# The Local Root of Wage Inequality

Hugo Lhuillier*

August 26, 2022

**PRELIMINARY AND INCOMPLETE.**

**Abstract**

I study the role of cities in shaping wage inequality. I document three novel facts that serve to motivate my analysis. First, high-paying jobs are much more spatially concentrated than low-paying jobs. Second, local search frictions generate city-specific job ladders, and the concentration of high-paying jobs in a few cities renders their ladders steeper. Third, high-paying jobs are concentrated in locations with larger firm-size wage premia. To rationalize these facts, I develop a framework of employer sorting across frictional local labor markets. Positive assortative matching prevails: productive jobs are located in larger cities, their higher willingness-to-pay leads to fiercer local competition, and productive employers offer high wages to attain their desired size. Altogether, firm sorting generates greater inequality in larger locations. I estimate the model and quantify the fraction of aggregate inequality caused by the spatial sorting of employers. Finally, I study the consequences of a place-based policy subsidizing job creation on local and national wage inequality.

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

In this paper I argue that cities matter for wage inequality. Most interactions between employers and workers take place within the city where both parties reside, and these interactions are shaped by the conditions of the local labor market. For instance, employers have to offer higher wages to poach workers from other firms in cities where the competition is fiercer, and workers have greater incentives to search for jobs in places where good job opportunities are numerous. These spatially differentiated interactions shape local and overall wage inequality.

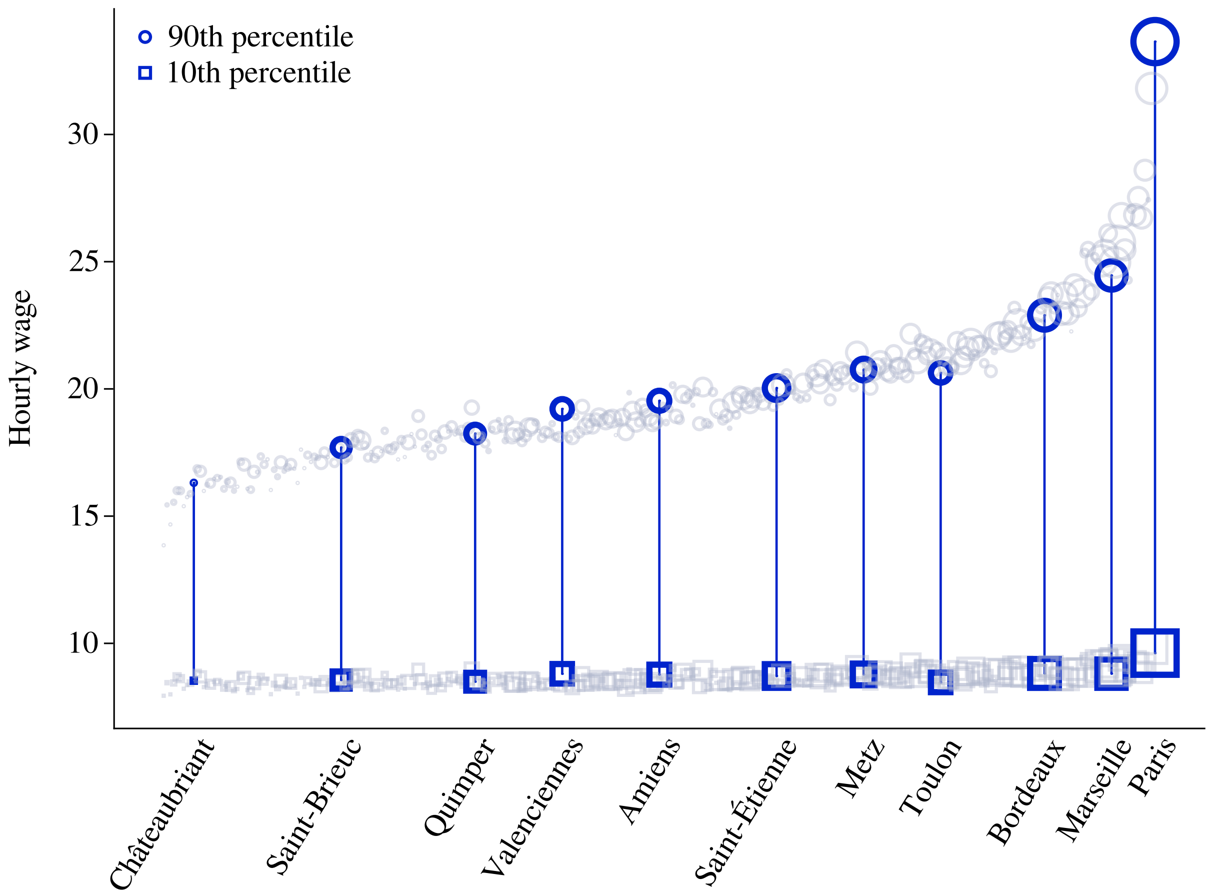
In fact, wage inequality varies substantially across cities. Figure 1 displays the distribution of hourly wages for the 300 French commuting zones. For instance, workers at the 90$^{\text{th}}$ percentile of the wage distribution in Paris are paid 70% more than workers at the 90$^{\text{th}}$ percentile in Saint-Étienne. The between-city difference in wage for workers at the 10$^{\text{th}}$ percentile is only 10%. This heterogeneous effect of space on wages across the distribution implies sizeable differences in local inequality: high-paid workers earn wages three times larger than low-paid workers in Paris, while this gap is only around two in Saint-Étienne. Why do high wages vary drastically over space while low wages do not? To what extent do cities exacerbate aggregate inequality? And what are the implications of place-based policies on local and aggregate inequality?

In this paper, I answer these questions in two parts. In the first part, I leverage administrative data from France to document three novel facts about cities and inequality. First, using a two-way fixed effect model, I find that job heterogeneity matters to understand the spatial differences in wage inequality. In particular, spatial variation in job heterogeneity explains more than 40% of the between-city differences in inequality. I further show that these differences in job heterogeneity are driven by the concentration of high-paying jobs in a few locations while low-paying jobs are present throughout the country.

Second, I show that low-paying and high-paying jobs coexist in each city due to the presence of local search frictions which prevent the immediate reallocation of workers to better-paying jobs. Job ladders specific to each city emerge as a consequence of the local frictions, and these ladders resemble the wage distributions of Figure 1: their bottom rungs are identical everywhere but the ladders are steeper in cities with greater inequality. Quantitatively, workers in cities where the variance of log wages is one standard deviation higher experience wage gains upon a job transition that are half a percentage point bigger. For instance, a job transition in Paris and in Saint-Étienne increases wages by 2.4 and 1.4 percentage points respectively. After thirty years of experience in the job market, these differentiated dynamic gains imply that workers in Paris earn wages on average 10 percentage points higher simply from reallocating to better-paying jobs.

Finally, I document that high-paying jobs are concentrated in the cities where the firm-size wage premium is larger. I estimate that doubling the size of an employer leads to a wage increase of 10 and 4 percentage points in Paris and Saint-Étienne respectively. Even after controlling for unobserved firm heterogeneity and the sorting of workers across employers, I estimate the firm-size wage premium in Paris to be 2 percentage points larger than the size premium in Saint-Étienne.

Figure 1: Distributions of hourly wage by city in France




Each circles and squares represent a French commuting zone. The blue squares and circles correspond to the commuting zones highlighted on the $x$-axis. The squares and circles display respectively the 10th and 90th percentile of the hourly wage distribution in a particular commuting zones. The size of the markers is proportional to the size of the local labor market. Figure A.1 reproduces this figure with log hourly wage, while Figure A.2 displays the local wage distribution for workers not born in the commuting zone where they work. Figure A.3 plots the hourly wage distribution by metropolitan area in the United States.

These three empirical facts suggest that the difference in competition for workers across cities, together with local search frictions, shape the spatial variation in inequality.

In the second part of the paper, I build a framework that can account for these patterns. There are many locations that differ with regard to their population of homogeneous and immobile workers.[1] Employers with heterogeneous productivity decide where to locate. Once established in a location, employers compete for workers in the local frictional labor market à la Burdett and Mortensen (1998). Specifically, search frictions prevent employers from attaining their optimal size: to hire a large workforce, they must offer relatively high wages to poach workers from their local competitors. Job ladders specific to each city emerge as a result from this local poaching competition, and with them, within-city wage inequality.

[1]The city size distribution can be micro-founded by an appropriate distribution of amenities. Alternatively, most of the model's predictions hold if I replace spatial differences in size by differences in local productivity.

The location choice of employers determines in equilibrium the degree of competition in each labor market together with the local wage distributions. The marginal value of a worker for productive employers is relatively larger, and as result, productive employers want to be located in the city where they can maximize their workforce. However, search frictions introduce a tradeoff between the tightness of the labor market and the intensity of its competition. If all the productive jobs are indeed located in a slack market, the competition for workers is fierce as all employers have a high willingness to pay, and it is therefore costly for an employer to hire many workers. Consequently, downgrading to a tighter but less competitive location may be profitable. Despite this tension, I show that sorting prevails in equilibrium: productive jobs are disproportionately located in larger cities. However, in contrast to standard assignment models, the equilibrium does not always feature perfect assortative matching as some jobs are present in multiple cities.

Through the spatial sorting of jobs, larger cities thus feature labor market with greater competition, and this fiercer competition has wage implications that are empirically relevant. First, the size wage premium is larger since poaching a worker is relatively more costly. Second, as relatively unproductive employers hire most of their workforce from unemployment, the wage they offer is insensitive to the spatial differences in competition. On the contrary, workers employed at productive jobs in competitive cities can extract large rents as their employers must compete with other productive employers. This differential impact of competition on wages generates steeper job ladder in larger cities, and with it, more inequality.

The dispersion in rents that emerges in this framework highlights why cities matter for inequality. When competition in the labor market differs across locations and employers sort across space, less productive jobs in remote places are shielded from the competitive pressure of other employers. Meanwhile, since productive jobs are concentrated in a few locations, they face a greater degree of competition. Given that the rent extracted by workers depends on the competition their employer faces, it is not possible to properly quantify the dispersion in these rents without integrating cities in the analysis.[2]

**Related literature** This paper is related to three broad empirical literatures. The first strand studies the drivers of wage heterogeneity. Since the seminal work of Abowd et al. (1999), this labor literature has in particular focused on the relative importance of worker and job heterogeneity (e.g. Card et al., 2013; Bonhomme et al., 2019; Kline et al., 2020; Bonhomme et al., 2022). Most related to this project, Dauth et al. (2022) estimates a two-way fixed effect model and show that the average match quality between workers and firms is higher in larger cities. I contribute to this literature by highlighting that there is significant degree of job heterogeneity within cities, and that high-paying jobs are concentrated in a few locations.

[2] A third part, in which I estimate an extended version of the model and quantify the local root of wage inequality, is currently in progress. In this third part, I also use the present framework to evaluate the consequences of a place based policy that subsidizes job creation in remote areas on the spatial allocation of jobs as well as on local and aggregate inequality.

The second related literature investigates the spatial differences in inequality (Glaeser et al., 2009; Baum-Snow and Pavan, 2013; Eeckhout et al., 2014; Baum-Snow et al., 2018).[3] Thus far, this literature has focused on worker heterogeneity as the driver of local inequality. Glaeser et al. (2009) finds, for instance, that a third of the spatial variation in inequality is explained by differences in workers' skills. My project emphasizes that job heterogeneity is essential as it explains roughly half of the differences in local inequality. One exception in this literature is Papageorgiou (2022) which quantifies that the larger number of occupations in big cities explain a third of their greater local inequality.[4] Throughout my empirical analysis, I control flexibly for workers' skills and therefore consider an occupation as a job attribute. I also show that jobs continue to matter after controlling for the occupational composition of cities.

Finally, this paper interacts with the literature that studies the dynamics of wages across cities (see for instance Baum-Snow and Pavan, 2012 in the United States, D'Costa and Overman, 2014 in the United Kingdom, Roca and Puga, 2017 and Porcher et al., 2021 in Spain, and Eckert et al., 2022 in Denmark). I contribute to this literature by decomposing these gains into a learning and a job switching component, and find that the reallocation of workers to better-paying jobs constitutes between half and three-fourth of the extra wage growth experienced by job switchers.

On the theory side, this paper is first related to the literature that investigates the impact of search frictions on wage inequality. Started by Burdett and Mortensen (1998), recent quantitative applications were carried by Engbom and Moser (2021), Gouin-Bonenfant (2022), or Bilal and Lhuillier (2022). In these models, the amount of rent dispersion generated by search frictions depends on the competition faced by employers. I extend Burdett and Mortensen (1998) to a spatial context where the labor markets are local and the competition is endogenous. Secondly, this paper is related to Jarosch et al. (2021) and Berger et al. (2022) which also build frameworks with endogenous local monopsony power.[5] This project extends this literature by pointing out that spatial differences in competition have implications on local and aggregate inequality.[6]

Finally, several recent papers have integrated search frictions in spatial frameworks to study the existence of wage differentials across cities (Schmutz and Sidibé, 2019; Martellini, 2020; Heise and Porzio, 2022). My analysis is most closely related to the recent paper by Lindenlaub et al. (2022).

[3]A long list of economic geography papers has studied the contribution of worker heterogeneity in explaining between-city differences in average wages (e.g. Berry and Glaeser, 2005; Combes et al., 2008; Bacolod et al., 2009; Hendricks, 2011; Combes et al., 2012; Diamond, 2016; Giannone, 2017). These spatial average variations explain between 5% and 15% of the overall wage dispersion. On the contrary, this paper focuses on the between-city differences in within-city wage inequality.

[4]Gautier and Teulings (2009) also provide evidence that the pool of industries and occupations is more diverse in larger cities, and Brinkman (2014) and Rossi-Hansberg et al. (2019) study respectively the average wage implications of these compositional differences in industry and occupation across cities.

[5]Several recent papers have highlighted the empirical prevalence of monopsony power in the labor market, measured through employer concentration (Benmelech et al., 2020; Azar et al., 2020, 2022; Rinz, 2022), wage markdown (Brooks et al., 2021; Yeh et al., 2022), or heterogeneous pass-through rates (Chan et al., 2021).

[6]Neither of these models feature within-city wage inequality, and consequently, cannot be used to study the local root of wage inequality.

Like me, they construct a framework with employers sorting across cities and competing in frictional local labor markets with on-the-job search.[7] Our papers answer different questions, and as such, complement each other. While Lindenlaub et al. (2022) is interested in quantifying the effect of firm sorting on the wage gap between East and West Germany, my paper studies how employer sorting shape local and aggregate inequality.

The remainder of the paper is structured as follows. Section 1 presents the data and the empirical analysis. Section 2 builds a theory and characterizes its qualitative implications. The last section concludes the paper.

# 1 Descriptive evidence

This section presents evidence that job heterogeneity matters to understand the spatial differences in wage inequality. In Section 1.1, I describe the data. In Section 1.2, I quantify the respective contribution of worker and job heterogeneity on local wage inequality. I then study the between-city differences in wage gains from a job transition in Section 1.3, before analyzing the heterogeneity in firm-size wage premium in Section 1.4.

## *1.1 Data*

Quantifying empirically the contribution of jobs on local wages requires a dataset with two specificities. First, to separate the effect of jobs from that of workers, a worker panel with employer information is required. Second, to estimate precisely the heterogeneity of jobs across space, the data must have sufficient power at the city level. Datasets publicly available such as the American Community Survey or the Current Population Survey lack either of these requisites.

To understand the role of jobs in explaining the spatial differences in inequality, I therefore use employer tax records from France (DADS) between 1993 and 2007. The data comes in two flavors. To construct statistics at the city-level, I use repeated cross-sections that cover the universe of French firms and workers. To compute individual-level outcomes, I use a 4% representative panel that tracks workers throughout their labor market histories.

The primary measure of wage used is hourly wage, defined as net salary (net of all taxes but the income tax) divided by the number of hours worked.[8][9] To minimize measurement errors, I impose several restrictions on the sample. First, to limit movements in and out of the labor force that are hard to capture with administrative data, I restrict the sample to prime age workers (25 to 55 years old) with at least three years of experience. Second, I keep only in the sample full-time workers employed in non-temporary jobs to limit measurement error in the number of hours worked. Finally, I exclude from the analysis industries that take mostly place outside of cities (agriculture and the

[7] Bilal (2020) also builds a framework with employer sorting in frictional local labor markets. However, his theory does not include on-the-job search.

[8] In the panel, the data is collected at the worker-establishment-occupation-year level. If a worker is employed at two jobs at the same time, I keep the job that delivers the highest salary.

[9] There is no spatial variation in income tax nor in unemployment benefits in France.

extraction industry) as well as the public sector.[10]

I define a city as a French commuting zone.[11] I focus the analysis on metropolitan commuting zones with more than 250 observations in the panel to minimize measurement errors, which leaves me with 298 cities. Finally, for computational feasibility, I construct ten population-weighted deciles of cities according to the cities' inequality, defined as the difference between the 90th and 10th percentile of the log hourly wage. For reference, the first decile contains 96 cities (amongst which Châteaubriant and Saint-Brieuc), Saint-Étienne is in the third decile, and the tenth decile is solely composed of Paris. Finally, I define a job as a pair establishment-occupation, where occupations are at the two digits level. [12]

### *1.2 The local importance of job heterogeneity*

Why are some cities more unequal than others? In particular, is Paris unequal because its workers are very diverse? Or rather, is it that some jobs in Paris pay their workers above and beyond the compensations offered by other employers?

To answer these questions, I estimate a two-way fixed effect model that break down workers' wages into a worker effect and a job effect. Specifically, I estimate

$$\omega_{i,t} = \underbrace{\alpha_i + \beta_{\ell(i,t)} \cdot x_{i,t}}_{\text{Worker contribution}} + \underbrace{\gamma_{j(i,t)}}_{\text{Job contribution}} + \varepsilon_{i,t}, \tag{1}$$

where $\omega_{i,t}$ is the log hourly wages of worker $i$ in year $t$, $\ell(i,t)$ returns the city decile in which worker $i$ works in year $t$, $j(i,t)$ indices the job at which worker $i$ is employed at time $t$, and $x_{i,t}$ is the experience of the worker. The worker contribution is the sum of two variables. The first component is static, measured by the worker fixed effect $\alpha_i$, and captures workers' observed and unobserved heterogeneity constant over time (e.g. sex, education, place and cohort of birth). The second component is dynamic, estimated by $\beta_{\ell(i,t)} \cdot x_{i,t}$, to capture that workers' skills evolve over time through on-the-job and off-the-job learning. Following Roca and Puga (2017), I estimate city-specific returns from experience to allow for spatially differentiated learning. Meanwhile, the job contribution is defined as the job fixed effect $\gamma_j$. This variable measures the extra wage that workers receive when working at job $j$ compared to the remuneration they would on average receive given their skills. Accordingly, I also refer to $\gamma_j$ as the job premium of job $j$.[13]

According to the empirical model (1), Paris is more unequal than Saint-Étienne for three potential reasons. First, workers may be more diverse in Paris. Second, the pool of job opportunities may be

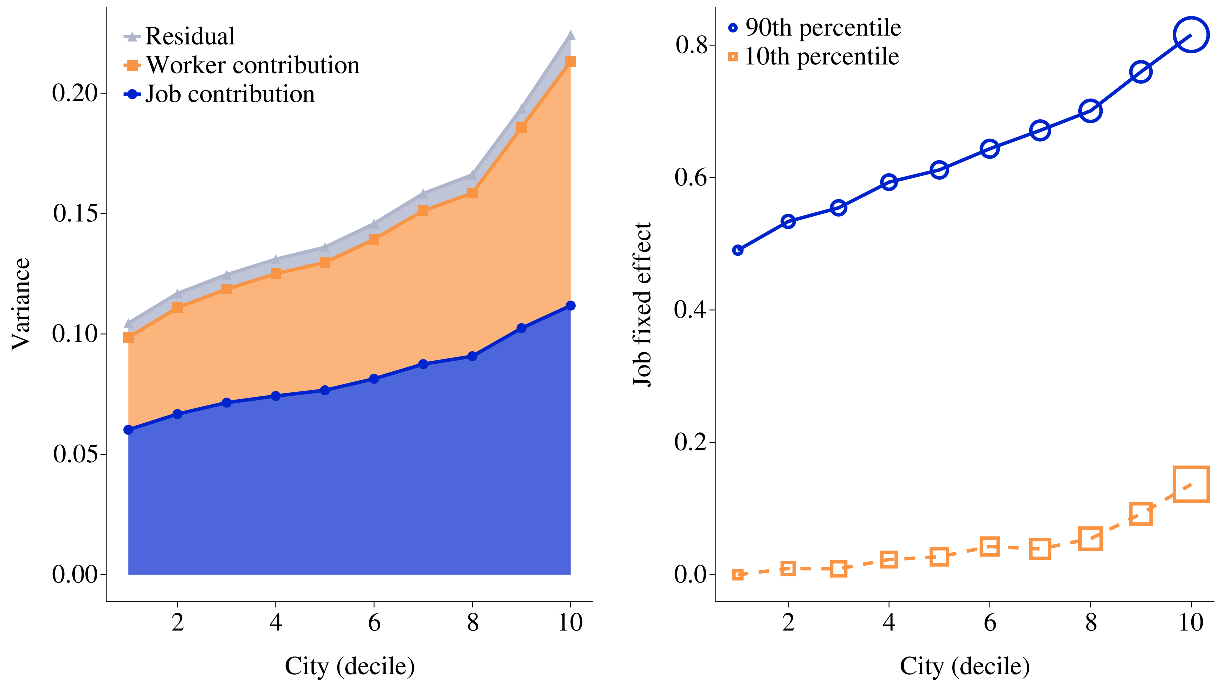
[10] The French public sector has a relatively flat income schedule over space in contrast to the private sector.

[11] The French statistical institute defines a commuting zone as an area where most of the residents work at jobs located in that same area. There are 328 such commuting zones which cover the entirety of France.

[12] Throughout the empirical analysis, I control for static and dynamic differences in workers' skills. Hence, differences in wages across occupations are referred to job rather than worker heterogeneity.

[13] To circumvent concerns regarding limited mobility bias, I follow Bonhomme et al. (2019) and group jobs in 1,000 bins according to their unconditional average wage. The estimates found in this section are robust to binning establishments into 250, 500, or 5,000 groups.

Figure 2: The role of job heterogeneity in explaining local wage inequality




(a) Variance decomposition across cities

(b) The local distributions of job fixed effect

Left panel: variance decomposition based on (13). A job is defined as a pair occupation-establishment. Jobs are grouped in 1,000 bins according to their unconditional expected wages. Blue area: the job contribution. Orange area: the worker contribution. Grey area: the residual. Right panel: the between-city differences in distributions of the job premia. Blue solid line: $90^{\text{th}}$ percentile of the city-specific job fixed effect distribution, normalized to zero in city decile 1. Orange dashed line: $10^{\text{th}}$ percentile of the city-specific job fixed effect distribution, normalized to zero in city decile 1.

more heterogeneous. Or finally, the residual may be more important. Figure 2a decomposes the variance of log wages into these three terms for each of the city deciles studied. As expected, Paris (top decile) is more unequal than Saint-Étienne ($3^{\text{rd}}$ decile): the variance of log wages equals 0.22 in the former city while it is only 0.12 in the latter. Figure 2a reveals that this greater inequality is largely caused by a greater degree of job heterogeneity in Paris. The variance of the job premia is indeed 0.11 in Paris, 0.04 log points higher than in Saint-Étienne. Generalizing to all cities, I find that places with a variance of log wages 0.037 higher (one standard deviation) have a variance of the job fixed effects that is on average 0.016 log points bigger; that is, the spatial variations in job heterogeneity account for 43% of the between-city differences in inequality.[14]

Figure 2a also reveals that workers are more diverse in unequal cities. An increase in the variance of log wages by 0.037 log points is indeed associated with an increase of the variance of the worker contribution by 0.019 log points. Overall, the job and worker contributions thus explain a similar

[14] In comparison, spatial variations in average job premia explain 40% of the between-city differences in average wages. However, this constitutes only 4% of the aggregate wage dispersion versus 51% for the within-city job heterogeneity (see decomposition (14) in Section A.1 for details). While the empirical facts and the theory that follows can speak to both between- and within-city differences in job heterogeneity, I mostly focus my attention on the second moment as the goal of the paper is to explain the contribution of cities to overall inequality.

fraction of the spatial differences in wage inequality. This paper focuses solely on the role played by job heterogeneity for two reasons. First, because spatial job heterogeneity has not been studied previously.[15] Second, and more importantly, because the spatial differences in job heterogeneity are intrinsically linked with the local root of wage inequality. To the extent that the premium an employer is willing to offer to its workers depends on the competition it faces on the local labor market, it is indeed impossible to properly estimate the contribution of job heterogeneity on aggregate wage inequality without integrating it in a spatial framework. On the other hand, for a given aggregate skill distribution, it is not necessary to know the spatial allocation of workers to quantify the impact of skill heterogeneity on aggregate inequality.

In the decomposition presented in Figure 2a, the covariance between the worker and job fixed effects is distributed equally in the job and worker contributions. I do so as this covariance term would be nil if job premia did not exist. That is, to understand why the covariance may differ across cities, it is first required to grasp why there exists job premia within cities. Nevertheless, Figure A.4 in Appendix plots the covariance term independently. I then find that the covariance of the job fixed effects and the covariance between the job and worker effects are equally important.[16] This result echoes that of Dauth et al. (2022) and shows that the covariance matters not only to explain spatial differences in average wages but also in inequality.[17]

Why are jobs more diverse in Paris? Looking back at Figure 1, spatial variations in the right tail of the local wage distributions were the source of the spatial differences in inequality. To see whether this holds at the job level, Figure 2b plots the local distributions of the job fixed effects. The orange dashed line and the blue solid line display respectively the 10th and 90th percentiles of the local job fixed effect distributions after normalizing the 10th percentile of the job distribution in city 1 to zero.[18]

Both low-paying and high-paying jobs offer relatively higher wages in unequal cities. However, as for wages, the spatial variation in high-paying jobs is relatively more pronounced, thus driving up inequality. For instance, high-paying jobs in Paris offer wages 0.26 log points higher than high-paying jobs in Saint-Étienne – half the wage gap between the 10th and 90th of the job distribution in Saint-Étienne. Meanwhile, this between-city difference is only of 0.12 log points for the low-paying jobs. On average, I find that cities with a variance of log wages 0.037 higher have high-paying and

[15] A rare exception is Papageorgiou (2022) which looks at the differences in number of occupations across cities. Figure A.5 presents the results of variance decomposition in which the job fixed effects estimated in (1) are further projected on occupations and industry terms. Occupation heterogeneity explains between 50% and 60% of the local variations in jobs.

[16] Specifically, places with a variance of log wages 0.037 higher have a variance of the job fixed effects and a covariance between the job and worker effects that are on average 0.008 log points bigger.

[17] Although I am not the first to estimate a two-way fixed effect model to then investigate the spatial property of its estimates, Dauth et al. (2022) use this model to study average differences in wages across cities while I focus on differences in inequality. To the best of knowledge, Figure 2 thus provides a novel fact.

[18] The aggregate level of the job fixed effect distribution is not identified from (1) hence the normalization. The city levels are identified since jobs are a pair establishment-occupation and there is not city fixed effect.

low-paying jobs offering wages 0.1 and 0.04 log points higher respectively. Finally, to generalize Figure 2b to the entire job distribution, I perform quantile regressions in which I project the job fixed effects estimated from (1) on local wage inequality. The estimated elasticities are displayed in Figure A.6 and support that high-paying jobs (the $70^{\text{th}}$ to $90^{\text{th}}$ percentiles of the local job distributions) are the ones with the greatest spatial variance.

Cities with a greater level of inequality are thus more unequal because they concentrate high-paying jobs whereas low-paying jobs are present throughout the country.[19] Why are workers willing to accept low-paying jobs while better paying ones exist? And why are high-paying jobs concentrated in a few cities? The next two subsections provides empirical answers to these questions.

### *1.3 The existence of city-specific job ladders*

How can workers be willing to work for low-paying jobs in Paris while there exists better employment opportunities – at home and throughout the country? This section shows that low-paying jobs exist in Paris due to the presence of search frictions. In particular, I argue that each city has its own job ladder and that these job ladders are steeper in more unequal cities. Workers are thus willing to be employed at low-paying jobs in Paris because they expect to move up to relatively better jobs in the future.

Search frictions are indeed sizeable and local. To see this, Table 1 reports average and local statistics on job transitions. Each year, one worker out of five switches jobs.[20] Most of these transitions are accompanied with a wage increase, and half of them occurs within occupation. Importantly, almost all job transitions take place within the labor market where the workers were already employed: only 5% of the job switchers transition to a new commuting zone. Finally, workers who actually change location tend to be more often unemployed than the average workers. A third of the migrants currently employed were indeed unemployed in the previous period compared to a tenth for the average French worker. Columns two and three, together with Figure A.8 in Appendix, report these statistics by city deciles to confirm that these patterns are present throughout the country. I take these frequent local job transitions as evidence of the existence of job ladders that are specific to each city.[21]

[19] Figure A.7 alternatively plots the share of low- and high-paying jobs in each city, where a high-paying (low-paying) job is defined as a job with a fixed effect above the 90th percentile (below the 10th percentile) of the aggregate distribution. The fraction of high-paying jobs in Paris is 26% against 4.6% in Saint-Étienne, whereas the fraction of low-paying jobs is 4.5% and 13.6% in each city respectively.

[20] A job switch is here defined as a transition between two jobs that takes less than 90 days. This definition is for instance consistent with that of the U.S. Census.

[21] For further structural evidence on the existence of local search frictions, see Schmutz and Sidibé (2019) which build a model with both search and spatial frictions.

Table 1: Job transition and migration rates

| | Aggregate | City deciles | |
|---|---|---|---|
| | | D1 | D10 |
| Job transition (%) | 19.3 | 16.6 | 23.3 |
| Fraction of job transitions (%) | | | |
| Associated with wage gains | 73.5 | 73.5 | 73.4 |
| Associated with migrations | 5.50 | 4.67 | 6.00 |
| Within occupation | 46.5 | 38.9 | 57.1 |
| Workers previously unemployed (%) | | | |
| Full sample | 10.3 | 10.1 | 11.8 |
| Migrants | 29.4 | 34.8 | 24.6 |

A job is defined as a pair occupation-establishment. A job transition is defined as a transition between two jobs where the end-date of the previous job was less than a quarter before the start-date of the current job. A migration is defined as a transition between two commuting zones of residence. A worker is defined as previously unemployed if the end-date of the last job was more than a quarter before the start-date of the current job.

These city-specific job ladders furthermore differ substantively from one another. To study their spatial heterogeneity, I estimate the following empirical model

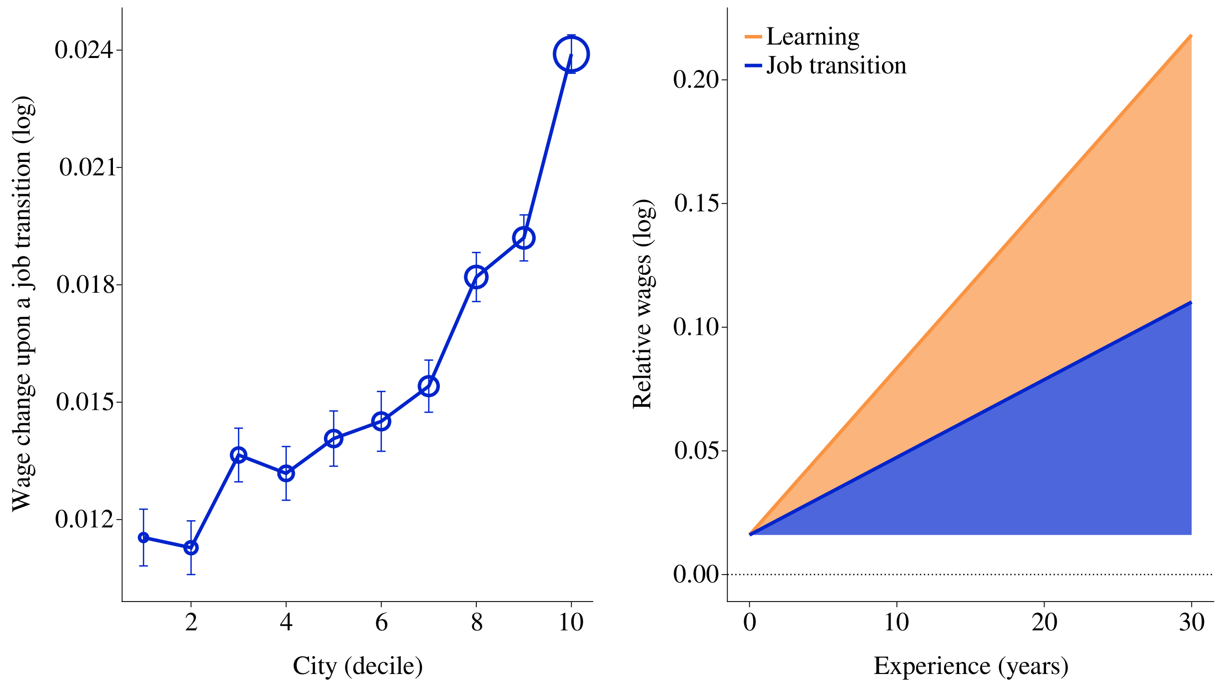
$$\omega_{i,t} = \alpha_{\ell(i,t)} + \beta_i + \delta_{\ell(i,t)} x_{i,t} + \gamma_{\ell(i,t)} \sum_{\tau=\underline{t}_i}^{t} \mathrm{J2J}_{i,\tau} + \varepsilon_{i,t}. \tag{2}$$

In specification (2), $\underline{t}_i$ is worker's $i$ year of entry in the labor market, $\mathrm{J2J}_{i,\tau}$ equals one if a job transition happened between period $\tau - 1$ and $\tau$, and $x_{i,t}$ is $i$'s experience at time $t$. The parameters $\beta_i$ are worker fixed effects while $\delta_\ell$ measures the city-specific returns to experience. The city fixed effects, $\alpha_\ell$, captures the wage boost (or penalty) that workers experience at the beginning of their career when employed in city $\ell$, holding the skill distribution fixed across cities. A high $\alpha_\ell$ thus means that the bottom rungs of the job ladder in city $\ell$ are relatively higher. Finally, the key parameter of interest is $\gamma_\ell$, which measures the wage gains following a job transition holding workers' skills constant. A larger $\gamma_\ell$ implies a steeper job ladder in city $\ell$.[22]

Figure 3a displays the estimated average wage gain upon a job transition for each city decile alongside their 95$^{\text{th}}$ confidence interval. First, a job ladder exist in each city as job transitions

[22]Heise and Porzio (2022) estimates a similar model in first difference but focus their analysis on job transition attached to a migration decision. Lindenlaub et al. (2022) also estimates a linear model with city-specific returns to job transition, but do not include worker fixed effects. As such, Figure 3 provides the first set of evidence on the spatial differences in wage gains upon a job transition at the worker level.

Figure 3: The role of job transitions for local wage growth




(a) Wage gains upon job transition by city

(b) Relative wages (Paris vs. Saint-Étienne)

Left panel: point estimates from specification (2). A job is defined as a pair occupation-establishment. A job transition is defined as a transition between two jobs where the end-date of the previous job was less than a quarter before the start-date of the current job. Solid blue line: change in log wages generated by a job transition, i.e. $\{\gamma_\ell\}_\ell$. Error bars represent the 95% confidence interval, with standard errors clustered at the worker level. The size of the markets is proportional to the size of the local labor market. Right panel: difference in wages between Paris and Saint-Étienne across the lifecycle, computed from the estimates obtained in (2). Blue area: relative wage growth from job transition, computed as $\gamma_\ell \mathbb{E}[\text{J2J} \mid \ell] - \gamma_{\ell'} \mathbb{E}[\text{J2J} \mid \ell']$. City-specific job transition rates presented in Figure A.8a. Orange area: relative wage growth from learning, computed as $\delta_\ell - \delta_{\ell'}$.

are a meaningful source of wage growth for workers.[23] Furthermore, the job ladders of unequal cities are much steeper. In Paris, a job transition yields on average a wage increase of 0.024 log points while it is only 0.014 log points in Saint-Étienne. For comparison, Figure A.9b and A.9c presents the estimated city-specific starting wages ($\{\alpha_\ell\}_\ell$) and learning rates ($\{\delta_\ell\}_\ell$). Consistently with Figure 2b, I find that there is very little spatial dispersion in the bottom rungs of the job ladders. For instance, the initial wage gap between Paris and Saint-Étienne is only of 0.016 log points, a gap thus filled after the first job transition in Saint-Étienne.[24] I also find that workers in Paris learn faster than those in Saint-Étienne, resulting in an extra wage growth of 0.004 log points per year spent in the city. That is, the reallocation of workers to better-paying jobs constitutes 71% of the total difference in wage growth between these two places for job switchers. Generalizing to all locations, workers in cities with a variance of log wages 0.037 higher experience wage gains upon a job switch 0.004 log points larger.

[23] Compared to the learning estimates $\{\delta_\ell\}_\ell$ (see Figure A.9c), a worker who switch jobs experience a yearly wage growth between 22% and 44% higher than a worker who does not switch.

[24] Furthermore, I find no statistical significant differences in starting wages for the city decile 1 to 8.

To address possible concerns regarding omitted variables, I perform several robustness exercises presented in Figure A.10. To verify that the patterns of Figure 3a are not caused by more frequent transitions through unemployment in low inequality cities misclassified as job transitions, I re-estimate (1) looking only at job transitions associated with a positive wage gain. The wage gains upon a job transition are then 0.017 log points higher in Paris than in Saint-Étienne.[25] To investigate whether the larger wage gains in Paris are due to occupation switching (Papageorgiou, 2022), I focus on transitions within occupation but across establishments and include occupation fixed effects in (2). The between-city differences in job ladder's steepness remain unchanged. Finally, Figure 3a continues to hold after including a quadratic term for experience, or adding as additional control variables age and tenure at the establishment.

Search frictions generate inequality across identical workers within cities, and the spatial differences in the ladders' steepness imply different level of local inequality. As some workers transition to better paying jobs while others do not, wage dispersion indeed appears within cities. Furthermore, as the wage increases upon a job transition are larger in Paris than in Saint-Étienne, so is the wage dispersion. Using the estimates from (2), Figure 3b helps visualize how spatial differences in job reallocations generate these between-city differences in inequality. Specifically, the figure displays the difference in wages across the lifecycle for workers in Paris versus those in Saint-Étienne. After thirty years on the job market, workers earn a wage on average 0.1 log points higher in Paris than in Saint-Étienne simply from the reallocation to better-paying jobs.[26] This represents 43% of the total wage premia that Parisian workers experience at the end of their career, and it is six times larger than the initial bonus earned by workers at the beginning of their career.[27]

Each city thus has its own job ladder. These ladders have common bottom rungs but some are much steeper than others. The differences in ladders' steepness imply significant spatial variations in wages for workers on the high rungs, and therefore differences in inequality across cities. Understanding why some ladders are steeper than others is thus essential to comprehend the local root of wage inequality, a question to which I now turn.

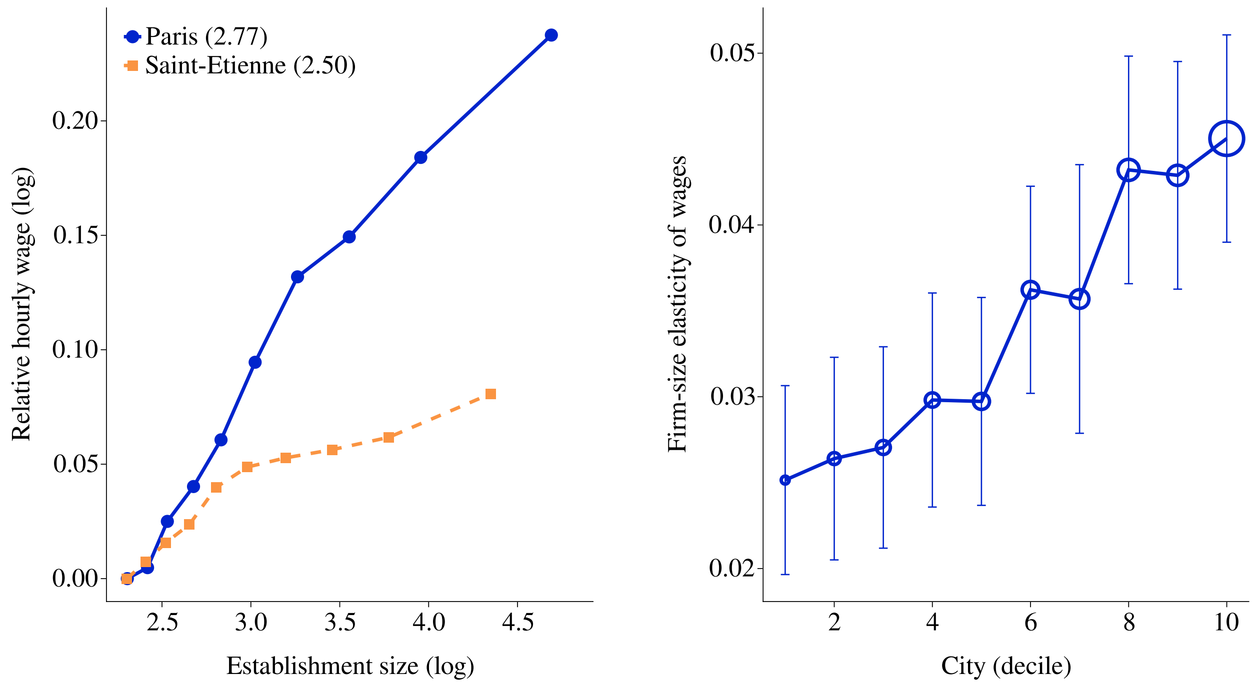
---

[25] I have also restricted the definition of a job switch to a transition between jobs with a gap lasting less than a month, and the point estimates were virtually unchanged.

[26] The difference in wage growth attributable to job transition between city $\ell$ and $\ell'$ is given by $\Delta^{\text{J2J}}_{\ell\ell'} \equiv \gamma_\ell \mathbb{E}[\text{J2J} \mid \ell] - \gamma_{\ell'} \mathbb{E}[\text{J2J} \mid \ell']$, where $\mathbb{E}[\text{J2J} \mid \ell]$ is the average job switching rate in city $\ell$. Workers in city $\ell$ may therefore experience faster wage growth due to either higher wage gains upon a job transition or more frequent job transitions. Rewriting the wage growth as $\Delta^{\text{J2J}} = (\gamma_\ell - \gamma_{\ell'})\mathbb{E}[\text{J2J}] + \{\gamma_\ell(\mathbb{E}[\text{J2J} \mid \ell] - \mathbb{E}[\text{J2J}]) - \gamma_{\ell'}(\mathbb{E}[\text{J2J} \mid \ell'] - \mathbb{E}[\text{J2J}])\}$, I find that spatially differentiated wage gains (first term) and more frequent transitions (term in curly bracket) explain respectively 63% and 37% of the higher growth.

[27] Combined with the initial wage boost of 0.016 log points, I thus estimate that workers at the end of their career earn a wage 0.11 log points higher in Paris than in Saint-Étienne. This is half of the between-city difference in the 90th percentile of the job premia distributions found in Figure 2b. This should not be surprising as the first estimate is an average effect. To the extent that not all workers start their careers on the lowest rungs of the ladder and that some transition more frequently across jobs than others, this generates further wage dispersion captured in Figure 2b but not in Figure 3b.

Figure 4: The firm-size wage premia across cities




(a) Cross-sectional differences in wages

(b) Within worker and firm estimates

Left panel: bin scatter plot of average hourly wage by deciles of the firm size distributions for Paris (blue solid line) and Saint-Etienne (orange dashed line). Right panel: point estimates from specification (3). Solid blue line: firm-size elasticity of wages, i.e. $\{\gamma_\ell\}_\ell$. Error bars represent the 95% confidence interval, with standard errors clustered at the establishment level. The size of the markets is proportional to the size of the local labor market. Sample selection for both panels: firms with more than 10 employees. The sample is then further winsorized at the 5% level. Finally, to estimate (3), the sample is restricted to growing establishments ($\Delta \log N_{j,t} \geq 0$) with a size of between 10 and 500 employees.

### *1.4 Between-city differences labor market competition*

Why is the job ladder steeper in Paris? In particular, why are employers willing to pay workers more than what is offered by their competitors present in the same labor market? In this section, I show that employers need to offer higher wages to attain a given size in more unequal cities – a fact I take as evidence of fiercer competition between employers.

Figure 4a displays the average wage per decile of the size distribution in Paris and Saint-Étienne.[28] To highlight the differences in wages across employers in each city, the average wage for the first decile of the size distribution has been normalized to zero. The first thing to notice is the overlap in the size distributions in each city. While the average establishment size is larger in Paris, there are small and large firms in both cities, with the 90$^{\text{th}}$ percentile of the firm-size distribution in Saint-Étienne being larger than the 80$^{\text{th}}$ percentile in Paris.

To reconcile the overlap in the size distributions with the spatial differences in high-paying jobs, it must therefore be that the size wage premium is larger in Paris. For instance, in Saint-Étienne, employers with 20 employees pay their workers wages 0.049 log points higher than employers with

[28] I define here an employer as an establishment. Its size is defined as the number of workers it employs.

10 employees. This gap is 0.095 log points in Paris. Said differently, an increase in size by 10 employees adds an extra 20% to the between-city wage gap that already existed across small firms. Generalizing to all cities, I find that doubling the size of an establishment implies an increase in wages by an extra 0.012 log points in cities with a variance of log wages 0.037 log points bigger. These differences in size gradient creates wage dispersion across workers. Furthermore, they suggest that employers in Saint-Étienne are shielded away from the competition of employers in Paris as they are able to attain the same size while offering lower wages.

However, employers of different sizes may differ from one another in other ways than the wage they offer to their workers, thus biasing the aforementioned estimate. First, firms compete in different industries. To the extent that the average size and the wage varies across sectors, differences in the industrial composition of cities would affect the spatial differences in size wage premia. Figure A.11b reproduces Figure 4a after taking out year and 4-digit industry fixed effects. While the relationship between firm-size and wage flattens up, in particular in Paris, there remains significant between-city differences. Nevertheless, even within narrowly defined industries, employers may still differ from one another in ways that impact wages. The main concern is in particular that large firms may be better at attracting qualified workers. To the extent that skilled workers are relatively more abundant in unequal cities, this concern would bias upward the between-city differences in size wage premia.

To control for the potential differences in worker skills across firms and cities, I thus turn to within-worker within-firm estimates of the size wage premium. Specifically, using only workers who stayed at the same employer between two periods, I estimate the following model

$$\Delta\omega_{i,t+1} = \alpha_{t,\ell(i,t)} + \delta_{\ell(i,t)}\Delta \log N_{e(i,t),t+1} + \varepsilon_{i,t}, \tag{3}$$

where the functions $\ell(i,t)$ and $e(i,t)$ are indexing the city and the establishment in which worker $i$ lives and works at time $t$, and $N_{e,t}$ is the number of workers employed by $e$ at time $t$. The time-by-city fixed effects $\{\alpha_{t,\ell}\}_{t,\ell}$ are present to control both for the spatial differences in learning as well as local shocks that would affect size and wages simultaneously. The parameters of interest are $\{\delta_\ell\}_\ell$. They measure the city-specific size elasticities of wages, controlling both for the allocation of workers' across firms and cities as well as the allocation of firms across locations.

Figure 4b displays the estimated size wage premia for each city decile alongside with their 95$^{\text{th}}$ confidence intervals. First, we see that even after controlling for differences in the worker composition of firms across cities, larger firms continue to pay higher wages. Second, and most importantly, there exits significant differences in firm-size wage premia across cities: a firm doubling its size must increase its wage by 0.045 log points in Paris versus 0.027 log points in Saint-Étienne. Generalizing, I find that doubling the size of an establishment implies an increase in wages by an extra 0.007 log points in cities with a variance of log wages 0.037 log points bigger. [29]

[29] Appendix A.3 discuses the possible threats to causality underlying Figure 4b, and Figure A.12 displays the results of several robustness exercises. In particular, adding occupation fixed effects or controlling for workers' experience and tenure at the establishment leave the results virtually unchanged.

Large employers must therefore offer relatively higher wages in Paris. To the extent that employers compete for workers through wages, this evidence suggests that the competition is fiercer in Paris. This fiercer competition pushes wages up at the top of the job ladder, leading to the concentration of high-paying jobs in a few locations. The spatial differences in local competition, together with their distributional effect on wages, constitute the local root of wage inequality. However, how much competition there is on the labor market is not a primitive of cities. As such, a full understanding of the role of cities in generating inequality is not possible without an internally coherent and empirically relevant framework that rationalizes the spatial differences in competition.

## 2 The local root of wage inequality: microfoundations

In this section, I build upon Burdett and Mortensen (1998) to construct a spatial model of job sorting in frictional local labor markets. Job sorting endogenously generates spatial differences in competition which, interacted with search frictions, creates local wage distributions resembling those of Section 1.

### *2.1 Setup*

The economy is constituted of two entities that are both infinitely lived. There is a unit mass of *ex-ante* homogeneous workers and a mass $M \leq 1$ of heterogeneous jobs. Workers and jobs meet within $L$ locations to produce an homogeneous good. This good is freely traded across cities and taken as the numeraire. Time is continuous, and the common discount rate is $\rho$. There is no aggregate shock, and I focus on the stationary equilibrium.

**Cities** Cities are indexed by $\ell$ with $\ell \in \{1, 2, \ldots, L\}$. They are identical in all respects except that some locations are larger than others. Specifically, the mass of workers residing and working in city $\ell$ is $\mathrm{n}_\ell$.[30] Cities are ordered by increasing size so that $\mathrm{n}_1 < \mathrm{n}_2 < \cdots < \mathrm{n}_J$. The force that drives job sorting is thus between-city differences in market size. An alternative motive could be differences in local productivity, for instance to reflect differences in average worker productivity. From the employers' point of view, both drivers are isomorphic. I focus here on city size because it highlights the role played by job sorting in generating spatial differences in wage inequality.[31]

**Workers** Workers are risk neutral, do not have access to any savings device, and derive utility from consuming the homogeneous good. Their lifetime utility thus equals their lifetime income

$$\mathbb{E}_0 \left[ \int_0^\infty \mathrm{e}^{-\rho t} c_t \,\mathrm{d}t \right].$$

Workers are endowed with one unit of labor that they supply inelastically when employed. However, the presence of search frictions generate unemployment. Specifically, unemployed and employed

[30] I assume for now that geographical frictions are large enough so that no workers desire to move. In the quantitative version of the model, I let workers choose their locations and assume that cities offer different amenities. In Appendix **??**, I further show that any worker allocation can be sustained by an appropriate vector of amenities.

[31] In any model with constant returns to scale, differences in city size alone do not affect directly wages.

workers receive job offers at rate $\lambda^u$ and $\lambda^e$ respectively from the *local* job offer distribution $F_\ell$. Employed workers are free to switch jobs at any point in time. Meanwhile, they also lose their job and fall back to unemployment at rate $\delta$, with $\delta < \lambda^e \leq \lambda^u$. Accordingly, the mass of unemployed and employed workers in city $\ell$ are respectively $u_\ell = \delta/(\delta + \lambda^u)\mathrm{n}_\ell$ and $e_\ell = \lambda^u/(\delta + \lambda^u)\mathrm{n}_\ell$. Unemployed workers earn $b$, which could come either from unemployment benefits or a backyard technology.

The assumption that workers receive offers only from local employers is aligned with the very few job transitions that take place across cities (Table 1).[32] Furthermore, motivated by the evidence that the heterogeneity in wage gains upon a job switch explains the majority of the dynamic gains from job transitions, search frictions are assumed exogenous and constant across locations. The quantitative extension relaxes this assumption by adding city-specific matching functions.

**Jobs** Jobs are *ex-ante* heterogeneous. Jobs are indexed by their productivity, $z$, with support $[\underline{z}, \bar{z}]$, $\bar{z} \leq \infty$, and distributed according to some aggregate probability measure $\Gamma$.[33] The revenues generated by a job with productivity $z$ are $R(z, n) = z \cdot n$, where $n$ is the number of workers hired at the job.[34] Jobs can be located (and relocated) anywhere at any time and the productivity of a job is constant throughout space. However, to locate a job in city $\ell$, employers need to pay a unit housing cost $r_\ell$.[35]. Employers then decide of the spatial allocation of jobs to maximize their profits. The local distribution of jobs are summarized by two objects: the mass of jobs in the city, $v_\ell$, and the distribution of job productivity, $\mathrm{d}\Gamma_\ell$. A feasible spatial job allocation is one where the number of jobs allocated across cities is no greater than the aggregate number of jobs available. That is, the allocation $\{v_\ell, \Gamma_\ell\}_{\ell=1}^{L}$ is feasible if, $v_\ell \geq 0$ for each $\ell$, and for each $z \in [\underline{z}, \bar{z}]$, the inequality

$$\sum_{\ell=1}^{L} v_\ell \mathrm{d}\Gamma_\ell(z) \leq M \mathrm{d}\Gamma(z) \tag{4}$$

holds. Given this notation, I suppose that housing prices are increasing in the number of jobs in the city, $r(v_\ell) = c \cdot v_\ell^\chi$.[36]

Once they have allocated their jobs over space, employers hire workers on the frictional local labor market. Following Burdett and Mortensen (1998), employers hire workers by posting a vacancy in

[32]For a quantitative model with wage posting and between-city search, see Heise and Porzio (2022).

[33]I assume that $\underline{z}$ is high enough so that all jobs are profitable in at least one city.

[34]An alternative version of this model would have heterogeneous firm consists of multiple occupations and decide the allocation of these occupations across space. A job would then be defined as a firm-location-occupation triplet, as in the data. To the extent that occupations are perfectly substitutable and labor markets are integrated across occupations, these two models are isomorphic.

[35]This assumption is for tractability as it turns the jobs sorting problem into a pseudo assignment model (see Section 2.4).

[36]The iso-elastic cost function can be micro-founded via profit maximizing local promoters subject to decreasing returns to scale. Suppose that a housing developer in each city owns the local land. Land is turned into commercial housing through the production function $L(x) = x^\rho$, where $x$ is a bundle of inputs with constant marginal cost $q$. Market clearing in the commercial housing market then requires $r_\ell \propto v_\ell^{(1-\rho)/\rho}$ where the constant of proportionality is common across locations.

the labor market where their job is located.[37] They possess all the bargaining power but commit to a constant wage throughout the duration of the job. The wage offered for the job is indicated on the vacancy. Vacancies are then randomly sent to the local workers, and employers cannot counteroffer when one of its workers receive an alternative job offer.

I solve this model in three steps. First, I characterize the local labor supply curve that employers face in each city (Section 2.2). I then derive the local wage distributions as a function of the spatial job allocation and obtain conditions under which the wage distributions match the patterns of Section 1 (Section 2.3). Finally, I solve for the spatial job allocation in Section 2.4 and show that it yields empirically relevant wage distributions.

## *2.2 The local labor supplies*

The local labor supply curves are obtained from the job-switching behavior of workers within city.[38] Let $U_\ell$ and $V_\ell(w)$ denote respectively the lifetime utility in city $\ell$ of an unemployed worker and a worker employed at wage $w$. These value functions satisfy

$$\rho U_\ell = b + \lambda^u \int \max\{V_\ell(w) - U_\ell, 0\} \mathrm{d}F_\ell(w),$$
$$\rho V_\ell(w) = w + \lambda^e \int \max\{V_\ell(w') - V_\ell(w), 0\} \mathrm{d}F_\ell(w') + \delta[U_\ell - V_\ell(w)].$$

The lifetime utility of unemployed workers in city $\ell$ is therefore composed of their contemporaneous earnings, $b_\ell$, and the expected value of searching for a job in that city. Similarly, the value of being employed at wage $w$ is composed of today's earnings, the expected value of receiving a better job offer, and finally the expected drop in utility that occurs when a worker becomes unemployed.

The utility of employed workers is increasing in their wage. Workers thus behave as income maximizers, climbing the local job ladder by accepting better-paying job offers while employed. When unemployed, workers tradeoff a higher search efficiency against lower present earnings, thus generating a city-specific reservation wage under which workers prefer to stay unemployed, $\underline{w}_\ell$. The reservation wage of workers in city $\ell$ is given by[39]

$$\underline{w}_\ell = b + (\lambda^u - \lambda^e) \int_{\underline{w}_\ell}^{\infty} \frac{\bar{F}_\ell(w)}{\delta + \lambda^e \bar{F}_\ell(w)} \mathrm{d}w. \tag{5}$$

Jobs with wages strictly under $\underline{w}_\ell$ cannot attract any worker. Since all jobs are worth allocating and because employers cannot hire workers outside of their market, $\underline{w}_\ell$ corresponds to the lower bound of the wage distribution in location $\ell$. The gap between the search efficiency off and on the job controls the spatial differences in the bottom rungs of the job ladders. In particular, if $\lambda^u \approx \lambda^e$, differences in high-paying jobs across cities are not reflected in the reservation wages.

[37]A framework aimed at explaining the empirical patterns of Section 1 must feature frictions to break the law of one price in the labor market. Search frictions with on-the-job search are relevant as they map into the job switching facts documented in Section 1.3. Finally, Burdett and Mortensen (1998) is a good starting because it contains at its core a relationship between employer size and wages, as in Section 1.4.

[38]This derivation is identical to that of the original Burdett and Mortensen (1998).

[39]The derivations of this section are contained in Appendix B.1 and B.2. It is already assumed that $\rho/\delta \approx 0$.

For wages above the reservation wage, the movement of workers up the job ladder determines the number of employees at a given wage. Let $G_\ell(w)$ be the fraction of workers employed at wages below $w$ in city $\ell$. Equating the inflows and outflows of workers for each wage interval yields

$$G_\ell(w) = \frac{F_\ell(w)}{1 + k\bar{F}_\ell(w)}, \tag{6}$$

where $k \equiv \lambda^e/\delta$ is 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 $F_\ell$ and the employment distribution $G_\ell$, the number of employed workers per wage offer $w$ in location $\ell$ is

$$n_\ell(w) = \frac{1}{\theta_\ell} \cdot \frac{1+k}{\left[1 + k\bar{F}_\ell(w)\right]^2}, \tag{7}$$

where $\theta_\ell \equiv v_\ell/e_\ell$ is the tightness of the labor market in city $\ell$. Equation (7) corresponds to the labor supply curve faced by employers in city $\ell$. This labor supply curve is in particular increasing in $w$: workers on high rungs of the job ladder stay employed relatively longer at the same employer since there are relatively fewer jobs to which they can transition.

Spatial differences in the supply curves are twofold. First, a relatively tighter market implies a lower size on each rung of the ladder. Second, employers in different cities face different degree of competition. These spatial differences in local competition are captured by the wage offer distribution. For instance, for two cities $\ell$ and $\ell'$, if $F_\ell \succ_{\text{FOSD}} F_{\ell'}$, a higher wage is needed in $\ell$ to achieve a given size. High-paying jobs are indeed more numerous in $\ell$. Hence, for any given wage, employers see their workers leave to better paying competitors more frequently.[40] Finally, search frictions, summarized by $k$, dampens or amplifies the spatial differences in competition. If frictions are high, workers transition at a slow pace across jobs, thus reducing the intensity of the between-employer poaching game. In the limit where frictions are enormous and $k \to 0$, employers hire workers only from unemployment and differences in competition become irrelevant.

### *2.3 The local job ladders*

The local wage distributions are given by employers' optimal wage posting behavior. I focus here on the employers' problem in steady state assuming that their discount rate is low. Employers then choose the location of their job and its wage to maximize their flow profits. The flow profits generated by a job with productivity $z$ are given by

$$\pi(z) = \max_\ell \pi_\ell(z) = \max_\ell \left\{ \max_{w,n} R(z,n) - w \cdot n - r_\ell \quad \text{s.t.} \quad n \leq n_\ell(w) \right\}. \tag{8}$$

As can be seen from the constraint in (8), employers internalize the labor supply curve of each city. In particular, upon choosing a location, a higher wage is necessary for an employer to hire more workers. I solve this problem in two stages. First, I derive the solution to the wage-posting game in each city for a given spatial job allocation. I then characterize the sorting patterns of jobs.

[40] From this point onward, I use $\succ$ to order distributions in the sense of first order stochastic dominance.

Within each city, the standard results from Burdett and Mortensen (1998) continue to hold for any feasible spatial job allocations. First, employers are always on their labor supply curve, $n = n_\ell(w)$, and the local wage offer distributions are continuous on an interval $[\underline{w}_\ell, \bar{w}_\ell]$. These results hold regardless of the shape of the local job distributions.[41] Furthermore, the technological complementarity in the revenue function between the productivity of the job and the number of workers employed guarantees that wages are strictly increasing in $z$ within each city: more productive employers desire to hire more workers to increase their revenues, and for that, offer higher wages. Finally, wages being strictly increasing in productivity, the rank of a job in the local wage offer distribution must correspond to its rank in the local job distribution. That is, if $w_\ell(z)$ denotes the optimal wage offered by job $z$ in city $\ell$, then the wage offer distribution must satisfy $F_\ell[w_\ell(z)] = \Gamma_\ell(z)$.

Combined with the employers' first order condition with respect to $w$, these results imply that job $z \in \operatorname{supp} \Gamma_\ell$ in city $\ell$ offers the wage

$$w_\ell(z) = \underline{w}_\ell \left( \frac{n_\ell(\underline{z}_\ell)}{n_\ell(z)} \right) + \int_{\underline{z}_\ell}^{z} \zeta \left( \frac{n'_\ell(\zeta)}{n_\ell(z)} \right) \mathrm{d}\zeta, \tag{9}$$

where $\underline{z}_\ell \equiv \inf \Gamma_\ell$ is the least productive job in city $\ell$ and $n_\ell(z)$ is the number of workers employed at that job,

$$n_\ell(z) \equiv n_\ell[w_\ell(z)] = \frac{1}{\theta_l} \cdot \frac{1+k}{\left[1 + k\bar{\Gamma}_\ell(z)\right]^2}.$$

To interpret equation (9), start from the bottom of the job distribution. The least productive job has the lowest willingness to pay for workers. This job thus locates itself at the bottom of the ladder and only hires unemployed workers by offering the reservation wage. More productive employers offer higher wages to poach workers from their low-paying competitors.[42] This poaching game builds up the competitive pressure from the bottom, and the wage paid by an employer depends on the productivity of all the other jobs on the lower rungs of the ladder. In particular, the most productive job in a city must compete with all the other jobs, making the wage that it offers more sensible to the spatial allocation of jobs.

Local search frictions thus generate city-specific job ladders, and differences in job distributions across cities lead to different ladders. In particular, the type of employers that decide to locate in a city acts as a local externality on wages. In cities where jobs are all relatively productive, wages at the top of the distribution are high not directly because they are offered by productive employers,

[41] For instance, even if some local job distribution has a hole, jobs with productivity just above the hole have no incentives in offering wages greater than the wage offered by jobs just below the hole as this would yield the same size.

[42] The wage increase between two adjacent employers is equal to the marginal profit of the job. Constant returns to scale in the revenue function imply that this marginal profit is independent of the local tightness of the labor market. Hence, two cities with different sizes but similar job distributions would have the identical wage offer distributions. This is a strength of the model as it highlights the role of the spatial variations in competition in generating between-city differences in inequality.

but rather because all the jobs located on the lower rungs are also more productive. Proposition 1 summarizes a sufficient condition on the spatial job allocation under which these the job ladders are empirically relevant.

**Proposition 1** (Between-city differences in job ladders)**.**
*Take two locations $\ell$ and $\ell'$ and suppose that jobs are relatively more productive in location $\ell$, $\Gamma_\ell \succ \Gamma_{\ell'}$. Then, in location $\ell$,*

1. *Wages are relatively higher, $G_\ell \succ G_{\ell'}$;*
2. *The top-to-bottom wage gap is larger.*

*In addition, if $\lambda^e \approx \lambda^u$, then*

3. *The wage gains upon a job transition are bigger for workers on low rungs of the ladder;*
4. *The size wage premium is larger.*

When jobs are relatively more productive in a location, employers' willingness to pay are relatively higher, leading to a fiercer poaching competition. For an employer to acquire a worker from one its competitor, a larger wage is necessary, and as a result, the size wage premium is bigger (Proposition 1.4 and Figure 4). The fiercer competition allows workers to extract on average larger rents, but in particular so for the workers at the top of the ladder whose employers have to compete with all the other jobs in the city. As a result, all jobs offer relatively higher wages but the inequality is also greater (Proposition 1.1 and 1.2 and Figure 2). Finally, as workers climb their local ladder from low-paying jobs that are relatively homogeneous over space to high-paying jobs that are concentrated in a few locations, they experience greater wage gains (Proposition 1.3 and Figure 3).

According to Proposition 1, the local wage distributions match the empirical patterns of Section 1 to the extent that the local job distributions can be ordered in term of first order stochastic dominance. I now turn to show that this is the case in equilibrium.

### *2.4 The spatial allocation of jobs*

Where are jobs allocated across space? Combining (8) and (9), the profit generated by a job with productivity $z$ in city $\ell$ is

$$\pi_\ell(z) = \underbrace{\frac{1}{\theta_\ell}}_{\text{Average size}} \underbrace{\left( z\eta_\ell(\underline{z}_\ell) + \int_{\underline{z}_\ell}^{z} (z-\zeta)\,\eta'_\ell(\zeta)\mathrm{d}\zeta \right)}_{\text{Net revenues per worker}} - \underbrace{(\underline{w}_\ell n_\ell(\underline{z}_\ell) + r_\ell)}_{\text{Local prices}}, \tag{10}$$

where $\eta_\ell(z)$ is the size of job $z$ in city $\ell$ net of the tightness of the labor market, $\eta_\ell \equiv \theta_\ell \cdot n_\ell$. Employers locate their job in the city that yields the greatest profit. That is, for each $z \in [\underline{z}, \bar{z}]$, the profit maximization condition

$$\pi_\ell(z) \geq \pi_{\ell'}(z) \text{ for all } \ell' \neq \ell \iff z \in \text{supp } \Gamma_\ell \tag{11}$$

must hold. An equilibrium spatial allocation of jobs is a tuple $\{v_\ell, \Gamma_\ell\}_{\ell=1}^{L}$ that jointly satisfies feasibility (4) and profit maximization (11).

The technological complementarity in the revenue function between job productivity and the number of employees hired remain present in equation (10).[43] More productive employers therefore want to establish their job in locations where they can attain a larger size. However, this size depends on two endogenous objects. First, the tightness of the local labor markets affect the average number of workers per job. Second, the degree of competition in each city determines how easy it is for a given employer to poach employees from its competitors. Furthermore, a local tension exists between tightness and competition: if all the productive employers locate their job in the least tight market, that city becomes very competitive, and most employers are eventually only able to hire a few workers.[44]

A tight characterization of the spatial job allocation remains possible even with this tension.[45] For this, I first solve for the job distribution in each city taking as given the vector of market tightness $\{\theta_\ell\}_{\ell=1}^{L}$. I then close the model by solving for the vector of market tightness.

**Proposition 2** (Local job distributions)**.**
*For two cities $\ell$ and $\ell'$, $\theta_\ell < \theta_{\ell'}$ if and only if $\Gamma_\ell \succ \Gamma_{\ell'}$. Furthermore, if $\theta_{\ell'} < \theta_\ell(1+k)^2$, jobs overlap in both cities.*

According to Proposition 2, sorting prevails in equilibrium: more productive jobs locate disproportionally to cities with less tight labor market. In particular, the local job distributions can be ranked in terms of first order stochastic dominance and Proposition 1 always holds. Yet, the equilibrium does not always feature pure positive assortative matching: if the search frictions are not too strong, some jobs are indeed present in multiple locations.[46]

Figure 5 helps visualize the overlap between the job distributions that can emerge in equilibrium. The left panel shows the number of employees per job as a function of the local productivity rank of the job for two cities, city 1 (orange) and city 2 (blue), and two levels of search frictions, low (solid) and high (dashed). The right panel displays the support of the local job distributions. When on-the-job search frictions are large, employers can only hire from unemployment. As a result, employers' size are independent of their wage offer and are proportional to the tightness of the labor market. High frictions attenuates the between-city differences in competition and perfect

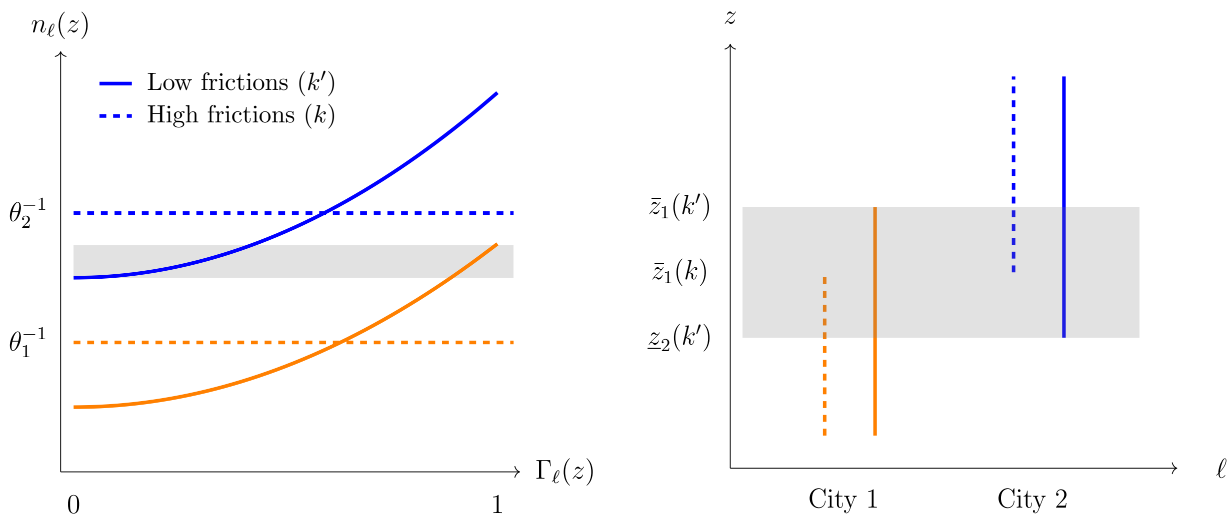
[43] In particular, the envelope condition guarantees that $\pi'_\ell(z) = n_\ell(z)$ almost everywhere.

[44] This tension is standard in models with directed search. In this theory, while job offers are randomly sent to workers within each city, employers can partially direct their search by choosing the location of their job.

[45] In equation (10), both the tightness the revenues per worker are affected by the allocation of jobs across space. This class of assignment models were coined matching problems with externality by Chade and Eeckhout (2020) and traditional optimal transport techniques cannot be used to solve for their solutions.

[46] In a relatively similar model, Lindenlaub et al. (2022) proves that perfect positive assortative always exist when search frictions are sufficiently large. I generalize this result by showing that sorting always prevails although it may not be perfect.

Figure 5: The effect of search frictions on the spatial job allocation




(a) Number of employees per job

(b) Job productivity distributions

sorting prevails, as visualized on Figure 5b. When frictions are smaller, high productivity employers offer high wages to poach workers from their competitors and the labor supply curves slop upward. When that is the case, an employer that offers a relatively high wage in a city with a tight labor market (city 1) attains the same size as an employer offering a relatively low wage in a city with a slack market (city 2). This overlap in the number of workers that a job can attract, visualized by the grey area on the left panel, leads several employers to be indifferent between the two locations – the grey area on the right panel.

Proposition 2 provides a characterization of the local job distributions in each city. To close the model, it is furthermore needed to solve for the number of vacancies in each locations, $\{v_\ell\}_{\ell=1}^L$. In equilibrium, as in standard assignment models, the sorting of jobs across places is sustained by a vector of housing prices that render the marginal employer indifferent between the two locations, $\pi_\ell(\underline{z}_\ell) = \pi_{\ell-1}(\underline{z}_\ell)$ for $\ell > 1$. This indifferent condition implies a difference equation

$$r_\ell - r_{\ell-1} = (\underline{z}_\ell - \underline{w}_\ell)n_\ell(\underline{z}_\ell) - (\underline{z}_{\ell-1} - \underline{w}_{\ell-1})n_{\ell-1}(\underline{z}_{\ell-1}) - \int_{\underline{z}_{\ell-1}}^{\underline{z}_\ell} n_{\ell-1}(\zeta)\mathrm{d}\zeta, \tag{12}$$

for $\ell > 1$, which, together with the consistency condition $\sum_{\ell=1}^L v_\ell = M$, yields a fixed point over the vector $\{v_\ell\}_{\ell=1}^L$. When congestion forces are large enough, this fixed point has a solution. Furthermore, the solution to this fixed point implies that cities with a thick labor market are the most attractive: they attract more employers and these employers are relatively more productive. Proposition 3 concludes the theoretical section by summarizing these two results.

**Proposition 3** (Existence)**.**
*If $\lambda^e \approx \lambda^u$ and $c \gg b$, there exists an equilibrium spatial allocation of jobs. In equilibrium, large cities attract more jobs and these jobs are relatively more productive:* $\mathrm{n}_\ell > \mathrm{n}_{\ell'} \iff v_\ell > v_{\ell'} \iff \theta_\ell < \theta_{\ell'}$.

Why are there differences in inequality across cities? Some cities have a comparative advantage in attracting employers because there are larger. Employers with relatively productive jobs locate in those cities to maximize their size. In each location, employers compete between one another in the local labor market to attract workers. In large cities where jobs are relatively more productive, this competition is fiercer as each employer has a higher willingness to pay. As a result, to attain their desired size, productive employers need to offer larger wages to poach workers from their competitors. The local fiercer competition thus leads to a steeper size wage premium, and with it, larger rent extraction for workers that are employed at productive jobs. On the other side of the job ladders, relatively low productivity employers hire most of their workforce from unemployment, and are thus not much affected by the spatial differences in competition. The local job ladders that emerge in equilibrium therefore have similar bottoms but largely differentiated tops, and when workers climb the ladder of large and competitive cities, they experience larger wage gains as they transition from one job to another. Altogether, the spatial sorting of jobs generates greater wage inequality in larger cities and rationalizes the three facts mentioned in Section 1.

# 3 Conclusion

In this paper, I have documented three novel facts on local wage inequality. First, cities with greater wage inequality feature a greater concentration of high-paying jobs while all locations offer low-paying jobs. Furthermore, in cities with a greater dispersion in job wage premium, the local job ladder is steeper and the firm-size wage premium is larger. I have rationalized these three facts through a theory of employer sorting in heterogeneous local frictional labor markets. In this framework, the sorting of employers generates endogenously fiercer competition in cities that attract the most productive jobs. The fiercer competition in turn implies a steeper local job ladder and with it a greater degree of inequality.

This paper has therefore laid out the microfoundations for the local root of wage inequality: to the extent that competition for workers is local and that inequality emerges out of that competition, cities contribute to aggregate wage inequality. The next step is therefore to estimate a quantitative version of the framework presented in Section 2 to quantify how much of the observed aggregate inequality is caused by the local root of wage inequality.

Finally, this framework could be extended in several fruitful directions. First, workers are here taken to be homogeneous. Yet, workers with different skills might react differently to firm sorting. A natural extension is therefore to study the interaction between firm and worker sorting in a frictional setting with local labor markets. Second, this framework could be used to study the effect of an increase in aggregate firm concentration on local concentration and wage inequality (Rossi-Hansberg et al., 2021; Autor et al., 2023).

## References


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

Autor, D., Patterson, C., and Van Reenen, J. (2023). Local and national concentration trends in jobs and sales: The role of structural transformation.

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

Azar, J., Marinescu, I., Steinbaum, M., and Taska, B. (2020). Concentration in us labor markets: Evidence from online vacancy data. *Labour Economics*, 66:101886.

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.

Benmelech, E., Bergman, N. K., and Kim, H. (2020). Strong employers and weak employees: How does employer concentration affect wages? *Journal of Human Resources*, pages 0119–10007R1.

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

Berry, C. R. and Glaeser, E. L. (2005). The divergence of human capital levels across cities. *Papers in regional science*, 84(3):407–444.

Bilal, A. (2020). The geography of unemployment.

Bilal, A. and Lhuillier, H. (2022). Outsourcing, inequality and aggregate output.

Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M., and Setzler, B. (2022). How much should we trust estimates of firm effects and worker sorting?

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

Brinkman, J. (2014). The supply and demand of skilled workers in cities and the role of industry composition.

Brooks, W. J., Kaboski, J. P., Li, Y. A., and Qian, W. (2021). Exploitation of labor? classical monopsony power and labor's share. *Journal of Development Economics*, 150:102627.

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

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

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

Chan, M., Salgado, S., and Xu, M. (2021). Heterogeneous passthrough from tfp to wages.

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., and Roux, S. (2012). Sorting and local wage and skill distributions in france. *Regional Science and Urban Economics*, 42(6):913–930.

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

D'Costa, S. and Overman, H. G. (2014). The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics*, 48:168–179.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.

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

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

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

Gautier, P. A. and Teulings, C. (2009). Search and the city. *Regional Science and Urban Economics*, 39:251–265.

Giannone, E. (2017). Skilled-biased technical change and regional convergence.

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

Gouin-Bonenfant, E. (2022). Productivity dispersion, between-firm competition, and the labor share.

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

Hendricks, L. (2011). The skill composition of us cities. *International Economic Review*, 52(1):1–32.

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

Kline, P., Saggio, R., and Sølvsten, M. (2020). Leave-out estimation of variance components. *Econometrica*, 88(5):1859–1898.

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

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

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

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

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

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

Rossi-Hansberg, E., Sarte, P.-D., and Schwartzman, F. (2019). Cognitive hubs and spatial redistribution.

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.

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

Yeh, C., Macaluso, C., and Hershbein, B. J. (2022). Monopsony in the us labor market.

# A Descriptive Evidence

Figure A.1: Distributions of log hourly wage by city in France

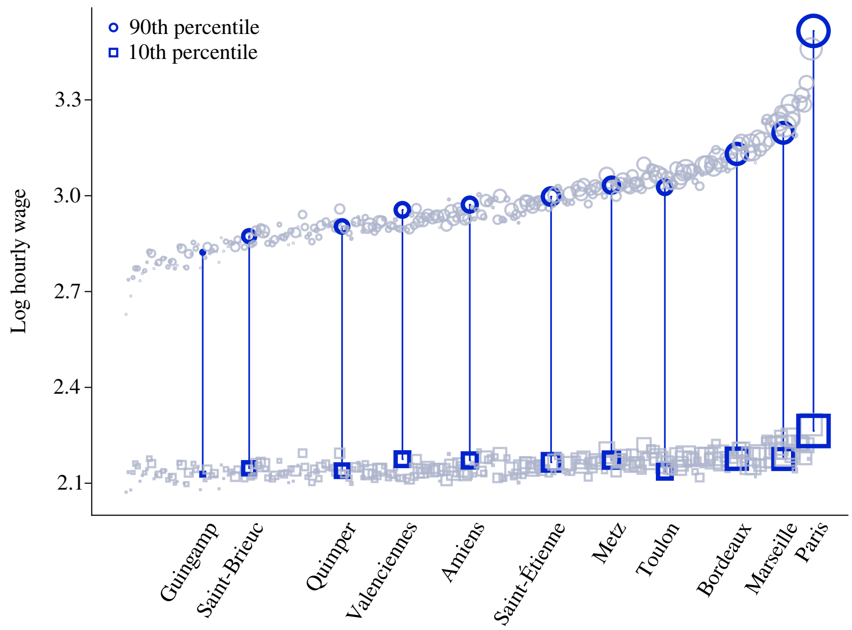
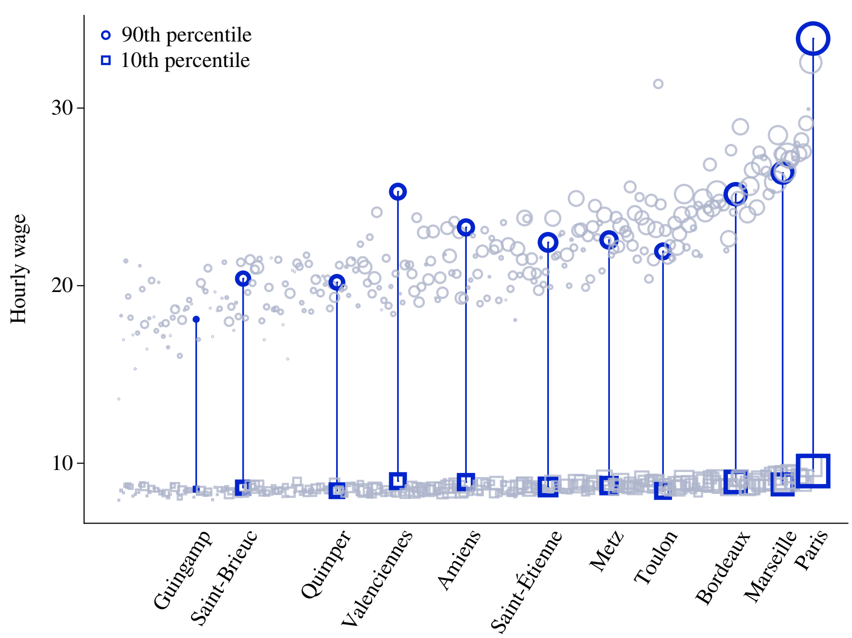
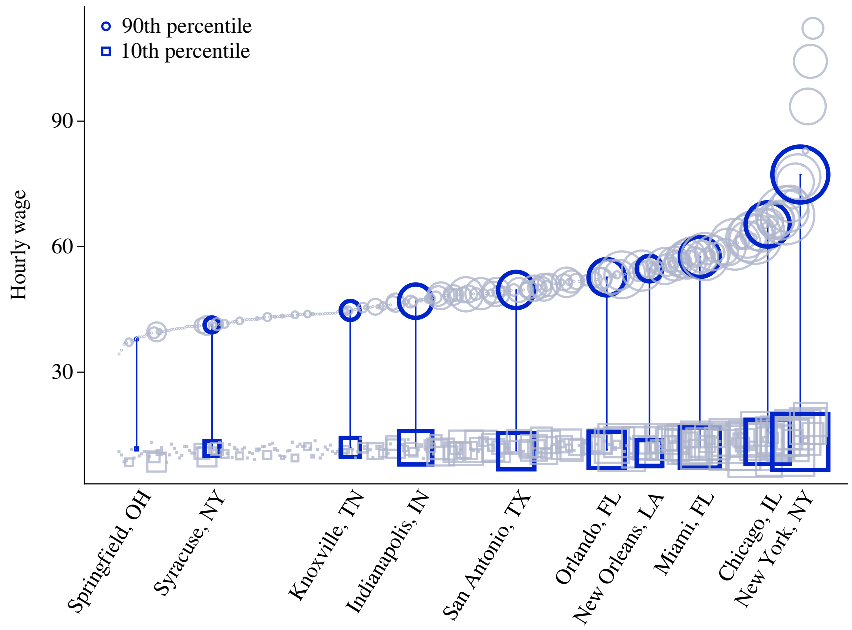



Figure A.2: Distributions of hourly wage by city in France for movers




Reproduction of Figure 1 for movers. Movers are defined as workers who work in commuting zones different from the ones they were born in.

Figure A.3: Distributions of hourly wage by city in the United States




Reproduction of Figure 1 in the United States. Each circles and squares represent a Metropolitan statistical area in the United States. Data from the American Community Survey (2011-2020).

## *A.1 The local importance of job heterogeneity*

Take the empirical model (1),

$$\omega_{i,t} = \alpha_i + \beta_{\ell(i,t)} \cdot x_{i,t} + \gamma_{j(i,t)} + \varepsilon_{i,t}.$$

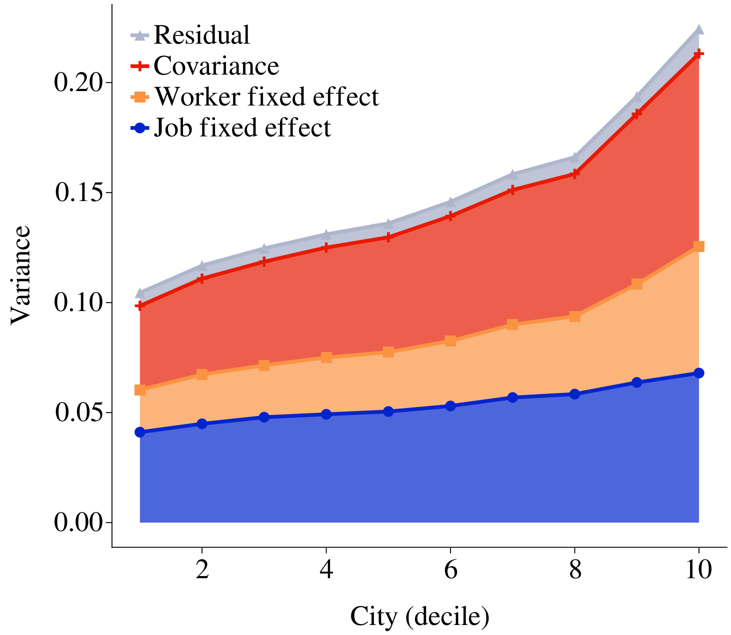
To limit concerns regarding limited mobility biases, I group jobs in 1,000 clusters according to their unconditional average wage. The above model is then estimated via OLS. Applying the variance operator on each side of the above equation conditionally on workers being located in city $l$ returns

$$\begin{aligned} V_l \equiv \mathrm{V}[\omega_{i,t} \mid \ell(i,t) = l] = \; & \underbrace{\mathrm{V}[\alpha_i + \beta_l \cdot x_{i,t} \mid l] + \mathrm{Cov}[\alpha_i + \beta_l \cdot x_{i,t}, \gamma_{j(i,t)} \mid l]}_{\text{Worker contribution } \equiv \, W_l} + \\ & \underbrace{\mathrm{V}[\gamma_{j(i,t)} \mid l] + \mathrm{Cov}[\alpha_i + \beta_l \cdot x_{i,t}, \gamma_{j(i,t)} \mid l]}_{\text{Job contribution } \equiv \, J_l} + \\ & \underbrace{\mathrm{V}[\varepsilon_{i,t} \mid l]}_{\text{Residual}}, \end{aligned} \tag{13}$$

where half of the covariance term between the worker and the job effects is allocated in the worker contribution, and the other half is in the job contribution. This is the variance decomposition used in Figure 2a. In Figure A.4, I separate the covariance term from the decomposition. Furthermore, overall wage dispersion can be decomposed in a between-city and within-city component,

$$\mathrm{V}[w] = \underbrace{\mathrm{V}[\mathbb{E}(w \mid \ell)]}_{\text{Between-city}} + \underbrace{\mathbb{E}[\mathrm{V}(w \mid \ell)]}_{\text{Within-city}}.$$

Figure A.4: Variance decomposition of wages across cities




Variance decomposition based on (13) after taking out the covariance terms from the worker and job contributions. A job is defined as a pair occupation-establishment. Jobs are grouped in 1,000 bins according to their unconditional expected wages. Blue area: the job contribution. Orange area: the worker contribution. Red area: covariance between the worker and job fixed effect. Grey area: the residual.

Plugging in the specification (1), we thus have

$$\begin{aligned} \mathrm{V}[w] = \mathrm{V}[\mathbb{E}(\alpha_{i,t} \mid \ell)] + \mathbb{E}[\mathrm{V}(\alpha_{i,t} \mid \ell)]\mathbb{E}[\mathrm{C}(\alpha_{i,t}, \gamma_{j(i,t)} \mid \ell)] + \mathbb{E}[\mathrm{V}(\varepsilon_{i,t} \mid \ell) \mid \ell)] + \\ \underbrace{\mathrm{V}[\mathbb{E}(\gamma_{j(i,t)} \mid \ell)]}_{\text{Job between-city}} + \underbrace{\mathbb{E}[\mathrm{V}(\gamma_{j(i,t)} \mid \ell)] + \mathbb{E}[\mathrm{C}(\alpha_{i,t}, \gamma_{j(i,t)} \mid \ell)]}_{\text{Job within-city}}, \end{aligned} \tag{14}$$

where $\alpha_{i,t} \equiv \alpha_i + \beta_{\ell(i,t)} \cdot x_{i,t}$ and I have assumed for simplicity that the average residual is minimal in each city. Footnote 14 uses this decomposition. In particular, the between-city differences in job explain 40% of the overall between-city differences in wages, $\mathrm{V}[\mathbb{E}(\alpha_{i,t} \mid \ell)] + \mathrm{V}[\mathbb{E}(\gamma_{j(i,t)} \mid \ell)]$, but 4% of the overall wage dispersion, $V[w]$.

In the main text, I report two elasticities that relate job fixed effects to local inequality. The first one is at the city-level and correlates the local variance of jobs with local inequality,

$$J_l = \varphi + \zeta V_l + \nu_l,$$

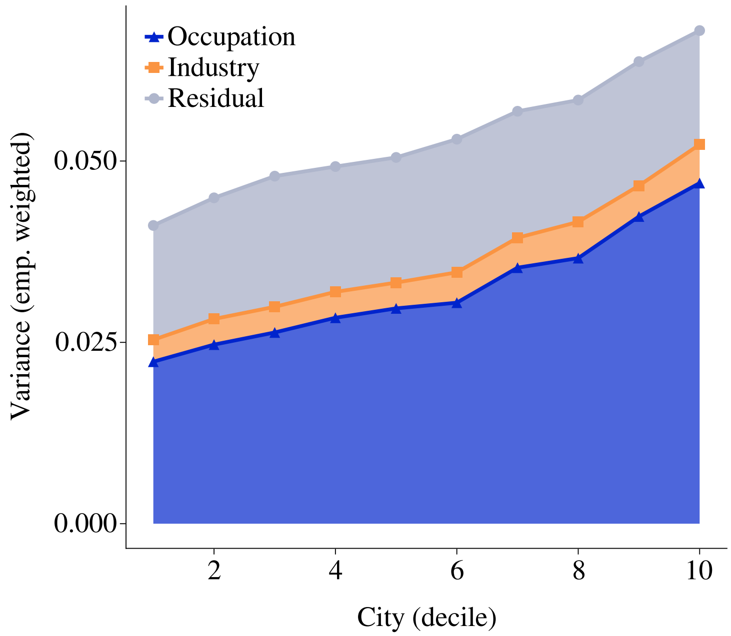
where the variables are defined in (13). The second one is at the individual level and correlates job fixed effects with local inequality. Specifically, I perform quantile regressions of job fixed effects on local wage inequality for various deciles. The results are reported in Figure A.6.

To understand the relative importance of occupation and industry heterogeneity in explaining job heterogeneity, I estimate the following model

$$\gamma_{j(i,t)} = \chi_{o(i,t)} + \phi_{s(i,t)} + \epsilon_{i,t}, \tag{15}$$

where the variable on the left-hand side corresponds to the job fixed effects estimated from (1). The variable $o(i,t)$ and $s(i,t)$ returns respectively the occupation and the sector in which worker $i$ workers at time $t$.

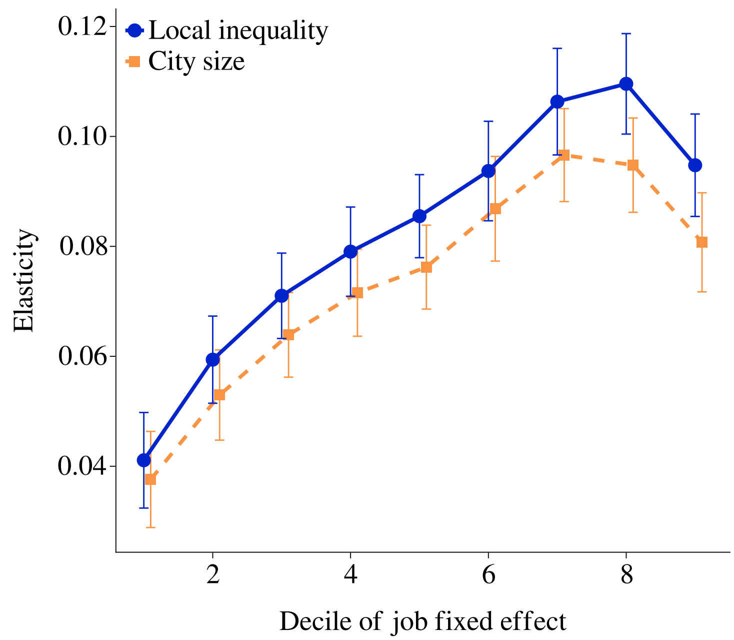
Figure A.5: Variance decomposition of job fixed effects across cities




Variance decomposition based on the empirical model (15). A job is defined as a pair occupation-establishment. Jobs are grouped in 1,000 bins according to their unconditional expected wages. Occupations are 2-digit occupations and industry are 4-digit industries. As in (13), the covariance term between the occupation and industry fixed effects is allocated equally to the occupation and industry contribution. Blue area: variance of the occupation fixed effect. Orange area: variance of the industry fixed effect. Grey area: variance of the residual.

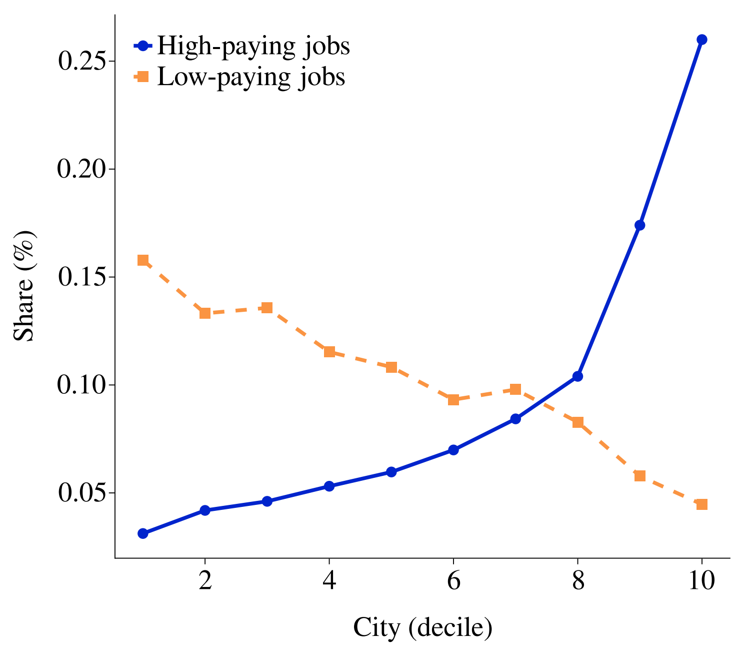
A decomposition similar to that in equation (13) can then be implemented. The results are discussed in footnote 15 and are plotted in Figure A.5.

Figure A.6: Estimated effects of local characteristics on the job distribution




Job fixed effects estimated from (15). A job is defined as a pair occupation-establishment. Jobs are grouped in 1,000 bins according to their unconditional expected wages. Effects of local characteristics on job fixed effect estimated for various deciles through quantile regressions. Blue solid line: estimates from quantile regression of job fixed effects on local inequality (standardized). Orange dashed line: estimates from quantile regression of job fixed effects on log city size (standardized).

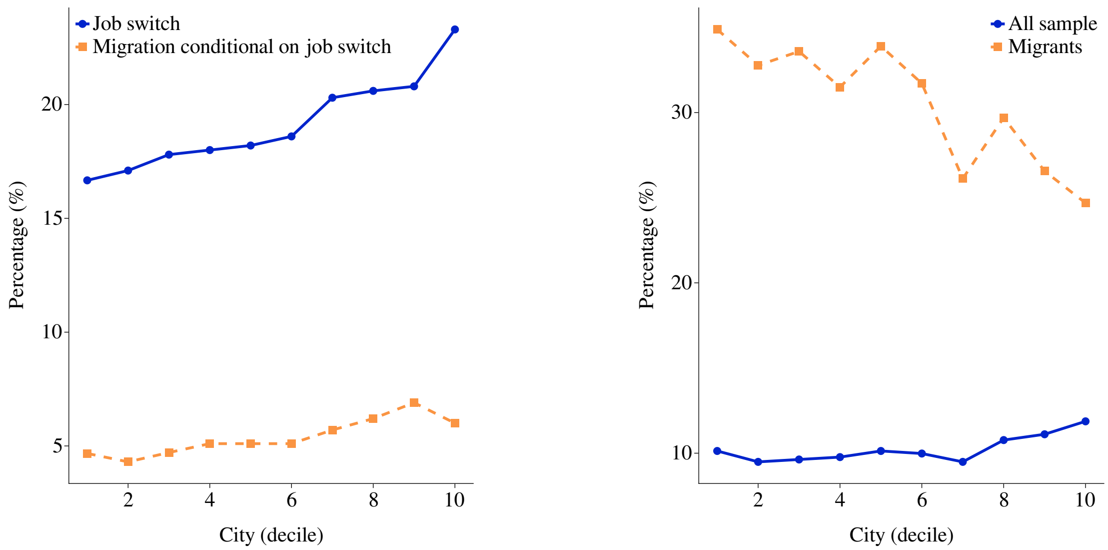
Figure A.7: Share of low- and high-paying jobs across cities




Job fixed effects estimated from (15). A job is defined as a pair occupation-establishment. Jobs are grouped in 1,000 bins according to their unconditional expected wages. A high-paying job defined as a job with a fixed effect above the 90th percentile of the aggregate distribution. A low-paying job defined as job with a fixed effect below the 10th percentile. Blue solid line: fraction of high-paying jobs in a city decile. Orange dashed line: fraction of low-paying jobs in a city decile.

## A.2 *The existence of city-specific job ladders*

Figure A.8: Job transition and migration rates by cities

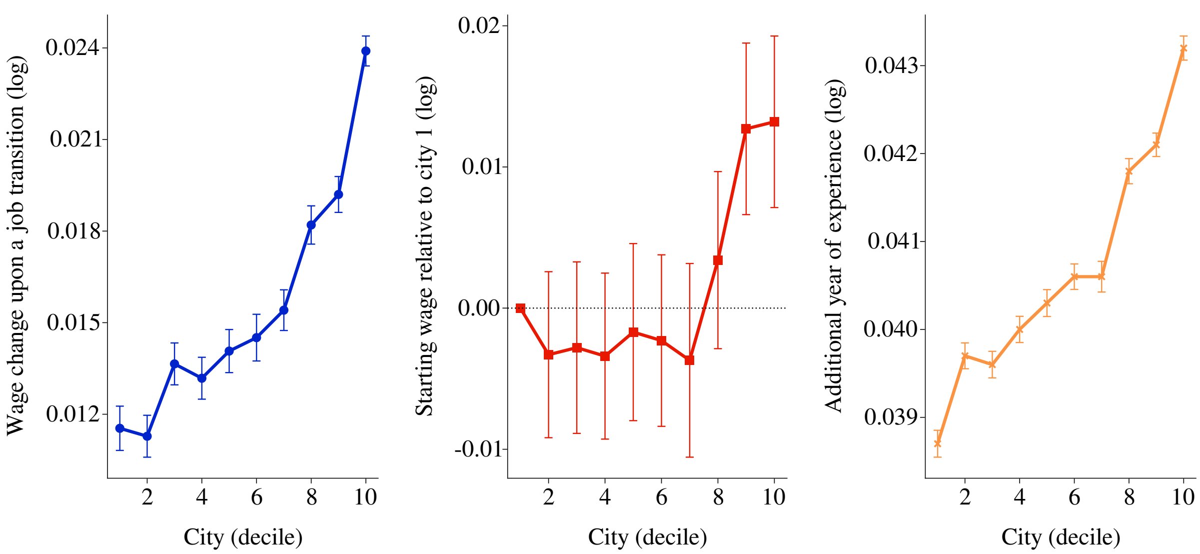



(a) Job switching and migration rate

(b) Workers previously unemployed

A job is defined as an occupation-establishment. A job transition is defined as a transition between two jobs where the end-date of the previous job was less than a quarter before the start-date of the current job. A migration is defined as a transition between two commuting zones of residence. A worker is defined as previously unemployed if the end-date of the last job was more than a quarter before the start-date of the current job.

Figure A.9: The role of job transitions and learning for local wage growth



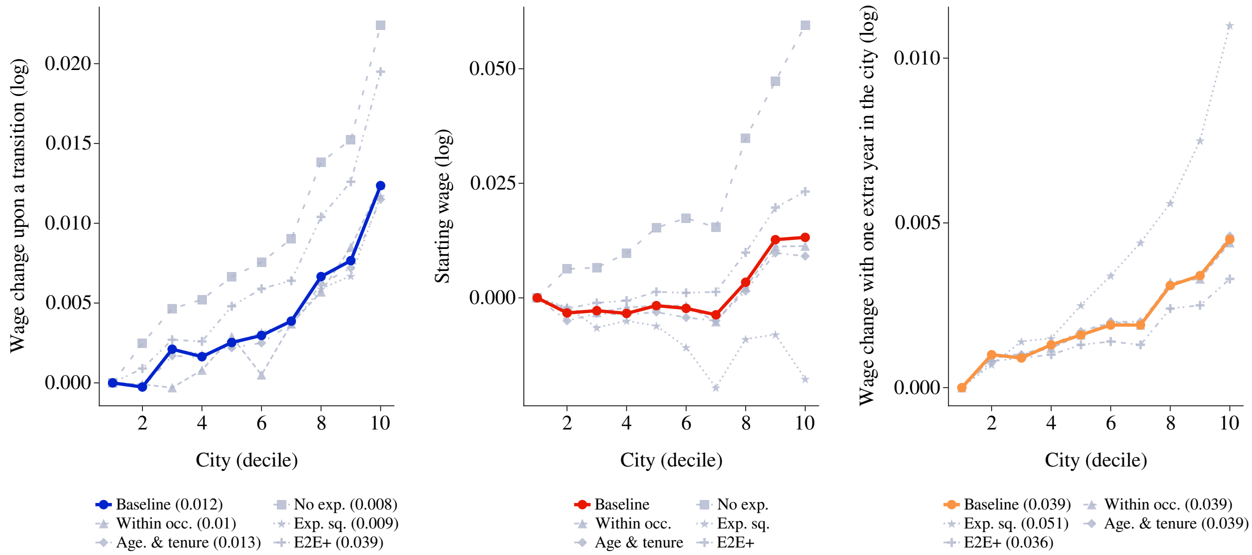

(a) Job transition

(b) Experience

(c) Experience

Point estimates from specification (2). A job is defined as an occupation-establishment. A job transition is defined as a transition between two jobs where the end-date of the previous job was less than a quarter before the start-date of the current job. Panel (a): change in log wages generated by a job transition, i.e. $\{\gamma_\ell\}_\ell$. Panel (b): change in log wages generated by one additional year of experience, i.e. $\{\delta_\ell\}_\ell$. Panel (c): starting wage in city $\ell$, i.e. $\{\alpha_\ell\}_\ell$. Error bars represent the 95% confidence interval, with standard errors clustered at the worker level.

Figure A.10: The role of job transitions and learning for local wage growth, robustness



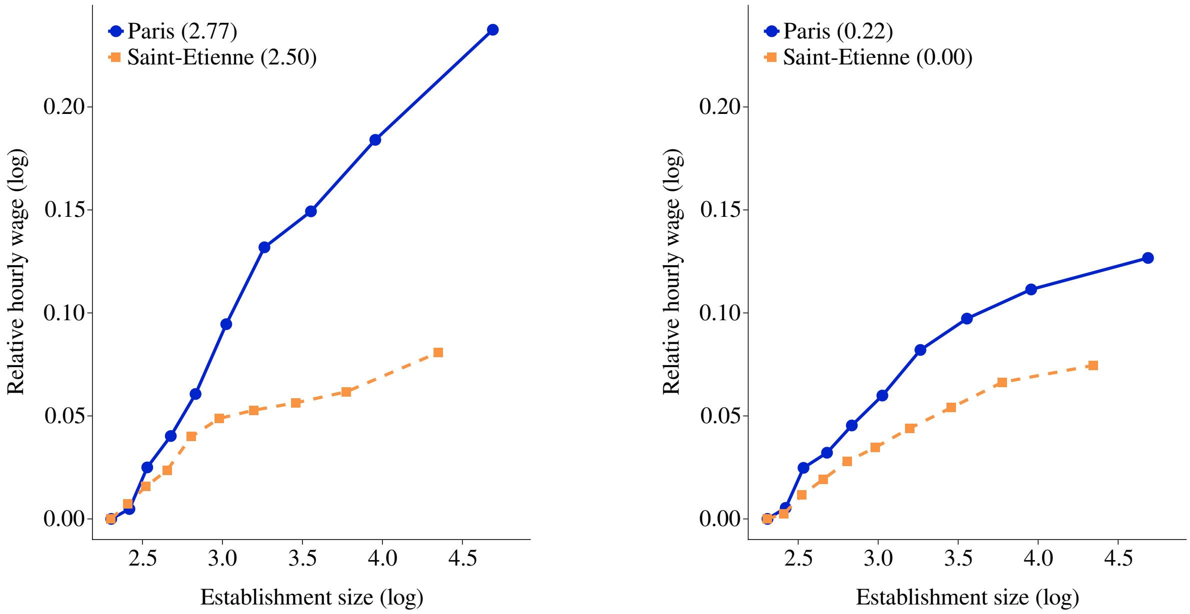

(a) Job transition (b) Starting wage (c) Learning

Point estimates in estimated from specification (2). A job is defined as an occupation-establishment. A job transition is defined as a transition between two jobs where the end-date of the previous job was less than a quarter before the start-date of the current job. Panel (a): change in log wages generated by a job transition, i.e. $\{\gamma_\ell\}_\ell$. Panel (b): change in log wages generated by one additional year of experience, i.e. $\{\delta_\ell\}_\ell$. Panel (c): starting wage in city $\ell$, i.e. $\{\alpha_\ell\}_\ell$. Error bars represent the 95% confidence interval, with standard errors clustered at the worker level. Squares: do not control for experience. Triangles: include 2-digit occupation fixed effects, and define a job transition as a switch across establishments within a 2-digit occupation. Stars: include a city-specific quadratic term for experience. Diamonds: control city-specific effects of age and within-establishments tenure in addition to the effects of experience. Crosses: job transition restricted to transitions associated with a positive wage change.

## *A.3 Between-city differences labor market competition*

The size elasticity $\{\delta\}_\ell$ estimated from (3) have a causal interpretation to the extent that we can rule out reverse causality and omitted variables concerns. Reverse causality is a concern if an increase in worker $i$'s productivity triggers a simultaneous increase in their wage and in the size of the firm. Meanwhile, if hiring better skilled workers increases the productivity of current workers through learning spillovers, not controlling for the changes in the average worker productivity would also bias the results. These concerns are of limited validity in the present setting. First, because I consider short time difference of one year. To the extent that it takes time for knowledge to spillover from one worker to another or from a worker to the firm's overall productivity, this limits the concerns about the simultaneity of the increase in workers' skills and in firm size. Second, as for Figure 4a, the sample is restricted to establishments with more than ten employees. This ensures that that any worker has a limited impact on other workers' and the firm's productivity.

Figure A.11: The firm-size wage premia across cities

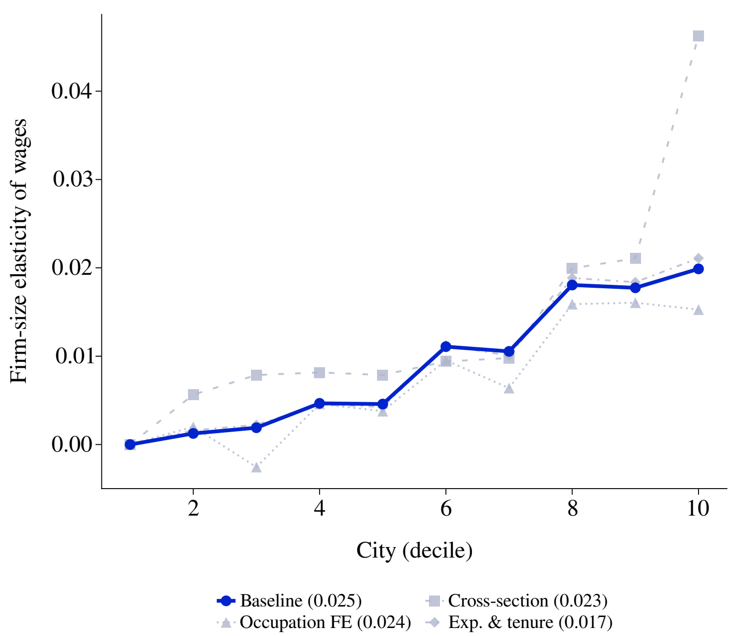



(a) Raw data

(b) Year and industry fixed effects

Left panel: bin scatter plot of average hourly wage by deciles of the firm size distributions for Paris (blue solid line) and Saint-Etienne (orange dashed line). Right panel: bin scatter plot of average hourly wage by deciles of the firm size distributions for Paris (blue solid line) and Saint-Etienne (orange dashed line) after controlling for year and 4-digit industry fixed effects. The within-city level of the firm size distribution is computed as the city fixed effect plus a weighted average of the industry fixed effects, where the weights are the employment share of the industries at the nationwide level. Sample selection for both panels: firms with more than 10 employees. The sample is then further winsorized at the 5% level.

Figure A.12: Firm-size elasticity of wages




Solid blue line: firm-size elasticity of wages, i.e. $\{\gamma_\ell\}_\ell$ in specification (3). Squares: firm-size elasticity of wages estimated in the cross-section and not in first difference. Triangles: include 2-digit occupation fixed effect. Diamonds: include controls for experience and tenure at the establishments. Sample selection: growing firms with more than 10 but less than 500 employees. Winsorization at the 5% level.

# B Proof

## B.1 Reservation wage

Omit the location indices. Suppose without loss of generality that $F$ admits a density. Differentiating the value of employment with respect to $w$ yields

$$V'(w) = \frac{1}{\rho + \delta + \lambda^e \bar{F}(w)}.$$

Integrating back with respect to $w$ returns

$$V(w) = U + \int_{\underline{w}}^{w} \frac{1}{\rho + \delta + \lambda^e \bar{F}(w')} \mathrm{d}w'.$$

Combining the above expression with the expression for the value of unemployment, I obtain

$$\rho U = b + \lambda^u \int_{\underline{w}} \frac{\bar{F}(w)}{\rho + \delta + \lambda^e \bar{F}(w)} \mathrm{d}w. \tag{16}$$

By definition, $V(\underline{w}) = U$, and therefore

$$(\rho + \lambda^u)U = b + \lambda^u \int_{\underline{w}} V(w) f(w) \mathrm{d}w,$$
$$(\rho + \lambda^e)U = \underline{w} + \lambda^e \int_{\underline{w}} V(w) f(w) \mathrm{d}w.$$

Together, these imply $\rho U = (\lambda^u \underline{w} - \lambda^e b)/(\lambda^u - lE)$, which plugged back into (16) yields (5) when $\rho/\delta \to 0^+$.

## B.2 Local supply curves

The local supply curves are first derived by obtaining $G_\ell(w)$, the fraction of workers in city $\ell$ with wage lower than $w$. Since there are no movements between cities, $G_\ell$ can be characterized solely by looking at the within-city worker flows. 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 $\lambda^u F_\ell(w) u_\ell = [\delta_\ell + \lambda^e \bar{F}_\ell(w)] e_\ell G_\ell(w)$. The left-hand side is the flow of workers into the wage interval $[\underline{w}_\ell, w)$, constituted of unemployed workers receiving a wage offer lower than $w$. The right-hand side is the flow of workers out of the wage interval, made of those workers employed at a job with wage under $w$ that either lose their job or find a better-paid job opportunity. Solving for $G_\ell(w)$ yields (6).

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

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

where $F_\ell(w^-) = \lim_{\varepsilon \to 0^-} F(w - \varepsilon)$. This is equivalent to equation (7) once mass points have been ruled out.

## B.3 Local wage offer distributions

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

I now show that $F_\ell[w_\ell(z)] = \Gamma_\ell(z)$. Since $n_\ell$ is strictly increasing in $w$, $\pi_\ell$ 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 (9). Since there are no mass point in the wage offer distribution, we can take the first-order conditions of (11) with respect to $w$ for almost every job productivity $z$,

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

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

$$w_\ell'(z) = \frac{n_\ell'(z)}{n_\ell(z)} \left(z - w_\ell(z)\right), \tag{17}$$

after using $R(z, n) = z \cdot n$. Integrating this ODE with respect to $w$ returns (9).

## *B.4 Proof of Proposition 1*

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

To prove Proposition 1.1 and 1.2, it is first needed to show that the reservation wage is increasing in $\Gamma_\ell$. Under the change of variable $w \to w_\ell(z)$, the reservation wage (5) becomes

$$\underline{w}_\ell = b + (k^u - k) \int_{\underline{z}_\ell} \frac{\bar{\Gamma}_\ell(z) w_\ell'(z)}{1 + k\bar{\Gamma}_\ell(z)} \mathrm{d}z = b + (k^u - k) O_\ell, \tag{18}$$

where $O_\ell$ is the option value of search. Using the wage FOC (17), the option value of search is

$$O_\ell = \frac{1}{1 + k} \left( \int z\bar{\Gamma}_\ell(z)\eta_\ell'(z)\mathrm{d}z - \frac{\underline{w}_\ell k}{1 + k} \right).$$

Combining this expression with (18) returns

$$O_\ell = \mathcal{K} \left( \int z\bar{\Gamma}_\ell(z)\eta_\ell'(z)\mathrm{d}z - \frac{bk}{1 + k} \right), \tag{19}$$

where $\mathcal{K}$ is a parametric constant.[47] Finally, using the change of variable $\Gamma_\ell(z) \to q$, the option value of

[47] Precisely,

$$\mathcal{K} = \frac{1}{1 + k} \left( 1 + \frac{k(k^u - k)}{(1 + k)^2} \right)^{-1}$$

search in city $\ell$ is

$$O_\ell = \mathcal{K}\left(\int \Gamma_\ell^{-1}(q)(1-q)\frac{\partial \eta^Q(q)}{\partial q}\mathrm{d}q - \frac{bk}{1+k}\right).$$

This expression is independent of $\theta_\ell$. It follows that, if $\Gamma_\ell \succ \Gamma_{\ell'}$, then $O_\ell > O_{\ell'}$ and $\underline{w}_\ell \geq \underline{w}_{\ell'}$, where the inequality is strict whenever $\lambda^e < \lambda^u$.

**FOSD ordering (1.1)** The wage paid by the employers at the $q$-th quantile of the wage distribution in city $\ell$ reads

$$\begin{aligned} w_\ell[\Gamma_\ell^{-1}(q)] &= \underline{w}_\ell\left(\frac{\eta^Q(0)}{\eta^Q(q)}\right) + \int_{\underline{z}_\ell}^{\Gamma_\ell^{-1}(q)} \zeta\left(\frac{\eta_\ell'(\zeta)}{\eta^Q(q)}\right)\mathrm{d}\zeta \\ &= \underline{w}_\ell\left(\frac{\eta^Q(0)}{\eta^Q(q)}\right) + \int_0^q \Gamma_\ell^{-1}(x)\left(\frac{\partial_x \eta^Q(x)}{\eta^Q(q)}\right)\mathrm{d}x, \end{aligned} \tag{20}$$

where the second equality follows from the change of variable $\Gamma_\ell(\zeta) \to x$. Take now two cities so that $\Gamma_\ell \succ \Gamma_{\ell'}$. We already know that $\underline{w}_\ell \geq \underline{w}_{\ell'}$. Hence, for all $w \in [\underline{w}_{\ell'}, \underline{w}_\ell]$, we have $F_{\ell'}(w) \geq F_\ell(w) = 0$. From (20), $w_{\ell'}[\Gamma_\ell^{-1}(q)] < w_\ell[\Gamma_\ell^{-1}(q)]$ for all $q \in (0,1]$. In particular, $\bar{w}_{\ell'} < \bar{w}_\ell$, so that $F_{\ell'}(w) = 1 > F_\ell(w)$ for all $w \in [\bar{w}_{\ell'}, \bar{w}_\ell)$. Finally, for any $w \in (\underline{w}_\ell, \bar{w}_{\ell'})$, (20) implies that $w_\ell[\Gamma_\ell^{-1}(q_\ell)] = w = w_{\ell'}[\Gamma_{\ell'}^{-1}(q_{\ell'})]$ if and only if $q_\ell > q_{\ell'}$. Hence, $F_{\ell'}(w) = F_{\ell'}\{w_{\ell'}[\Gamma_{\ell'}^{-1}(q_l)]\} = \Gamma_{\ell'}[\Gamma_{\ell'}^{-1}(q_{\ell'})] = q_{\ell'} > q_\ell = \Gamma_\ell[\Gamma_\ell^{-1}(q_\ell)] = F_\ell\{w_\ell[\Gamma_\ell^{-1}(q_\ell)]\} = F_\ell(w)$, where the second equality follows from the rank-preserving property of the job ladder. Concluding, we have $F_{\ell'}(w) \geq F_\ell(w)$ for all $w$ and $F_{\ell'}(w) > F_\ell(w)$ on the joint support of $F_\ell$ and $F_{\ell'}$, so that $F_\ell \succ F_{\ell'}$. From equation (6), it automatically follows that $G_\ell \succ G_{\ell'}$.

**Inequality (1.2)** Let $\Delta_{\ell\ell'} \equiv \bar{w}_{\ell'} - \underline{w}_\ell' - (\bar{w}_{\ell'} - \underline{w}_{\ell'})$ be the difference in top-to-bottom wage gap between city $\ell$ and $\ell'$. Using (9), this difference writes

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

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

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

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

**E2E gains (1.3)** The $\ell$ are dropped as long as comparisons across cities are absent. The expected wage growth upon a job-to-job transition in a given city is

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

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

$$\Delta^{\mathrm{EE}}[F^{-1}(q)] = \frac{1}{1-q}\int_{F^{-1}(q)} \frac{w'}{F^{-1}(q)}\mathrm{d}F(w'). \tag{21}$$

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

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

If $\Gamma_\ell \succ G_{\ell'}$, then $F_\ell \succ F_{\ell'}$, and therefore $\Delta_\ell^{\text{EE}}(\underline{w}_\ell) < \Delta_{\ell'}^{\text{EE}}(\underline{w}_{\ell'})$. Since the inequality is strict and (21) is continuous in $q$, this holds for workers on low rungs of the ladder.

**Size wage premium (1.4)** From the first-order condition (17), the size wage premium of some employer $z$ is

$$\text{FP}(z) \equiv \frac{\partial_z w(z)}{\partial_z n(z)} = \frac{z - w(z)}{n(z)}$$

The employment-weighted average size wage premium is therefore

$$\text{FP} \equiv \int \text{FP}(z) n(z) \mathrm{d}\Gamma(z) = \frac{1}{3} \int z \left(1 + \frac{2}{(1 + k\bar{\Gamma}(z))^3}\right) \mathrm{d}\Gamma(z) - c_{\underline{w}} \underline{w}$$

for $c_{\underline{w}}$ some constant.[48] If $\lambda^e \approx \lambda^u$, then $\underline{w} \approx b$, and it follows that if $\Gamma_\ell \succ \Gamma_{\ell'}$ then $\text{FP}_\ell > \text{FP}_{\ell'}$.

## *B.5 Proof of Proposition 2*

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

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

*Proof.* From (10), $\pi'_\ell(z)$ is differentiable. Differentiating with respect to $z$ returns $\pi'_\ell(z) = n_\ell(z)$. Since $F_\ell$ has no mass point, $n_\ell(z)$ is continuous in $z$. □

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

*Proof.* Suppose not. That is, for some city $\ell$, there exists at least one hole in the support of $\Gamma_\ell$. 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 (11) then requires

1. $\pi_\ell(z') \geq \pi_{\ell'}(z')$ for all $z' \in [\underline{z}_\ell, z]$ and all cities $\ell'$,
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 $\ell'$.

The first two conditions occur only if $\pi'_\ell(z') < \pi'_{\ell^\star}(z')$ for $z' \in B^+(z)$.[49] Using the Envelope condition, this rewrites $n_\ell(z') = n_\ell(z) = \pi'_\ell(z) < \pi'_{\ell^\star}(z') = n_{\ell^\star}(z')$, where the first equality follows from $n_\ell$ being

[48]Specifically, we have

$$c_{\underline{w}} \equiv \frac{3 + k(3+k)}{3(1+k)^2}$$

[49]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 $x$.

constant outside of the support of $\Gamma_\ell$. It follows that $\pi'_\ell(z') < \pi'_{\ell^\star}(z')$ for all $z' \in (z, z+\varepsilon)$, and therefore $\pi_\ell(z+\varepsilon) < \pi_{\ell^\star}(z+\varepsilon)$, a contradiction with the third condition. $\square$

**Lemma 3** (First order stochastic dominance ordering)**.**
*For two cities $\ell$ and $\ell'$ and a vector of tightness that is equilibrium compatible, $\theta_\ell < \theta_{\ell'}$ if and only if $\Gamma_\ell \succ \Gamma_{\ell'}$.*[50]

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

$$\pi'_\ell(z) = \frac{1}{\theta_\ell}\frac{1+k}{(1+k\bar{\Gamma}_\ell(z))^2} \geq \frac{1}{\theta_\ell}\frac{1+k}{(1+k\bar{\Gamma}_{\ell'}(z))^2} > \frac{1}{\theta_{\ell'}}\frac{1+k}{(1+k\bar{\Gamma}_{\ell'}(z))^2} = \pi'_{\ell'}(z). \tag{22}$$

If $\Gamma_\ell(z), \Gamma_{\ell'}(z) \in (0,1)$, Lemma 2 implies that $\pi_\ell(z) = \pi_{\ell'}(z)$. However, from equation (22), $\pi'_\ell(z') > \pi'_{\ell'}(z')$ for $z' \in B^+(z)$, and therefore $\pi_\ell(z') > \pi_{\ell'}(z')$. Lemma 2 then implies $z = \bar{z}_{\ell'}$ and $z < \bar{z}_\ell$, or $\bar{\Gamma}_{\ell'}(z) = 0 < \bar{\Gamma}_\ell(z)$, a contradiction.

If $\Gamma_{\ell'}(z) = 0$ and $\Gamma_\ell(z) \in (0,1)$, then it must from Lemma 2 that $\pi_{\ell'}(z) \leq \pi_\ell(z)$. Furthermore, the positive measure of jobs locating in city $\ell'$ requires the existence of $\underline{z}_{\ell'} \geq z$ such that $\pi_{\ell'}(\underline{z}_{\ell'}) \geq \pi_\ell(\underline{z}_{\ell'})$. If $\underline{z}_{\ell'} = z$, then it must be that $\pi_{\ell'}(\underline{z}_{\ell'}) = \pi_\ell(\underline{z}_\ell)$. However, (22) implies that $\pi'_{\ell'}(z') < \pi'_\ell(z')$ for all $z' \in B^+(\underline{z}_{\ell'})$, and therefore $\pi_{\ell'}(z') < \pi_\ell(z')$ for $z' \in B^+(\underline{z}_{\ell'})$, which contradicts Lemma 2. Hence, it must be that $\underline{z}_{\ell'} > z$. But (22) implies $\pi'_{\ell'}(z') < \pi'_\ell(z')$ for all $z' \in (z, \underline{z}_{\ell'})$, and therefore $\pi_{\ell'}(\underline{z}_{\ell'}) < \pi_\ell(\underline{z}_{\ell'})$, a contradiction with the definition of $\underline{z}_{\ell'}$. It follows that $\underline{z}_{\ell'}$ does not exist, a contradiction.

Finally, suppose $\Gamma_\ell(z) = 1$, so that $z \geq \bar{z}_\ell$. If $\Gamma_{\ell'}(z) \in (0,1)$, then $\bar{z}_{\ell'} > \bar{z}_\ell$. Meanwhile, it must also be that $\pi_{\ell'}(\bar{z}_\ell) \leq \pi_\ell(\bar{z}_\ell)$. Finally, for all $z' \in [\bar{z}_\ell, \bar{z}_{\ell'})$, we have $\Gamma_{\ell'}(z') < \Gamma_\ell(z') = 1$, and therefore $\pi'_{\ell'}(z') < \pi'_\ell(z')$ from (22). However, this implies $\pi_{\ell'}(z') < \pi_\ell(z')$ for all $z' \in [\bar{z}_\ell, \bar{z}_{\ell'})$, which contradicts $\bar{z}_{\ell'} > \bar{z}_\ell$ from Lemma 2. Hence, suppose that $\Gamma_{\ell'}(z) = 0$. However, in this case, a similar argument as in the previous paragraph shows that there does not exist a $\underline{z}_{\ell'}$, a contradiction. Hence, the only possibility is $\Gamma_{\ell'}(z) = 1$, a contradiction of $\Gamma_{\ell'} \succ \Gamma_\ell$.

The other direction automatically follows. Suppose not $\theta_\ell < \theta_{\ell'}$, such that $\theta_\ell \geq \theta_{\ell'}$. If $\theta_\ell > \theta_{\ell'}$, then we have already shown that $\Gamma_{\ell'} \succ \Gamma_\ell$, or not $\Gamma_\ell \succ \Gamma_{\ell'}$. If $\theta_{\ell'} = \theta_\ell$, it must be that $\Gamma_{\ell'}(z) = \Gamma_\ell(z)$ for all $z$. To see this, suppose for the sake of contradiction that $\bar{\Gamma}_{\ell'}(z) > \bar{\Gamma}_\ell(z)$ for some $z$. Note that, as in (22), this implies $\pi'_\ell(z) > \pi'_{\ell'}(z)$, and it is possible to re-use the same arguments as above to create a contradiction. By symmetry, $\bar{\Gamma}_\ell(z) > \bar{\Gamma}_{\ell'}(z)$ for some $z$ also creates a contradiction. $\square$

**Lemma 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 $\ell$,*

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

*Then, if $z \in$ supp $\Gamma_\ell$, the relative measure of $z$-jobs in $\ell$ is*

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

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

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

*Proof.* Take any $z \in \text{supp } \Gamma$. If there exists a city $\ell$ so that $\pi_\ell(z) > \pi_{\ell'}(z)$ for all other $\ell'$, then all $z$ employers locate in $\ell$. The feasibility condition (4) then requires $v_\ell \mathrm{d}\Gamma_\ell(z) = M\mathrm{d}\Gamma(z)$. Otherwise, let $\mathcal{L}$ be the set of city for which $\pi_\ell(z) = \pi(z)$ for $\ell \in \mathcal{L}$. Unless $z \in \{\underline{z}_\ell, \bar{z}_\ell\}$ – a measure zero event, then this condition must also hold for a neighborhood around $z$. Hence, $n_\ell(z) = \pi'_\ell(z) = \pi'_{\ell'}(z) = n_{\ell'}(z)$ for any $(\ell,\ell') \in \mathcal{L}$. Differentiating $n_\ell(z) = n_{\ell'}(z)$ with respect to $z$ implies

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

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

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

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

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

**Lemma 5** (Overlapping supports)**.**
*Take two cities $\ell$ and $\ell'$ 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 3 implies $\bar{z}_{\ell'} \le \underline{z}_\ell$. Furthermore, $\pi_{\ell'}(\underline{z}_\ell) \le \pi_\ell(\underline{z}_\ell)$ and $\pi_{\ell'}(z) > \pi_\ell(z)$ for $z \in B^-(\bar{z}_{\ell'})$. Hence, $\pi_\ell$ and $\pi_{\ell'}$ must cross at least once in $[\bar{z}_{\ell'}, \underline{z}_\ell]$ and $\pi_\ell$ crosses $\pi_{\ell'}$ from below; that is, there exists a $z^\star \in [\bar{z}_{\ell'}, \underline{z}_\ell]$ so that $\pi_\ell(z^\star) = \pi_{\ell'}(z^\star)$ and $\pi'_\ell(z) > \pi'_{\ell'}(z)$ for $z \in B^-(z^\star)$. By continuity, we therefore have $\pi'_\ell(z^\star) \ge \pi'_{\ell'}(z^\star)$. Furthermore, it must be that $\Gamma_{\ell'}(z^\star) = 1$. Combining these elements and using the Envelope theorem, this inequality writes

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

where the first and last equality follows from $n_\ell$ 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 $\theta_{\ell'} \ge (1+k)^2\theta_\ell$. 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 2 then implies supp $\Gamma_\ell \cap$ supp $\Gamma_{\ell'} = [\underline{z}_\ell, \bar{z}_{\ell'}]$ with $\underline{z}_\ell < \bar{z}_{\ell'}$. Employers in $[\underline{z}_\ell, \bar{z}_{\ell'}]$ must be indifferent between the two cities, $\pi_\ell(z) = \pi_{\ell'}(z)$ for all $z \in [\underline{z}_\ell, \bar{z}_{\ell'}]$, and therefore $\pi'_\ell(z) = \pi'_{\ell'}(z)$. By continuity, this must hold at $\underline{z}_\ell$, or using the Envelope theorem,

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\frac{1}{\theta_\ell}\left(\frac{1+k}{1+k\bar{\Gamma}_\ell(\underline{z}_{\ell'})^2}\right) = \frac{1}{\theta_{\ell'}}\left(\frac{1}{1+k}\right). \tag{23}$$

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

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

Combined with equation (23), these two equations returns $\underline{z}_{\ell'}$.

**end**

**end**

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

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

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

**end**

**Algorithm 1:** Boundaries solver

## *B.6 Proof of Proposition 3*

The proof of Proposition 3 is broken up in several lemmas. Lemma 6 proves that, in equilibrium, cities must necessarily have different market tightness. Lemma 7 then derives the expression for the local prices that sustain the spatial job allocation. Lemma 8 then orders cities based on their fundamentals, and Lemma **??** concludes by showing the existence of an equilibrium.

**Lemma 6** (No symmetric equilibria)**.**
*In equilibrium, no two cities can have the same labor market tightness.*

*Proof.* Take two cities, $\ell$ and $\ell'$ with $e_\ell > e_{\ell'}$ (wlog). Suppose for the sake of contradiction that $\theta_\ell = \theta_{\ell'}$. Then, it must be that $v_{\ell'}/v_\ell = e_{\ell'}/e_\ell$, or $v_{\ell'} < v_\ell$. It follows that $r_{\ell'} < r_\ell$. At the same time, Lemma 3 implies that $\Gamma_{\ell'} = \Gamma_\ell$, and therefore $\underline{z}_{\ell'} = \underline{z}_\ell = \underline{z}$. Furthermore, from equation (5), $\underline{w}_{\ell'} = \underline{w}_\ell \equiv \underline{w}$. Finally, employers must be indifferent between both locations. In particular, $\pi_{\ell'}(\underline{z}_{\ell'}) = (\underline{z} - \underline{w})n(\underline{z}) - r_{\ell'} = (\underline{z} - \underline{w})n(\underline{z}) - r_\ell = \pi_\ell(\underline{z}_\ell)$ must hold. However, this requires $r_{\ell'} = r_\ell$, a contradiction. □

**Lemma 7** (Local prices)**.**
*Fix* $\boldsymbol{\theta}$ *and re-arrange cities so that* $\theta_\ell$ *is decreasing in* $\ell$. *Then, the spatial job allocation is sustained by a vector of local prices that satisfies the difference equation*

$$\kappa_{\ell+1} - \kappa_\ell = \underline{z}_{\ell+1} n_{\ell+1}(\underline{z}_{\ell+1}) - \underline{z}_\ell n_\ell(\underline{z}_\ell) - \int_{\underline{z}_\ell}^{\underline{z}_{\ell+1}} n_\ell(\zeta)\mathrm{d}\zeta, \tag{24}$$

*for* $\kappa_1$ *given. Furthermore, for two cities* $\ell$ *and* $\ell'$, $\theta_{\ell'} < \theta_\ell \iff \kappa_\ell > \kappa_{\ell'}$.

*Proof.* For any city $\ell$, take the city $\ell'$ such that $\theta_{\ell'} > \theta_\ell$ and there is no other third city $l$ so that $\theta_{\ell'} > \theta_l > \theta_\ell$. If $z \in \operatorname{supp} \Gamma_\ell$ and $z \in \operatorname{supp} \Gamma_{\ell'}$, then it must be that $\pi_\ell(z) = \pi_{\ell'}(z)$ by profit maximization (11). By definition, $\underline{z}_\ell \in \operatorname{supp} \Gamma_\ell$. Hence, if $\underline{z}_\ell \in \operatorname{supp} \Gamma_{\ell'}$, then $\pi_\ell(\underline{z}_\ell) = \pi_{\ell'}(\underline{z}_\ell)$. For the sake of contradiction, suppose $\underline{z}_\ell \notin \operatorname{supp} \Gamma_{\ell'}$; that is, $\bar{z}_{\ell'} < \underline{z}_\ell$ from Lemma 2. Feasibility (4) then requires that there exists at least one city $l$ for which $\pi_{\ell'}(z) < \pi_l(z)$ for $z \in B^+(\bar{z}_{\ell'})$, and thus $1 = \Gamma_{\ell'}(z) > \Gamma_l(z) > \Gamma_\ell(z) = 0$. However, from Lemma 3, this is possible in equilibrium only if city $l$ satisfies $\theta_{\ell'} > \theta_l > \theta_\ell$, a contradiction. We therefore conclude that $\pi_\ell(\underline{z}_\ell) = \pi_{\ell'}(\underline{z}_\ell)$.

To prove the second statement, I start by showing that $\theta_\ell < \theta_{\ell'}$ implies $\kappa_\ell > \kappa_{\ell'}$. For that, take two cities $\ell$ and $\ell'$ that are adjacent in the $\boldsymbol{\theta}$-space with $\theta_\ell < \theta_{\ell'}$. Since $n_\ell$ is a strictly increasing function, equation (24) implies

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

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

**Lemma 8** (City ordering)**.**
*Suppose that* $\lambda^e \approx \lambda^u$ *and* $c \gg b$. *Then, in equilibrium,* $e_\ell > e_{\ell'} \iff v_\ell > v_{\ell'} \iff \theta_\ell < \theta_{\ell'}$.

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

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

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