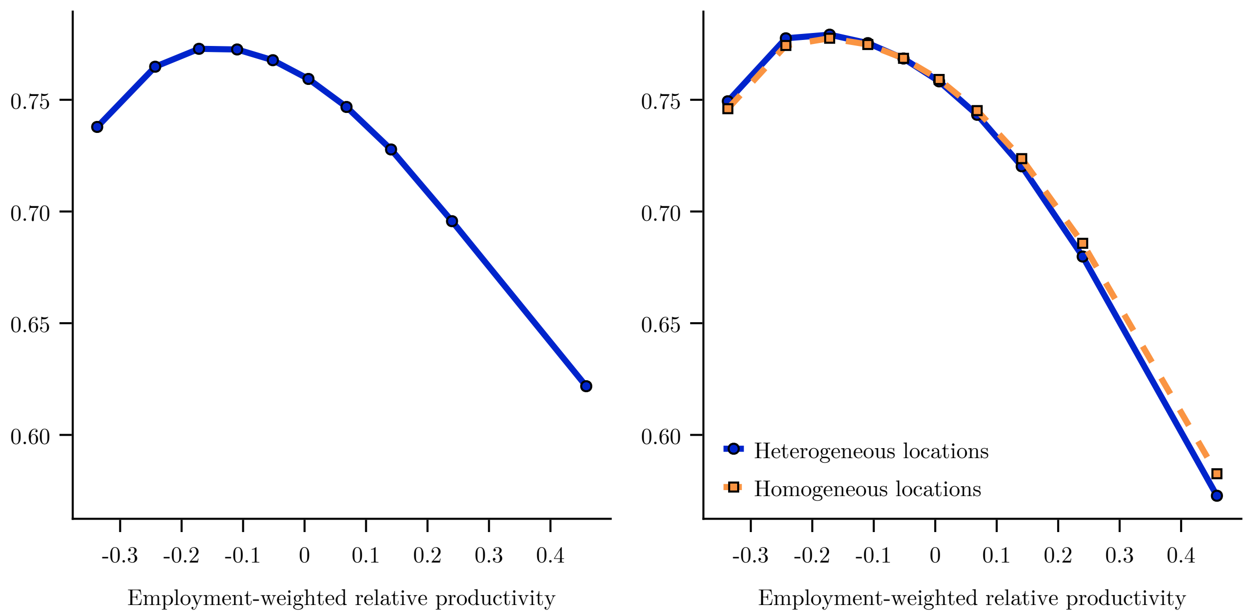
Figure D.20: Wage standard deviation



Blue circles depict the equilibrium, within-city wage standard deviation. The orange rectangles show the counterfactual wage standard deviations if employers were to price using their average markdowns, $w_\ell(z) = z\bar{\mu}(z)$.

Figure D.21: Markdowns distribution

(a) Employment-weighted (b) Unweighted



Left-panel shows the (employment-weighted) average markdown by decile of (employment-weighted) productivity. Right-panel shows in blue the unweighted average markdown ($\bar{\mu}(z) = L^1 \sum_\ell \mu_\ell(z)$) by decile of (employment-weighted) productivity. Orange line depicts the equilibrium markdown in a counterfactual economy with homogeneous locations.

Figure D.22: Search frictions and the spatial distribution of activity

(a) Lifetime earnings (b) Average productivity

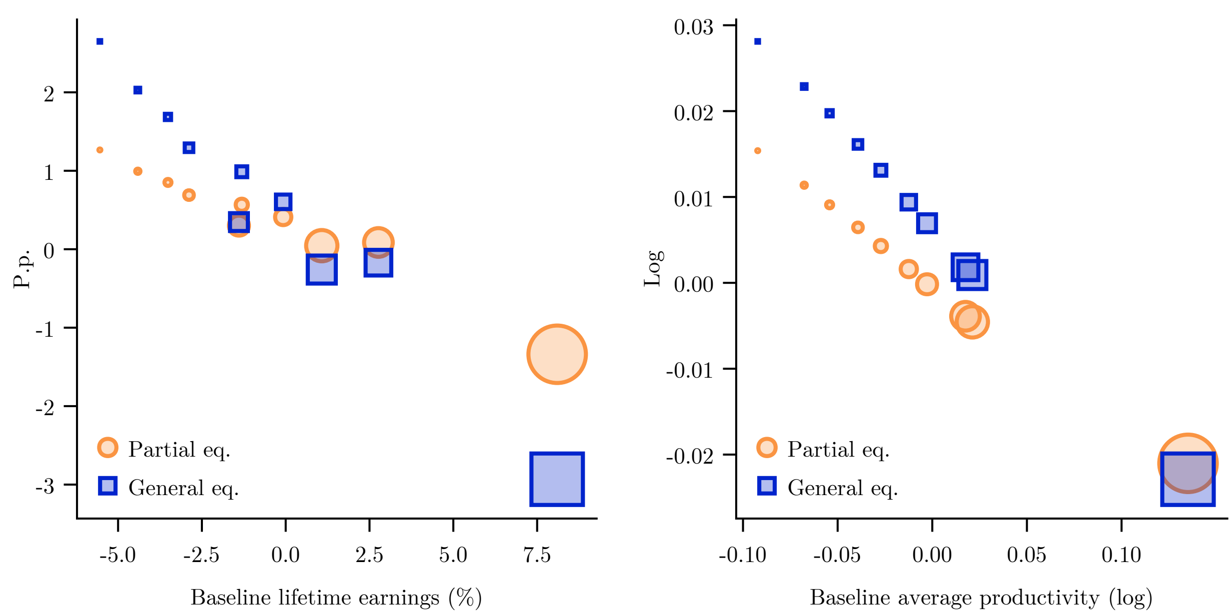


Figure D.23: The consequences of search frictions on workers

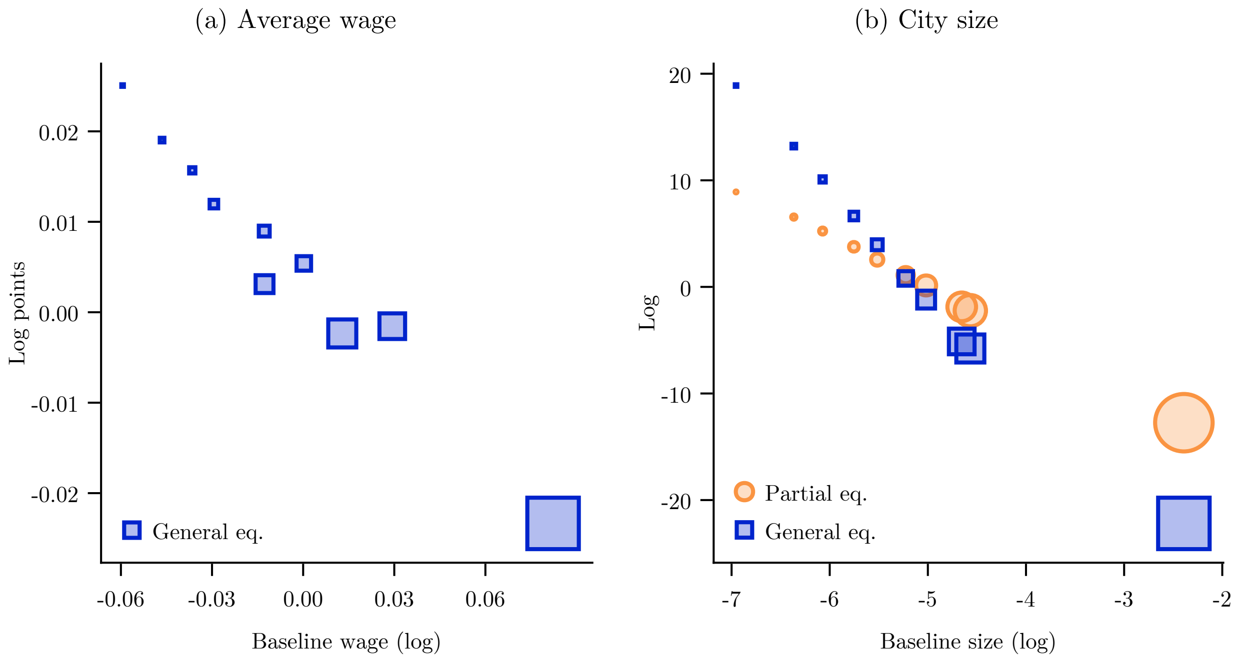


Figure D.24: The consequences of search frictions on employers