The Geography of Unemployment*

Adrien Bilal
Princeton University

Job Market Paper

January 7, 2020

(Click here for latest version)

Abstract

Unemployment rates differ widely across local labor markets. I offer new empirical evidence that high local unemployment emerges primarily because of elevated local job losing rates, even for observationally identical workers. I then propose a theory in which spatial differences in job loss arise endogenously. Highly productive employers prefer to fill vacancies rapidly, as waiting implies foregoing high profits. Therefore, they sort into high wage locations with few vacancies per job seeker while less productive employers sort into tight labor markets with low wages. Jobs at more productive employers are endogenously more stable, and spatial gaps in job losing rates arise. In contrast, the equilibrium response of reservation wages results in flatter job finding rates across locations. Due to labor market frictions, productive employers over-value locating close to each other. Thus, the optimal policy incentivizes productive employers to relocate to areas with high job losing rates, providing a rationale for commonly used place-based policies. After structurally estimating the model on French administrative data, I show that it accounts for over 90% of the cross-sectional dispersion in unemployment rates, as well as for the respective contributions of job losing and job finding rates. Employers' inefficient location choices amplify spatial unemployment differentials five-fold. Finally, I show that both real-world and optimal place-based policies yield sizable welfare gains at the local and aggregate level.

*Email: abilal@princeton.edu; website: `https://sites.google.com/site/adrienbilal`. I thank my advisors Gregor Jarosch, Richard Rogerson, Esteban Rossi-Hansberg and Gianluca Violante for invaluable guidance and support. This paper has benefited from discussions with Mark Aguiar, Fabian Eckert, Nik Engbom, Cécile Gaubert, Gene Grossman, Adrien Matray, Steve Redding, Chris Tonetti and Owen Zidar. I gratefully acknowledge financial support from the Simpson Center for Macroeconomics, the International Economics Section, and the Industrial Relations Section at Princeton University. Specials thanks to Xavier Ragot and the Observatoire Français des Conjonctures Économiques at Sciences Po Paris for offering physical access to the data.

Introduction

Unemployment rates vary enormously across local labor markets. In 2017 in Versailles, an affluent French city close to Paris, only five out of a hundred workers were unemployed. In southern Marseille, that ratio exceeded twelve out of a hundred. Such differences can be found in most developed countries, including the United States.[1] Despite their magnitude, these spatial gaps persist over decades. Being exposed to unemployment in a strong labor market is undoubtedly harmful, but spending a lifetime in a distressed labor market can have dire consequences for workers. While local governments devote billions of dollars every year to attract jobs, there is but scant guidance as to the determinants of spatial unemployment differentials. Why is the unemployment rate persistently high in some places, while it remains low in others? What are the welfare implications of this spatial dispersion for workers? Can place-based policies improve the prospects of local residents as well as labor market conditions in the aggregate economy?

This paper proposes answers to these questions with four contributions. First, I offer new empirical evidence showing that spatial unemployment differentials result from spatial gaps in the rate of job loss, tied to employers rather than to workers. Second, I propose a theory of the location choice of employers with labor market frictions that accounts for spatial differences in job stability. Third, I structurally estimate the framework on French administrative data. Fourth, I quantify the local and aggregate welfare gains from place-based policies in general equilibrium. This paper thus consists of four parts, one for each contribution, which I now describe in more detail.

In the first part of the paper, I examine how local labor market flows differ between locations. I use French administrative matched employer-employee data to assess whether differences in unemployment rates across commuting zones reflect differences between inflow (job losing) versus outflow (job finding) rates. Strikingly, differences in job losing rates emerge as the primary source of spatial unemployment differentials, accounting for 86% of the variation. In contrast, job finding rates are close to constant across locations. Controlling for industry and worker composition does not affect these results, which also hold in publicly available data in the United States. The dominant role of the job losing rate indicates that locations have high unemployment because workers repeatedly lose their job there, not because finding a job is particularly hard. This result contrasts with aggregate unemployment fluctuations, as well as with existing models of spatial unemployment, that have focused on the job finding rate.[2] The limited role of worker composition suggests that spatial gaps in the rate of job loss arise because of systematic differences in job stability between employers.

In the second part of the paper, I propose an analytical theory to account for spatial gaps in job losing rates. Workers choose freely where to live and work, and employers choose where to open jobs.[3] They meet in frictional local labor markets, with housing in limited supply. Jobs differ in their initial productivity. Job productivity subsequently fluctuates due to idiosyncratic shocks, leading to endogenous job loss. As a result, initially more productive jobs are more stable. The distinct interaction of the location choice of employers and labor market frictions gives rise to labor market pooling complementarities that lie at the heart of the model's implications. Employers trade off higher wages against high vacancy

[1] In 2017, the unemployment rate was 5.4% in Boston, Massachusetts. It was 13% in Flint, Michigan. See the OECD (2005) report for more OECD countries.

[2] Changes in the job finding rate have been found to be the dominant force in aggregate unemployment fluctuations along the business cycle. See Shimer (2005), Hall (2005), Fujita and Ramey (2009) and Krusell et al. (2017).

[3] As is common in the search literature, there is no difference between employers, firms and jobs in my framework.

contact rates across locations. More productive employers make higher profits when operating. Thus, they forego relatively more than unproductive employers while waiting for a worker: productive employers have a higher opportunity cost of time. Hence, they prefer locating in slack labor markets where they fill vacancies rapidly but wages are high. In contrast, unproductive employers are priced out where wages are high but forego lower profits where the vacancy contact rate is low. Thus, they self-select into low wage areas where the labor market is tight and vacancies are filled slowly. As a result, sorting emerges in spatial equilibrium.

Labor market flows reflect the spatial sorting of employers. The job losing rate is high where employers are unproductive. Crucially, job finding rates depend on two components. The first is the rate at which workers contact employers, which rises when there are more vacancies relative to unemployed workers. But not all contacts result in a viable job. The second component of job finding rates is the conditional probability that a contact is successful. I argue that both components offset each other. In locations with many unemployed workers, there are also more employers since labor is cheaper. On net, the labor market is tighter, workers contact many employers, and employers contact few workers. Only unproductive employers locate in high-unemployment locations. There, expected wages are close to reservation wages. Thus, the probability that a contact results in a viable job is low, offsetting the higher worker contact rate. When both forces closely balance, the job finding rate is flat across locations.

I then show that the spatial equilibrium features misallocation. Because of labor market pooling complementarities, productive employers over-value the benefits from locating close to each other. Labor market frictions prevent productive employers from attracting as many workers as would be socially optimal, should they enter in a location with low productivity employers. Hence, productive employers find it privately optimal to concentrate too much in top locations with a larger pool of workers relative to vacancies. A utilitarian planner thus chooses an optimal policy that incentivizes productive employers to relocate towards high unemployment areas. A profit subsidy that rises with the local job losing rate implements the optimal allocation, providing a rationale for commonly used place-based policies that target high unemployment locations.

The third part of the paper develops and structurally estimates a quantitative version of the framework. The main additions are local productivity and amenity differences; migration frictions that hinder workers' ability to move; and heterogeneity in human capital, which grows while workers accumulate experience at work, and depreciates while they are unemployed. To isolate the importance of labor market pooling externalities, I abstract from other sources of agglomeration. Despite its richness, the quantitative model produces estimating equations that allow for transparent identification leveraging the many dimensions of the French administrative data. In particular, the estimation directly targets neither the cross-sectional variance of local unemployment rates nor its breakdown into job losing and job finding rates.

The estimated model accounts for the primary margins of spatial unemployment differentials. It generates over 90% of the cross-sectional variance of local unemployment rates in the data. It also closely replicates the respective contributions of job losing and job finding rates. 86% stems from the job losing rate in the data, against 85% in the model. Pooling externalities are crucial to rationalize the location choice of employers, and hence job losing rate differences. Shutting down pooling externalities reduces the spatial variation in unemployment rates by over 80%. Over-identification checks support these results by highlighting that the model can match a number of non-targeted moments.

I then propose a set of validation exercises to support the core structure of the theory. First, I verify the key link between labor productivity differences and job losing rates. Leveraging firm-level balance sheet data for the near universe of French businesses, I show that labor productivity is 37% lower in locations with a one percentage point higher job losing rate. This gap rises by an additional 43% for establishments less than two years old, that have a higher proportion of new jobs. In contrast, labor productivity growth is flat across locations. Second, I use survey data to unpack how job finding rates reflect worker contact rates and contact-to-job probabilities. Consistent with the view that more productive employers locate where there are few vacancies per job seeker, I find that worker contact rates rise with the unemployment rate. In contrast, the proportion of contacts resulting in a viable job falls with the unemployment rate.

The fourth and last part of the paper conducts two policy counterfactuals. The first exercise explores the local and aggregate effects of economy-wide place-based policies. I contrast the optimal policy with the French Enterprise Zones (EZ) program – which consisted in heavy subsidies for businesses opening jobs in high unemployment areas.[4] In both cases, the policy is funded at the federal level. Qualitatively, the optimal policy resembles the French EZ policy. Quantitatively, however, it is much larger. By massively relocating productive jobs towards high job losing rate areas, the optimal policy cuts the local unemployment rate by over 10 percentage points in the most afflicted locations. It also achieves over 20% welfare gains for their residents. While half of those gains follow directly from the unemployment rate reduction, the other half are indirect amplification effects due to human capital improvements. Although the optimal policy primarily redistributes jobs across locations, it ameliorates aggregate outcomes: the aggregate unemployment rate falls by 0.5 percentage points and utilitarian welfare rises by 5.1%. In contrast, the model indicates that the French EZ program reduced unemployment only in targeted areas by 1 percentage point, leading to 3% local welfare gains but little aggregate effects.

The second policy counterfactual concludes the paper by assessing whether locally funded, discretionary place-based policies can achieve local outcomes similar to federal programs. Specifically, it examines how attracting a "Million Dollar Plant" (MDP) affects a location. In the model, a MDP calibrated to the estimates in Greenstone et al. (2010) lowers commuting-zone level unemployment by 0.2 percentage points, largely driven by improved job stability. Gross welfare of residents rises by 1.15%. In practice, attracting a MDP comes at a cost. To discipline that trade-off, I also compute the optimal MDP subsidy in the model financed by non-distortionary local taxes. The results indicate that locally funded policies can achieve as much as 20% net welfare gains in high unemployment locations.

Literature. This paper adds to four strands of literature. First and most closely related is the body of work that examines persistent spatial unemployment differentials.[5] Kline and Moretti (2013), Şahin et al. (2014) and Marinescu and Rathelot (2018) study spatial variants of the Diamond (1982), Mortensen (1982), and Pissarides (1985) model. These papers focus on the role of the job finding rate and abstract from job losing rate differentials. Kline and Moretti (2013) find that subsidies to high unemployment areas reduce welfare. In contrast, I stress that a different theory is needed after documenting that job losing rate gaps are the key empirical determinant of spatial unemployment differentials. As a result, I find

[4] The *"Zone Franches Urbaines"*.

[5] Blanchard and Katz (1992)'s seminal work found little evidence of state-level unemployment persistence between 1975 and 1985. In contrast, Kline and Moretti (2013) and Amior and Manning (2018) show that unemployment and labor force participation differentials between US commuting zones are highly persistent after 1980. Amior and Manning (2018) specifically focus on long-run adjustments to persistent labor demand shocks and the slow response of migration. The empirical analysis and model therein abstract from unemployment and worker flows.

that subsidies to high unemployment areas raise welfare, reconciling theory with real-world place-based policies.

Second, this paper adds to the large literature that studies the location decisions of agents. A first subset thereof has focused on workers' location decisions based on income prospects (Roback, 1982, Kennan and Walker, 2011, Desmet and Rossi-Hansberg, 2013, Bilal and Rossi-Hansberg, 2018).[6] A second set of papers studies firms' location choices (Combes et al., 2012, Gaubert, 2018). Both literatures abstract from unemployment, while I show that inclusion thereof has distinct policy implications. A final strand of the literature proposes theoretical assignment models to study the sorting between workers and employers (Sattinger, 1993, Shimer and Smith, 2000, Davis and Dingel, 2014, Eeckhout and Kircher, 2018), which the present paper builds on.

Third, this paper adds to the body of work that studies the efficiency properties of search models (Hosios, 1990, Mortensen and Pissarides, 1994).[7] The labor market pooling externality can be seen as a spatial analogue of Acemoglu (2001). He shows that when high and low productivity jobs coexist in the labor market, too many low productivity jobs open in the aggregate labor market because they fail to internalize that they divert workers away from productive jobs. In my model, a similar force pushes less productive jobs to inefficiently locate in places that are too productive for them. In contemporaneous work, Brancaccio et al. (2019) emphasize a similar mechanism in the context of transport markets.

Finally, this paper is closely tied to the large literature on agglomeration and congestion externalities. Going back to at least Marshall (1920), externalities operating at the local level have formed the basis for place-based policies. Recent empirical analyses of the latter have found mixed employment effects across several countries (Glaeser and Gottlieb, 2008, Hanson, 2009, Neumark and Simpson, 2014, Mayer et al., 2015, Slattery and Zidar, 2019). Several recent papers propose spatial models with either or both worker and firm mobility to analyze place-based policies but all abstract from unemployment (Ossa, 2017, Fajgelbaum et al., 2018, Fajgelbaum and Gaubert, 2018, Slattery, 2019). In many cases, agglomeration economies call for subsidies to high income locations, which contrasts with many real-world place-based policies. While the overall net policy should account for that largest possible set of agglomeration and congestion externalities, I highlight and quantify a particular mechanism whereby labor market pooling externalities favor subsidies to low income locations. The idea that redistributing a given set of jobs across heterogeneous local labor markets can improve aggregate outcomes even in the absence of technological spillovers goes back at least to Bartik (1991), and has been recently revived by Austin et al. (2018). This paper proposes a theory of frictional local labor markets that makes this idea precise.

The remainder of the paper is structured as follows. Section 1 presents the data and empirical analysis. Section 2 builds a simple model of spatial unemployment differentials with endogenous job loss and characterizes the spatial equilibrium. Section 3 lays out the quantitative extensions and the estimation strategy. Section 4 describes the estimation results and proposes additional validation exercises. Section 5 presents the policy counterfactuals. The last section concludes. Proofs and additional details can be found in the Appendix.

[6]For structural change over time, see Diamond (2016), Giannone (2017) Caliendo et al. (2017), Glaeser et al. (2018), and Couture et al. (2019).

[7]See Jarosch (2015) and Mangin and Julien (2018) for recent contributions.

1 Descriptive evidence

This section first describes the data. Next, I highlight that spatial unemployment gaps are large and persistent. Then, I show that spatial unemployment gaps are primarily driven by spatial differences in job losing rates, independently from local worker and industry composition. My main analysis focuses on France where I can exploit the richness of administrative data, but I also confirm the main findings in the United States.

Data

Worker flows in and out of unemployment are central components of labor market studies. Aggregate time series exercises typically break down the contribution of job losing and job finding rates in accounting for the unemployment rate. While they are jointly determined equilibrium variables, separating their contributions is a useful diagnostic device that informs the underlying economic mechanisms.

Adapting this approach to a geographic setting is challenging. On the one hand, large repeated cross-sections like the Census or the American Community Survey (ACS) are ill-suited for the measurement of worker flows. On the other hand, surveys with a short panel dimension such as the Current Population Survey (CPS) typically have a much smaller cross-section. This limitation leads to measurement error concerns, particularly for the outflow from unemployment, and prevents any compositional split.[8] In addition, panel surveys often stop tracking movers who change location.

To circumvent these difficulties, I turn to administrative matched employer-employee data from France. I use a combination of the DADS and of the French Labor Force Suvey (LFS) between 1997 and 2007.[9] The DADS have two advantages. First, they are a representative dataset covering almost one million individuals in any cross-section. Second, it is a panel that consists of the entire work history of individuals, with rich demographic, geographic and firm-level information. Thus, the DADS are well-suited to study the employment versus non-employment status of individuals across cities. The sample size allows to break down the analysis by city, industry, and finely disaggregated worker groups to control for composition.

One drawback of the DADS is that it only allows to discriminate between employment and non-employment.[10] To address this limitation, I first restrict my sample to males between 30 and 52 years old. This group has a high and stable labor force participation rate, which allows to abstract from life-cyle changes therein. Second, I complement the DADS with the LFS. I compute conditional transition probabilities between employment, non-employment and unemployment in the LFS, by broad city and worker group. I then use those conditional transition probabilities from the LFS to probabilistically discriminate between non-employment and unemployment in the DADS.[11] In practice, this imputation has a limited impact on the results because non-employment and unemployment are closely related in

[8]Once broken down by city or commuting zone, the CPS data has about one hundred individuals by city, and thus only about five unemployed individuals in any cross-section.

[9]DADS: *"Déclaration Aministrative de Données Sociales."* The LFS is the *"Enquête Emploi."*

[10]Consistent with the International Labour Office's definition, I define an employed individual as one who has a job. A non-employed individual is one who is not working for a wage. An unemployed individual is one who is not working but is actively looking for a job and available to start work within two weeks.

[11]This imputation exercise resembles Blundell et al. (2008) who use the Panel Study of Income Dynamics to complement consumption categories in the Consumption Expenditure Survey. For instance, if an individual goes through an employment to non-employment transition in the DADS, I define her employment status after the transition (unemployment or no-employment) based on the LFS transition probabilities.

the cross-section of locations. Therefore, the results are very similar if I use only the unemployment information in the LFS. I aggregate the resulting sample at the quarterly frequency. Table 5 in Appendix A.1 shows that aggregate statistics in this sample and in the LFS alone are similar.

I complement these datasets with several other data sources. To compute city-level and establishment-level variables, I use a repeated cross-section version of the DADS that covers the universe of French workers. For some over-identifying exercises in Section 4.4, I use firm-level balance sheet data covering the near universe of French business for the same period. I also use a single cross-section of housing prices from an online realtor, MeilleursAgents.com.

I define a location as a commuting zone as defined by the French statistical institute INSEE.[12] A commuting zone is an area where most of the residents work at jobs located in that same area. There are 328 commuting zones, and they partition the French territory. This definition is most natural as a spatial notion of a local labor market. In what follows, location, commuting zone and city are used interchangeably. I construct a measure of skill from occupation and age data because the main DADS panel dataset does not have education data. Skill is defined as the average age and occupation wage premium for a worker, derived from a Mincer regression. Appendix A.1 provides more details.

For the United States, I use the CPS. I define a location as a metropolitain statistical area, and use a similar definition of skill as in France.[13] I focus on white males between 30 and 52 years old that are household heads, and use the CPS's definition of unemployment.

1.1 Dispersion and persistence of spatial unemployment differentials

I start by showing that local unemployment rates are widely dispersed and highly persistent across locations in France. Figure 1 (a) maps commuting-zone level unemployment rates in mainland France. Darker shades of blue encode higher unemployment rates. Figure 1 (a) highlights that commuting zones with unemployment rates above 10% or below 5% can be found throughout the country. The cross-sectional standard deviation is 2.3 percentage points, amost twice as much as the the time-series standard deviation of the aggregate unemployment rate (1.3 percentage points).

To assess the persistence of spatial unemployment differentials, I then split the sample in two subperiods, 1997-2001 and 2002-2007. Figure 1 (b) plots the local unemployment rate in the second subperiod against the unemployment rate in the first subperiod for every city. The blue circles represent a city, with the size proportional to population. Figure 1 (b) reveals that local unemployment rates are highly persistent, as they line up closely around the orange 45 degree line. The 5-year autocorrelation is 0.86.[14]

Figure 1 confirms earlier findings from Kline and Moretti (2013) and Amior and Manning (2018) for the United States. I now turn to the main empirical contribution of this paper: unpacking how worker flows in and out of unemployment differ between commuting zones.

1.2 Worker flows in and out of unemployment

Inflows from local employment, from non-participation and in-migration from other locations all contribute to local unemployment. Similarly, outflows into local employment, into non-participation and

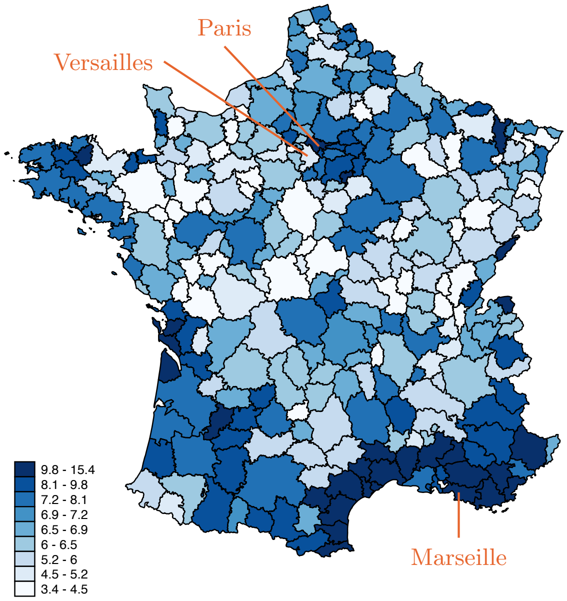
[12] *"Institut National de la Statistique et des Etudes Economiques"*.

[13] I also check that using education to define skill in the CPS leaves the results unchanged.

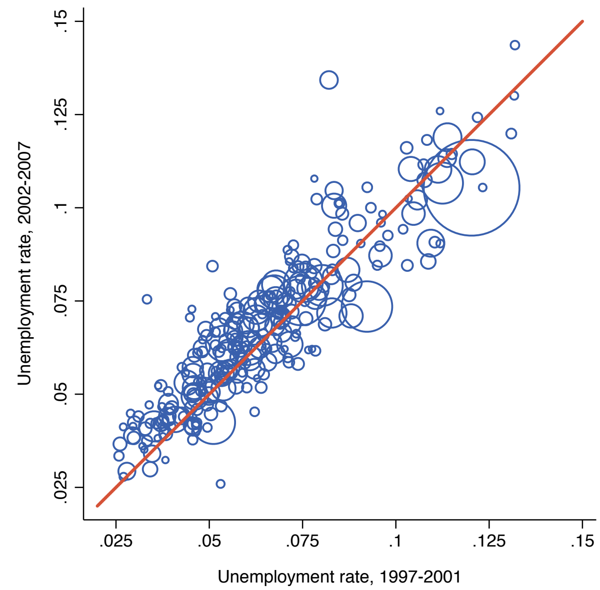
[14] In Appendix A, I show that controlling for economy-wide industry business cycles increases local persistence, with a conditional autocorrelation of 0.99.

Figure 1: Unemployment rates in France, by commuting zone and over time

(a) Local unemployment rates, 1997-2007 averages



(b) Persistence of local unemployment rates



out-migration reduce the number of unemployed workers. In what follows, I borrow standard terminology from the literature and call the the rate at which employed workers flow into unemployment the job losing rate. Similarly, I call the rate at which unemployed workers flow into employment the job finding rate.

To guide the analysis, start with a simple two-state accounting model. Suppose that employed workers in city c face a constant job losing rate s_c per unit of time (i.e. separation rate to unemployment), and that unemployed workers face a constant job finding rate f_c per unit of time. Abstract from movements in and out of the labor force and migration for now. In steady state, the local equilibrium unemployment rate u_c satisfies

$$\log \frac{u_c}{1-u_c} = \log s_c - \log f_c \; . \tag{1}$$

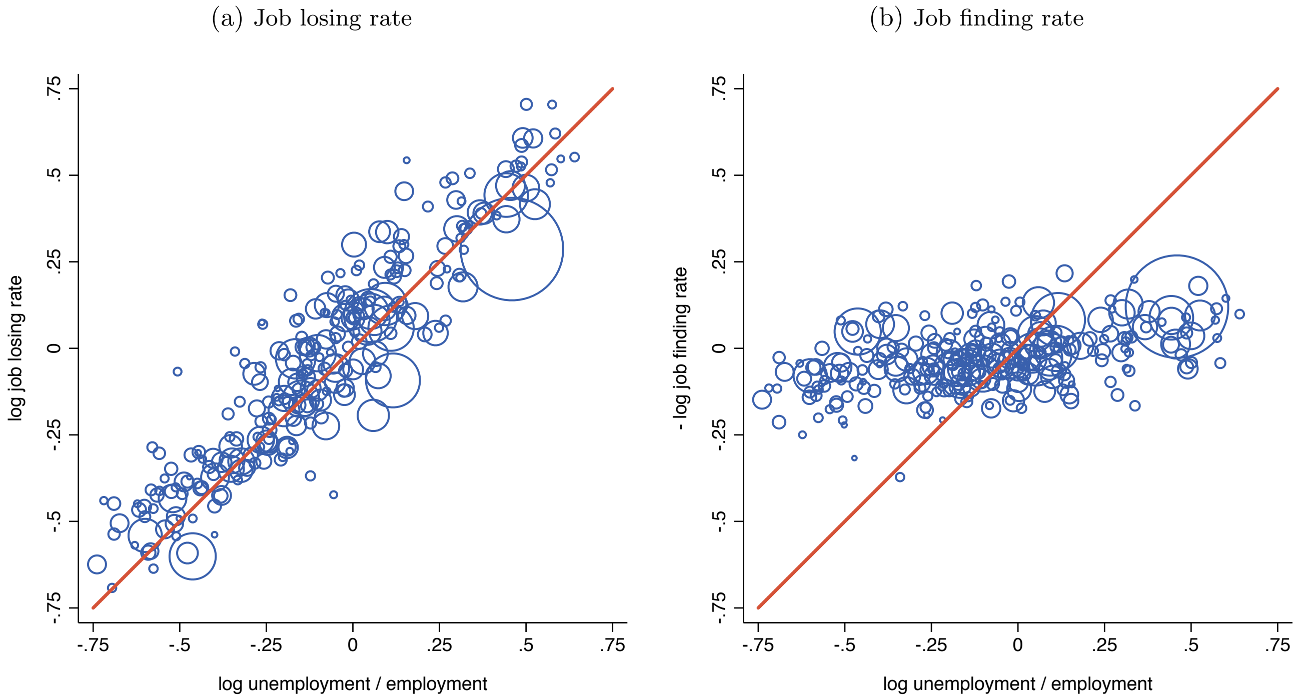
Both s_c and f_c can be directly measured in the data using transition probabilities between employment and unemployment.[15]

To examine the respective contributions of the job losing and job finding rates to local unemployment, panel (a) in Figure 2 plots the logarithm of the job losing rate s_c against the logarithm of the unemployment-to-employment ratio $\frac{u_c}{1-u_c}$ for France. The data align closely along the 45 degree line in orange, indicating that local job losing rates are the primary determinants of spatial unemployment differentials. Panel (b) in Figure 2 plots the logarithm of the job finding rate against the logarithm of the unemployment-to-employment ratio. In contrast to the job losing rate, the job finding rate appears almost flat across locations.[16]

[15]In principle, these quarterly transition *probabilities* must be corrected for time aggregation in order to obtain instantaneous transition *rates*. In practice, Figure 17 in Appendix A.3 shows that the time-aggregation correction leaves the variance share of the job losing share virtually unchanged in France and the United States.

[16]Similarly, Figure 15 in Appendix A.3 shows that the job-to-job mobility rate is not systematically associated with the local unemployment rate.

Figure 2: Local job losing and finding rates against unemployment-to-employment ratios in France.



In practice, movements in and out of the labor force, migration and local transitional dynamics could introduce a wedge between the left-hand-side and right-hand-side of equation (1). They account for part of the dispersion around the 45 degree line in Figure 2 (a). To assess exactly how much variation stems from the job losing rate, Appendix A.3 extends equation (1) to a three-state model with labor force participation, which can then be used to for an exact variance decomposition. In France, the job losing rate accounts for 86% of the cross-sectional variation in spatial unemployment rate. The job finding rate accounts for almost all of the remaining 14%. Figure 16 in Appendix A.3 replicates the same exercise in the United States. There, the job losing rate contributes 73%.

1.3 Worker and industry composition

In principle, differences in the local industry mix and worker skill mix may account for some or all of the differences in unemployment and job losing rates in Figure 2. To unpack the contribution of worker and industry composition, I estimate econometric models of the following form:

$$Y_{i,t} = \alpha_{C(i,t)} + \beta_{J(i,t)} + \gamma_{S(i)} + e_{i,t} \ , \tag{2}$$

where C denotes a city, J denotes a 3-digit industry, S denotes a skill group, i denotes a worker identifier, and t is a quarter. Y is an outcome of interest, for instance the local unemployment, job losing or finding rates, or some functional transformation thereof. α_C is a city effect, β_J an industry effect, and γ_S a skill effect. $e_{i,t}$ is a conditionally mean zero residual. I first estimate linear probability models with 232 industry fixed effects and 300 skill fixed effects. Then, I replicate the exercise from Figure 2 with the estimated city fixed effects $\hat{\alpha}_c$.[17]

[17]For variable Y, I impose that the city fixed effects have a mean equal to the unconditional mean of Y.

Figure 3: Local job losing and finding rates against unemployment-to-employment ratios in France. City fixed effects net of local industry and worker composition.

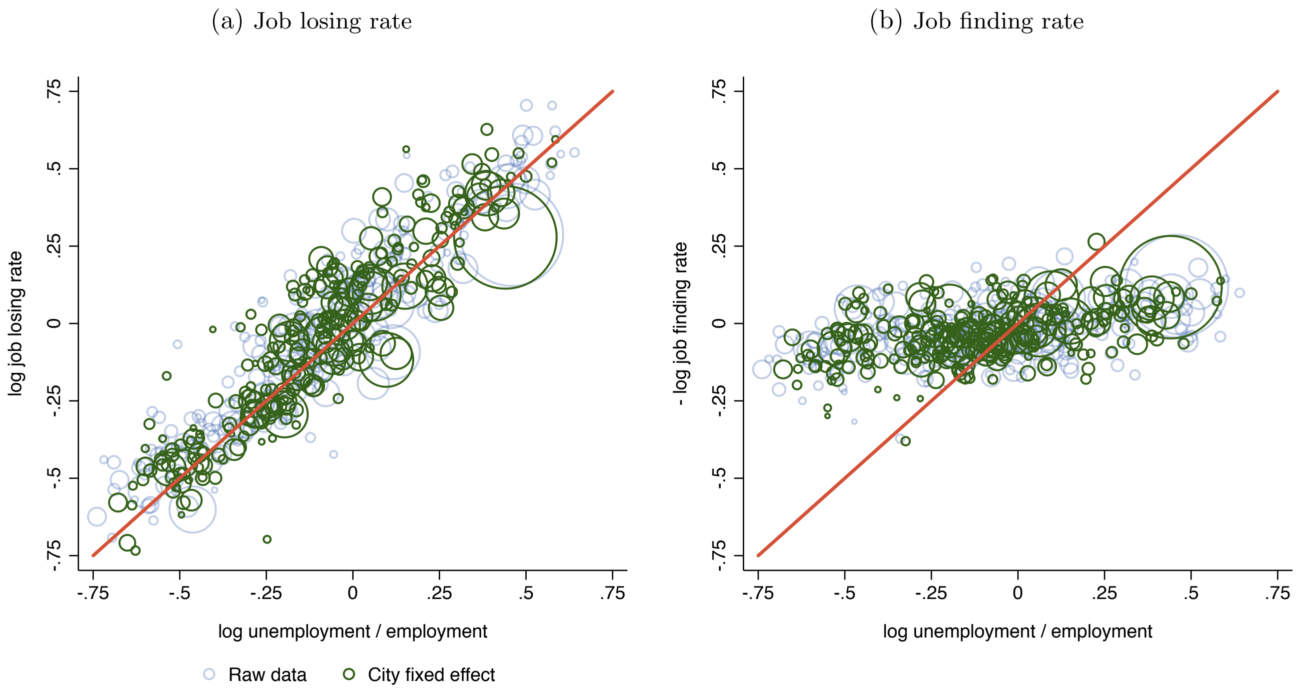


Figure 3 reveals that industry and worker composition do not contribute significantly to spatial unemployment differentials. In addition, Figure 3 shows that, even after controlling for local composition, the job losing rate remains the dominant source of spatial unemployment gaps. In fact, its variance share increases to 91% when using the estimated city fixed effects.

To make sure that my results are not driven by small sample biases or functional form assumptions, I also estimate probit models as well as correlated random effect models in Figures 19 and 20 in Appendix A.4. In all cases, industry and worker composition do not account for more than 20% of spatial differences in job losing or job finding rates.

One remaining concern is that unobserved worker heterogeneity within the 300 skill groups drives part of the spatial differences in job losing rates. I test for that possibility by also estimating equation (2) with worker effects in the French data.[18] Figure 21 in Appendix A.4 indicates that over 70% of the spatial variation still stems from city effects in that case.

1.4 Conditional correlations

I briefly conclude this section by discussing two additional dimensions of the spatial distribution of worker flows that will inform the quantitative features of the model. First, I describe conditional correlations between local labor market flows and two local observables: commuting zone wages and population density. Table 6 in Appendix A.5 indicates that spatial gaps in job losing rates are negatively correlated with local wages. Job losing rates appear uncorrelated with local population density conditional on wages and compositional controls. Second, Figure 22 in Appendix A.6 shows that spatial gaps in job losing rates are

[18]This procedure imposes stronger data requirements. Mobility conditional on city, industry and worker effects must be random. See Card et al. (2013) for details.

starkest during the first two years of a job, and stabilize thereafter.

Overall, the results in this section indicate that spatial differences in job losing rates are by far the largest contributor to spatial unemployment rate differentials in France and in the United States. In addition, these spatial differences are not explained by the local industry mix or composition of the workforce, suggesting systematic differences on the employer side. These findings are, to the best of my knowledge, new to the literature. They elude existing models of local unemployment that have focused on the job finding rate. In contrast, the job losing rate lies at the heart of the theory I propose below.

2 A model of spatial unemployment differentials

This section develops an analytical theory of spatial unemployment differentials. I build on Kline and Moretti (2013)'s model of frictional unemployment in spatial equilibrium. I add two key ingredients. First, heterogeneous employers decide where to locate. More productive employers prefer filling vacancies faster in locations with slack labor markets where hiring is easy. Hence, spatial sorting emerges in equilibrium. Second, job loss is endogenous and tied to employers. Thus, spatial gaps in job losing rates arise. Distinct normative implications follow, that I present after describing the structure of the model in more detail.

2.1 Setup

Time is continuous. There is a single final good used as the numeraire and freely traded across locations.

Geography. There is a continuum of ex-ante heterogeneous locations endowed with one unit of housing. Locations differ in productivity ℓ with cumulative distribution function F_ℓ on a connected support $[\underline{\ell}, \overline{\ell}]$, with density F_ℓ'. Thus, a location is characterized by its productivity ℓ rather than its particular name. Local productivity differences ℓ are useful to fix ideas and provide a natural ordering of locations, but are not necessary for the main mechanism.

Workers. There is a unit mass of infinitely-lived homogeneous workers. Their preferences over streams of consumption of the final good c_t and housing services h_t are

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

with $\omega > 0$. Workers consume their income each period. They only search when unemployed.[19] Workers are freely mobile across locations.

Employers and jobs. As is common in the search literature, the productive unit is an employer-worker match. Thus, the notions of firms, employers and jobs are interchangeable in the model.[20] An employer

[19] I do not incoporate is job-to-job mobility, for two reasons. First, job-to-job moves do not directly affect the unemployment rate as they relocate workers from one job to another. Second, I show in Figure 15 in Appendix A.3 that, just like the job finding rate, the job-to-job mobility rate is not systematically correlated with the unemployment rate. Together, these remarks suggest that including job-to-job moves would have small quantitative effects on counterfactuals affecting spatial unemployment differentials.

[20] The model can also be seen as one in which there are large, constant-returns-to-scale firms that open many jobs at cost c_e per job. For models with a well-defined notion of firm size through decreasing returns to scale and search frictions, see Bilal et al. (2019), Schaal (2017) and Elsby and Michaels (2013). I explore the role of the boundary of the firm in the data

pays a fixed cost c_e to open a new job. After paying c_e, the employer draws a job quality (or expected productivity) z that informs its initial productivity draw. The population distribution of quality z is F_z, with connected support $[\underline{z}, \overline{z}]$ and density F'_z. After observing job quality z, employers choose a location ℓ to open their job and search for workers by posting a single vacancy in the local labor market. After they match with a worker, they draw their initial match productivity y_0 from a conditional distribution $G_0(y_0|z)$ that depends on the employer's quality z. Drawing a higher z implies that the job will be more productive on average, in a sense made precise in Assumption 1 below. After observing this initial draw, the matched pair decides to start producing together or not. If not, the worker returns to unemployment, and the job disappears.

An active job with productivity y_t in a location ℓ produces $y_t\ell$: there is a technological complementarity between the local productivity ℓ and the job's idiosyncratic productivity y_t. Over time, the productivity of the job fluctuates according to a geometric Brownian motion

$$d \log y_t = -\delta dt + \sigma dW_t \ , \tag{3}$$

where $\delta > 0$: productivity depreciates on average. This assumption is required for endogenous separations to take place, as well as for a well-defined steady-state distribution to arise.[21] σ is the volatility of shocks. A geometric Brownian motion is the continuous-time analogue of an otherwise standard random walk with drift. Importantly, the productivity process is identical in all locations, so that any spatial differences in job loss must originate from differences between employers. For values to remain finite, I impose that $\rho + \delta > \frac{\sigma^2}{2}$. If the match breaks up, the job disappears.

Local labor markets. Unemployed workers search for jobs only in the location where they live, and employers search for workers only in the location where their job is open. Workers randomly meet vacancies in a single labor market in each location. Meetings occur according to a Cobb-Douglas matching function $\mathcal{M}(\mathcal{U}(\ell), \mathcal{V}(\ell)) = m\mathcal{U}(\ell)^\alpha \mathcal{V}(\ell)^{1-\alpha}$, where $\mathcal{U}(\ell)$ is the local number of unemployed workers, and $\mathcal{V}(\ell)$ is the local number of vacancies in that market.[22]

Local market tightness is $\theta(\ell) = \mathcal{V}(\ell)/\mathcal{U}(\ell)$. Workers' local contact rate is then $f(\theta(\ell)) = m\theta(\ell)^{1-\alpha}$ while the vacancy contact rate for employers is $q(\theta(\ell)) = m\theta(\ell)^{-\alpha}$. The contact rate may differ from the realized finding rate if some contacts do not result in a new job. Denote by $f_R(\ell)$ and $q_R(\ell)$ the realized rates.

Flow value of unemployment. Unemployed workers in location ℓ consume $b\ell$. This specification captures the idea that unemployment benefits are a constant replacement rate of past wages, because wages will scale with local productivity ℓ. It also helps with analytical tractability.[23]

in Section 4.4.

[21] $\delta < 0$ reflects the difference between parameters governing productivity growth at new jobs relative to incumbent jobs in endogenous growth models such as Engbom (2018).

[22] The main results extend to the case of a non-Cobb-Douglas constant returns to scale matching function if its elasticities are bounded away from 0 and 1.

[23] The specification can also be seen as home production or self-employment with the same production function as firms, but with an efficiency b. Because the model features aggregate constant returns to scale in production, defining unemployment benefits to be directly a constant replacement rate of past wages leads to multiplicity.

Wage determination. Workers and employers bargain à la Rubinstein (1982), with worker bargaining power β. For simplicity, I assume that renegotiation occurs each instant.[24]

Ownership. A representative mutual fund owns housing and claims to employers' profits. The mutual fund rents land to workers at equilibrium rents $r(\ell)$. It also collects profits from employers. For simplicity, I assume in this section that absentee owners receive the profits from housing rents and firms.[25]

2.2 Value functions and wage determination

In what follows, the economy is in steady-state.

Unemployment and employment. Let U be the value of unemployment. Because unemployed workers are freely mobile, their value must be equalized across all locations that they populate. The Inada property of the matching function ensures that any populated location must have some unemployed workers. Thus, the value of unemployment is equal across all populated locations.[26]

To keep the exposition simple in the main text, I consider wage functions $w^*(y,\ell)$ that only depend on productivity y and the location ℓ. As shown in Appendix B.1, this restriction is without loss of generality. Let $V^E(y,\ell)$ be the value of employment at wage $w^*(y,\ell)$ in location ℓ. U and V^E satisfy the recursions

$$\begin{aligned} \rho U &= b\ell r(\ell)^{-\omega} + f(\ell)\mathbb{E}_\ell\big[\max\{V^E(y_0,\ell) - U, 0\}\big] & (4) \\ \rho V^E(y,\ell) &= w^*(y,\ell)r(\ell)^{-\omega} + (L_y V^E)(y,\ell)\ . & (5) \end{aligned}$$

The first term on the right-hand-side reflects workers' flow value when unemployed or employed. Because of Cobb-Douglas preferences their housing choice, workers spend a constant share ω of their income on housing. Hence, workers' flow value is income adjusted by local housing prices $r(\ell)$. The second term on the right-hand-side of equation (4) reflects unemployed workers' future employment opportunities. At rate $f(\ell)$, they meet potential employers. The latter then draw the initial productivity y_0. Provided it is sufficiently high, the worker is hired and the matched pair starts producing together. Because initial productivity y_0 is unknown prior to meetings, the value of employment opportunities reflects the expected value from employment conditional on the pool of employers in location ℓ. The second term on the right-hand-side of equation (5) reflects the expected continuation value of employment due to productivity shocks. Given the geometric Brownian motion assumption (3), the functional operator L_y is defined by

$$L_y V^E = \left(\frac{\sigma^2}{2} - \delta\right) y \frac{\partial V^E}{\partial y} + \frac{\sigma^2}{2} y^2 \frac{\partial^2 V^E}{\partial y^2}\ .$$

[24] As in Bilal et al. (2019), wage setting protocols such as renegotiation by mutual consent (Cahuc et al., 2006, Postel-Vinay and Turon, 2010) leave the surplus of matched pairs unchanged and the resulting allocations coincide with those under renegotiation each instant. Wage dynamics differ.

[25] Alternatively, the proceeds from land rents and profits can be rebated to workers as a flat earnings subsidy. In that case the cross-sectional implications are unchanged. To keep the focus on the efficiency properties of the location choice of employers and abstract from distributional considerations between owners and workers, I use the flat earnings subsidy rebate in the quantitative exercises.

[26] Specifically, the Inada property reads $\lim_{\mathcal{U}\downarrow 0} \frac{\partial \mathcal{M}}{\partial \mathcal{U}}(\mathcal{U},\mathcal{V}) = +\infty$.

Employers. The value of a matched employer with productivity y in location ℓ solves

$$\rho J(y,\ell) \quad = \quad y\ell - w^*(y,\ell) + (L_y J)(y,\ell) \ .$$

Employers value flow profits $y\ell - w^*(y,\ell)$ as well as the contribution of future productivity changes.

Joint surplus and wage determination. A common solution method in search models is to focus on the joint surplus from an employed worker and her employer. The wage then drops out, which allows to use the standard solution to Rubinstein (1982)'s split-the-pie game. In the present geographic setting, workers consume local housing with heterogeneous prices across locations. Therefore, workers' marginal valuation of a dollar (or euro) depends on local housing prices. In contrast, employer's marginal valuation of a dollar does not. Therefore, adding up the worker's and the firm's surplus does not have a well-defined interpretation.

Despite this apparent complication, I show in Lemma 4 in Appendix B.3 that the bargaining game still delivers a simple solution for wages, because both parties have differing but constant marginal valuations of a dollar. The only requirement is to put each side's value on a common numeraire scale. This property obtains because wage determination in the alternating offers game depends on each side's valuation a dollar relative to their own outside option, not relative to the other side's valuation.

Lemma 4 greatly simplifies the analysis and extends otherwise standard bargaining results. It allows to restrict attention to a single object, that I dub the adjusted surplus. It is defined as

$$S(y,\ell) = J(y,\ell) + r(\ell)^{\omega} \cdot \big(V^E(y,\ell) - U\big) \tag{6}$$

It is independent from wages. Appendix B.1 shows that it follows a recursion similar to that of employers. Lemma 4 then implies that wages adjust so that workers and employers each receive a constant adjusted share of the adjusted surplus, Namely,

$$r(\ell)^{\omega} \cdot \big(V^E(y,\ell) - U\big) = \beta S(y,\ell) \ , \qquad\qquad J(y,\ell) = (1-\beta) S(y,\ell) \ . \tag{7}$$

In particular, both sides agree when to break up the match when the adjusted surplus drops to zero. In that case, a separation occurs. Existing matches therefore solve a forward-looking optimal stopping problem, which is detailed in Appendix B.3. There, I characterize its solution which is described in the following lemma, useful for future reference.

Lemma 1. *(Adjusted surplus)*
There exists a unique solution adjusted surplus, given by

$$\rho S(y,\ell) \quad = \quad \ell\big(b + v(\ell)\big)\mathcal{S}\left(\frac{y}{\underline{y}(\ell)}\right)$$

for $y \geq \underline{y}(\ell)$*, and* $S(y,\ell) = 0$ *for* $y \leq \underline{y}(\ell)$*, where*

$$\rho\frac{\underline{y}(\ell)}{\underline{y}_0} = b + v(\ell) \ , \qquad v(\ell) = \frac{f(\ell) r(\ell)^{\omega} \mathbb{E}_\ell[\max\{V^E(y_0,\ell) - U, 0\}]}{\ell} \ , \qquad \mathcal{S}(Y) = \frac{\tau Y + Y^{-\tau}}{1+\tau} - 1 \ ,$$

and $\tau, \underline{y}_0$ *are transformation of* ρ, δ, σ *given in Appendix B.3.*

The local endogenous separation cutoff $\underline{y}(\ell)$ increases as the worker's local value of unemployment $b + v(\ell)$ rises, which is itself the equilibrium outcome of local market tightness $\theta(\ell)$ and the local mix of employers. The adjusted surplus S is an increasing function of current productivity y relative to the local endogenous cutoff $\underline{y}(\ell)$. The nonlinearity in the function $\mathcal{S}$ arises because of the option value of separation, which rises as productivity y approaches the cutoff $\underline{y}(\ell)$. Hence, the adjusted surplus S satisfies both the value matching and smooth-pasting conditions at the cutoff: $S(\underline{y}(\ell), \ell) = \frac{\partial S}{\partial y}(\underline{y}(\ell), \ell) = 0$.[27]

It is also useful to define workers' reservation wage $\underline{w}(\ell)$ in each location, in efficiency units of local productivity ℓ. It satisfies

$$\underline{w}(\ell) = \underline{w}_0 \underline{y}(\ell) \ , \tag{8}$$

where $\underline{w}_0 = (1 - \beta)\rho/\underline{y}_0 + \beta$. When the local separation threshold is higher, matches break up at higher productivity levels because workers value more the option to search for a different job in the same local labor market. Thus, the local reservation wage is higher.

Given reservation wages $\underline{w}(\ell)$, the free mobility condition takes a simple form,

$$U = \frac{\ell \underline{w}(\ell)}{\underline{w}_0 \underline{y}_0 r(\ell)^{\omega}} \ . \tag{9}$$

Across locations, higher housing prices compensate either higher productivity or a higher local reservation wage. Employed workers do not move because their value exceeds the common value of unemployment. With those results at hand, it is now possible to characterize the location choice of employers.

2.3 The location choice of employers

An employer with a quality z contemplates the expected value from entering in each location, and chooses the location that delivers the highest payoff. When it matches, the employer receives a share $1 - \beta$ of the adjusted surplus. The employer's expected payoff in each location $\bar{J}(z, \ell)$ then follows from integrating over the job's initial productivity distribution $G_0(y_0|z)$, adjusted for the vacancy contact rate $q(\ell)$:

$$\rho \bar{J}(z, \ell) = q(\ell)(1 - \beta) \int S(y_0, \ell) dG_0(y_0|z) \tag{10}$$

To facilitate the exposition, I assume that the starting distribution G_0 is Pareto in the main text. I show that the Pareto assumption is empirically plausible in Section 4.4. Nonetheless, I also provide more general distributional conditions under which my results hold in Appendix B.4.

Assumption 1. *(Initial productivity distribution)*
Assume that the conditional starting distribution is Pareto with support $[Y, +\infty)$*,*

$$G_0(y_0|z) = 1 - \left(\frac{Y}{y_0}\right)^{\frac{1}{z}} \quad , \quad z \in (0, 1) \ .$$

[27] The term $Y^{-\tau}$ rises as Y approaches 1 from above. When an adverse productivity shocks pushes the match below the cutoff, both parties are better off separating rather than producing at below cutoff productivity, thereby insuring the pair against negative shocks. As productivity approaches the cutoff from above, the probability of productivity dropping below the cutoff rises, and so must the option value of separation.

Under Assumption 1, Lemma 1 implies that the expected payoff of job z in location ℓ satisfies

$$\log\left(\left(\bar{\rho}\bar{J}(z,\ell)\right)^{\frac{z}{1-z}}\right) = \frac{z}{1-z}\log\bar{S}(z) + \frac{z}{1-z}\log\ell + \frac{z}{1-z}\log q(\ell) - \log\underline{w}(\ell) \tag{11}$$

where $\bar{\rho} = \rho + \frac{\beta}{1-\beta}\underline{y}_0$ and $\bar{S}(z) = (Y/\underline{w}_0)^{1/z}\frac{z}{1-z}\frac{\tau z}{\tau z+1}$.

The four terms on the right-hand-side of equation (11) reveal four forces that shape how employers value different locations. The first term encodes the absolute advantage of employers according to their job quality z. High quality jobs draw from a better starting distribution, have higher productivity on average and earn higher profits regardless of their location. This term does not affect the location choice of employers.

The second term reflects standard technological complementarities in production. From the production function, more productive employers benefit relatively more from high local productivity ℓ. As a result, they value locating in more productive locations relatively more than unproductive employers. While local productivity differences ℓ are useful to fix ideas because they define a natural ordering of locations, they are not the central ingredient of the model. In fact, I show in Corollary 1 below that one can think of productivity ℓ being identical to 1 in every location, and ℓ simply indexing geographically distinct but otherwise ex-ante identical locations.

In contrast, the third term in equation (11) lies at the core of the mechanism this paper proposes. It reveals that more productive employers value relatively more locations where hiring is easy – where the vacancy contact rate $q(\ell)$ is high. Because more productive employers generate higher profits, waiting longer until they meet a worker and start producing is relatively more costly for them. Higher foregone profits translate into a higher opportunity cost of time for more productive employers. Importantly, some contacts do not result in a viable match, so that the vacancy filling rate and the vacancy contact rate differ. The probability that a contact results in a match, $\left(Y/\underline{y}(\ell)\right)^{1/z}$, depends on both the employer type z (first term) and on local reservation wages $\underline{w}(\ell)$ through the separation threshold $\underline{y}(\ell)$ (last term).

The vacancy contact rate $q(\ell) = m\theta(\ell)^{-\alpha}$ is an equilibrium object that depends on local market tightness $\theta(\ell)$. Ultimately, it depends on the pool of employers and workers who choose to locate in ℓ. Therefore, I follow Marshall (1920)'s terminology and call the complementarity between the employer's productivity z and the location's vacancy contact rate $q(\ell)$ a pooling complementarity. In contrast to technological complementarities which can be found in the assignment literature without frictions, the pooling complementarity arises at the confluence of the location choice of heterogeneous employers and frictional local labor markets.

Finally, the fourth term in equation (11) reflects the expected cost of labor in a particular location ℓ, which can be summarized by the reservation wage $\underline{w}(\ell)$. All employers prefer locations with low labor costs where the reservation wage is low.

In equilibrium, local reservation wages are related to local vacancy contact rates though labor market tightness $\theta(\ell)$. Therefore, employers face a trade-off between local vacancy contact rates and local wages. From the pooling complementarity, more productive employers value high vacancy contact rates relatively more. As a result, productive employers are willing to pay more for locating in places with a slack labor market and a high vacancy contact rate. In contrast, unproductive employers are priced out in high wage locations, while they forego lower profits by waiting for workers in locations with tight labor markets.

The differential valuation of locations by different employers plays the role of a single-crossing condition. The value of employers reflects their forward-looking optimal separation decision, which is fully determined by current productivity y relative to the endogenous threshold $\underline{y}(\ell)$ as per Lemma 1. From the Pareto distribution, the probability of drawing a below-threshold starting productivity decays more slowly for high quality employers.[28] Thus, both the Pareto assumption and the structure of the forward-looking optimal separation problem give rise to the single-crossing property. Appendix B.4 proposes more general distributional conditions under which employer payoffs exhibit the single-crossing property. For instance, it also arises if the starting distribution is a mass point at $y_0 = z$ and the starting productivities z are sufficiently far from the largest cutoff $\max_\ell \underline{y}(\ell)$.

An employer with quality z thus solves

$$\ell^*(z) \;=\; \underset{\ell}{\operatorname{argmax}} \;\; \frac{z}{1-z}\log \ell + \frac{z}{1-z}\log q(\ell) - \log \underline{w}(\ell) \tag{12}$$

Although every active job faces a dynamic optimal stopping problem in each location, the explicit solutions in Lemma 1 allow to simplify the location choice problem to one that shares many features with standard static assignment problems. Examples thereof can be found in Sattinger (1993) and Davis and Dingel (2014).[29] Apart from the underlying dynamic production decision, a distinction with those studies arises. Traditional assignment problems resolve the sorting between two-sided markets with exogenously given quantities. In contrast, in the present model, local labor markets clear through the adjustment of labor market tightness $\theta(\ell)$. The latter in turn feeds back into the vacancy contact rate, thereby adding an additional layer of general equilibrium effects to the payoffs that determine the assignment. This feedback acts as an agglomeration force, with two implications. First, cities with different ex-post characteristics emerge in equilibrium even in the absence of ex-ante heterogeneity.[30] Second, well-known multiplicity issues may arise.[31]

I define an assignement pair as a pair of functions $\ell \mapsto (z(\ell), \underline{w}(\ell))$, where $z(\ell)$ is the assignment function of employers to locations. It is the inverse of $\ell^*(z)$. In this paper, I call $z(\ell)$ the assignment function, while $\mathcal{M}$ is the matching function that determines contacts in the labor market. $\underline{w}(\ell)$ is the equilibrium reservation wage that supports this location choice. To facilitate the exposition in the main text, I restrict attention to assignments that exhibit weak positive assortative matching, i.e. for which z is increasing. In Appendix B.4, I show that this is only a mild restriction, for two reasons. First, when the matching function elasticity α is not too large, only assignments with positive assortative matching can exist. Second, any other potential steady-state assignment is dynamically unstable for any value of α, in a sense made precise in Appendix B.4.

Proposition 1. *(Sorting)*
Suppose that Assumption 1 holds. Fix the equilibrium value of unemployment U and the mass of new jobs M_e. There exists a unique solution $\ell \mapsto (z(\ell), \underline{w}(\ell))$ to (12) among all possible assignments with increasing z. There exists a threshold $\underline{\alpha} > 0$ such that for all $\alpha \in [0, \underline{\alpha}]$, this solution is unique among all possible assignments. z and $\underline{w}$ are strictly increasing functions.

[28]Indeed, $\frac{\partial^2 \log(1-G_0(y_0|z))}{\partial z \partial \log y_0} > 0$.

[29]For an in-depth exposition of the underlying theory, see Topkis (1998), Villani (2003) and Galichon (2016).

[30]See Gaubert (2018) for a similar idea when employers' technology directly depends on city population.

[31]See Grossman and Rossi-Hansberg (2012) for an example of multiple equilibria in a spatial context with agglomeration economies and exogenous differences across locations.

Proposition 1 establishes uniqueness of the assignment with positive assortative matching between local productivity ℓ and firm quality z: more productive jobs go to more productive locations. It also shows that this assignment is the only possible one when the matching function elasticity α is not too large. Proposition 9 in Appendix B.4 extends this result to all dynamically stable assignments and under more general distributional conditions for G_0.

The equilibrium response of local reservation wages $\underline{w}(\ell)$ to the location choice of employers sustains the positive assignment. Reservation wages adjust up to the point where the marginal employer is indifferent between locations ℓ and $\ell + d\ell$. This adjustment reflects two forces. First, reservation wages reflect expected future wages conditional on starting work, which depend on equilibrium employer quality $z(\ell)$. Therefore, reservation wages rise with ℓ, but less than one-for-one relative to wages of employed workers due to discounting. Second, reservation wages also reflect the job finding rate $f_R(\ell)$ and labor market tightness $\theta(\ell)$. As employers sort across locations, more workers locate in places with high expected wages and high employer quality $z(\ell)$. In response, labor market tightness $\theta(\ell)$ falls there, reducing the job contact rate $f(\theta(\ell))$. Since the value of search $v(\ell)$ reflects both the rising expected wages conditional on work and the falling job contact rate $f(\theta(\ell))$, reservation wages $\underline{w}(\ell)$ rise with ℓ, but again less than one-for-one relative to $z(\ell)$. By characterizing the allocation of heterogeneous jobs to locations, these results deliver predictions for spatial unemployment differentials. I turn thereto in the next section.

2.4 Endogenous job loss and unemployment

In every location, the job losing rate depends on three forces: the average starting productivity at new jobs, the productivity separation threshold, and how fast productivity depreciates from the starting productivity down to the threshold. The productivity depreciation rate is governed by the productivity process (3) and is constant across locations by assumption. Therefore, any differences in local job losing rates must arise because of differences in the ratio between the starting productivity and the separation threshold. Both are related to the equilibrium assignment function $z(\ell)$ and reservation wage $\underline{w}(\ell)$.

To make this argument precise and determine how many workers lose their job per unit of time, it is necessary to solve for the invariant distribution of employment across productivities in each location ℓ. Denote $g(y, \ell)$ its density function. In steady-state, $g(y, \ell)$ solves the Kolmogorov Forward Equation (KFE),

$$0 = (L_y^* g)(y, \ell) + n(\ell) g_0(y, \ell) \ , \qquad y > \underline{y}(\ell) \ , \tag{13}$$

where $g_0(\cdot, \ell)$ is the density associated with the entry distribution $G_0(y_0|z(\ell))$, which in turn depends on the equilibrium quality of jobs $z(\ell)$ that open in location ℓ. $n(\ell)$ is the endogenous inflow of unemployed workers into employment. The operator L_y^* encodes how productivity shocks shape the distribution.[32] Under the geometric Brownian motion assumption (3), it is given by

$$(L_y^* g)(y) = -\left(\frac{\sigma^2}{2} - \delta\right) \frac{\partial}{\partial y}\Big(y g(y, \ell)\Big) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial y^2}\Big(y^2 g(y, \ell)\Big) \ .$$

By construction, the density g must integrate to unity in each location: $1 = \int_{\underline{y}(\ell)}^{\infty} g(y, \ell) dy$. Because of

[32] It is the formal adjoint of the operator L_y. See Appendix B.5 for a heuristic derivation of the KFE (13), and Oksendal (1992) for a formal derivation.

Brownian shocks, the distribution must satisfy the additional boundary condition $g(\underline{y}(\ell), \ell) = 0$.[33] There always exists a closed-form solution to the KFE (13) with the boundary condition. To facilitate exposition in the main text, Lemma 2 describes that solution under Assumption 1. The generalized solution is given in Appendix B.5.

Lemma 2. *(Employment distribution)*
Let $\kappa = \frac{2\delta}{\sigma^2}$. Under Assumption 1, the solution to the KFE (13) with $g(\underline{y}(\ell), \ell) = 0$ satisfies

$$g(y, \ell) = \frac{\kappa}{\kappa z(\ell) - 1} \left[\left(y/\underline{y}(\ell) \right)^{-\frac{1}{z(\ell)}} - \left(y/\underline{y}(\ell) \right)^{-\kappa} \right] \quad , \ \forall y \geq \underline{y}(\ell) \ .$$

The steady-state distribution has two components. The first component reflects the productivity distribution of new jobs. The invariant productivity distribution inherits the right tail from the starting distribution $1/z(\ell)$. The right tail is thicker in locations with high quality $z(\ell)$. The second component reflects the productivity process. When the negative drift δ is higher, κ is higher, implying that the distribution is more left-skewed as productivity depreciates faster. When volatility σ is higher, κ is lower and the distribution is more right-skewed: more jobs receive large positive shocks, while large negative shocks are truncated due to endogenous job loss. Finally, the entry rate n does not appear because it simply scales the overall mass of employed workers, as in Hopenhayn and Rogerson (1993).

Having solved for the invariant distribution in each location ℓ, it is possible to determine the endogenous job losing rate $s(\ell)$ (or separation rate into unemployment). Given the steady-state distribution, the local endogenous job losing rate depends on how many workers are close to the cutoff. In Appendix B.5, I show that it is

$$s(\ell) = \frac{\sigma^2}{2} \frac{\partial g}{\partial y}(\underline{y}(\ell), \ell) \ . \tag{14}$$

Recall that close to the cutoff, only workers who receive a negative shock become unemployed. Because Brownian shocks dominate in any small time period and are symmetric, only half of shocks result in job loss, hence the division by two. Because $g(\underline{y}(\ell), \ell) = 0$, the number of job losers follows from the second order contribution of the mass of workers close to the cutoff, which is $\frac{\partial g}{\partial y}(\underline{y}(\ell), \ell)$.

Expression (14) for the local job losing rate is useful when combined with the explicit solution to the distribution in Lemma 2. Together, they produce a simple solution to the local endogenous job losing rate as well as for labor market flows at the local level. Again to facilitate exposition in the main text, only the expression under the Pareto assumption is presented in the main text. Appendix B.5 describes the general solution.

Proposition 2. *(Spatial unemployment differentials)*
Under Assumption 1, the local job losing, finding and unemployment rates in location ℓ are

$$s(\ell) = \frac{\delta}{z(\ell)} \ , \qquad f_R(\ell) = f(\theta(\ell)) \times \left(\frac{Y}{\underline{y}(\ell)} \right)^{1/z(\ell)} \ , \qquad u(\ell) = \frac{s(\ell)}{s(\ell) + f_R(\ell)} \ .$$

[33] In a small time period, the Brownian motion shocks dominate the negative drift. Because these shocks are symmetric, half of the workers close to the cutoff are pushed into unemployment in any small time period. Compounded over a non-zero time interval, this process leaves no workers at the cutoff. A formal proof is provided in Appendix B.5

In addition, the job losing rate is decreasing in ℓ.

The Pareto case is particularly transparent. When the negative drift δ is higher, productivity depreciates faster everywhere and the endogenous job loss rate increases in all locations.[34] In low ℓ locations, local jobs are of low quality $z(\ell)$. Hence, they draw from a left-skewed distribution and enter close to the endogenous threshold. Thus, they fall below the threshold early on and the local job losing rate is high. In high ℓ locations, local jobs are of high quality $z(\ell)$ and hence enter highly productive. Because of both discounting and the general equilibrium adjustment of labor market tightness, reservation wages $\underline{w}(\ell)$ rise slower than the assignment function $z(\ell)$ across locations. Hence, the endogenous separation threshold $\underline{y}(\ell)$ increases less than one-for-one across locations relative to $z(\ell)$. As a result, the ratio between the average starting productivity and the threshold $\underline{y}(\ell)$ is larger and productivity takes more time to fall below the local threshold $\underline{y}(\ell)$. Therefore, the job losing rate is low in high ℓ locations. Overall, positive assortative matching between firm quality z and local productivity ℓ implies that the job losing rate is decreasing in local productivity.

By contrast, the job finding rate is the outcome of two opposing forces. It is the product of the worker contact rate and the probability that a given contact results in a job, the contact-to-job probability. First, the worker contact rate depends positively on labor market tightness $\theta(\ell)$. As more productive employers $z(\ell)$ benefit more from higher vacancy contact rates $q(\theta(\ell))$, the worker contact rate $f(\theta(\ell))$ is negatively correlated with $z(\ell)$. However, the contact-to-job probability $\left(Y/\underline{y}(\ell)\right)^{1/z(\ell)}$ pushes in the other direction. In locations with more productive employers $z(\ell)$, contacts are more likely to result in a job because new matches draw from a better productivity distribution, and because the endogenous separation threshold $\underline{y}(\ell)$ rises less than one-for-one with local employer productivity. Both forces need not offset each other exactly, but when they almost do, the job finding rate is close to flat across locations.

Recall that, if anything, the job finding rate is moderately negatively correlated with the job losing rate in the data (Figure 2). In the model, pooling complementarities incentivize employers with stable jobs to locate where there are few vacancies per job seeker. As a result, the worker contact rate is positively correlated with the job losing rate. Thus, the model can rationalize the moderate negative correlation between job losing and job finding rates only if the contact-to-job probability more than offsets the direct correlation with the worker contact rate. In Section 4.4, I validate this implication of the model. Using survey data, I show that the worker contact rate is indeed positively correlated with the job losing rate while the probability that a contact results in a job pushes the other way more than one-for-one.

2.5 Equilibrium and comparative statics

Having characterized how the location choice of employers shapes spatial unemployment differentials, I close the economy in the decentralized equilibrium. Local housing and labor markets must clear in each

[34] Perhaps surprisingly, the volatility of the productivity process σ does not affect the job losing rate with Pareto entry. This reflects two opposing forces. When volatility σ increases, matches receive larger *negative* shocks, which may push them into breaking up more frequently – the direct volatility channel. But matches are also subject to larger *positive* shocks, which raises the option value of producing and lowers the cutoff – the option value channel. In general, this second channel operates through the cutoff $\underline{y}$ that appears in the general expression for the job losing rate in Appendix B.5. When firms enter with a Pareto distribution, the direct volatility channel and the option value channel exactly offset each other and changes in volatility do not affect the endogenous job losing rate in steady-state.

location ℓ,

$$r(\ell) = \omega L(\ell)\Big(u(\ell)b\ell + (1-u(\ell))\overline{w}(\ell)\Big) \ , \qquad\qquad \theta(\ell) = \frac{M_e F_z'(z(\ell))z'(\ell)}{u(\ell)L(\ell)F_\ell'(\ell)} \ , \tag{15}$$

where $L(\ell)$ is population in location ℓ, and $\overline{w}(\ell) = \int w^*(y,\ell)g(y,\ell)dy$ is the average wage in location ℓ.

Local housing prices reflect local expenditures on housing. The labor market clearing condition simply states that labor market tightness is the ratio between the number of vacancies and the number of unemployed workers in locations with productivity ℓ. The number of unemployed workers is the unemployment rate times total population across the $F_\ell'(\ell)d\ell$ locations with productivity in $[\ell, \ell + d\ell)$. The number of vacancies in a location reflects the total number of new jobs, M_e, but also the spatial sorting of employers. There are fewer employers in locations where the assignment function z is steep. In that case, a given mass of employers is stretched across a wider set of locations.

Finally, employers enter freely each period, so that the cost of entry is equal to the expected value from entering, and total population in the economy must add up to unity,

$$c_e = \int \bar{J}(z, \ell^*(z))dF_z(z) \ , \qquad\qquad 1 = \int L(\ell)dF_\ell(\ell) \ . \tag{16}$$

A decentralized equilibrium is comprised of a mass of entering employers M_e, a value of unemployment U, an assignment function $z(\ell)$, a reservation wage function $\underline{w}(\ell)$, wages of employed workers $w^*(y,\ell)$, an employment distribution $g(y,\ell)$, a distribution of unemployment $u(\ell)$ and market tightness $\theta(\ell)$, housing prices $r(\ell)$, and a population distribution $L(\ell)$, such that (4), (5), (7), the definitions in Lemma 1, (8), (9), (12), (13), (14), (15), and (16) hold. The following proposition characterizes existence and uniqueness of the decentralized equilibrium.

Proposition 3. *(Existence and uniqueness)*
Under Assumption 1, there exists a decentralized steady-state equilibrium with weak positive assortative matching. There exist $d_z, d_\ell > 0$ such that, for $|\overline{z} - \underline{z}| < d_z$ and $|\overline{\ell} - \underline{\ell}| < d_\ell$, the equilibrium is unique.

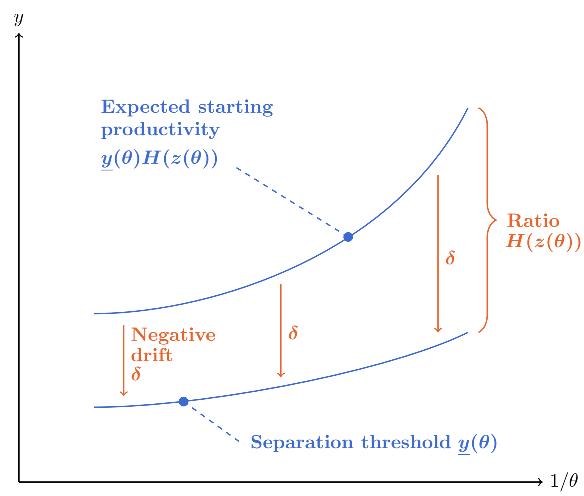
Proposition 3 guarantees that there exists a unique steady-state equilibrium with weak positive assortative matching, when there is not too much dispersion in spatial and productivity primitives. Although this result only proves uniqueness in a particular region of the parameter space, I check in simulations of the estimated model that my algorithm selects an equilibrium that continuously converges to one in the local uniqueness region as I shrink the dispersion in F_ℓ, F_z to zero. To facilitate exposition, Proposition 3 focuses on the Pareto case in Assumption 1. In the Appendix, I extend these results to more general entry distributions G_0 as well as to all dynamically stable steady-states.

With sufficient conditions for existence and uniqueness at hand, it is possible to study comparative statics across equilibria. To shed further light on how spatial unemployment differentials depend on the labor market pooling complementarity, it is useful to consider a particular limiting economy in which ex-ante spatial differences in ℓ become arbitrarily small. In that case, only the pooling complementarity may determine sorting as well as any ex-post differences across locations. Corollary 1 below shows that spatial differentials in job losing and unemployment rates arise even in the absence of any substantial ex-ante heterogeneity between locations.

Corollary 1. *(Non-vanishing spatial differences with ex-ante identical locations)*
Suppose that the conditions in Proposition 3 hold and that the matching function elasticity α is strictly positive. Then the variance of local job losing and unemployment rates remain strictly positive and bounded above zero as the variance in exogenous differences ℓ goes to zero.

This results highlights that the pooling complementarities suffice to sustain sorting in equilibrium, irrespectively of technological complementarities.[35] When technological differences ℓ vanish, locations are ex-ante identical and ex-post differences emerge endogenously. In particular, job losing and unemployment rates differ across locations. This is possible because congestion in local housing markets allows for differences in reservation wages across locations. Figure 4 depicts the structure of the equilibrium in that case. In the limit, locations can be re-indexed by labor market slackness $1/\theta$, which is on the x-axis. The y-axis shows the endogenous separation threshold $\underline{y}(\theta)$ as a function of labor market tightness, as well as expected starting productivity. From the solution to the KFE in Proposition 2, the ratio between the average starting productivity and the separation threshold is $H(z(\theta)) = \frac{\kappa}{(\kappa-1)(1-z(\theta))}$. Consistently with Proposition 2, it rises with the assignment function $z(\theta)$. Thus, it also rises with market slackness $1/\theta$. In contrast, if housing played no role $\omega = 0$, all locations would become ex-post identical.

Figure 4: Spatial equilibrium with ex-ante identical locations.



Corollary 1 indicates that labor market pooling complementarities combined with congestion in local housing markets have the potential to lead to large differences in job losing and unemployment rates by incentivizing heterogeneous employers to self-select in space. I now assess whether the resulting location choices are efficient.

[35] At a more formal level, taking the limit of arbitrarily small differences selects one particular equilibrium in the limit without any exogenous spatial heterogeneity. When exogenous spatial differences are exactly zero, there is an infinity of equilibria because locations can be arbitrarily reshuffled. However, there are only two possible spatial distribution of equilibrium outcomes: the mixing distribution in which all locations are identical, and the separating distribution in which locations differ due to sorting. The separating distribution survives because of labor market pooling complementaries. Taking the limit under vanishing spatial heterogeneity always selects the separating distribution. In addition, the mixing distribution is trembling-hand unstable.

2.6 Efficiency and planning allocation

To fix ideas, recall that in a single-location model of the labor market such as Mortensen and Pissarides (1994), the only sources of inefficiency are the overall entry and separation margins. These arise because of a single missing price: the price of labor market tightness. Both margins are efficient only when employers are compensated for opening and shutting down jobs by exactly as much as they congest the matching function. This is the case when the Hosios (1990) condition $\alpha = \beta$ holds. The same logic carries through to the model with many locations for the overall entry and separation decisions.

With geography, employers must make an additional decision: the location choice. It introduces an additional margin of inefficiency. There are many labor markets to choose from, but there is still no price for market tightness in any local labor market. Thus, there is not one, but many missing markets. Efficiency requires that employers are compensated by exactly as much as they congest the matching function in each location. However, due to the spatial heterogeneity in profitability and the spatial sorting, the congestion effect on the matching function varies across space.

To understand the nature of this spatial externality, consider two locations $\ell_1 < \ell_2$. Each location is populated with jobs $z_1 = z(\ell_1) < z(\ell_2) = z_2$. Consider a marginal job $z \in (z_1, z_2)$ contemplating opening in locations ℓ_1 or ℓ_2. If job z enters in location ℓ_2, it is worse than the average local job. Due to labor market frictions however, it does meet as many workers as its more productive competitors. By opening in location ℓ_2, job z exerts a negative externality on all other open jobs there because it diverts workers away from them. This externality is also socially harmful, as workers are redirected towards a less productive job, $z < z_2$. Symmetrically, the marginal job z exterts a negative externality on other jobs in location ℓ_1 if it enters there. However, the externality is socially beneficial in this case, as workers are redirected towards a more productive job, $z > z_1$. In both cases, the magnitude of the externality depends on the quality of local jobs z_1 or z_2, and on the quality of the newcomer z.

On net, the marginal job has an incentive to free-ride the favorable hiring conditions in location ℓ_2, because its vacancy contact rate does not reflect that it is worse than average there. Wages are bargained ex-post and thus do not fully price contact rates. As a result, employers will concentrate too much in the best labor markets relative to the social optimum. This inefficiency trickles down across locations and generates misallocation throughout the economy.

The externality thus emerges at the confluence of three features of the model. First, geography creates many labor markets. Second, employers are heterogenous and choose where to locate. Third, labor markets are frictional and matches are formed with some degree of randomness. The externality arises because heterogeneous employers would be pooled in the same matching function, should they deviate off equilibrium play. Thus, I call it a labor market pooling externality. To make these arguments precise, I now define the planner's problem.

Planning problem. A utilitarian planner maximizes a possibly weighted sum of values of all individuals in the economy, taking the search frictions as given. The decentralized equilibrium is inefficient when there exists no set of utilitarian weights such that the allocations under the decentralized equilibrium and under the planning solution coincide. Otherwise, the decentralized equilibrium is efficient. Because the planner can freely reallocate the final good across locations while workers can only consume their income in the decentralized equilibrium, only one set of utility weights delivers planning allocations that may coincide with the decentralized equilibrium. These weights are defined in equation (49), Appendix B.8.

The planner controls where to send unemployed and possibly employed workers to search for jobs. The planner also decides when to break up matches, and is subject to the same search frictions as in the decentralized equilibrium. Because idiosyncratic productivity shocks are persistent, the planner must take the entire distribution of employment across productivities and locations as a state variable. If the planner does not know this distribution, she may not break up matches optimally. This distribution is an infinite-dimensional object. Nevertheless, a well-defined planner problem can be established with carefully chosen functional spaces for the distribution, described in Appendix B.8. To do so, I build on the work of Moll and Nuño (2018) who propose a method to solve planning problems with infinite-dimensional heterogeneity.[36] Because it involves additional notation, I relegate the formal definition of the planning problem to Appendix B.8 and simply characterize it in the main text. Denote with SP supercripts variables in the planning solution, and with DE superscripts variables in the decentralized equilibrium.

Proposition 4. *(Planning solution)*

- *With utility weights from* (49)*, sorting (Proposition 1), local labor market flows (Proposition 2), and existence and uniqueness (Proposition 3) results extend to the planning solution under the same conditions.*
- *The decentralized equilibrium is inefficient for all values of $\alpha, \beta \in (0, 1]$.*
- *Suppose $\beta = \alpha$ and that the supports of F_ℓ, F_z are not too large as in Proposition 3. Then for all ℓ:*
    - $z^{SP}(\ell) \geq z^{DE}(\ell)$ *with equality if and only if $\ell \in \{\underline{\ell}, \overline{\ell}\}$.*
    - $\frac{\partial \log \underline{w}^{DE}}{\partial \ell}(\ell) > \frac{\partial \log \underline{w}^{SP}}{\partial \ell}(\ell)$
- *The planning allocation coincides with the allocation in a decentralized equilibrium in which search is directed.*

Proposition 4 first establishes that the basic sorting, labor market flows, existence and uniqueness properties of the decentralized equilibrium also hold in the planning solution. Second, it formalizes the discussion above by stressing that the decentralized equilibrium is always inefficient, even when the Hosios (1990) condition $\alpha = \beta$ holds. To illustrate the externality, I compare the private value of jobs entering a particular location in the decentralized equilibrium, to the planner's value of sending the same job to the same location. Conditional on the same separation threshold $\underline{y}(\ell)$, these values satisfy

$$\left(\frac{\bar{J}^{DE}(z,\ell)}{\bar{J}^{SP}(z,\ell)}\right)^{\frac{1-\alpha}{\alpha}} = \frac{\bar{S}(z^{DE}(\ell))}{\bar{S}(z)} \cdot \left(Y/\underline{y}(\ell)\right)^{\frac{1}{z^{DE}(\ell)} - \frac{1}{z}} . \tag{17}$$

The planner's valuation of opening job z in location ℓ only depends on the quality of that particular job, z. In contrast, the private value from entering in the same location ℓ for job z also depends on the quality other local jobs $z^{DE}(\ell)$. This difference exactly encodes the labor market pooling externality, acting though the vacancy contact rate. Because $z^{DE}(\ell)$ is increasing, employers over-value opening jobs in locations where other employers are productive. As the planner considers all possible assignment

[36] My approach departs from theirs in two aspects. First, because of endogenous separations, their results do not directly apply and I must start from first principles. Second, minor modifications are needed for the general methodology to be fully consistent. For instance, the distribution must actually lie in a Sobolev-Strichartz space rather than a Lebesgue space. I provide details in Appendix B.8. I thank Ben Moll and Galo Nuño for useful related discussions.

functions $z^{SP}(\ell)$, it internalizes that mixing different jobs in the same location is not optimal. In contrast, deviating away from sorting is a viable alternative for employers in the decentralized equilibrium.

The comparison between the assignments z^{DE} and z^{SP} follows. For any location ℓ, the local employer is not productive enough in the decentralized equilibrium relative to the planner's choice. Indeed, the more productive employers are too concentrated in high productivity locations $\ell' > \ell$ in the decentralized equilibrium. As a result, shadow reservation wages rise too fast in the decentralized equilibrium.

Finally, recall that the spatial externality arises because there is no price for labor market tightness in any location. I outline an alternative setup with directed search in Appendix B.8. The key assumptions are that firms are able to commit to fully state-contingent contracts and that workers can perfectly allocate between submarkets within each location should they offer different contracts. Employers then internalize that by entering in a local labor market with higher quality than their own, they depress their contact rate as workers direct their search away towards the more productive jobs. As a result, they post wage contracts that exactly price congestion effects and the decentralized equilibrium is efficient. Whether search is directed or random is ultimately an empirical question with data requirements that go beyond the scope of this paper. In principle, reality is likely to lie between both models.

Nevertheless, I propose two checks to lend credibility to this paper's welfare implications. First, I allow employers to post many vacancies in the extended model of Section 3. More productive employers post more vacancies than less productive ones. Thus, they contact relatively more workers, mitigating the strength of the externality, akin to directed search. The vacancy cost elasticity then determines where the model lies between random and directed search. At the estimated cost, I find large welfare effects from place-based policies. Second, Table 12 in Appendix E shows that re-estimating the model under the directed search assumption delivers too little dispersion in local unemployment rates relative to the data and misses the variance decomposition into job losing and finding rates described in section 1. Conditional on the rest of the model and in this spatial context, the data thus supports the random search assumption among those two extreme cases.

2.7 Optimal policy

Given that the the decentralized equilibrium does not attain the first best, a natural question is whether it can be restored using standard policy instruments. An optimal policy should achieve the following. First, it should correct the pooling externality by incentivizing employers to open jobs in low profitability locations. Second, it should enforce the Hosios (1990) condition. I introduce place-based policies into the model in Appendix B.9 and show in Proposition 5 that they can be used to bring the economy back to its first-best.

Proposition 5. *(Optimal policy)*
Constrained efficiency is restored with a combination of place-based policies:

- *A labor subsidy increasing in local productivity ℓ if and only if $\beta < \alpha$.*
- *A profit subsidy decreasing in local productivity ℓ.*
- *A lump-sum transfers to owners.*[37]

[37] Alternatively, if there are no absentee owners and profits are rebated to workers with a flat earnings subsidy, then a flat earnings tax replaces the lump-sum tax on owners.

The labor subsidy implements the Hosios (1990) condition. As in Kline and Moretti (2013), spatial variation in workers' value of search makes that policy place-specific. Similarly to their results, labor needs to be taxed more heavily in low productivity locations on the empirically relevant side of the Hosios (1990) condition $\beta < \alpha$.[38] Because this particular trade-off has been extensively studied, I focus primarily on the externality in the location choice of jobs.

The spatial misallocation that results from the labor market pooling externality results in an optimal profit subsidy that resembles real-world place-based policies. The Empowerment Zone program in the United States and its French equivalent both grant large effective profit subsidies for firms opening jobs in distressed areas – in practice, they guarantee tax exemptions relative to a baseline tax rate. In the model, the profit subsidy corrects the labor market pooling externality that equation (17) obviates. Subsidies must rise as local productivity ℓ diminishes, and thus rise with the local job losing rate as per Proposition 2. From Section 1, those locations have high unemployment in the data. Provided the model can tie together high job losing rates and high unemployment rates, it propose a structural justification for subsidizing high unemployment areas: high productivity employers fail to internalize their positive labor market spillovers there. To the best of my knowledge, this is the first paper to propose a structural justification for such policies based on frictional labor markets and two-sided mobility of workers and employers. Finally, lump-sum transfers to owners balance the government's budget.

So far the spatial and individual heterogeneity in the model has remained minimal. To quantitatively account for local labor market flows and the welfare effects of place-based policies, I enrich this baseline framework in Section 3 below.

3 Extended model and estimation

In this section, I first describe the extensions of the model. I then establish how the results from Section 2 extend to the richer environment. Finally, I detail the estimation strategy.

3.1 Quantitative setup

Geography. There is ample empirical evidence that locations differ in residential amenities. Better amenities attract more workers which may congest the labor market. Incorporating amenities thus allows to capture the joint spatial variation in population, wages, and unemployment. Hence, I now assume that locations differ both in productivity p and amenities a. Locations are indexed by productivity-amenity pairs $\ell = (p, a)$, and are exogenously distributed with cumulative function F_ℓ on a connected support.

Housing supply. The magnitude of welfare gains from place-based policies that attract jobs and workers depends on how much local congestion offsets the direct gains from the policy. To better capture this force, I introduce perfectly competitive land developers using the final good to produce housing on a unit endowment of land with an isoelastic production function. It results in a local housing supply given by $H(r(\ell)) = H_0 r(\ell)^\eta$.

Migration frictions. The externality that arises due to labor market pooling complementarities relies on the general equilibrium response of worker mobility to changes in local economic conditions. Thus,

[38] Kline and Moretti (2013) deem this conclusion to be "rather counter-intuitive".

the migration elasticity of workers crucially affects the welfare gains from place-based policies. Instead of being freely mobile, workers now receive the opportunity to move at Poisson rate $\mu \geq 0$. When hit by this "moving opportunity", they receive a set of preference shocks for locations $\{\varsigma_\ell\}_\ell$ that are Frechet-distributed with shape parameter $1/\varepsilon$, and choose where to locate.[39] Those shocks stay constant until the next moving opportunity arrives.

Preferences. The flow utility function becomes $u(c, h, a, \varsigma) = \left(\frac{c}{1-\omega}\right)^{1-\omega} \left(\frac{h}{\omega}\right)^{\omega} a\varsigma$.

Non-participation. Workers stochastically exit the labor force at Poisson rate $\Delta > 0$. When they do, they are replaced by a single new worker. Entry and exit from the labor force stabilizes the human capital distribution described below.

Learning and human capital. An important channel through which unemployment harms workers above and beyond direct earnings losses is by hindering their ability to accumulate labor market experience. When out of work, not only do individuals fail to accumulate valuable knowledge, but their human capital tends to depreciate over time. In a spatial context with limited worker mobility, these scarring effects in high unemployment areas produce pockets of low human capital labor. There, high quality jobs may be less likely to open, further worsening local labor market conditions and magnifying spatial disparities. Thus, learning effects and localized unemployment interact through the location choice of employers and may amplify welfare gains from place-based policies.[40]

To parsimoniously capture this idea, I assume that workers now differ in their human capital k. When employed, workers' human capital grows at rate $\lambda \geq 0$. When unemployed, their human capital grows at rate $\lambda - \varphi$. $\varphi \geq 0$ encodes the relative depreciation rate of human capital for unemployed workers. Consistently with the idea that young workers enter the labor force with human capital that reflects the average human capital in the economy, I assume that the distribution of human capital of new workers k_t also shifts at rate λ: the rescaled distribution $k_t e^{-\lambda t}$ for new workers does not depend on calendar time t, and is denoted F_k.[41] I also assume that workers with different human capital in the same location search in the same labor market: potential employers cannot discriminate between workers with different human capital prior to meeting with them.

Production. I allow employers to use housing in production, to capture the idea that local congestion due to higher population affects production costs. Filled jobs with idiosyncratic productivity y in a location with local productivity p thus use housing h and human capital k of their employee to produce, with production function $(ypk)^{\frac{1}{1+\psi}} h^{\frac{\psi}{1+\psi}}$.

[39] The shifter is normalized to $1/\Gamma(1-\varepsilon)$, where Γ is Euler's Gamma function, because it is not separately identified from amenities a.

[40] Human capital differences also allow the model to jointly account for the role of sorting in wage and unemployment differentials. Wages reflect human capital, the sorting of which thus contributes to spatial wage differentials directly. In contrast, because human capital is transferable between jobs, the separation decision is independent of human capital. As a result, the local mix of human capital does not directly affect spatial job loss differentials.

[41] This assumption can be understood as young workers learning from older workers prior to entry in the labor force. The economy is therefore on a balanced growth path determined by λ. In levels, the distribution of knowledge of new workers is a "travelling wave with constant shape". I also assume that F_k has a density with full support equal to $\mathbb{R}_+$. These assumptions also help with tractability.

Recruiting intensity. Finally, I let employers adjust their recruiting efforts. This channel potentially mitigates the strength of labor market pooling externalities. Thus, employers with open jobs are now allowed to post many vacancies v at cost $\frac{c_v}{1+1/\gamma}v^{1+1/\gamma}$.

3.2 Characterization

The extensions preserve the basic structure of the location choice of employers. I show in Appendix C that when migration opportunites are rare enough $\mu \ll 1$ and the relative depreciation rate of human capital is small enough $\varphi \ll 1$, the location choice of job z in equation (12) becomes

$$\underset{(p,a)=\ell}{\text{argmax}} \ \frac{z}{1-z}\left\{ \log\left(\left[p^{\mathcal{Q}}a^{-\psi\mathcal{P}}\right]\cdot C\big(\underline{w}(\ell), z(\ell)\big)^{\psi\mathcal{P}}\right) + \log q(\ell) + \log\left(\bar{h}(u(\ell))^{\mathcal{Q}}\right)\right\} - \log \underline{w}(\ell) \tag{18}$$

where $\mathcal{P} = \frac{1}{\omega+\psi+\varepsilon(1+\eta+\psi)}$ and $\mathcal{Q} = \frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}$, and the average human capital in location ℓ is a decreasing function of the local unemployment rate $u(\ell)$. It is equal to $\bar{h}(u(\ell)) = \frac{\Delta}{\Delta+\varphi u(\ell)}$ up to a general equilibrium constant. Recall that $\ell = (p, a)$ now indexes productivity-amenity pairs. The function C is defined in Appendix C and is increasing in each argument.

Equation (18) first highlights that in the extended model, technological complementarities depend on a combination of productivity p and amenities a. Higher productivity p makes locations more lucrative for jobs, but higher local amenities reduce profitability. Higher amenities bring in more workers, raising housing prices and driving up production costs. This housing price channel explains why the amenity contribution enters with an elasticity ψ. Anticipating a result showing that this single index is a local sufficient statistic for the model's outcomes, I identify a pair $\ell = (p, a)$ with the combined index of local advantage

$$\ell(p, a) \equiv p^{\mathcal{Q}}a^{-\psi\mathcal{P}} \ . \tag{19}$$

In addition, local expenditures on housing also depend on local wages, as captured by $C(\underline{w}(\ell), z(\ell))$.

Second, equation (18) shows that labor market pooling complementarities remain unchanged and still depend only on the local vacancy contact rate. Similarly, the expected cost of labor continues to be summarized by local reservation wages $\underline{w}(\ell)$.

Third, equation (18) reveals the contribution of learning at the workplace for the location choice of employers. Average local human capital $\bar{h}(u(\ell))$ falls as the local unemployment rate rises as workers are more frequently scarred by unemployment. When scarring effects φ are stronger relative to how frequently the workforce turns over (Δ), a given unemployment rate is associated with worse average local human capital.

Because the structure of the location choice of employers in equation (18) closely resembles its more stylized version in equation (12), virtually all the analytical results from Section 2 carry through.

Proposition 6. *(Characterization of the extended model)*
When the migration rate μ and the scarring effects of unemployment φ are not too large, Propositions 1, 2, 3, 4, 5 and Corollaries 1, 2 obtain in the extended framework under the same conditions, with three modifications. First, replace the local unemployment rate by $u(\ell) = \frac{s(\ell)+\mu+\Delta}{s(\ell)+\mu+\Delta+f_R(\ell)}$. Second, replace ℓ

with the combined index of local advantage $\ell(p,a)$. Third, population depends on both $\ell(p,a)$ and on a conditional on $\ell(p,a)$: $L(p,a) \equiv L(\ell(p,a),a)$.

Population cannot be summarized solely by the local advantage index $\ell(p,a)$ because workers value amenities directly, while employers value amenities only through local housing prices. As a result, amenities generate variation in population even conditional on the local advantage index $\ell(p,a)$. Thus, in the model as in the results from Section 1.4, local unemployment correlates negatively with local wages, but not with population conditional on wages. I provide more details in Appendix C.1.7. Having laid out the structure of the extended framework, I turn to the structural estimation.[42]

3.3 Estimation strategy and identification

Despite its rich structure, the quantitative model is transparent enough to produce estimating equations for almost all key parameters. In particular, no simulation will be required until the last step, which estimates the entry cost. To make this argument precise, I discuss how each parameter can be recovered recursively given the data I choose. A proposition at the end of this subsection summarizes the formal identification of the model. Different specific estimators are used for different parameters, but all can be nested into an overarching Generalized Methods of Moments (GMM) estimator. In total, there are 19 parameters to be estimated: $\rho, \Delta, \omega, \psi, \delta, \sigma, \beta, b, Y, \eta, \mu, \varepsilon, \alpha, \gamma, c_v, m, \lambda, \varphi$; together with two distributions $F_z, F_{p,a}$. I do not specify functional forms in the main estimation as they can be recovered non-parametrically.

The 19 parameters can be divided into three groups. Parameters in the first group – $\rho, \Delta, \omega, \psi, \mu, b, c_v, m$ – directly map into empirical counterparts or can be normalized, thus only requiring simple Minimum Distance Estimators (MDE). Parameters in the second group – $\delta, \sigma, \beta, Y, \eta, \varepsilon, \alpha, \gamma, \lambda, \varphi$ – require more involved estimating equations, together with different estimators. The third group of parameters only contains the entry cost c_e, which is estimated by numerical search (Method of Simulated Moments). Finally, I parametrize the distributions $F_z, F_{p,a}$ after the estimation for simulation purposes, leading to a fourth group of parameters. Before describing how to estimate each group of parameters, I briefly discuss the data used to construct empirical targets.

Data. I use data from France for all years between 1997 to 2007. I choose a quarter as the baseline time period $[t, t+1)$. For most of the estimation, I use averages over the entire period. For some parameters I break down the sample into two subperiods, and use averages for 1997-2001 and for 2002-2007. I index locations (cities) in the data by c. I use aggregate data for expenditure shares on housing for households. I measure expenditures on real estate for firms in the firm-level balance sheet data. Using the DADS-LFS combination, one obtains measures of local unemployment rates u_c, local job losing rates for stayers s_c, local job finding rates for stayers f_{Rc}, local average wages W_c, and population shares L_c. The DADS-LFS combination also delivers measures of aggregates such as the geographic mobility rate of workers and the average job offer acceptance probability. Finally, the DADS-LFS allow finer disaggregation of job losing rates and wages by tenure and location which will be useful to estimate several parameters in the second group. The last data source is the online realtor MeilleurAgents.com, from which I construct commuting zone housing prices r_c.

[42] Appendix C.2 discusses additional possible extensions and how they may affect the results: industry heterogeneity, labor market segmented by skill, technological and amenity spillovers, and trade costs.

First group (8 parameters). The moving opportunity rate μ is directly identified from the geographic mobility rate for individuals transitioning into unemployment at the same time.[43] Δ then follows from the flow equation for unemployment: $\Delta = \sum_c L_c \left(\frac{\bar{u}}{1-\bar{u}} f_{Rc} - s_c - \mu\right)$, where $\bar{u}$ is the aggregate unemployment rate. The interest rate identifies ρ through the effective discount rate of individuals $\rho + \Delta$. Next, household's expenditure share on housing ω can be directly equated to the value reported by INSEE (23%).[44] Similarly, the expenditure share on real estate out of value added by employers ψ is equated to my estimate of 11%.[45] The remaining parameters can be normalized $b = c_v = m = 1$.[46]

Second group (10 parameters).

Productivity process $\boldsymbol{\delta}$ and $\boldsymbol{\sigma}$. To estimate (δ, σ), I use data on job losing rates and wage growth by tenure. To that end, I leverage a closed-form solution to the time-dependent KFE equation derived in Appendix D.2. This solution delivers an explicit expression for the time-aggregated job losing rate in the first year in each location in the model. Given the measured average job losing rate s_c in city c, the job losing rate in the first year of a job in city c is $s_1(s_c, \hat{\delta})$, where s_1 is an explicit decreasing function of $\hat{\delta}$ given μ, Δ, and is specified in Appendix D.2. Intuitively, if the volatility σ is much larger than the drift δ, many separations occur at early tenure. Denoting s_{1c} the measured job losing rate in the first year in city c, I recover $\hat{\delta}$ directly by estimating

$$s_{1c} = s_1(s_c, \hat{\delta}) \tag{20}$$

with Non-Linear Least Squares (NLLS), treating residuals as measurement error.

Given the estimated ratio $\hat{\delta} = \frac{\delta}{\sigma}$, the same solution to the time-dependent KFE allows to explicitly compute wage growth by tenure when β is not too large. Appendix D.3 shows that it identifies the common scale of δ, σ. Intuitively, when productivity depreciates faster, wages at continuing jobs fall behind wages at new jobs at a faster pace. Thus, a NLLS regression similar to (20) estimates δ. When β is large, the tenure profile of wages must be computed numerically, and (β, σ) must be estimated jointly. At the estimated bargaining power β the difference is negligible.

Bargaining power $\boldsymbol{\beta}$. The labor share in location c is $\beta + \frac{1-\beta}{H(s_c)}$, where H only depends on δ and σ. I target the aggregate labor share net of debt servicing expenditures to identify β by simple MDE.[47]

Learning rates $\boldsymbol{\lambda}$ and $\boldsymbol{\varphi}$. Wage changes for workers coming out of unemployment reflect human capital losses, that grow with unemployment duration. Appendix D.5 shows that, for worker i who loses

[43] In the model, unemployed and employed workers always change location and enter unemployment when they receive the moving opportunity at rate μ. That rate must be time-aggregated quarterly.

[44] INSEE's calculations reflect both renters and homeowners.

[45] Balance sheet data lists all rental expenditures, as well as the book value of land, building and structures owned by the firm. I annuitize the value of those properties using a 5% annual interest rate, and add the annuitized value to the rental expenditures. This defines expenditures on real estate.

[46] The unemployment income parameter b is not separely identified from productivity ℓ. The shifter of the vacancy cost function c_v and the matching function efficiency are not separately identified from the entry cost c_e.

[47] I net out debt payments from value added before computing the labor share. This adjustment is correct under the assumption of a national frictionless non-housing capital market rented from households, with a Cobb-Douglas production function in efficient units of labor, non-housing capital, and housing. In that case, the labor share is 0.85 in my sample.

her job at time t_0 and finds a new job at time t_1, wages satisfy

$$\log W_i(t_1, \ell) = (\lambda - \varphi)(t_1 - t_0) + \Phi(\ell) + \log W_i(t_0, \ell) + \hat{W}_i(\ell) \ , \tag{21}$$

where $\hat{W}_i(\ell)$ is a mean-zero random variable that reflects draws from the local new job distribution, and $\Phi(\ell)$ is a location fixed effect. Because productivity draws are independent from unemployment duration, $\hat{W}_i(\ell)$ does not depend on $t_1 - t_0$. Hence, OLS consistently estimate $\lambda - \varphi$ using equation (21). λ can then be directly obtained from aggregate real wage growth (see Appendix D.5). Thus, I recover φ.

In practice, mechanisms left out from the model may generate endogeneity issues.[48] To address those concerns, Appendix D.5 proposes several other specifications with more flexible controls (for instance, industry fixed effects, worker fixed effects, past wage controls, employed workers as control group). The point estimate of φ remains stable around 1% per quarter and statistically significant across specifications.

Local quality and cutoff. For the remainder of the estimation, I recover estimates of the local job quality $z(\ell)$ and the local productivity cutoff $\underline{y}(\ell)$ in each city. They are endogenous outcomes, not fixed primitives of the economy. Given the estimate for δ, local job losing and finding rates directly identify job quality and the threshold in each city as per Proposition 2,

$$z_c = \frac{\delta}{s_c} \ , \qquad\qquad \underline{y}_c = \frac{b\underline{y}_0}{\hat{\rho}} \frac{\beta f_{Rc} \bar{S}(z_c)}{\hat{\rho} - \beta f_{Rc} \bar{S}(z_c)} \ . \tag{22}$$

where $\hat{\rho} = \rho + \Delta + \mu + \varphi - \lambda$, and $\underline{y}_0$ and the function $\bar{S}$ can be calculated from known parameters.

Lower bound of initial productivity draws $\boldsymbol{Y}$. With an estimate of $\underline{y}_c$ at hand, I use data on job search behavior from the LFS to identify Y. In Appendix D.6, I show how to use this information to recover the contact-to-job probability conditional on meeting. I estimate it to be 20.6%.[49] From the model, the contact-to-job probability in city c is $\left(BY/\underline{y}_c\right)^{1/z_c}$, where B is a known constant. Y is estimated by MDE between the average acceptance probability across locations in the model, and the empirical target of 20.6%.

Housing elasticity $\boldsymbol{\eta}$. At this stage, it is possible to construct demand for housing in each city in the model. Appendix D.8 derives a known function r_0 such that $\log r_c = r_1 + \frac{1}{1+\eta} \log r_0(W_c, L_c, u_c, z_c, \underline{y}_c)$. I then obtain η with OLS, assuming that measurement error is the only residual.[50]

Migration elasticity $\boldsymbol{1/\varepsilon}$. Migration shares by destination $\pi(\ell)$ satisfy

$$\log \pi(\ell) = \pi_0 + \frac{1}{\varepsilon} \log \bar{U}(\ell) + \log a \tag{23}$$

where $\bar{U}(\ell) = \frac{\bar{w}(\ell)}{(1-\beta+\beta H(s(\ell))r(\ell)^{\omega}}$ can now be computed in the model, and π_0 is a general equilibrium constant. Because unobserved amenities a are correlated with $\bar{U}(\ell)$, I split the sample into two subperiods

[48] Permanent differences across workers may correlate with their human capital and productivity draws, or with their probability of finding a new job. Workers coming out of unemployed may draw starting productivities that depend on unemployment duration. The new productivity draws may be correlated with current human capital.

[49] Data reported in Faberman et al. (2017) suggest an acceptance probability of 29.6% in the United States.

[50] Omitted factors like heterogeneous housing supply elasticities may be a source of endogeneity. With repeated cross-sections of housing prices, difference-in-difference specifications using shift-share shocks as instruments could be used to correct for endogeneity. With only one cross-section, these approaches are not possible.

0 and 1 and first-difference equation (23). Then, I use local productivity shocks based on shift-share projections of economy-wide industry shocks as instruments for the change $\log \frac{\bar{U}(\ell_1)}{\bar{U}(\ell_0)}$. I thus estimate $1/\varepsilon$ with Two Stage Least Squares (2SLS) using (23) in first differences. The identification assumption is that economy-wide industry-level shocks are orthogonal to local changes in amenities. I further discuss how to map industry-level shocks into the model and the identification assumption in Appendix D.9.

Non-parametric distributions of local productivity, amenities and job quality. At this stage I need to recover non-parametric estimates of local productivity and amenities (p_c, a_c) in each city, as well as the density function of job qualities f_z. Equation (67) in Appendix D.7 shows that local productivity p_c follows from inverting the model's predictions for local wages. Given the migration elasticity estimate, inverting the population equation (55) in Appendix C.1.4 then delivers an estimate of local amenities a_c in each city. Together, the estimates (p_c, a_c) provide a non-parametric estimate of the distribution $F_{p,a}$.[51] Finally, Appendix D.10 shows that the density function of job losing rates across locations identifies f_z using (22).

Matching function and vacancy cost elasticities $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$. To estimate α and γ, I express local job finding rates as a function of estimated market tightness and employers' values in equation (68) in Appendix D.11, together with more details. I use the same shift-share approach in first differences to estimate α, γ jointly with 2SLS.

Together with the details in Appendix D, the previous arguments prove identification of the 15 parameters that need not be normalized, together with the distributions of fundamentals in the economy. All the previous estimators can be formally collected into an overarching GMM estimator.

Proposition 7. *(Identification)*
When μ, Δ and β are not too large, the parameters $\mu, \Delta, \rho, \omega, \psi, \delta, \sigma, \beta, \lambda, \varphi, \eta, \varepsilon, Y, \alpha, \gamma$, as well as the distribution of firms qualities F_z, the joint distribution of local productivities and amenities $F_{p,a}$, are exactly identified by the GMM estimator. The other parameters can be normalized except the entry cost.

Third group (1 parameter). After estimating those 15 parameters, a numerical search estimates the entry cost c_e by targeting the aggregate unemployment rate.[52] For simulations purposes, I impose parametric functional forms for the distributions, to which I turn now.

Fourth group (7 parameters). I estimate a joint lognormal distribution for local amenities and productivities, with respective standard deviations σ_a, σ_ℓ and correlation $c_{\ell,a}$. I estimate a Beta distribution for the distribution of employer quality. Its shape parameters are g_1, g_2 and its support is $[\underline{z}, \overline{z}]$.

The next section discusses the results from the estimation and reports additional over-identification and validation exercises.

[51] Altenatively, amenities could be obtained as residuals from the migration share equation (23). Because the estimation relies on observed population shares, I choose to match population rather than migration shares. In practice, they are highly correlated.

[52] Given the estimator for Δ, this procedure searches for the entry cost c_e that delivers the right average job finding rate.

4 Results

This section presents the estimation results. First, I discuss the parameter estimates. Second, I show that the model can quantitatively account for spatial unemployment differentials. Third, I conduct a set of over-identification checks to support the model's predictions. Finally, I propose direct evidence to validate the model's core structure.

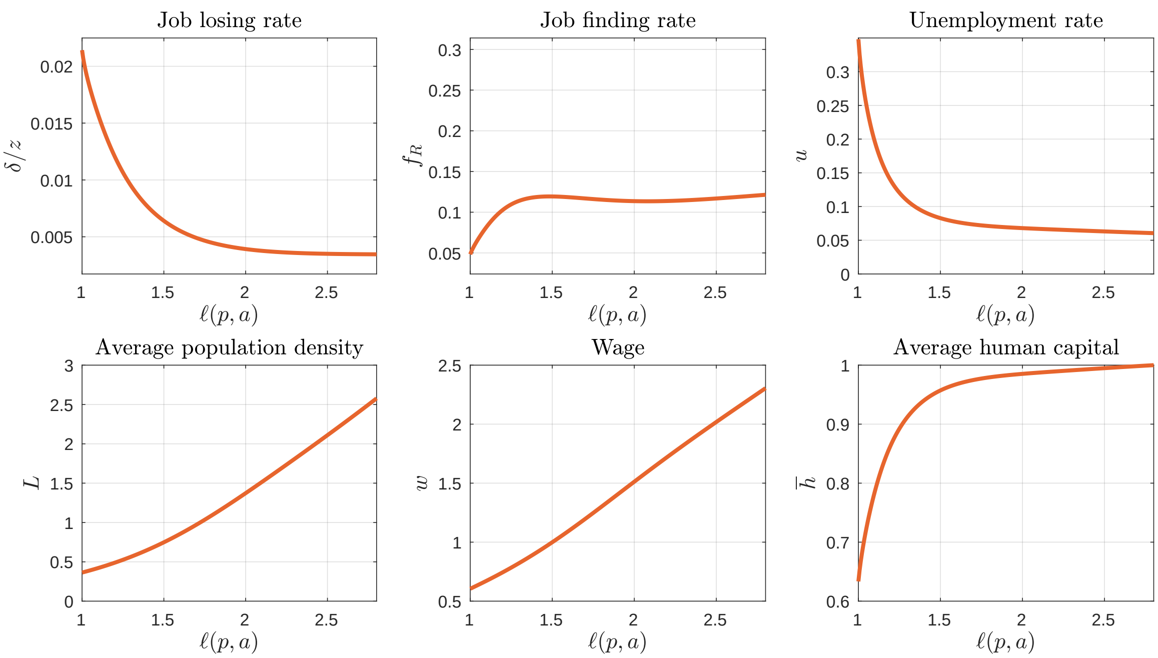
4.1 Parameter estimates

Table 1: Parameter estimates

Parameter	Interpretation	Target	Estimator	Estimate
ρ	Discount rate	Annual interest rate	MDE	0.009
Δ	Labor force exit rate	Aggregate unemployment rate	MDE	0.003
ω	Housing share (workers)	Expenditures on housing	MDE	0.23
ψ	Housing share (firms)	Expenditures on housing	MDE	0.11
δ	Drift of productivity	Job losing rate by tenure	NLLS	0.001
σ	Volatility of productivity	Wage growth by tenure	NLLS	0.004
β	Bargaining power	Labor share	MDE	0.10
Y	Lower bound of init. prod.	Job acceptance probability	MDE	0.88
η	Housing elasticity	Housing prices	OLS	3.49
μ	Migration rate	Migration rate	MDE	0.001
$1/\varepsilon$	Migration elasticity	Migration shares	2SLS	1.65
α	Matching function elasticity	Local job finding rates	2SLS	0.47
γ	Vacancy cost elasticity	Local job finding rates	2SLS	1.31
λ	Learning rate	Aggregate growth	MDE	0.002
φ	Depreciation rate	Unemployment scar	OLS	0.01
$\underline{z}$	Lowest job quality	Local job losing rates	MDE	0.07
$\overline{z}$	Highest job quality	Local job losing rates	MDE	0.41
g_1	Shape of job quality distrib.	Local job losing rates	MDE	1.27
g_2	Shape of job quality distrib.	Local job losing rates	MDE	1.80
σ_p	St.d. of local productivity	Local wages	MDE	0.18
σ_a	St.d. of local amenities	Local population	MDE	0.77
$c_{p,a}$	Correlation prod.–amenities	Local wages and population	MDE	0.17

Table 1 reports the parameter estimates. Overall, they are close to values found in the literature. The housing shares for workers $\omega = 0.23$ is close to the commonly used value of 0.3 for the United States. Similarly, the housing share for firms $\psi = 0.11$ is in the range of estimates reported in Desmet et al. (2018). While the negative drift δ of the worker-level productivity process is close to the quarterly value of 0.5% implied by the estimates in Engbom (2018), the volatility σ is somewhat smaller. The bargaining power $\beta = 0.10$ is close to the estimate in Hagedorn and Manovskii (2008) and references therein. The housing supply elasticity η implies a price-to-population elasticity of 0.22, which is within the range of estimates reported in Saiz (2010) for the United States. The key driver of steady-state population adjustments is the shape parameter of the idiosyncratic preference shock distribution $1/\varepsilon$, which also coincides with the migration elasticity. Its value is 1.65, well within the values reported in the literature between 0.5 and

Figure 5: Model's solution in the decentralized equilibrium.



3.[53] The matching function elasticity α is 0.47, in the middle of the range reported in Petrongolo and Pissarides (2001), and the vacancy cost elasticity parameter γ implies that the cost function is close to quadratic, also in line with existing estimates. Finally, the estimate of the unemployment scar φ implies a 4% wage loss for workers who spent a year unemployed – roughly the average duration of unemployment – relative to workers who remained employed throughout the year. This value is somewhat conservative relative to the value implied by the estimate of 10% in Jarosch (2015).[54]

4.2 Spatial job loss differentials and unemployment

With the estimated model at hand, I start by describing the decentralized equilibrium in Figure 5. Locations are indexed by their local advantage index $\ell(p, a)$ relative to the lowest value thereof.[55]

Consistent with Proposition 6, the job losing rate declines as the local advantage index increases due to rising job quality $z(\ell)$. Because of the opposing forces highlighted in Proposition 2, the finding rate is non-monotonic in the local advantage index in the decentralized equilibrium. Although it declines in the low $\ell(p, a)$ locations, in proportional terms it falls by about half while the job losing rate soars more than four-fold. As a result, the unemployment rate largely follows the declining pattern of the job losing rate.

Tracking the rise in job quality, wages grow as locations become better suited for production. Thus,

[53]The estimate for μ implies an annual migration rate of 0.4%. This is lower than the overall migration rate in my sample which is about 3%. This discrepancy is due to the fact that many migrants are employed workers moving with a job at hand. However, the quantitative results are similar if I target the overall migration rate rather than the migration rate into unemployment. In steady-state, the migration elasticity is the key driver of population movements, not the migration rate.

[54]Jarosch (2015) estimates the long-run effect of an initial job loss, whereas I estimate the elasticity of wage losses to unemployment duration. In the data and in Jarosch (2015)'s model, current job loss begets future job losses, thereby increasing the long-run effect of job loss on human capital relative to my estimate.

[55]The lognormal distribution of productivity and amenities used in the estimation has an unbounded support. For simulation purposes, numerical lower and upper bounds must be chosen. I pick the 0.1% and 99.9% quantiles of the $\ell(p, a)$ distribution. Because the supply of locations left out is then small enough, the equilibrium outcomes are virtually unchanged when widening out those bounds.

Table 2: Aggregate and local unemployment rates in the decentralized equilibrium.

	Model			Data
	Baseline	No pooling	Identical loc.	
Aggregate unemployment rate	0.076	0.071	0.075	0.076
St. dev. unemployment rate	0.022	0.004	0.010	0.023
St. dev. log unemp. / emp.	0.281	0.045	0.158	0.335
Job losing rate	85 %	-180 %	133 %	86 %
Job finding rate	15 %	280 %	-33 %	14 %

average population density $L(\ell)$ also rises. Recall that conditional on the advantage index ℓ, there is residual variation in amenities a across locations. Conditional on ℓ, there are thus locations with higher or lower population density than the average $L(\ell)$. Finally, mirroring the falling unemployment rate as per (18), average human capital $\bar{h}(u(\ell))$ steeply rises across locations. The somewhat conservative estimate of φ still implies human capital gaps over 30% between residents of the best and worst locations, due to the interaction of spatial unemployment differentials and scarring effects of job loss.

I now connect the model's outcomes back to the motivating evidence. The model generates predictions for the spatial variation in local unemployment rates. By non-parametrically recovering the distribution of job quality f_z from local job losing rates, the estimation imposes some restriction on the spatial variation in job losing rates. This moment results from (a) the distribution f_z, but also (b) the equilibrium assignment of heterogeneous employers to locations. In equilibrium, the assignment of employers to locations is an endogenous object. Importantly, it is not constrained by the estimation and may freely adjust in counterfactuals. In addition, the estimation does not restrict the spatial variation in job finding rates apart from the two coefficients in equation (68) that identify α and γ. Therefore, both the spatial variation in the unemployment rate and its split into the job losing and job finding contributions are useful moments to assess the model's ability to speak to spatial unemployment differentials.

The first and fourth columns of Table 2 reveal that the model accounts for over 90% of the cross-sectional variance of the unemployment rates. The standard deviation is 0.022 in the model against 0.023 in the data. Table 2 also highlights that the model closely replicates the contribution of job losing rates to spatial unemployment differentials, which is 85% in the model and 86% in the data. Table 2 finally reports the (targeted) aggregate unemployment rate in the model and the data.

I now assess the relative importance of labor market pooling complementarities versus technological complementarities in shaping spatial unemployment differentials. To that end, I conduct two exercises. First, I shut down the labor market pooling externality in the decentralized equilibrium. This exercise only removes the off-equilibrium-play incentive for employers to choose locations with too high an $\ell(p, a)$, and hence isolates the contribution of technological differences. In the second exericse, I remove technological differences in productivity p and amenities a across locations. The economy then behaves as the limiting one with ex-ante identical locations in Section 2.5, and isolates the role of pooling complementarities.

The second column of Table 2 reveals that, without pooling complementarities, the standard deviation of local unemployment rates falls from 0.022 to 0.004, which represents only 18% of the standard deviation in the baseline model. This drop is accompanied by a complete reversal of the role of the job finding and job losing rates. In contrast, the third column of Table 2 indicates that even without technological differences across locations, the model can generate 45% of the spatial dispersion in unemployment rates. In that case, the model also qualitatively matches the contribution of job losing and finding rates.

Overall, Table 2 shows that pooling complementarities are the primary mechanism generating empirically relevant dispersion in job losing rates across space. Quantitatively, pooling complementarities account for about half the spatial dispersion in unemployment rates. Technological differences contribute one sixth, and the interaction between pooling and technological complementarities contributes a third.

In addition, Table 12 in Appendix E shows that re-estimating the model under alternative assumptions such as the Hosios (1990) condition or directed search yield predictions that are at odds with the data. In particular, none of the alternative assumptions can account jointly for the large cross-sectional variation in unemployment rates and the respective contributions of the job losing and job finding rates in the data.

Having characterized the model's ability to account quantitatively for the margins of spatial unemployment differences, I now propose additional over-identification exercises.

4.3 Over-identification exercises

This subsection proposes a set of over-identifying exercises. The goal is to support the identification of key parameters using non-targeted moments.

Productivity process. I start by discussing three exercises that lend credibility to the estimates of the productivity process δ and σ. First, Figure 6 (a) shows that, despite relying on a single degree of freedom (δ/σ) to predict job losing rates in the first year as per (20), the model closely fits the full cross-sectional variation between cities. Second, I construct job losing rates at all tenures across cities in the model.[56] Figure 6 (b) shows that the model closely accounts for the estimated tenure profiles from Figure 22, both across different tenures and across different cities.

The third exercise offers direct evidence supporting the estimate of the common scale δ. I use balance sheet data to compute firm-level labor productivity growth relative to aggregate labor productivity growth, which should be close to δ. Focusing on large and high labor productivity firms to minimize survival selection bias, I obtain a relative decline of 0.5% annually, close to the annualized estimate of δ, 0.4%.[57]

Bargaining power. To estimate β, I target the aggregate labor share. The model's labor share equation also predicts a negative correlation between wages and local labor shares. This negative correlation mirrors the fact that reservation wages rise less than one-for-one relative to employer productivity. Table 8 in Appendix D.12 shows that the regression coefficient of labor share on local wages is -0.19. In the estimated model, it is -0.11. Despite the simplicity of the bargaining protocol, the model replicates prominent characteristics of rent-sharing across locations.

Amenities. A natural check of the non-parametric amenity estimates a_c is to correlate them with local characteristics that should affect the value of living in a particular location. I regress the estimated log amenities on the log of sun hours per month, as well as a the log density of residential service

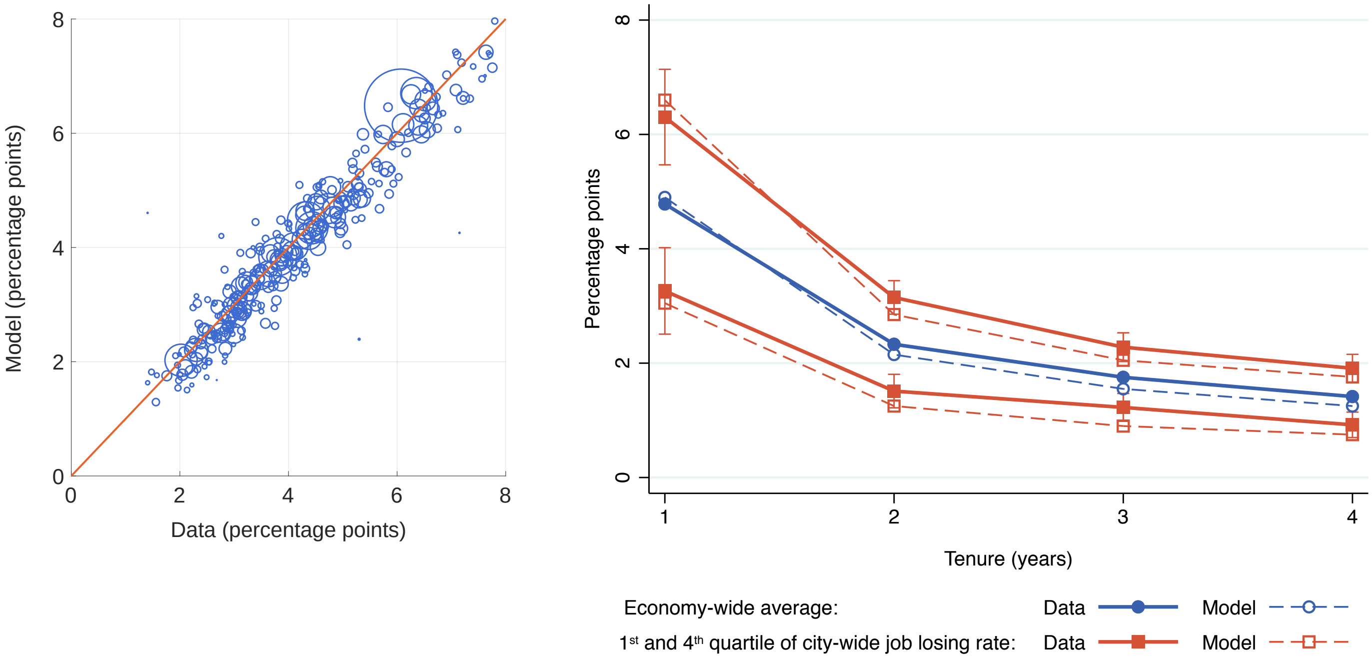
[56] The estimation targets the average job losing rate in each city, and the average job losing rate in the first year economy-wide. The structure of the model then fills in the gaps.

[57] These firms are the least likely to exit in the data. They are also least likely to exit according to theories of firm dynamics with frictional labor markets as Bilal et al. (2019).

Figure 6: Over-identification: job losing rates across cities and tenure

(a) First year job losing rate: model vs data

(b) Job losing rate by tenure and city groups: model vs. data.



establishments of various kinds.[58] Table 9 in Appendix D.12 shows that more sun hours and a higher density of health, education or commercial services are all positively associated with higher amenities.[59] While these results cannot be interpreted as causal, they support the view that the estimated amenities capture salient features of a location's residential attractiveness.

Housing elasticity. To assess how well the estimated housing supply elasticity accounts for cross-sectional dispersion in housing prices, Figure 7 (a) plots housing prices in the model against housing prices in the data. The estimation targets a single moment, the correlation between local house prices and local income. While there is some residual dispersion, the model's predictions are centered around the 45 degree line in orange.

Positive assortative matching. The final over-identification exercise tests the positive assortative matching prediction from Proposition 1. Using the non-parametric estimates of local job quality $z(\ell)$ and the local cutoff $\underline{y}(\ell)$, I check whether locations with higher cutoffs also have higher job quality. While this increasing relationship is not enforced in the estimation, Figure 7 (b) shows that the data supports it.

4.4 Model validation

Having proposed over-identification checks for parameter estimates, I now turn to a last set of validation exercises. The goal of this subsection is to lend credibility to four crucial mechanisms of the model before turning to counterfactuals. To that end, I propose direct evidence using firm-level balance sheet data that

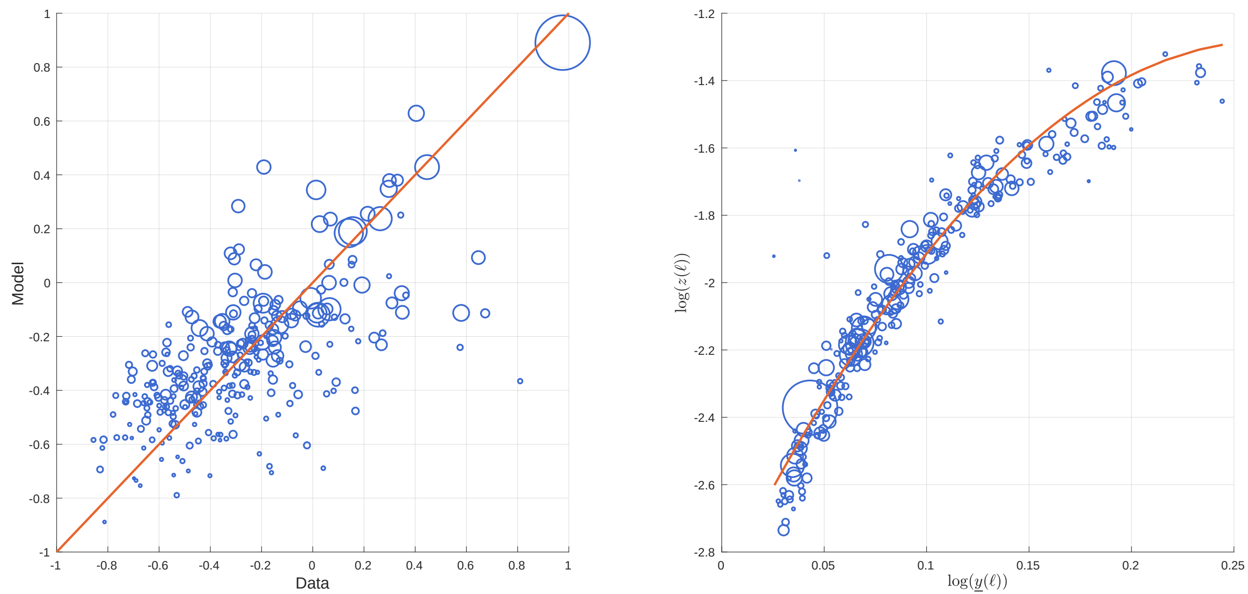
[58]The number of establishments is adjusted by the area of the commuting zone.

[59]For instance, a 10% increase in the number of sun hours per month raises the amenity value of a location by 3.3%. A 10% rise in the density of health establishments increases amenities by 2.2%.

Figure 7: Over-identification: housing prices and positive assortative matching.

(a) Housing prices: model vs data

(b) Estimated job quality against estimated cutoff



is not used in the estimation.

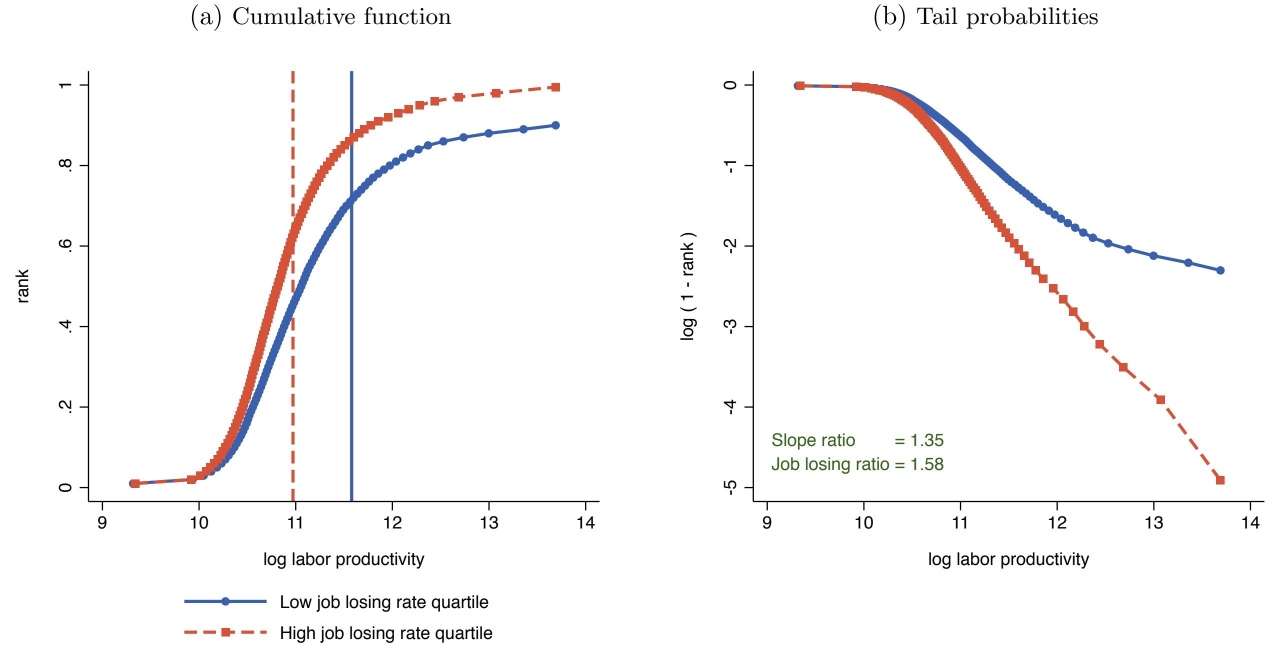
Labor productivity across and within locations. The first validation exercise emphasizes the link between labor productivity differences and job losing rate differentials. Using the solution to the productivity distribution from Lemma 2, the model produces testable implications that tie labor productivity with job losing rates.

1. *Average labor productivity is higher in locations with lower job losing rates.*
2. *The labor productivity distribution in low job losing rate locations first-order stochastically dominates the distribution in high job losing rate locations.*
3. *The labor productivity distribution has a Pareto tail with index $1/z(\ell)$ in each location.*
4. *The ratio of Pareto tails indices between locations is equal to the ratio of job losing rates between the same locations.*

Implications 1 and 2 are not special to the Pareto case, while implications 3 and 4 are closely tied to that particular functional form. To test implications 1 to 4, I compute labor productivity in single-establishment firms in the balance sheet data. I then group commuting zones into four job losing rate quartiles, and compute the employment-weighted labor productivity distribution in each group.

Figure 8 (a) displays the labor productivity distribution in the bottom and top quartiles of commuting zones, ranked by their job losing rate. The vertical lines are the local averages. Consistent with the first implication of the model, average labor productivity is higher in locations with low job losing rates. Furthermore, the cumulative distribution function of labor productivity in low job losing rate locations is always below the cumulative function in high job losing rate locations. Therefore, the data also supports the second, finer implication that the labor productivity distribution first-order stochastically decreases with the job losing rate.

Figure 8: Labor productivity distribution in high and low job losing rate commuting zones.



Panel (b) zooms into the right tail of the productivity distribution by showing the log tail probability as a function of log labor productivity. In both groups of locations, the log tail probability is approximately linear, consistent with the third implication of a Pareto tail. The fourth implication of the model imposes a strong link between the local job losing rate and the shape of the right tail of the labor productivity distribution. I estimate the ratio between the tail indices in each group of locations to be 1.35. It is close to the ratio of group averages of job losing rates, which is 1.58. Together, these results support the structure of the model that ties the heterogeneous productivity of employers to job losing rates, as well as the Pareto assumption.

Labor productivity by employer age. The transparent link between employer productivity and job losing rates obtains as the result of two assumptions in the model. First, the productivity process does not differ across locations. Second, differences in local job losing rates all stem from differences in new job productivity. To support these assumptions, I compute labor productivity in level and growth rate, both for entrant firms (less than two years old), and incumbent firms (at least two years old). All jobs are new at a firm that just entered. While it is unclear what fraction of jobs are new at an incumbent firm, it is arguably less than at an entrant firm. The model then predicts that (a) the negative correlation between labor productivity and job losing rate should be more negative for entrants. The model also predicts that (b) labor productivity growth should not correlate with the job losing rate. Table 10 in the Appendix shows that the data supports both implications (a) and (b). A one percentage point increase in the local job losing rate is associated with 37% lower labor productivity on average across locations. For entrant firms, labor productivity is lower by an additional 43%. In contrast, labor productivity growth rises by an economically and statistically insignificant 0.01.[60]

Given this association between employer's life-cycle and job losing rates, it is natural to ask how much

[60] Any spatial gaps in δ/σ would also introduce a location-specific residual in equation (20). The close fit in Figure 6 indicates that any spatial variation in δ/σ is likely to be small.

spatial differences in establishment-level exit rates contribute to spatial gaps in job losing rates. Recall that the model does not take a stand on the boundaries of the firm, and hence may be interpreted at the firm, establishment or job level. Figure 25 in Appendix D.13 shows that job loss at surviving establishments is the dominant source of geographical variation in the job losing rate. Surviving establishments account for 89% of the cross-location variance in the job losing rate, while exiting establishment contribute only 11% of the cross-location variation. This finding is not particularly surprising given that establishment-level exit rates rarely exceed 10-15%. Thus, locations have different job losing rates primarily because local establishments lay off workers more frequently without exiting.

The role of firm-specific job instability. I now ask how much of spatial gaps in job stability are tied to the firms that open establishments in particular locations, rather than within-firm differences across establishments. To do so, I use firms that open multiple establishments in different cities in the same year.[61] The idea is to compare the job losing rate for new establishments within a given firm to the average job losing rate of their location, both unconditionally and conditionally on the firm's identity. The structure of the model allows to interpret those moments and delivers three predictions. First, a positive correlation between establishment-level job losing rates and local job losing rates should arise – implication (a). To the extent that firms present large differences in their average job quality, conditioning on firms' identity should substantially weaken that correlation, implication (b). The correlation should not disappear if there is residual within-firm job quality dispersion, implication (c).[62]

The main test is thus whether the within-firm correlation between establishment job losing rates is lower than the between-firm correlation. If this is not the case, then it is likely that local factors unrelated to the location choice of heterogeneous employers affect spatial job losing rate differences. Table 11 in Appendix D.13 indicates that the data supports all three implications. It reports conditional correlations of establishment-level job losing rates with the average local job losing rate. First, the R-squared of the regression increases from 0.14 to 0.64 with firm fixed effects, suggesting that there are subtantial differences in job losing rates across firms common to their many establishments. Second, consistent with implication (a), the regression coefficient without firm fixed effects is 1.34 and statistically significant: a one percentage point increase in the average local job losing rate is associated with a 1.34 increase in new establishments' job losing rates in the first year.[63] When including firm fixed effects, the regression coefficient drops from 1.34 to 0.30. This 78% reduction indicates that there are substantial differences in job stability inherent to firms which contribute significantly to spatial job instability gaps, supporting implication (b). The remaining 22% are due to either within-firm dispersion in job quality – implication (c) – or to local factors.

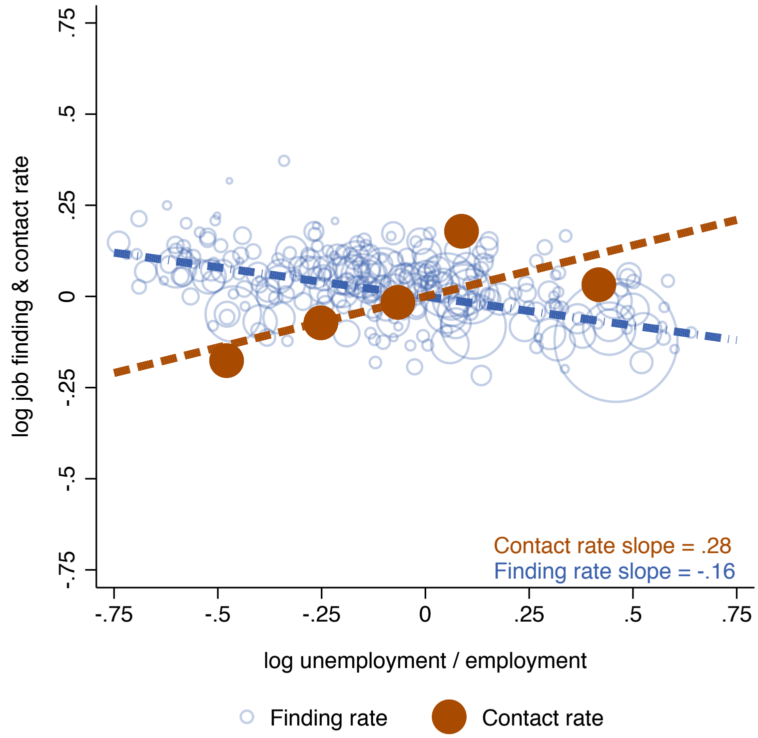
Job finding rate vs. worker contact rate. Having investigated the determinants of job losing rates across space, I conclude this validation section by unpacking the countervailing effects of worker contact rates and contact-to-job probabilities in shaping job finding rates. In the model, sorting arises largely because more productive employers prefer to locate where filling vacancies is easy. Hence, unemployment

[61] I use the DADS *Postes* that have firm and establishment identifiers for the universe of French workers.

[62] Assume that a firm f is a discrete set of heterogeneous jobs j, $\{z_{fj}\}_j$, with mass m_{fj} for each j. The model then predicts that firm f opens at least one establishment per distinct value z_{fj}. Firm f sorts its establishments and jobs across locations: establishments with high z_{fj} jobs locate in low job losing rate locations, while establishments with low z_{fj} jobs locate in high job losing rate locations.

[63] Consistent with Figure 22, the coefficient is larger than 1, reflecting larger job losing rates at new jobs.

Figure 9: Local job finding and worker contact rates against local unemployment-to-employment ratio.



rates should correlate positively with worker contact rates. I test this implication using the French Labor Force Survey. Replicating the approach outlined in Section 3.3, I estimate both worker contact rates and contact-to-job probabilities across locations. I bin commuting zones into 5 unemployment quintiles to reduce small sample measurement error.[64] Figure 9 shows that the worker contact rate is indeed positively correlated with the unemployment rate in the data, the opposite of the job finding rate. Hence, the data indicate that the contact-to-job probability more than offsets the positive correlation between contact rates and unemployment rates.

5 Policy counterfactuals

With the estimated model at hand, this final section presents two policy counterfactuals. The first exercise investigates the general equilibrium welfare gains from both optimal and real-world economy-wide place-based policies that are financed at the federal level. The second exercise studies the effects of a widespread type of discretionary place-based policy financed at the local level. Namely, it studies the effects of subsidies to attract a large, productive plant to a particular location – a "Million Dollar Plant" as in Greenstone et al. (2010)

[64] If anything, increasing the number of bins results in a steeper positive slope for the contact rate.

5.1 The aggregate effects of place-based policies

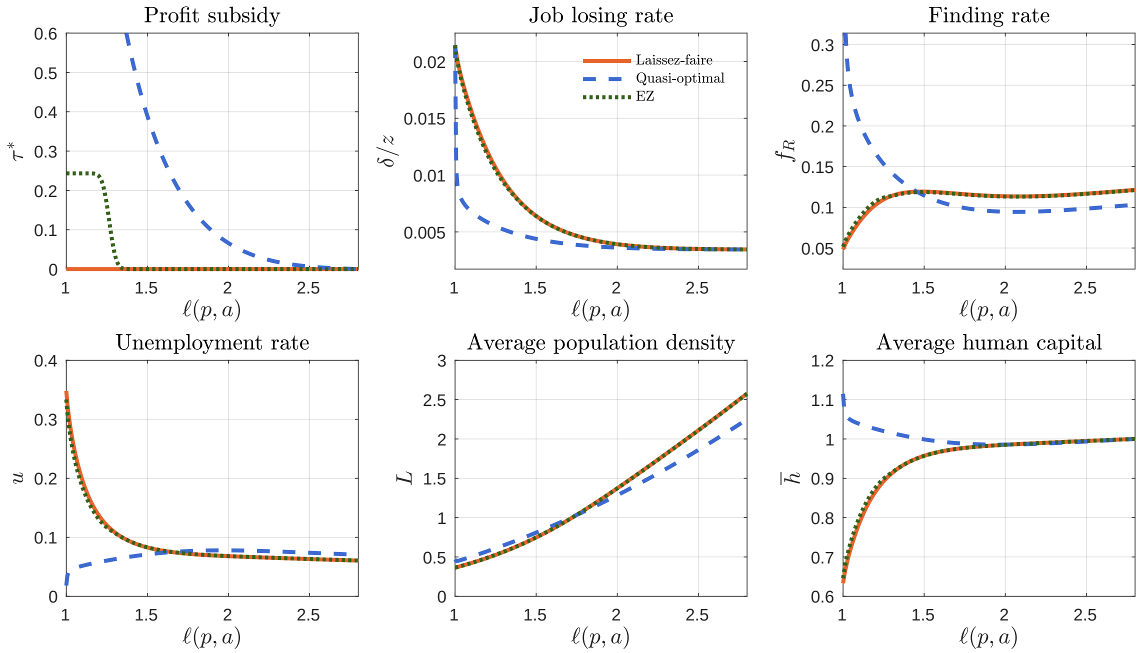
The first policy exercise evaluates the local and aggregate effects of economy-wide place-based policies. As per Proposition 6, the optimal policy subsidizes high job losing rate locations, which also tend to have high unemployment. I focus on the location choice of employers, and start by examining the quasi-optimal policy that corrects the labor market pooling externality. Under this policy, the economy is not fully efficient since the Hosios (1990) condition needs not hold. The quasi-optimal policy takes the form of a profit subsidy, financed with a non-distortionary tax. To compute welfare gains without taking a stand on distributional issues between owners and workers, I use the alternative, equivalent formulation in which profits and rents are redistributed to workers with a non-distortionary flat earnings subsidy, while the policy is subsidized with a flat earnings tax.

I contrast the effects of the quasi-optimal policy with a real-world example of an economy-wide set of place-based policies. Federal programs such as the Empowerment Zones program in the United States proposed considerable tax breaks for firms opening jobs in high unemployment areas. In France, a similar Enterprise Zones (EZ) program was rolled out in 1996 and subsequently expanded. The labor market pooling externality provides a theoretical basis for such policies. By changing employers' incentives to open jobs in various locations, the policies effectively relocate jobs across space and affect the general equilibrium of the economy. Ideally, the structural estimation would account for the policy during the sample period. However, reliable estimates of local policy expenditures are hard to obtain, making it difficult to net out the effects of the EZ policy in the estimation. In practice, the policy is small and has modest local and general equilibrium effects. Therefore, it is unlikely to substantially affect the parameter estimates and the counterfactuals.

Figure 10 displays the cross-sectional patterns of the equilibrium under the quasi-optimal policy and under a budget-equivalent version of the French EZ program. Figure 10 reveals that the EZ program subsidy is much smaller than the quasi-optimal one. However, it shares the same qualitative pattern: to incentivize high productivity employers to open jobs in high unemployment locations. Because of its smaller magnitude, it induces only few jobs to change location and has minor effects on the job losing rate. In contrast, the quasi-optimal policy massively relocates productive jobs towards initially high unemployment locations. This results in large drops in local unemployment rates at the bottom of the geography spectrum, that can exceed 10 percentage points. Consequently, spatial unemployment differentials plummet. Finally, local unemployment reductions are mirrored in substantial human capital gains for residents.

Figure 11 depicts the welfare gains for residents in all locations. Locations are ordered by their unemployment rate in the laissez-faire equilibrium, and grouped into population-weighted quantiles to reflect how many workers experience a given welfare increase. Figure 11 reveals that the quasi-optimal policy achieves large welfare gains in initially high unemployment locations. Appendix C.1.9 derives an exact welfare decomposition in the model, which corresponds to the different colored areas in Figure 11. The blue area shows that direct gains to the average resident unemployed worker steadily rise with pre-policy local unemployment, and exceed 10% in the most distressed areas. Importantly, in steady-state, unemployed workers all have the same expected utility due to compensating differentials in the form of preference shocks. Therefore, the blue area that describes the welfare gains to unemployed workers

Figure 10: Model's solution in the decentralized equilibrium, the quasi-optimal policy and the French EZ program.



conditional on a given preference shock $\varsigma = 1$ and conditional on human capital $k = 1$.[65] The blue area also correponds to each location's contribution to the aggregate welfare gains for unemployed workers in the economy. The green area represents the additional welfare gains to the average employed worker. Thus, the welfare gains to an employed worker with $k = 1$ is the sum of the blue and the green area. Under the quasi-optimal policy, employed workers gain slightly more that unemployed workers. Finally, human capital accumulation in orange benefits both unemployed and employed workers. It amplifies welfare gains more than two-fold in originally high unemployment locations. Because the quasi-optimal policy relocates jobs away from the best locations, residents there experience welfare losses. In contrast, the EZ program has more modest effects, with welfare gains peaking around 3% and concentrated in the targeted high unemployment areas.

To highlight the spatial distribution of these local welfare gains, Figure 12 (a) maps the gains from the quasi-optimal policy across all French commuting zones. Because welfare gains are strongly correlated with the local unemployment rate, the southern Mediterranean coast benefits most. In suburban areas close to Paris, several high unemployment commuting zones also benefit substantially. Figure 12 (b) shows that local welfare gains are accompanied by substantial TFP improvements. Finally, note that residential amenities a are fixed in counterfactuals. To the extent that other factors such as local crime rates or locally provided public services respond to lower unemployment rates, the welfare gains in the model are likely to be a lower bound relative to welfare gains inclusive of such adjustments.

Leveraging the structure of the model, I aggregate the local welfare gains and compute the aggregate welfare gains from the quasi-optimal policy and the EZ program. Table 3 first highlights that the quasi-optimal policy removes the pooling externality and reduces spatial unemployment differentials five-fold as in Table 2. This change follows a large relocation of high productivity jobs towards poorer locations.

[65] Alternatively, the blue are is equal to the steady-state welfare gains of an unemployed worker who never received the moving opportunity and so stayed in the same location, with $k = 1$.

Figure 11: Welfare gains from the quasi-optimal policy and the French EZ program.

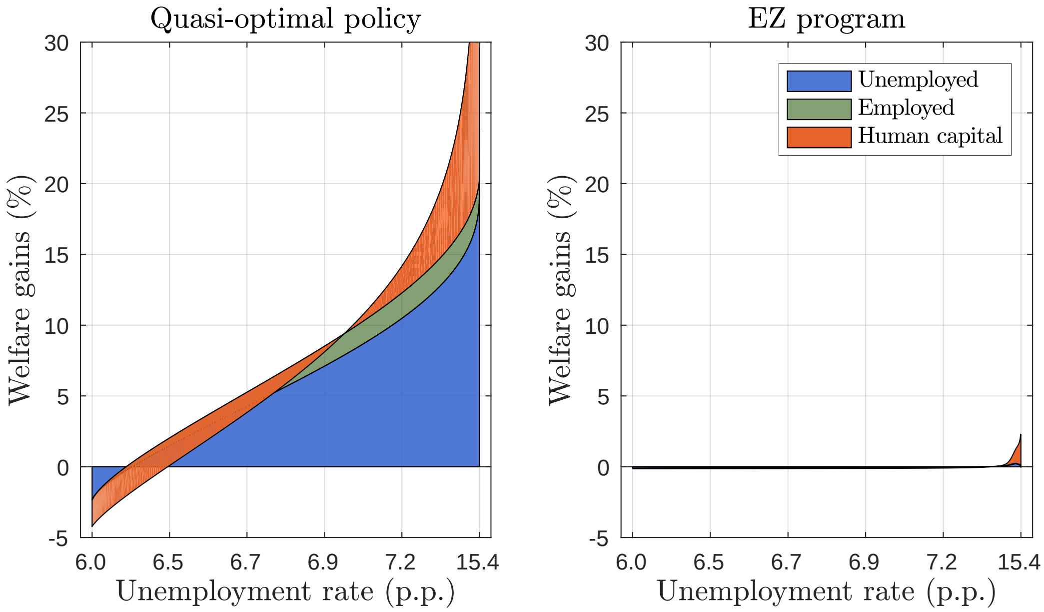


Figure 12: Local gains from the quasi-optimal policy

(a) Welfare (%)

(b) TFP (%)

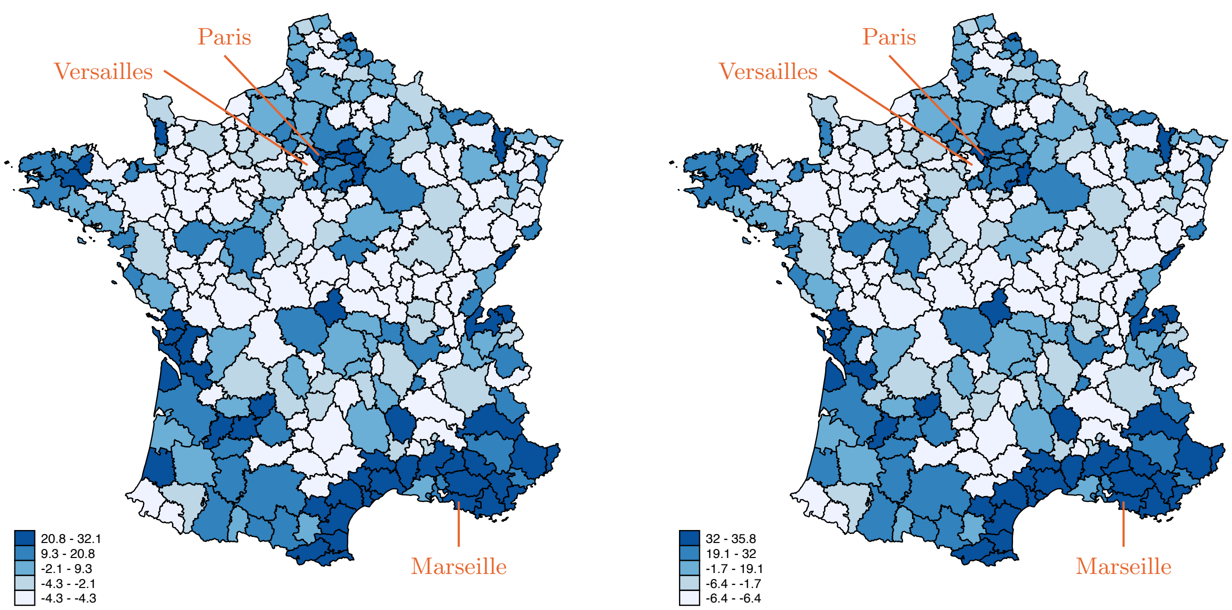


Table 3: Aggregate gains from place-based-policies

	Laissez-faire	EZ program	Quasi-optimal
Aggregate unemployment rate	0.076	0.076	0.071
St. dev. unemployment rate	0.022	0.020	0.004
Aggregate welfare gains (%)		0.23	5.08
Unemployed		0.09	2.93
Employed		0.01	1.05
Human capital		0.13	1.10
Redistribution (% of GDP)		0.04	5.05

As the aggregate efficiency of the economy rises, the aggregate unemployment rate falls by 0.5 percentage points, a 7% drop. The quasi-optimal policy achieves over 5% aggregate welfare gains. Overall, the quasi-optimal policy largely ameliorates opportunities for unemployed workers. They become shielded against the risk of remaining trapped in distressed labor markets.

Second, Table 3 reveals that despite its relatively small size, the EZ program reduced spatial unemployment differentials by 10%. While it had virtually no impact on the aggregate unemployment rate, it raised aggregate welfare by 0.23%. Most of these gains stem from better human capital accumulation in high unemployment areas. Interestingly, the relative contribution of human capital accumulation is lower under the quasi-optimal policy than under the EZ program. Because it is smaller, the EZ program harmlessly punctures high productivity jobs from the entire economy and helps residents of high unemployment locations. The quasi-optimal policy involves more drastic relocations. Therefore, human capital gains in poorer locations are partly offset by losses in richer locations.

It is not surprising that the EZ policy delivers smaller gains than the quasi-optimal policy. The EZ policy consists in a much smaller subsidy scheme as shown in Figure 10. Aggregate expenditures on the EZ policy represent redistributing 0.04% of Gross Domestic Product (GDP). Expenditures under the quasi-optimal policy are over 100 times larger. If scaling up the redistribution-efficiency ratio of the EZ policy was possible, welfare would rise by 5.75% for every percent of GDP redistributed. The redistribution-efficiency ratio of the quasi-optimal policy is close to 1, indicating that decreasing returns rapidly kick in. Indeed, one should expect the planner's problem to be concave in the profit subsidy around the quasi-optimal policy. Thus, the largest gains for a marginal increase in the profit subsidy should arise close to the laissez-faire.

5.2 The local effects of a "Million Dollar Plant"

The previous analysis focused on federally funded place-based policies. This last section explores the effects of locally funded place-based policies. Specifically, I examine the local employment and welfare effects of attracting a large, productive plant – a "Million Dollar Plant" (MDP). As documented by Greenstone et al. (2010) and Slattery (2019), local governments allocate considerable resources to discretionary subsidies to firms that seek a location for a new plant. Local governments are largely motivated by potential

Table 4: Employment and welfare gains from the calibrated MDP.

Moment	Greenstone et al.	Model
MDP employment share (%)	1.70	1.78
MDP output share (%)	2.00	2.00
TFP gains at other employers (%)	1.80	1.91
Job multiplier		-0.55
Unemp. rate change (p.p.)		-0.21
Job losing		-119%
Job finding		+19%
Population change (%)		0.58
Welfare gains (%)		1.15

employment expansions, both due to the direct jobs created by the MDP, but also due to potential spillover effects. These two channels are sometimes called direct and indirect job creation in this context.

From the perspective of the theory, attracting a large productive plant with subsidies resembles the optimal policy from Proposition 5, although in the MDP case the policy is locally funded. To quantify the potential employment and welfare gains from such a policy, I identify a MDP as a mass point $\bar{m}$ of high productivity jobs $\bar{z}$ in the model. I start with a simple positive exercise that abstracts from the cost of attracting a MDP. To that end, I select a location with a 10% unemployment rate, I target the MDP's employment and output share in the location.[66] Then, I compare the steady-state outcomes in the location before and after the MDP.[67]

Table 4 displays the results. Interestingly, the model produces Total Factor Productivity (TFP) gains at employers other than the MDP that are close to those estimated by Greenstone et al. (2010). Although the model abstracts from technological spillovers, TFP gains arise because of labor market pooling complementarities. The MDP improves the local job quality mix. Workers throughout the economy in-migrate to benefit from improved earnings prospects, leading to a half percent increase in population. The labor market pooling complementarity then kicks in as hiring conditions improve, thereby attracting better jobs even beyond the MDP. While technological spillovers are a common interpretation for the effects of a MDP, the model indicates that labor market pooling complementarities alone can also rationalize the ensuing TFP gains.

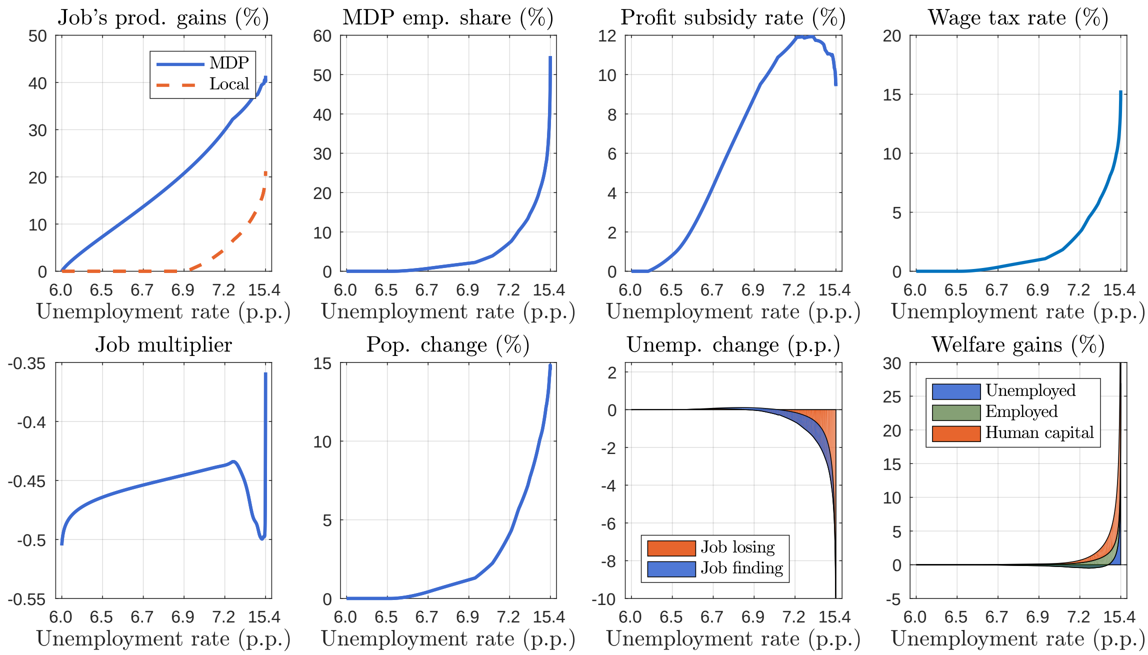
Together, the MDP and the better other employers contribute to a 0.2 percentage point reduction in the local unemployment rate. The drop in the job losing rate is the primary driver of this reduction, as jobs at the MDP and at other employers are more stable. As a result, gross welfare of local residents rises by 1.15%. Despite these welfare gains, the calibrated MDP experiment brings in few jobs above and beyond what the MDP directly contributes. Every new MDP job displaces on average half a pre-existing job, in line with existing evidence.[68]

[66] Their estimates are at the county level, while the model is at the commuting zone level. I map their estimates into commuting zone-level targets by deflating their estimates by the average number of counties per commuting zone in the United States.

[67] While local transitional dynamics can be computed, a full discussion thereof is beyond the scope of this paper.

[68] Slattery and Zidar (2019) find little support for systematic indirect job creation effects of MDPs. One caveat is that

Figure 13: Local gains from the optimal "Million Dollar Plant"



The exercise so far abstracted from the cost of attracting the MDP. In practice, these plants enjoy considerable subsidies often financed with local taxes rather than an economy-wide tax as in Section 5.1. As pointed out by Slattery and Zidar (2019), measuring the exact collection of subsidies that MDPs receive is challenging. To circumvent this difficulty, I discipline the cost of a MDP inside the model. For each location, I compute the optimal MDP subsidy that maximizes the welfare of its residents. To do so, I assume that every local government believes it acts in isolation – a complete evaluation of between-location tax competition is beyond the scope of this paper. The MDP benefits from a profit subsidy, which is financed through a local earnings tax.

Figure 13 displays the local effects from the optimal MDP in the cross-section of locations ordered by pre-policy unemployment rate. In line with evidence from Slattery (2019), the model indicates that high unemployment locations spend more on subsidies to MDPs. These locations offer higher profit subsidies, sometimes over 10%. They optimally attract more productive MDPs which can take over as much as 50% of local employment, in turn bringing in more productive jobs through labor market pooling complementarities. Local unemployment drops by 10 percentage points in the most distressed areas, largely driven by reductions in the job losing rate. On net, welfare rises by over 10% there. However, optimal discretionary spending comes at a subtantial fiscal cost that can exceed a 10% earnings tax. Given political economy or credit constraints, local government may not be able to take advantage of the large gains depicted in Figure 13.

the model does feature neither input-output linkages with trade costs nor technogological spillovers. Both may in principle amplify the indirect job creation effects, although the calibrated MDP already matches estimated TFP gains.

Conclusion

This paper proposes an alternative view of spatial unemployment differentials. I have shown that high localized unemployment arises because workers repeatedly lose their job, not because finding a job is particularly hard. Differences in job losing rates emerge as employers with unstable jobs self-select into similar locations, while employers with stable jobs locate in others. I have developed a theory in which labor market pooling complementarities are the central driver of the location choice of heterogeneous employers. As a result, employers with stable jobs over-value locating close to each other due to labor market pooling externalities. This view implies that redistributing from low unemployment locations towards high unemployment locations is welfare improving.

Of course, the idea that pooling complementarities result in too much concentration in the best options available to workers and employers is more general than the particular spatial context put forward in this paper. For instance, investigating the implications of pooling externalities for the allocation of workers and employers across occupations and industries could lead to interesting policy insights. Indeed, in the spatial context alone, pooling externalities quantitatively account for the lion's share of differences in unemployment across locations.

Consequently, the view of this paper emphasizes that spatial unemployment differentials are not an immutable characteristic of the economic landscape. Instead, place-based policies have the potential to drastically reshape the spatial distribution of unemployment, and ameliorate employment prospects at the aggregate level. While a long tradition of research has found that agglomeration economies call for taxes on poor locations, the implications thereof have remained at odds with a wide range of real-world spatial policies. The view that labor market pooling externalities lie at the heart of the location decisions of employers helps reconcile theory with policymakers' intuition that incentivizing businesses to open in distressed areas may help rather than harm individuals. Yet, the inherently local nature of many economic interactions gives rise to many other externalities. Therefore, any policy recommendation should account for as many sources of agglomeration and congestion as possible. As individuals who grow up and live and different places seem to face increasingly divergent economic opportunities, place-based policies appear more relevant than ever.

References

Abowd, John M., Kramarz, Francis, and Margolis, David N. (1999). "High Wage Workers and High Wage Firms". *Econometrica* 67.2, pp. 251–333.

Acemoglu, Daron (2001). "Good Jobs versus Bad Jobs". *Journal of Labor Economics* 19.1, pp. 1–21.

Amior, Michael and Manning, Alan (2018). "The Persistence of Local Joblessness". *American Economic Review* 108.7, pp. 1942–70.

Austin, Benjamin A, Glaeser, Edward L, and Summers, Lawrence H (2018). *Jobs for the Heartland: Place-Based Policies in 21st Century America*. Working Paper 24548. National Bureau of Economic Research.

Bartik, Timothy (1991). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute for Employment Research.

Bilal, Adrien and Rossi-Hansberg, Esteban (2018). *Location as an Asset*. Working Paper 24867. National Bureau of Economic Research.

Bilal, Adrien G., Engbom, Niklas, Mongey, Simon, and Violante, Giovanni L. (2019). "Firm and Worker Dynamics in a Frictional Labor Market". *Working Paper*.

Blanchard, Olivier Jean and Katz, Lawrence F. (1992). "Regional Evolutions". *Brookings Papers on Economic Activity* 23.1, pp. 1–76.

Blundell, Richard, Pistaferri, Luigi, and Preston, Ian (2008). "Consumption Inequality and Partial Insurance". *American Economic Review* 98.5, pp. 1887–1921.

Bonhomme, Stphane, Lamadon, Thibaut, and Manresa, Elena (2019). "A Distributional Framework for Matched Employer Employee Data". *Econometrica* 87.3, pp. 699–739.

Borovičková, Katarína and Shimer, Robert (2017). *High Wage Workers Work for High Wage Firms*. Working Paper 24074. National Bureau of Economic Research.

Brancaccio, Giula, Kalouptsidi, Myrto, Rosaia, Nicola, and Papageorgiou, Theodore (2019). *Efficiency in Decentralized Transport Markets*. Working Paper.

Cahuc, Pierre, Postel-Vinay, Fabien, and Robin, Jean-Marc (2006). "Wage Bargaining with On-the-Job Search: Theory and Evidence". *Econometrica* 74.2, pp. 323–364.

Caliendo, Lorenzo, Opromolla, Luca David, Parro, Fernando, and Sforza, Alessandro (2017). *Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement*. Working Paper 23695. National Bureau of Economic Research.

Card, David, Heining, Jrg, and Kline, Patrick (2013). "Workplace Heterogeneity and the Rise of West German Wage Inequality". *The Quarterly Journal of Economics* 128.3, pp. 967–1015.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, Puga, Diego, and Roux, Sbastien (2012). "The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection". *Econometrica* 80.6, pp. 2543–2594.

Couture, Victor, Gaubert, Cecile, Handbury, Jessie, and Hurst, Erik (2019). *Income Growth and the Distributional Effects of Urban Spatial Sorting*. Working Paper 26142. National Bureau of Economic Research.

Şahin, Ayşegül, Song, Joseph, Topa, Giorgio, and Violante, Giovanni L. (2014). "Mismatch Unemployment". *American Economic Review* 104.11, pp. 3529–64.

Davis, Donald R. and Dingel, Jonathan I. (2014). *The Comparative Advantage of Cities*. NBER Working Papers 20602. National Bureau of Economic Research, Inc.

Desmet, Klaus and Rossi-Hansberg, Esteban (2013). "Urban Accounting and Welfare". *American Economic Review* 103.6, pp. 2296–2327.

Desmet, Klaus, Nagy, David, and Rossi-Hansberg, Esteban (2018). "The Geography of Development". *Journal of Political Economy*.

Diamond, Peter A. (1982). "Wage Determination and Efficiency in Search Equilibrium". *Review of Economic Studies* 49.2, pp. 217–227.

Diamond, Rebecca (2016). "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000". *American Economic Review* 106.3, pp. 479–524.

Eeckhout, Jan and Kircher, Philipp (2018). "Assortative Matching With Large Firms". *Econometrica* 86.1, pp. 85–132.

Elsby, Michael W. L. and Michaels, Ryan (2013). "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows". *American Economic Journal: Macroeconomics* 5.1, pp. 1–48.

Engbom, Niklas (2018). *Firm and Worker Dynamics in an Aging Labor Market*. Tech. rep.

Faberman, R. Jason, Mueller, Andreas I, ahin, Ayegl, and Topa, Giorgio (2017). *Job Search Behavior among the Employed and Non-Employed*. Working Paper 23731. National Bureau of Economic Research.

Fajgelbaum, Pablo and Gaubert, Cecile (2018). *Optimal Spatial Policies, Geography and Sorting*. Working Paper 24632. National Bureau of Economic Research.

Fajgelbaum, Pablo D, Morales, Eduardo, Surez Serrato, Juan Carlos, and Zidar, Owen (2018). "State Taxes and Spatial Misallocation". *The Review of Economic Studies* 86.1, pp. 333–376.

Fujita, Shigeru and Ramey, Garey (2009). "The Cyclicality of Separation and Finding Rates". *International Economic Review* 50.2, pp. 415–430.

Galichon, Alfred (2016). *Optimal Transport Methods in Economics*. Princeton University Press. ISBN: 9780691172767.

Gaubert, Cecile (2018). "Firm Sorting and Agglomeration". *American Economic Review* 108.11, pp. 3117–53.

Giannone, Elisa (2017). *Skill-Biased Technical Change and Regional Convergence*. 2017 Meeting Papers 190. Society for Economic Dynamics.

Glaeser, Edward L and Gottlieb, Joshua D (2008). *The Economics of Place-Making Policies*. Working Paper 14373. National Bureau of Economic Research.

Glaeser, Edward L, Kim, Hyunjin, and Luca, Michael (2018). *Measuring Gentrification: Using Yelp Data to Quantify Neighborhood Change*. Working Paper 24952. National Bureau of Economic Research.

Greenstone, Michael, Hornbeck, Richard, and Moretti, Enrico (2010). "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings". *Journal of Political Economy* 118.3, pp. 536–598.

Grossman, Gene M. and Rossi-Hansberg, Esteban (2012). "Task Trade Between Similar Countries". *Econometrica* 80.2, pp. 593–629.

Hagedorn, Marcus and Manovskii, Iourii (2008). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited". *American Economic Review* 98.4, pp. 1692–1706.

Hall, Robert E. (2005). "Employment Fluctuations with Equilibrium Wage Stickiness". *American Economic Review* 95.1, pp. 50–65.

Hanson, Andrew (2009). "Local employment, poverty, and property value effects of geographically-targeted tax incentives: An instrumental variables approach". *Regional Science and Urban Economics* 39.6, pp. 721–731.

Hopenhayn, Hugo and Rogerson, Richard (1993). "Job Turnover and Policy Evaluation: A General Equilibrium Analysis". *Journal of Political Economy* 101.5, pp. 915–938.

Hosios, Arthur J. (1990). "On the Efficiency of Matching and Related Models of Search and Unemployment". *The Review of Economic Studies* 57.2, pp. 279–298.

Jarosch, Gregor (2015). "Searching for Job Security and the Consequences of Job Loss". *Working Paper.*

Kennan, John and Walker, James R. (2011). "The Effect of Expected Income on Individual Migration Decisions". *Econometrica* 79.1, pp. 211–251.

Kline, Patrick and Moretti, Enrico (2013). "Place Based Policies with Unemployment". *American Economic Review* 103.3, pp. 238–43.

Kline, Patrick, Saggio, Raffaele, and Solvsten, Mikkel (2019). *Leave-out Estimation of Variance Components*. Working Paper 26244. National Bureau of Economic Research.

Krusell, Per, Mukoyama, Toshihiko, Rogerson, Richard, and ahin, Ayegl (2017). "Gross Worker Flows over the Business Cycle". *American Economic Review* 107.11, pp. 3447–76.

Luenberger, David G. (1997). *Optimization by Vector Space Methods*. 1st. New York, NY, USA: John Wiley & Sons, Inc. ISBN: 047118117X.

Luttmer, Erzo G. J. (2007). "Selection, Growth, and the Size Distribution of Firms*". *The Quarterly Journal of Economics* 122.3, pp. 1103–1144.

Mangin, Sephorah and Julien, Benoit (2018). "Efficiency in Search and Matching Models: A Generalized Hosios Condition". *Becker Friedman Institute for Research in Economics Working Paper No. 2018-10.*

Marinescu, Ioana and Rathelot, Roland (2018). "Mismatch Unemployment and the Geography of Job Search". *American Economic Journal: Macroeconomics* 10.3, pp. 42–70.

Marshall, Alfred (1920). *Principles of Economics*. London: MacMillan.

Martellini, Paolo and Menzio, Guido (2018). *Declining Search Frictions, Unemployment and Growth*. NBER Working Papers 24518. National Bureau of Economic Research, Inc.

Mayer, Thierry, Mayneris, Florian, and Py, Loriane (2015). "The impact of Urban Enterprise Zones on establishment location decisions and labor market outcomes: evidence from France". *Journal of Economic Geography* 17.4, pp. 709–752.

Moll, Benjamin and Nuño, Galo (2018). "Social Optima in Economies with Heterogeneous Agents". *Review of Economic Dynamics* 28, pp. 150–180.

Mortensen, Dale T. (1982). "The Matching Process as a Noncooperative Bargaining Game". In: *The Economics of Information and Uncertainty*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 233–258.

Mortensen, Dale T. and Pissarides, Christopher A. (1994). "Job Creation and Job Destruction in the Theory of Unemployment". *The Review of Economic Studies* 61.3, pp. 397–415.

Neumark, David and Simpson, Helen (2014). *Place-Based Policies*. Working Paper 20049. National Bureau of Economic Research.

OECD (2005). "How Persistent Are Regional Disparities in Employment? The Role of Geographic Mobility". *OECD Employment Outlook 2005 Edition*.

Oksendal, Bernt (1992). *Stochastic Differential Equations (3rd Ed.): An Introduction with Applications*. Berlin, Heidelberg: Springer-Verlag. ISBN: 3-387-53335-4.

Osborne, Martin J. and Rubinstein, Ariel (1994). *A Course in Game Theory*. Vol. 1. MIT Press Books 0262650401. The MIT Press. ISBN: ARRAY(0x448056b8).

Ossa, Ralph (2017). *A Quantitative Analysis of Subsidy Competition in the U.S.* Working Paper 20975. National Bureau of Economic Research.

Petrongolo, Barbara and Pissarides, Christopher A. (2001). "Looking into the Black Box: A Survey of the Matching Function". *Journal of Economic Literature* 39.2, pp. 390–431.

Pham, Huyn (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications*. 1st. Springer Publishing Company, Incorporated. ISBN: 3540894993, 9783540894995.

Pissarides, Christopher A. (1985). "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages". *The American Economic Review* 75.4, pp. 676–690.

Postel-Vinay, Fabien and Turon, Hlne (2010). "On-the-Job Search, Productivity Shocks, and the Individual Earnings Process". *International Economic Review* 51.3, pp. 599–629.

Roback, Jennifer (1982). "Wages, Rents, and the Quality of Life". *Journal of Political Economy* 90.6, pp. 1257–1278.

Rubinstein, Ariel (1982). "Perfect Equilibrium in a Bargaining Model". *Econometrica* 50.1, pp. 97–109.

Saiz, Albert (2010). "The Geographic Determinants of Housing Supply". *The Quarterly Journal of Economics* 125.3, pp. 1253–1296.

Sattinger, Michael (1993). "Assignment Models of the Distribution of Earnings". *Journal of Economic Literature* 31.2, pp. 831–880.

Schaal, Edouard (2017). "Uncertainty and Unemployment". *Econometrica* 85.6, pp. 1675–1721.

Shimer, Robert (2005). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies". *American Economic Review* 95.1, pp. 25–49.

Shimer, Robert and Smith, Lones (2000). "Assortative Matching and Search". *Econometrica* 68.2, pp. 343–369.

Slattery, Cailin and Zidar, Owen (2019). "Evaluating State and Local Business Tax Incentives". *Journal of Economic Perspectives (in preparation).*

Slattery, Calin (2019). "Bidding for Firms: Subsidy Competition in the U.S". *Working Paper.*

Tao, Terence (2006). *Nonlinear Dispersive Equations: Local and Global Analysis.* American Mathematical Society and Conference Board of the Mathematical Sciences.

Topkis, Donald M. (1998). *Supermodularity and Complementarity.* Princeton University Press. ISBN: 9780691032443.

Villani, Cedric (2003). *Topics in Optimal Transportation Theory.* Vol. 58. American Mathematical Society.

ONLINE APPENDIX

This Appendix is organized as follows. Section A provides additional details on the data, sample construction and descriptive evidence. Section B contains proofs for the baseline model. Section C contains proofs for the quantitative framework, as well as a discussion of additional extensions. Section D provides details for the estimation and identification proof, as well as validation exercises.

A Data and descriptive evidence

A.1 Data

DADS panel. The central dataset is the 4% sample of the DADS panel, between 1993 and 2007. Once a worker enters the dataset in any year after 1976, all her subsequent employment spells are recorded. The dataset provides start and end days of each employment spell, the job's wage, the residence and workplace zipcodes of the individual, four-digit occupation and industry, as well as establishment and firm tax identifiers that can be linked to administrative balance-sheet data.

In addition to the sample restrictions described in the main text, I exclude from the sample individuals during the first year that they appear in it. This restriction ensures that aggregate fluctuations in non-employment are not driven by higher entry in the sample in a particular year, given that individuals are first observed when they have a job. I also drop individuals from the sample two years after their last job. I keep only the years after 1997 because the entry in the panel is noisier in the initial years 1993-1996. I stop in 2007 to avoid both an important classification changes in 2008 and the Great Recession in 2009.

DADS cross-section. The DADS *Postes*, are used by the French statistical institute to construct the DADS *Panel*. They cover the universe of French workers, but in the version available to researchers, worker identifiers are reshuffled every two years. The DADS Postes allow to compute employment, wages, occupational mix as well as exit rates and job losing rates for the near universe of French establishments, which can be located at the zipcode level.

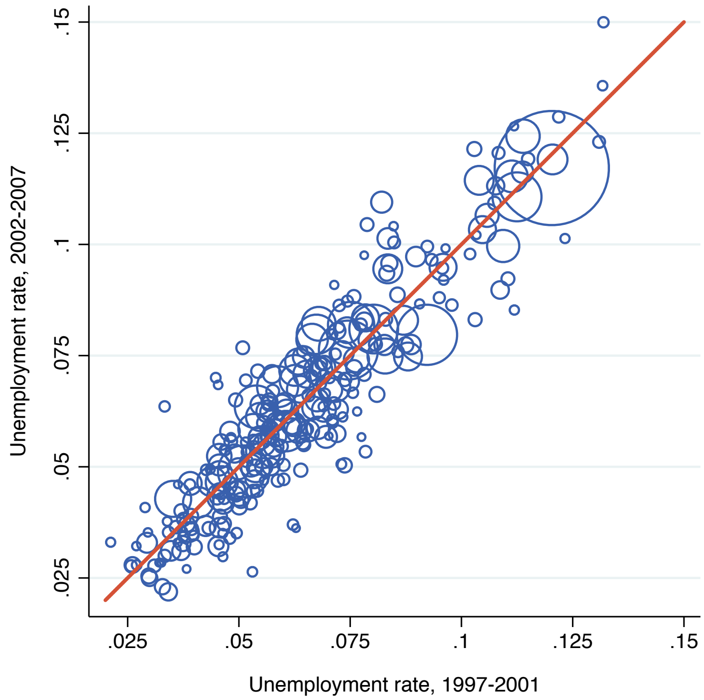
LFS. I complement the DADS panel with the LFS. I use the LFS starting in 2003 due to a large survey change in 2002. The LFS is quarterly and tracks individuals for six consecutive quarters. The LFS reports whether an individual is working, unemployed or out of the labor force. As in many surveys, the LFS drops individuals if they move between quarters, which is why the DADS panel is particularly useful. I apply the same demographics restrictions as in the DADS panel. I use the LFS to discriminate between unemployment and non-employment in the DADS panel. To that end, I estimate cell-level quarterly transition probabilities between employment, unemployment and non-participation in the LFS. A cell is an occupation and age group - city group bin. Occupation and ages are binned into 4 groups based on their average wage. Similarly, cities are binned into 4 groups based on their unemployment rate. With the estimated transition probabilities at hand, I probabalistically impute the non-participation vs. unemployment status of individuals in the DADS panel. Table 5 shows that the DADS panel and the LFS have similar aggregate statistics.

Table 5: Summary statistics in the DADS and the LFS

	DADS	LFS
Unemployment rate	0.077	0.071
Implied unemp. rate from losing and finding	0.057	0.055
Participation rate	0.911	0.903
E-to-U probability	0.011	0.015
U-to-E probability	0.180	0.261

Skill definition. Because the DADS panel does not have education data, I construct a measure of skill based on workers' occupation and age, I run a Mincer regression of worker wages on basic demographics (age and occupation fixed effects), industry and city fixed effects. I retrieve the age and occupation fixed effects, average them over the individual's work history. Then I rank thoses averages between workers, and define that rank as skill. I check that several alternative definitions of skill do not alter the results.

Figure 14: Five-year change in local unemployment rate. Aggregate industry cycle controls.



More precisely, I run the following Mincer regression:

$$\log w_{it} = \underbrace{\alpha_{O(i,t)}}_{\text{Occupation}} + \underbrace{\alpha_{Y(t)}}_{\text{Year}} + \underbrace{\alpha_{C(i,t)}}_{\text{City}} + \underbrace{\alpha_{A(i,t)}}_{\text{Age bin}} + \varepsilon_{it}$$

for employed workers i in quarter t. Age is binned into 5-year groups, and occupations are at the 3-digit level.[69] Then define skill as average occupation and age premium

$$\hat{S}_i = \frac{1}{N_{i,O}} \sum_{k=1}^{N_{i,O}} (\hat{\alpha}_{O(i,t)} + \hat{\alpha}_{A(i,t)})$$

Results are similar when including worker fixed effects in the Mincer regression and in the skill measure, or when including industry fixed effects.

Firm-level balance sheet data. For several over-identification exercises, I use firm-level balance sheet data. I use the FICUS data (*"Fichier Complet Unifié de Suse"*) which covers the near universe of nonfarm French businesses. The unit of observation is a firm-year. I link the firm identifier to the DADS postes, which allows to identify all workers in the different establishments of the firm. Except to compute the real estate expenditure share and to examine the location choices of multi-establishment firms, I restrict the analysis to single-establishment firms in order to have a well-defined notion of location. In the sample of single-establishment firms, I use firm age and industry. I can also compute value added per worker (labor productivity), average worker skill at a firm along with other variables used in the over-identification exercises.

A.2 Persistence

Figure 14 shows persistence in local unemployment rates after netting out country-wide industry cycles. The autocorrelation is 0.99. To remove the contribution of industry cycles at the country level, I first compute country-wide change in employment at the 3-digit industry level ΔE_j between bot subperiods 0 (1997-2001(and 1 (2002-2007). Then, I construct a predicted employment change at the commuting zone level by projecting the predicted

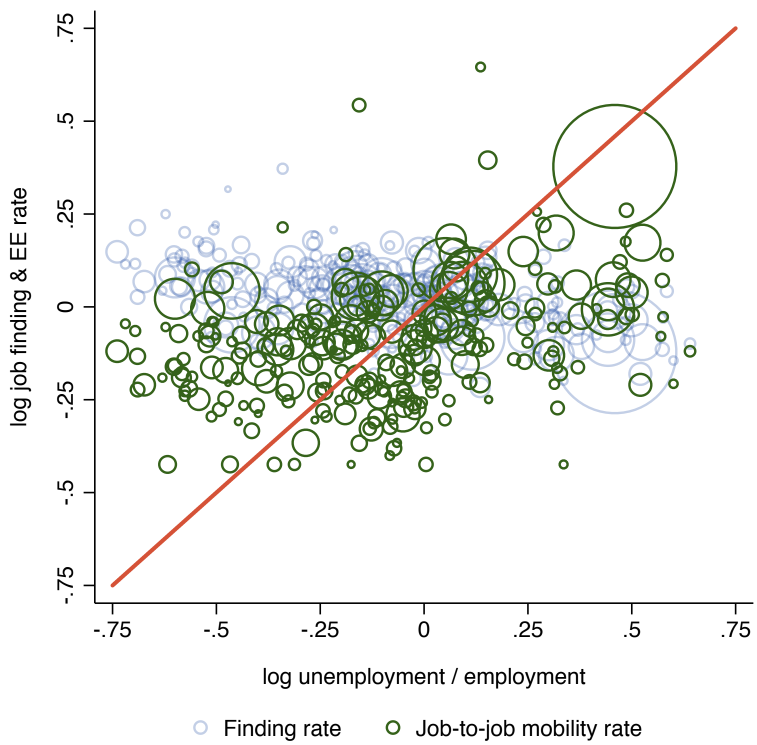
[69] In some years only 2-digit occupations are available. For workers who have at least one job in years with a 3-digit occupation, I use that occupation. For a few workers I only have 2-digit occupation information, in which case I use the 2-digit occupation wage premium.

industry employment changes ΔE_j at the local level using industry employment shares in each location in the 1997-2001 subperiod $w_{c,j,0}$: $\Delta E_c = \sum_j w_{c,j,0} \times \Delta E_j$. Next, I regress changes in local unemployment rates on this predicted change in employment $\Delta u_c = \beta_0 + \beta_1 \Delta E_c + \Delta \tilde{u}_c$. Finally, I extract the residuals from this regression $\Delta \tilde{u}_c$ and construct a measure of local unemployment net of industry cycles in the second subperiod as $\hat{u}_{c,1} = u_{c,0} + \Delta \tilde{u}_c$. Figure 14 plots $\hat{u}_{c,1}$ against $u_{c,0}$.

A.3 Transition rates

A.3.1 Job-to-job rate

Figure 15: Local job-to-job mobility rate against unemployment-to-employment ratios. France.



A.3.2 United States

See Figure 16.

A.3.3 Time-aggregation

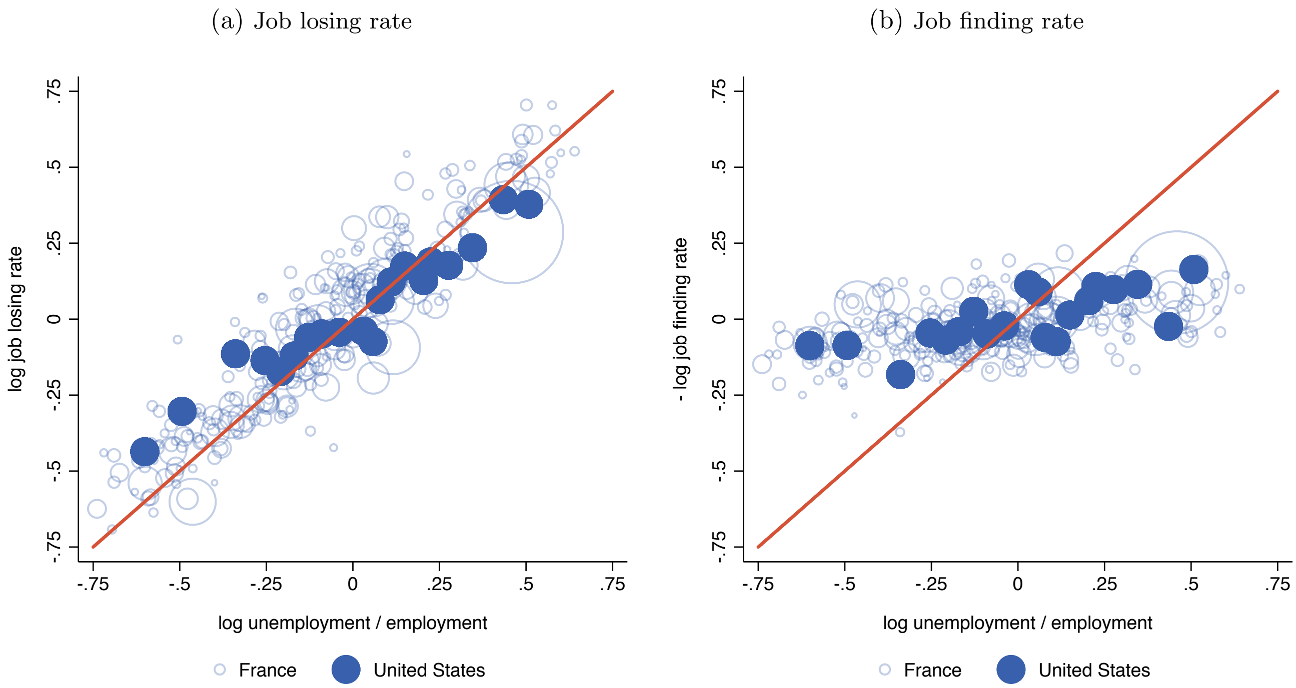
Two-state continuous time model. Consider first the case in which each city is isolated and workers never leave or enter the labor force which size is normalized to 1. Assume constant job losing and finding rates s, f. Then unemployment and employment in each city evolves according to the ODE system

$$\dot{u} = se - fu \quad ; \quad \dot{e} = fu - se \quad ; \quad e = 1 - u$$

This system has a simple solution

$$u(t) = u_\infty + (u_0 - u_\infty)e^{-(s+f)t} \quad ; \quad e(t) = e_\infty + (e_0 - e_\infty)e^{-(s+f)t}$$

Figure 16: Local job losing and finding rates against unemployment-to-employment ratios. France and United States.



where $u_\infty = \frac{s}{s+f}$ and $e_\infty = \frac{f}{s+f}$. Therefore, the transition probabilities in any given time interval $[0, t]$ are

$$\mathbb{P}_t[E \to U] = u(t)\big|_{u_0=0} = \frac{s(1 - e^{-(s+f)t})}{s+f} \quad ; \quad \mathbb{P}_t[U \to E] = e(t)\big|_{u_0=1} = \frac{f(1 - e^{-(s+f)t})}{s+f}$$

Hence, the instantaneous quarterly transition rates can be recovered from time-aggregated transition probabilities from

$$s = \mathcal{T} \times \mathbb{P}_1[E \to U] \quad ; \quad f = \mathcal{T} \times \mathbb{P}_1[U \to E]$$

where one quarter is the interval $[t, t+1)$, and the time aggregation factor is

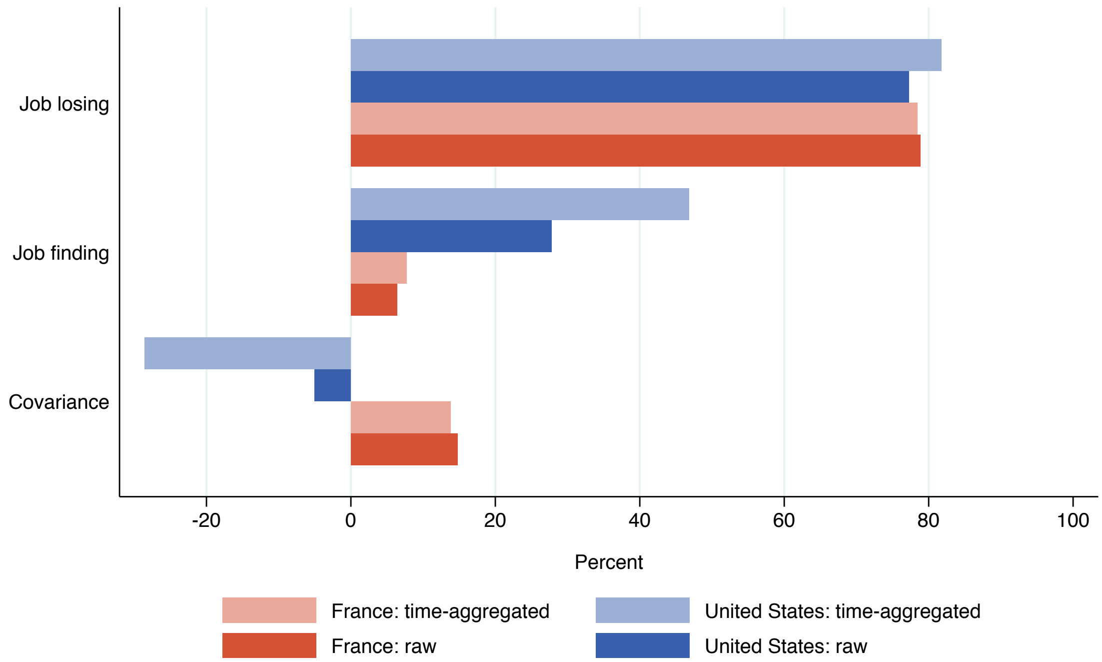
$$\mathcal{T} = \frac{\log\Big(1 - \mathbb{P}_1[E \to U] - \mathbb{P}_1[U \to E]\Big)}{\mathbb{P}_1[E \to U] + \mathbb{P}_1[U \to E]}$$

Time-aggregation in the data. To assess the importance of time-aggregation in the data, Figure 17 shows the variance decomposition of the log unemployment-to-employment ration into using transition probabilities and time-aggregated transition rates, for France and the United States. The job losing shares remains stable across all specifications and countries. In France, the time-aggregation correction also does not change the job finding and the covariance share. In the United States, these are more sensitive to time aggregation, without changing the main results. There are at least two explanations for this difference. First, the CPS data for the United States contains potentially substantial measurement error in the local job finding probabilities due to small sample issues. This inflates the variance share of job finding flows, which is in turn magnified by the time aggregation correction. Second, gross labor market flows are larger in the United States, leading to potentially more time aggregation bias.

Three-state model. I now consider a three-state version of the model, still with isolated locations and the total number of individuals normalized to 1 in each lcation. Denote now by $n(t)$ the number of individuals out of the labor force, so that $u(t) + e(t) + n(t) = 1$. There are transitions between all states, such that

$$\dot{u} = se - fu + rn - du \quad ; \quad \dot{n} = s_n e - f_n n - rn + du$$

Figure 17: Variance decomposition of predicted log unemployed-to-employed ratio at city level into job losing and job finding rates. France and United States. With and without time-aggregation correction.



where s_n is the separation rate into non-participation, f_n the finding rate out of non-participation, r the re-entry rate (NU) and d the drop-out rate (UN). In steady-state,

$$du - rn = se - fu \qquad ; \qquad du - rn = f_n n - s_n e$$

Finally, the unemployment rate u_R is $u_R = \frac{u}{e+u} = \frac{u}{1-n}$. Using $e = 1 - u - n$ and combining both equations,

$$u_R = \frac{s(f_n + r) + rs_n}{f_n(d + f + s) + r(f + s_n + s)}$$

and so

$$\frac{u_R}{1 - u_R} = \frac{s(f_n + r) + rs_n}{f_n(d + f) + rf}$$

Then define

$$p = \frac{s(f_n + r) + rs_n}{f_n(d + f) + rf} - \frac{s}{s + f}$$

Therefore, with flows in and out of the labor force within isolated cities c, equation (1) becomes

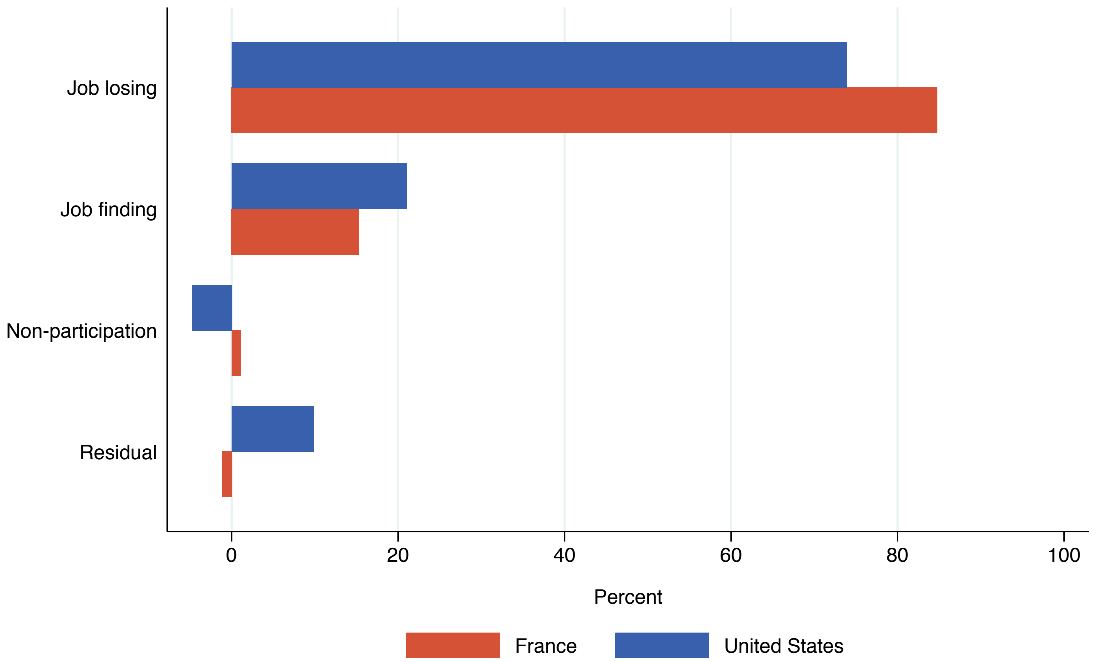
$$\log \frac{u_c}{1 - u_c} = \log s_c - \log f_c + \log p_c + e_c \tag{24}$$

where e_c is a residual that captures migration flows, local dynamics and measurement error. The exact variance decomposition of the log unemployment-to-employment ratio writes

$$\begin{aligned}\mathbb{V}\text{ar}\left[\log\frac{u_c}{1-u_c}\right] &= \mathbb{C}\text{ov}\left[\log\frac{u_c}{1-u_c},\log s_c\right]+\mathbb{C}\text{ov}\left[\log\frac{u_c}{1-u_c},-\log f_c\right]\\ &+ \mathbb{C}\text{ov}\left[\log\frac{u_c}{1-u_c},\log p_c\right]+\mathbb{C}\text{ov}\left[\log\frac{u_c}{1-u_c},e_c\right]\end{aligned}$$

Figure 18 reports the results from this variance decomposition.

Figure 18: Variance decomposition of local unemployment-to-employment ratio.



A.4 Composition

Equation (2) is useful to attribute the spatial variation in Y to city, industry and skill or worker characteristics. An exact variance decomposition follows from taking expectations conditional on city c on each side, and breaking up the resulting variance:

$$\mathbb{V}\text{ar}\big[\bar{Y}_c\big] = \mathbb{C}\text{ov}\big[\bar{Y}_c,\alpha_c\big]+\mathbb{C}\text{ov}\big[\bar{Y}_c,\mathbb{E}_c[\beta_j]\big]+\mathbb{C}\text{ov}\big[\bar{Y}_c,\mathbb{E}_c[\gamma_i]\big] \tag{25}$$

where $\bar{Y}_c = \mathbb{E}_c[Y_{c,j,i}]$ denotes the local average of Y. The first term on the right-hand-side of equation (25) is the contribution of city-specific heterogeneity to the spatial variation in Y. The second term is the contribution of industry heterogeneity, and the third of skill or worker heterogeneity. These two last terms are zero if $\mathbb{E}_c[\beta_j] = \mathbb{E}_c[\gamma_i] = 0$, which occurs when there is no systematic sorting of industries or skills across cities.

Fixed effects regressions estimate the decomposition in equations (2)-(25), but at at a fine level of disaggregation with over 300 cities, 220 industries and 300 skill groups, well-known small-sample biases in the covariances may arise. Small sample biases may arise similarly to those in worker and firm effects models as in Abowd et al. (1999). Therefore, I also estimate a correlated random effects structure following Borovičková and Shimer (2017), which provides an unbiased estimate of the variance-covariance matrix at the cost of distributional restrictions. The key idea is to use a leave-out estimator. I posit that the random effects follow a jointly normal distribution. Finally, I estimate linear and probit models for robustness to functional form assumptions. See also Kline et al. (2019) as well as Bonhomme et al. (2019) for alternative approaches. I now describe my correlated random effect estimator in more detail.

Consider the de-meaned job losing probability $EU_{c,j,i}$. Assume that

$$EU_{c,j,i} = \alpha_c + \beta_j + \gamma_i + e_{c,j,i} \tag{26}$$

Suppose that $(\alpha_c, \beta_j, \gamma_i)$ is jointly normally distributed in the employed population, with mean zero and covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_c^2 & \rho_{cj}\sigma_c\sigma_j & \rho_{ci}\sigma_c\sigma_i \\ \rho_{cj}\sigma_c\sigma_j & \sigma_j^2 & \rho_{ji}\sigma_j\sigma_i \\ \rho_{ci}\sigma_c\sigma_i & \rho_{ji}\sigma_j\sigma_i & \sigma_i^2 \end{pmatrix}$$

Suppose that $\varepsilon_{c,j,i}$ has mean zero conditional on (c, j, i), is normally, identically, and independently distributed across triplets. Given an estimate of Σ, the variance decomposition in (25) follows from conditional Gaussian distributions:

$$\begin{aligned} \mathbb{C}\mathrm{ov}\big[\bar{Y}_c, \alpha_c\big] &= \sigma_c V \\ \mathbb{C}\mathrm{ov}\big[\bar{Y}_c, \mathbb{E}_c[\beta_j]\big] &= \rho_{cj}\sigma_j V \\ \mathbb{C}\mathrm{ov}\big[\bar{Y}_c, \mathbb{E}_c[\gamma_i]\big] &= \rho_{ci}\sigma_i V \\ V &\equiv \sigma_c + \rho_{cj}\sigma_j + \rho_{ci}\sigma_i \end{aligned}$$

To estimate Σ, I use conditional second moments. I first outline the strategy in large samples. Then, I describe how I correct for small sample biases in practice.

Large samples. The random effects structure implies

$$\begin{aligned} \mathbb{E}_c[EU_{c,j,i}] &= \left(1 + \rho_{cj}\frac{\sigma_j}{\sigma_c} + \rho_{cs}\frac{\sigma_s}{\sigma_c}\right)\alpha_c \\ \mathbb{E}_j[EU_{c,j,i}] &= \left(1 + \rho_{cj}\frac{\sigma_c}{\sigma_j} + \rho_{sj}\frac{\sigma_s}{\sigma_j}\right)\beta_j \\ \mathbb{E}_i[EU_{c,j,i}] &= \left(1 + \rho_{ci}\frac{\sigma_c}{\sigma_s} + \rho_{ji}\frac{\sigma_j}{\sigma_s}\right)\gamma_i \end{aligned}$$

Thus, the random effects correlations can be directly estimated from correlations between conditional means:

$$\begin{aligned} \mathbb{C}\mathrm{orr}[\mathbb{E}_c[EU_{c,j,i}], \mathbb{E}_s[EU_{c,j,i}]] &= \rho_{cs} \\ \mathbb{C}\mathrm{orr}[\mathbb{E}_c[EU_{c,j,i}], \mathbb{E}_j[EU_{c,j,i}]] &= \rho_{cj} \\ \mathbb{C}\mathrm{orr}[\mathbb{E}_i[EU_{c,j,i}], \mathbb{E}_j[EU_{c,j,i}]] &= \rho_{ji} \end{aligned} \tag{27}$$

Given estimates of the pairwise correlations, recover variances from the 3x3 linear system:

$$\begin{aligned} \sqrt{\mathbb{V}\mathrm{ar}[\mathbb{E}_c[EU_{c,j,i}]]} &= \sigma_c + \rho_{cj}\sigma_j + \rho_{ci}\sigma_i \\ \sqrt{\mathbb{V}\mathrm{ar}[\mathbb{E}_j[EU_{c,j,i}]]} &= \sigma_j + \rho_{cj}\sigma_c + \rho_{ji}\sigma_i \\ \sqrt{\mathbb{V}\mathrm{ar}[\mathbb{E}_i[EU_{c,j,i}]]} &= \sigma_i + \rho_{ci}\sigma_c + \rho_{ji}\sigma_j \end{aligned} \tag{28}$$

which can be solved as

$$\begin{aligned} \begin{pmatrix} \sigma_j \\ \sigma_c \\ \sigma_s \end{pmatrix} &= \frac{1}{D}\begin{pmatrix} 1-\rho_{ci}^2 & \rho_{ci}\rho_{ij}-\rho_{cj} & \rho_{cj}\rho_{ci}-\rho_{ij} \\ \rho_{ci}\rho_{ij}-\rho_{cj} & 1-\rho_{ij}^2 & \rho_{cj}\sigma_{ij}-\rho_{ci} \\ \rho_{cj}\rho_{ci}-\rho_{ij} & \rho_{cj}\rho_{ij}-\rho_{ci} & 1-\rho_{cj}^2 \end{pmatrix} \cdot \begin{pmatrix} \sqrt{\mathbb{V}\mathrm{ar}[\mathbb{E}_j[EU_{c,j,i}]]} \\ \sqrt{\mathbb{V}\mathrm{ar}[\mathbb{E}_c[EU_{c,j,i}]]} \\ \sqrt{\mathbb{V}\mathrm{ar}[\mathbb{E}_i[EU_{c,j,i}]]} \end{pmatrix} \\ D &\equiv 1 - \rho_{cj}^2 - \rho_{ci}^2 - \rho_{ji}^2 + 2\rho_{cj}\rho_{ci}\sigma_{si} \end{aligned}$$

Replace $EU_{c,j,i}$ with the inverse probit transformation of EU for the probit model. When estimating with skill groups and worker effects, impose that the correlation between worker effects and skill effects is zero.

Small samples. In practice, "naive" estimators of the variances in (28) and the correlations in (27) are subject to small sample biases. First, variance estimates must be adjusted using the Bessel correction factor. Second, consider the "naive" correlation estimator that simply correlates estimated conditional means, In that case, the common observation in conditional means creates a positive bias .[70] To circumvent this difficulty given the i.i.d. assumption of the residual, it suffices to remove all common observations from the estimated conditional means. This strategy is often called a leave-out estimator.

Finally, because the relevant distribution is the employment-weighted distribution, it remains to specify how mobility is correlated with the fixed effects. To make progress, I follow Card et al. (2013) and Borovičková and Shimer (2017) and assume conditional random mobility. More precisely, I assume that the number of quarters of worker i in city c and industry j, $n_{c,j,i}$, is uncorrelated with α_c and β_j conditional on γ_i. This is admittedly a strong assumption, although it will be satisfied in the model. Similarly, I assume that $n_{c,j,i}$ is uncorrelated with γ_i and α_c conditional on β_j, and that it is uncorrelated with γ_i and β_j conditional on α_c.

In practice using cities, industries and skill groups, using the leave-out estimator or not does not affect the results because the common observation bias is small. Thus, the results are robust to relaxing the conditional random mobility assumption. However, when using cities, industries, skills and workers effects, the conditional random mobility must be imposed: there are only few different cities and industries for each worker.

With three groups, the leave-out estimation procedure imposes additional data requirements relative to a two-group situation. To facilicate exposition, I outline the estimator with two groups. The logic extends directly to three groups. Suppose there are only cities c and workers i.

Within variance. Following Borovičková and Shimer (2017), start by estimating the within-worker variance $\mathbb{V}\text{ar}_i[EU_{ijt}]$. First, an unbiased estimator of the conditional mean is

$$\overline{EU}_i = \frac{1}{N_i}\sum_{c,t} n_{i,c,t}EU_{ict}$$

where $N_i = \sum_{c,t} n_{ict}$ is the number of quarters for which individual i is observed. Now,

$$EU_{ict} \mid i \sim \mathcal{N}\left(\gamma_i + \mathbb{E}_i[\alpha_c], \sigma^2_{W,i}\right)$$

where by definition, $\sigma^2_{W,i}$ is the within-worker variance. An unbiased estimator of $\sigma^2_{W,i}$ is then

$$\widehat{\sigma^2_{W,i}} = \frac{N_i}{N_i - 1}\cdot\frac{1}{N_i}\sum_{c,t} n_{ict}(EU_{ict} - \overline{EU}_i)^2$$

The average within-worker variance is $\sigma^2_W \equiv \mathbb{E}[\mathbb{V}\text{ar}_i[EU_{ict}]]$, for which an unbiased estimator is then

$$\widehat{\sigma^2_W} = \frac{1}{\bar{N}}\sum_{i=1} N_i\widehat{\sigma^2_{W,i}} = \frac{1}{T}\sum_{i=1}\frac{N_i}{N_i - 1}\cdot\sum_{c,t} n_{ict}(EU_{ict} - \overline{EU}_i)^2$$

where $N = \sum_{i=1}^{\bar{N}} N_i$, and $\bar{N}$ is the number of workers in the sample.

Variance of conditional mean. To estimate the variance of the conditional mean, use the law of total variance

$$\mathbb{V}\text{ar}[\mathbb{E}_i[EU_{ict}]] = \mathbb{V}\text{ar}[EU_{ict}] - \mathbb{E}[\mathbb{V}\text{ar}_i[EU_{ict}]]$$

where the unconditional variance can be estimated with the standard variance estimator.

Covariance of conditional means To get an unbiased estiamator of the covariance between conditional means, compute conditional means leaving out common terms. Denote $\mathcal{N}_i$ the set of quarters t for which individual i is observed. Denote $\mathcal{N}_{i,-c}$ the set of quarters for which individual i is observed, but in a city different than c. Denote $N_{i,-c}$ the number of quarters in $\mathcal{N}_{i,-c}$. Symmetrically, denote $\mathcal{N}_{c,-i}$ the set of individual-quarters pairs

[70]This positive bias is distinct from the negative bias that arises in small sample fixed effect estimators. For fixed effects in samples, the negative bias arises because positive measurement error in one fixed effect is mechanically transmitted as negative measurement error into the other fixed effect of a given pair.

that are in city c outside of worker i, and $N_{c,-i}$ its cardinality. Then compute

$$
\begin{aligned}
& \mathbb{C}\mathrm{ov}\left(\frac{1}{N_{i,-c}} \sum_{(t,\ell)\in\mathcal{N}_{i,-c}} n_{i\ell t} EU_{ikt}, \frac{1}{N_{c,-i}} \sum_{(t,k)\in\mathcal{N}_{c-i}} n_{kct} EU_{kct}\right) \\
= \quad & \mathbb{C}\mathrm{ov}\left(\gamma_i + \frac{1}{N_{i,-c}} \sum_{(t,\ell)\in\mathcal{N}_{i,-c}} n_{i\ell t}(\alpha_k + e_{ikt}), \alpha_c + \frac{1}{N_{c,-i}} \sum_{(t,k)\in\mathcal{N}_{c-i}} n_{kct}(\gamma_k + e_{ikt})\right) \\
= \quad & \mathbb{C}\mathrm{ov}(\gamma_i, \alpha_c)
\end{aligned}
$$

which, adjusted by the estimated variances, delivers an estimate of the correlation. These estimators can be directly extended tow three groups, provided all common observations between one group and the two others are removed from the conditional mean estimators.

Results. Figure 19 displays the results of the variance decomposition for the job losing rate from equation (25) for France and the United States. In practice, the results are very close across specifications and countries. Industry and skill composition of cities account for no more than 10-15% of the spatial variation in the job losing rate, while city-specific effects account for over 80% of it. Figure 20 indicates that the results are similar for the job finding rate.

Finally, figure 21 the decomposition after including worker effects. Estimating that specification requires to restrict the sample to movers between industry and cities, as well as assuming that worker effects have mean zero conditional on the skill effect. For comparison, Figure 21 reports the results from similar decompositions without the worker effects, but on the same restricted sample of workers. The variance share of city effects modestly diminishes to about 75%, but mostly as a result of the sample selection. Including worker effects leaves cities' contribution essentially unchanged. It affects the contributions of industries and skills to spatial job losing rate differentials.

Figure 19: Variance decompositions of local job losing rate into city, industry and skill contributions. France and United States.

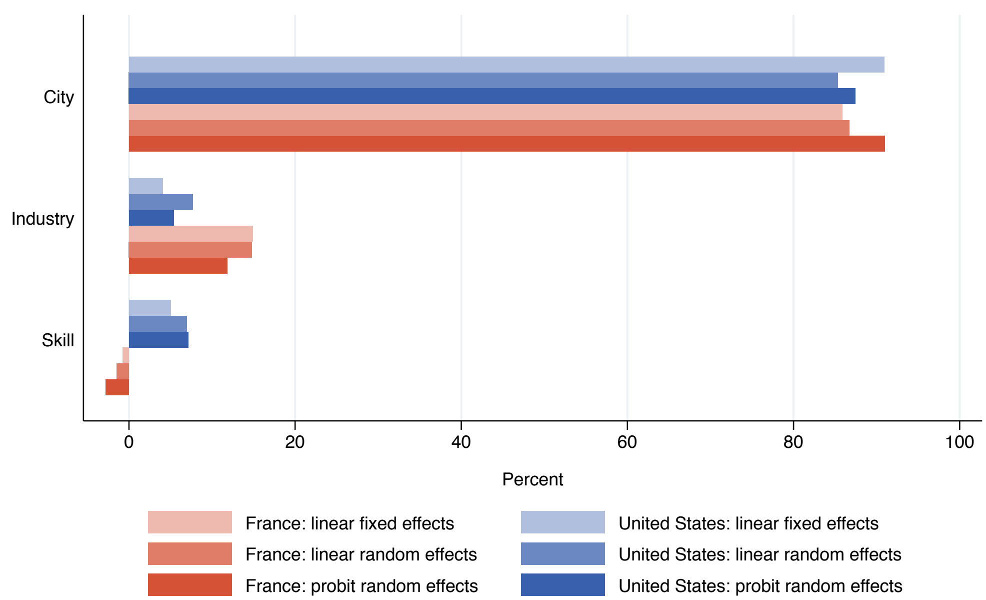


Figure 20: Variance decompositions of local job finding rate into city, industry and skill contributions. France and United States.

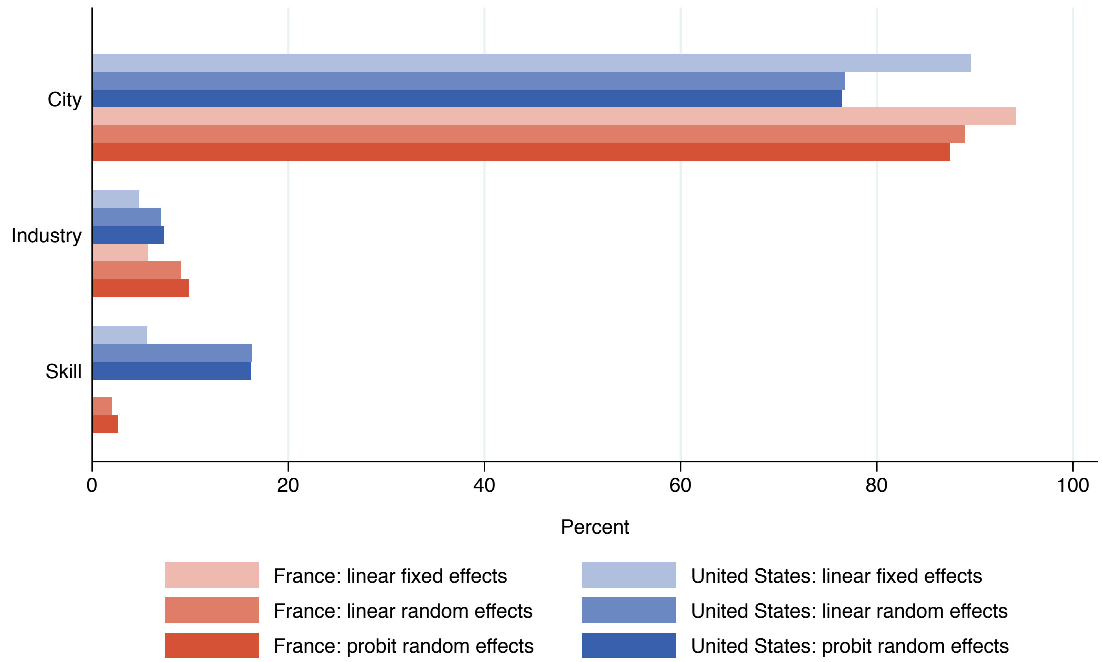
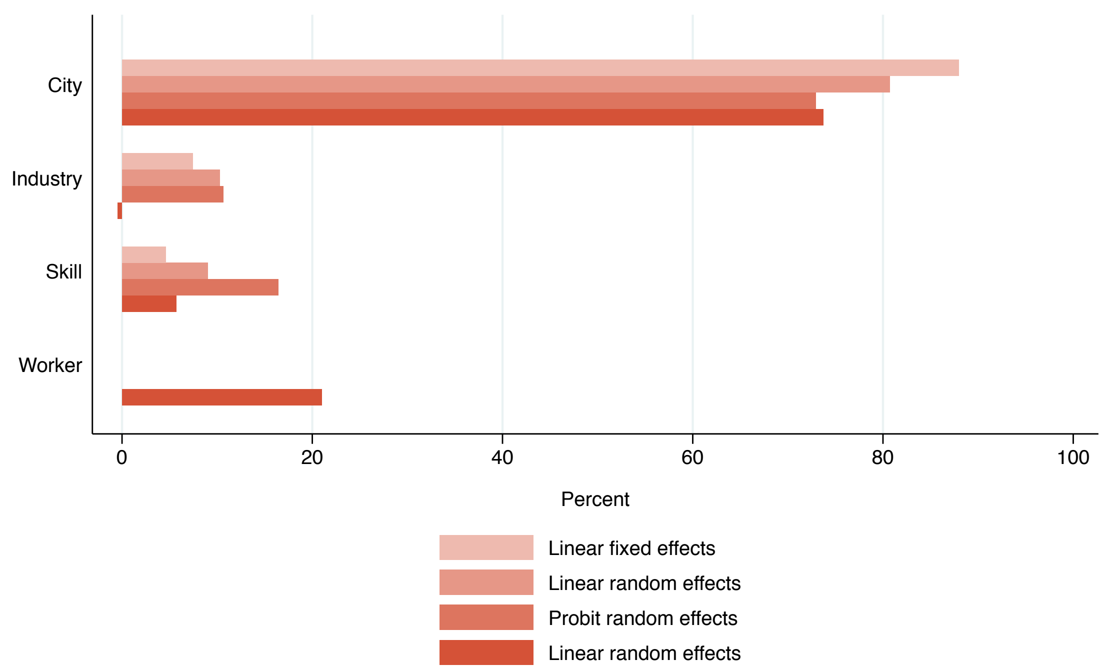


Figure 21: Variance decompositions of local job losing rate into city, industry, skill and worker contributions. France only.



A.5 Conditional correlations

See Table 6.

Table 6: OLS regressions of worker-level unemployment, job loss and job finding probabilities

	Unemployment			Job loss			Job finding		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log city wage	-1.44***	-0.79*	-0.62***	-0.30***	-0.21***	-0.16*	-0.66***	-0.64***	-0.73***
	(0.40)	(0.32)	(0.15)	(0.06)	(0.05)	(0.07)	(0.14)	(0.13)	(0.17)
Log city pop.	1.63***	1.40***	0.50*	0.23***	0.21***	0.12+	-0.04	0.01	-0.23
	(0.40)	(0.33)	(0.21)	(0.06)	(0.06)	(0.06)	(0.15)	(0.13)	(0.16)
Worker skill		-3.07***			-0.47***			0.55***	
		(0.18)			(0.02)			(0.12)	
Fixed Effects									
Year	✓			✓			✓		
Industry-Year		✓	✓		✓	✓		✓	✓
Worker			✓			✓			✓
Obs.	3005929	3005919	3005306	2699433	2699426	2697645	306496	306436	296336
R^2	0.006	0.070	0.361	0.001	0.015	0.105	0.002	0.017	0.232
W.-R^2	0.001	0.010	0.000	0.000	0.001	0.000	0.000	0.000	0.000

Standard errors in parenthesis, two-way clustered by city and 3-digit industry. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
City population by km2 (density). Quarterly frequency, 1997-2007. Movers only.
Left-hand-side variables in 100 percentage points. Right-hand-side variables standardized.

A.6 Tenure profile

It is well known that job losing rates tend to be highest at early job tenures, and subsequently decline. A natural question is then whether spatial gaps in job losing rates also occur at early job tenures.

To estimate the tenure profile of job loss, I run the following linear probability model

$$EU_{iq} = \alpha_{C(i,q),\tau(i,q)} + \beta_{J(i,q),\tau(i,q)} + \gamma_{S(i),\tau(i,q)} + \delta_{Y(q)} + e_{iq}$$

where i indexed workers, c cities, j 3-digit industries and q quarters. $\tau(i,q)$ denotes worker i's years of job tenure in quarter q. $Y(q)$ is a year fixed effect. I then retrieve all fixed effects. The solid blue line in Figure 22 is then the economy-wide average of all the fixed effects:

$$A_\tau = \frac{1}{N_\tau} \sum_{\substack{i,q \mid \\ \tau(i,q)=\tau}} \left(\hat{\alpha}_{C(i,q),\tau(i,q)} + \hat{\beta}_{J(i,q),\tau(i,q)} + \hat{\gamma}_{S(i),\tau(i,q)} \right)$$

It numerically coincides with the outcome of a simple OLS regression $\hat{A}_\tau$

$$EU_{iq} = \delta_{Y(q)} + A_{\tau(i,q)} + e_{iq}$$

The dashed orange lines are the city premia. They correspond to

$$P_{g,\tau} = A_\tau + \frac{1}{N_{g,\tau}} \sum_{\substack{i,q \mid \\ \tau(i,q)=\tau \\ (i,q)\in\mathcal{N}_g}} \left(\hat{\alpha}_{C(i,q),\tau(i,q)} - A_\tau \right)$$

where g denotes city groups. Cities are binned into four population-weighted raw job losing rate quartiles. $\mathcal{N}_{g,\tau}$ is the number of worker-quarter pairs in group g with annual job tenure τ. $N_{g,\tau}$ denotes the cardinality of $\mathcal{N}_{g,\tau}$.

Figure 22: Job loss rate by job tenure, aggregate and by city. France.

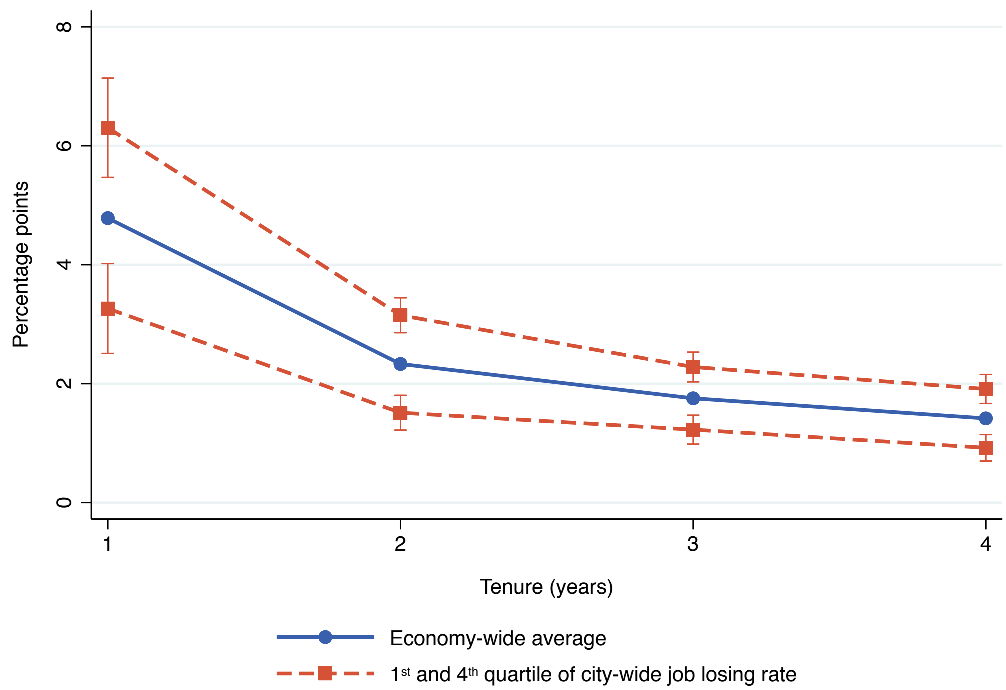


Figure 23: Job loss rate by job tenure, aggregate and by city. Difference between top and bottom quartile of cities ranked by job losing rate. France.

Figure 22 first reveals that average job losing probabilities decline with tenure in all cities. Most of the changes occur in the first two years. Second, it shows that the gap between high and low job losing cities is largest at early job tenures – in the first two years of the job – and stabilize thereafter. Figure 23 shows that the pattern is similar after controlling for worker fixed effects.

B Baseline model

B.1 Value functions

In this Appendix I solve the more general model without assuming that wages depend only on productivity and the location.

Values. When the wage needs not depend on (y, ℓ) but only follows a Markov process, workers' values become

$$\rho U = b\ell r(\ell)^{-\omega} + f(\ell)\mathbb{E}_\ell[V^E(w^*(y_0, \ell), \ell) - U] \qquad ; \qquad \rho V^E(w, \ell) = wr(\ell)^{-\omega} + (L_w V^E)(w, \ell)$$

where the expectation is taken over the starting productivity y_0 in location ℓ. L_w is the integro-differential infinitesimal generator that encodes the continuation value of employment due to wage changes. It needs not be explicitly specified at this stage.

Worker surplus. Workers' surplus from being employed $V^E - U$ solves

$$\rho\big(V^E(w, \ell) - U\big) \quad = \quad r(\ell)^{-\omega}\Big(w - (b + v(\ell))\ell\Big) + L_w\big(V^E - U\big)(w, \ell)$$

where I denote $v(\ell)\ell = f(\ell)\mathbb{E}[V^E(w^*(y_0, \ell)) - U]$ the efficiency value of search in location ℓ.

Employers. The value of a filled job paying wage w with productivity y in location ℓ solves

$$\rho J(w, y, \ell) \quad = \quad y\ell - w + (L_y J)(y, w, \ell) + (L_w J)(y, w, \ell)$$

B.2 Bargaining

To characterize wages and values, it is useful to define the adjusted surplus

$$S(y, \ell) = J(y, w, \ell) + r(\ell)^\omega \cdot \big(V^E(y, w, \ell) - U\big)$$

which is indepedent from wages, and solves the recursion

$$\rho S(y, \ell, a) = \ell \cdot \Big(y - b - v(\ell)\Big) - L_y S \tag{29}$$

for continuing matches. Renegotiation every instant means that employers and workers bargain over flow surpluses

$$r(\ell)^{-\omega}(w - (b + v(\ell))\ell) \qquad ; \qquad y\ell - w$$

Without loss of generality, these flow surpluses can be written as values

$$W(w) = W_0 w - W_1 \qquad ; \qquad F(w) = F_1 - w$$

To solve for wages in the bargaining game, the idea is now to make use Proposition 122.1 p.122, Chapter 7, of Osborne and Rubinstein (1994). The setup of the bargaining game is as follows. There is a parallel time for bargaining, in which the worker and the firm have linear flow preferences over a wage w given by $W(w), F(w)$, and discount the future. Denote by δ_F the discount factor of the worker in the bargaining space-time, and δ_F that of the firm.

Disagreement and admissible wages. If bargaining breaks down, each side gets 0. The admissible bargaining set is all w such that

$$B^F \equiv \frac{W_1}{W_0} \leq w \leq F_1 \equiv B^W$$

where B^W, B^F denote the worker's and firm's best agreement, respectively. Finally, define the Pareto frontier as the set of wages w such that there is no other wage w' such that both parties prefer w' to w in the initial round: $F(w') > F(w)$ and $W(w') > W(w)$. Because of the linearity of flow values, the Pareo frontier is exactly equal to the set of admissible wages. I now need check Assumptions (A1-A4) p.122 in Osborne and Rubinstein (1994).

(A1) – For no two distinct wages $w \neq w'$, it is the case that $W(w) = W(w')$ and $F(w) = F(w')$. Each party's objective is strictly monotonic in the chosen wage w, so (A1) is satisfied.

(A2) – Getting the other party's best agreement in the second round is the same as getting in the first round, i.e. $F(B^W) = \delta_F F(B^W)$ and $W(B^F) = \delta_W F(B^F)$. Since $F(B^W) = W(B^F) = 0$, (A2) is satisfied.

(A3) – The Pareto frontier is strictly monotome: for any efficient/admissible wage w, there is no other wage $w' \neq w$ such that each side weakly prefers w'. This again directly follows from linearity of payoffs.

(A4) – There is a unique pair of wages (w^W, w^F) such that $\delta^W(w^W) = W(w^F)$ and $\delta^F F(w^F) = F(w^W)$, and both (w^W, w^F) are efficient. I write down the system of equations

$$\begin{aligned} \delta^W[W_0 w^W - W_1 a] &= W_0 w^F - W_1 \\ \delta^F[F_1 - w^F] &= F_1 - w^W \end{aligned}$$

Use the second equation to obtain:

$$w^F = \frac{w^W}{\delta^F} - \frac{1-\delta^F}{\delta^F} F_1$$

Substituting into the first equation:

$$w^W = \beta^W F_1 + (1-\beta^W)\frac{W_1}{W_0}$$

where

$$\beta^W = \frac{1-\delta^F}{1-\delta^F\delta^W} \in (0,1)$$

and so

$$w^F = \beta^F F_1 + (1-\beta^F)\frac{W_1}{W_0}$$

where

$$\beta^F = \frac{(1-\delta^F)\delta^W}{1-\delta^F\delta^W} \in (0,1)$$

Finally, w^W, w^F are automatically on the Pareto frontier because they are admissible and payoffs are linear, which concludes the proof to the bargaining solution.

Without loss of generality, suppose that the worker moves first. Then the worker's effective bargaining power is $\beta = \beta_W$. Finally, note that the bargaining solution solves

$$\frac{W(w^*)}{W_0} = \beta \cdot \left(F(w^*) + \frac{W(w^*)}{W_0}\right) \quad ; \quad F(w^*) = (1-\beta)\cdot\left(F(w^*) + \frac{W(w^*)}{W_0}\right)$$

Therefore, it is enough to define an adjusted surplus $F(w) + \frac{W(w)}{W_0}$ which does not depend on wages. Then rescaled values split this adjusted surplus. This argument prooves the following Lemma.

Lemma 3. *(Bargaining solution)*
Suppose a worker and an employer play an alternating offer game à la Rubinstein (1982) with static surpluses $W(w) = W_0 w - W_1$ and $F(w) = F_1 - w$, and worker effective bargaining power β. Define the adjusted surplus $S(w) = F(w) + \frac{W(w)}{W_0}$. Then

- *The adjusted surplus is independent from wages $S(w) \equiv S$*
- *The equilibrium wage w^* solves*

$$\frac{W(w^*)}{W_0} = \beta S \quad ; \quad F(w^*) = (1-\beta)S$$

B.3 Adjusted surplus

Using Lemma 3, the solution to the dynamic bargaining problem immediately follows.

Lemma 4. *(Bargaining solution)*
Equilibrium wages $w^*(y,\ell)$ split the adjusted surplus into constant shares:

$$J(y,w^*(y,\ell),\ell) = (1-\beta)S(y,\ell) \qquad ; \qquad V^E(y,w^*(y,\ell),\ell) - U = \beta r(\ell)^{-\omega}\cdot S(y,\ell)$$

Because of static renegotiation, wages for continuing matches can then be immediately calculated

$$w^*(y,\ell) = \Big[(1-\beta)(b+v(\ell)) + \beta y\Big]\ell \tag{30}$$

However, all matches eventually break up. Thus, the adjusted surplus S solves an optimal stopping problem, and thus a Hamilton-Jacobi-Bellman-Variational-Inequality (HJB-VI):[71]

$$0 = \max\left\{\Big(y - \big(b+v(\ell)\big)\Big)\ell + (L_yS)(y,\ell) - \rho S(y,\ell)\ ,\ S(y,\ell)\right\} \quad , \quad \forall\, y \geq 0 \tag{31}$$

The structure of the HJB-VI (31) has two implications: first, there exists a continuation region in which the HJB (29) holds. As will become clear, the joint surplus is strictly increasing in this continuation region. Thus, it takes the form of an interval $[\underline{y}(\ell),+\infty)$ in each location: there is a cutoff productivity $\underline{y}(\ell)$ below which the match breaks up. Then, at that cutoff, the surplus must be zero: $S(\underline{y}(\ell),\ell)=0$. This condition is sometimes called the value-matching condition. Second, because the cutoff is chosen optimally, a first-order-condition with respect to the cutoff must hold, implying $\frac{\partial S}{\partial y}(\underline{y}(\ell),\ell)=0$. This condition is sometimes called the smooth-pasting condition.[72] In addition, the joint surplus must be smaller than the surplus of a match without any outside option, which is $\frac{y\ell}{\rho+\delta-\sigma^2/2}$.[73]

Together, the HJB (29), the value-matching, smooth-pasting conditions and the upper bound determine the value $S(y,\ell)$ and the endogenous separation cutoff $\underline{y}(\ell)$, which I summarize as

$$\begin{aligned}\rho S(y,\ell) \quad &= \quad \Big(y-\big(b+v(\ell)\big)\Big)\ell + (L_yS)(y,\ell) \quad , \quad \forall y \geq \underline{y}(\ell) \\ \text{s.t.} \quad & S(\underline{y}(\ell),\ell)=0\ ,\ \frac{\partial S}{\partial y}(\underline{y}(\ell),\ell)=0\ ,\ S(y,\ell) \leq \frac{y\ell}{\rho+\delta-\sigma^2/2}\end{aligned} \tag{32}$$

Lemma 1 then displays the solution to problem (32), with the constants

$$\tau = \frac{2\delta}{\sigma^2}\left\{\sqrt{1+\frac{2\rho\sigma^2}{\delta^2}}-1\right\} \qquad ; \qquad \underline{y}_0 = \frac{1+\tau}{\tau}\frac{\rho}{\rho+\delta-\sigma^2/2}$$

The solution method follows the arguments in Luttmer (2007). To make notation lighter, I drop location indices ℓ and solve without loss of generality

$$\begin{aligned}\rho S(y) \quad &= \quad y - c + L_yS \quad , \quad \forall y \geq \underline{y} \\ \text{s.t.} \quad & S(\underline{y})=0\ ,\ S'(\underline{y})=0\ ,\ S(y) \leq \frac{y}{\rho+\delta-\sigma^2/2}\end{aligned}$$

First re-express the problem in logs $x=\log y$:

$$\begin{aligned}\rho V(x) \quad &= \quad e^x - c - \delta V'(x) + \frac{\sigma^2}{2}T''(x) \quad , \quad \forall x \geq \underline{x} \\ \text{s.t.} \quad & V(\underline{x})=0\ ,\ V'(\underline{x})=0\ ,\ V(x) \leq \frac{e^x}{\rho+\delta-\sigma^2/2}\end{aligned}$$

[71] See Pham (2009) for a formal derivation of the HJB-VI from the sequential formulation.

[72] See Pham (2009) for a formal derivation of the interval property and of the smooth-pasting condition.

[73] Formally, from the sequential formulation, the joint surplus can be expressed as $S(y,\ell) = \ell\mathbb{E}_0\left[\int_0^\tau e^{-\rho t}(y_t-(b+v))dt|y_0=y\right]$ where τ is the stopping time. Taking an upper bound, the surplus must be bounded above by the aforementioned expression.

The homogeneous equation. Look for a solution $V(x) = e^{-\tau x}$ to $\rho V(x) = -\delta V'(x) + sV''(x)$ where $s = \sigma^2/2$. This delivers a second-order equation

$$\rho = \delta\tau + s\tau^2$$

Denote $\kappa = \mu/s$ and $\eta = \rho/s$, so that the equation re-writes $\tau^2 + \kappa\tau - 1 = 0$. The assumption on parameters implies $\eta > 1 + \kappa$. The discriminant is $D = \kappa^2 + 4\eta > 0$. The equation hence has two solutions in general:

$$\tau_{\pm} = \frac{-\kappa \pm \sqrt{\kappa^2 + 4\eta}}{2}$$

Both roots can be bounded. First, $\tau_- > 0$. Second, $-\tau_+ > 1$. Indeed, since $\eta > 1 + \kappa$, $-\tau_+ = \frac{\sqrt{\kappa^2+4\eta}-\kappa}{2} > \frac{\sqrt{\kappa^2+4\kappa+4}-\kappa}{2} = \frac{\sqrt{(\kappa+2)^2}-\kappa}{2} = \frac{|\kappa+2|-\kappa}{2} \geq \frac{\kappa+|+2-\kappa}{2} \geq 1$.

Therefore, the homogeneous solution with τ_+ violates the upper bound on the value function. The solution with $\tau \equiv \tau_-$ is thus the only possible homogeneous solution.

Thus, slightly abusing notation, the homogeneous equation subject to the upper bound has solutions

$$V_H(x) \quad = \quad Ae^{-\tau x},\ A \in \mathbb{R}$$

Inhomogeneous solution. Now look for solutions:

$$V(x) = Ae^{-\tau x} + Be^x - C$$

Substituting in the HJB, the homogeneous term drops out and we find:

$$\begin{aligned}
\rho B &= 1 - \delta B + sB \to B = \frac{1}{\rho + \delta - s} \\
-\rho C &= -c \to C = \frac{c}{\rho}
\end{aligned}$$

Notice that Be^x is the value if thematch continues forever, $-C$ is the annuitized option value. The term $Ae^{-\tau x}$ then captures the endogenous separation decision.

Because $e^{-\tau x}$ solves the homogeneous equation, A is not determined form the HJB. I am left with $(A, \underline{x})$ to determine, with the two boundary conditions $V(\underline{x}) = 0,\ V'(\underline{x}) = 0$:

$$\begin{aligned}
Ae^{-\tau\underline{x}} + Be^{\underline{x}} &= C \\
-A\tau e^{-\tau\underline{x}} + Be^{\underline{x}} &= 0
\end{aligned}$$

leading to $(\tau + 1)Be^{\underline{x}} = \tau C$, and hence

$$\begin{aligned}
e^{\underline{x}} &= \frac{\tau}{\tau+1}\frac{C}{B} = \frac{\tau}{\tau+1} \cdot \left(1 - \frac{1+\kappa}{\eta}\right) \cdot d \\
A &= \frac{B}{\tau} e^{(1+\tau)\underline{x}}
\end{aligned}$$

The solution finally writes $V(x) = \frac{B}{\tau} e^{-\tau x + (1+\tau)\underline{x}} + Be^{x-\underline{x}}e^{\underline{x}} - C = Be^{\underline{x}}\Big\{e^{x-\underline{x}} + \alpha^{-1}e^{-\tau(x-\underline{x})}\Big\} - C$, i.e.

$$V(x) = \frac{e^{\underline{x}}}{\rho + \delta - s} \cdot \Big\{e^{x-\underline{x}} + \tau^{-1}e^{-\tau(x-\underline{x})}\Big\} - \frac{c}{\rho}$$

Going back to $y = e^x$ and re-arranging delivers the expression in Lemma 1.

B.4 Sorting

Given the bargaining solution and the adjusted surplus, the value of an employer z in location ℓ then satisfies

$$\rho J(z, \ell) = (1 - \beta)q(\ell)\ell(b + v(\ell))\bar{S}(z, \underline{y}(\ell)) \tag{33}$$

where I denote

$$\bar{S}(z,\underline{y}) = \int \mathcal{S}\left(\frac{y_0}{\underline{y}}\right) G_0(dy_0|z)$$

Under Assumption 1, the integral can be explicitly computed and equation (33) becomes

$$\rho J(z,\ell) = (1-\beta)q(\ell)\ell(b+v(\ell))^{1-\frac{1}{z}}\bar{S}_0(z) \tag{34}$$

where

$$\bar{S}_0(z) = (\rho Y/\underline{y}_0)^{\frac{1}{z}} \frac{z}{1-z}\frac{\tau z}{\tau z+1}$$

Expressing $b + v(\ell) = \frac{\underline{w}(\ell)}{1-\beta+\beta\underline{y}_0/\rho}$, I obtain

$$\rho J(z,\ell) = \frac{1-\beta}{1-\beta+\beta\underline{y}_0/\rho} q(\ell)\ell\underline{w}(\ell)^{1-\frac{1}{z}}(1-\beta+\beta\underline{y}_0/\rho)^{1/z}\bar{S}_0(z)$$

Defining

$$\begin{aligned}\bar{S}(z) &= \bar{S}_0(z)(1-\beta+\beta\underline{y}_0/\rho)^{1/z} \\ &= (Y/\underline{w}_0)^{\frac{1}{z}} \frac{z}{1-z}\frac{\tau z}{\tau z+1}\end{aligned}$$

and raising to a power $\frac{z}{1-z}$ delivers (11). I now turn to the proof of Proposition 1.

B.4.1 Proof of Propostion 1

To make notation lighter, denote $\zeta = 1/z$. New jobs ζ solve

$$\max_{\ell} \; [\zeta - 1]\log\frac{1}{b+v(\ell)} + \log\big(\ell q(\ell)\big)$$

This is a non-standard assignment problem, where labor costs $v(\ell)$ enter both in the return to a location and as part of the endogenous price that adjusts to mediate the matching. It is useful to consider the inverse functions $\ell(v), q(v)$ rather than $\ell(v)$, and view the problem as

$$\max_{\ell} \; [\zeta - 1]\log\frac{1}{b+v} + \log\big(\ell(v) q(v)\big)$$

where now $\ell(v), q(v)$ act as the endogenous price that adjust to sustain the matching. The first part of the objective is decreasing and convex in v, and so we expect $\ell(v)$ to be increasing in equilibrium.

Continuum property. To use first-order conditions (FOC) to characterize the assignment, I first show that a closed interval of v's exists in equilibrium. Suppose for a contradiction that there is a "hole" in the distribution of equilibrium v's. Denote $v_1 < v_2$ the lim-sup before the jump and the lim-inf after the jump. Then firms who locate right above the jump would have an incentive to deviate down because the distribution of ℓ is continuous. Thus, there can be no "hole" in the distribution of v's in equilibrium. In this case, the continuuous sorting case is the only relevant one for the interior. Therefore, I can treat v as a continuous variable in the sorting problem.

Interval property. Suppose that there is only one v in equilibrium. Then all firms would locate at a corner since locations are heterogeneous. This cannot be an equilibrium as house rents would be infinite there.

Sorting. Because of the supermodularity between ζ and $\log\frac{1}{b+v}$, the solution must feature a one-to-one assignment function between ζ and v.[74] From the second-order condition, the matching function that maps $\log\frac{1}{b+v}$ to ζ must be increasing, so the matching function $\zeta(v)$ must be decreasing: $\zeta'(v) < 0$.

First-order condition. The FOC is

$$-(\zeta(\ell)-1)\frac{v'(\ell)}{b+v(\ell)} + \frac{1}{\ell} + \frac{q'(\ell)}{q(\ell)} = 0$$

[74] See Galichon (2016).

In what follows, I denote $\bar{\mathcal{S}}(\zeta) = \bar{S}(1/\zeta)$ where $\zeta = 1/z$. Now, from the definition of the worker's value,

$$\begin{aligned} v(\ell)\ell &= \rho^{-1}\beta f(\ell)\ell\big(b+v(\ell)\big)\left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)}\bar{\mathcal{S}}(\zeta(\ell)) \\ &= \rho^{-1}\beta m^{\frac{1}{\alpha}}q(\ell)^{-\frac{1}{\alpha}}\cdot \ell q(\ell)\big(b+v(\ell)\big)\left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)}\bar{\mathcal{S}}(\zeta(\ell)) \end{aligned} \tag{35}$$

Differentiating this identity and using the FOC for the envelope theorem,

$$\frac{1}{\ell}+\frac{v'(\ell)}{v(\ell)} = -\frac{1}{\alpha}\frac{q'(\ell)}{q(\ell)}+\left(\frac{\bar{\mathcal{S}}'}{\bar{\mathcal{S}}}+\log\frac{B}{b+v}\right)\zeta'(\ell) \tag{36}$$

Substitute back into the FOC to obtain

$$-\frac{v'(\ell)}{v(\ell)}\left[\alpha+\frac{v(\ell)}{b+v(\ell)}(\zeta(\ell)-1)\right]+\frac{1-\alpha}{\ell}+\alpha\left(\frac{\bar{\mathcal{S}}'}{\bar{\mathcal{S}}}+\log\frac{B}{b+v}\right)\zeta'(\ell)=0 \tag{37}$$

Now re-index in terms of v to use $\zeta'(v)<0$:

$$-\frac{1}{v\ell'(v)}\left[\alpha+\frac{v}{b+v}(\zeta(v)-1)\right]+\frac{1-\alpha}{\ell(v)}+\alpha\left(\frac{\bar{\mathcal{S}}'(\zeta(v))}{\bar{\mathcal{S}}(\zeta(v))}+\log\frac{B}{b+v}\right)\frac{\zeta'(v)}{\ell'(v)}=0$$

and therefore,

$$\frac{1}{v\ell'(v)}\left\{\alpha+\frac{v}{b+v}(\zeta(v)-1)+\alpha v\underbrace{\left(\frac{\bar{\mathcal{S}}'(\zeta(v))}{\bar{\mathcal{S}}(\zeta(v))}+\log\frac{B}{b+v}\right)}_{<0}\cdot\underbrace{(-\zeta'(v))}_{>0}\right\}=\frac{1-\alpha}{\ell(v)}$$

When $\alpha=0$, the bracket on the left-hand-side is always positive. In this case, $\ell'(v)>0$, which implies $z'(\ell)>0$: there is positive assortative matching (PAM). Therefore, there exists a region of the parameter space where α is small and positive assortative matching obtains.[75]

B.4.2 Generalization: starting productivity distribution

I now state the general set of assumptions required for positive sorting to obtain in equilibrium.

Assumption 2. *(Initial productivity distribution)*
Assume that

- $\frac{\partial \log \underline{y}\bar{S}(z,\underline{y})}{\partial \underline{y}} < \alpha$
- $\frac{\partial \log \bar{S}(z,\underline{y})}{\partial z} > 0$
- $\frac{\partial^2 \log \bar{S}(z,\underline{y})}{\partial \underline{y}\partial z} > 0$

These assumptions allow to generalize the sorting results.

Proposition 8. *(Sorting 2)*
Suppose that Assumption 2 holds. Then all the implications of Proposition 1 hold.

Proof. The structure of the proof closely follows Appendix B.4.1. The differences to check are supermodularity and the FOC. First, the location choice becomes

$$\max_{\ell}\log(\ell q(\ell))+\log\underline{y}(\ell)\bar{S}(z,\underline{y}(\ell)) \tag{38}$$

[75]Formally, this statement anticipates that the general equilibrium conditions involve only continuously differentiable fixed point functionals.

Because $\bar{S}(z, \underline{y})$ is log-supermodular in $(z, \underline{y})$, PAM between z and $\underline{y}$ obtains. The FOC is

$$\frac{1}{\ell} + \frac{q'(\ell)}{q(\ell)} + \frac{\underline{y}'(\ell)}{\underline{y}(\ell)} \left(1 + \frac{\underline{y}(\ell)\bar{S}_{\underline{y}}(z(\ell), \underline{y}(\ell))}{\bar{S}(z(\ell), \underline{y}(\ell))}\right) = 0 \tag{39}$$

Re-arranging the worker's value of search,

$$q(\ell) \propto \left(\frac{\underline{y}(\ell) - \underline{y}_1}{\underline{y}(\ell)\bar{S}(z(\ell), \underline{y}(\ell))}\right)^{-\frac{\alpha}{1-\alpha}}$$

where $\underline{y}_1 = b\underline{y}_0/\rho$. Thus,

$$\frac{1-\alpha}{\alpha}\frac{q'(\ell)}{q(\ell)} = \frac{\underline{y}'(\ell)}{\underline{y}(\ell)} \left(1 + \frac{\underline{y}(\ell)\bar{S}_{\underline{y}}(z(\ell), \underline{y}(\ell))}{\bar{S}(z(\ell), \underline{y}(\ell))} - \frac{\underline{y}(\ell)}{\underline{y}(\ell) - \underline{y}_1}\right) + \frac{\bar{S}_z(z(\ell), \underline{y}(\ell))}{\bar{S}(z(\ell), \underline{y}(\ell))} z'(\ell)$$

Substituting into the FOC,

$$0 = \frac{1-\alpha}{\ell} + \frac{\underline{y}'(\ell)}{\underline{y}(\ell)} \left(1 + \frac{\underline{y}(\ell)\bar{S}_{\underline{y}}(z(\ell), \underline{y}(\ell))}{\bar{S}(z(\ell), \underline{y}(\ell))} - \alpha\frac{\underline{y}(\ell)}{\underline{y}(\ell) - \underline{y}_1}\right) + \alpha\frac{\bar{S}_z(z(\ell), \underline{y}(\ell))}{\bar{S}(z(\ell), \underline{y}(\ell))} z'(\ell) \tag{40}$$

Thus, when $\alpha = 0$, $\underline{y}'(\ell) > 0$, which concludes the proof. □

A special case. Suppose that the starting distribution is degenerate at $y_0 = z > \max_\ell \underline{y}(\ell)$. In that case,

$$(1+\tau)\underline{y}\bar{S}(z, \underline{y}) = \tau z + \underline{y}^{1+\tau} z^{-\tau} - (1+\tau)\underline{y}$$

Then

$$\begin{aligned}
\frac{\partial \log \underline{y}\bar{S}(z, \underline{y})}{\partial z} &= \tau\frac{1 - z^{-\tau-1}\underline{y}^{1+\tau}}{\underline{y}\bar{S}(z, \underline{y})} && > \quad 0 \\
\frac{\partial \log \underline{y}\bar{S}(z, \underline{y})}{\partial \underline{y}} &= \frac{1+\tau}{\underline{y}}\frac{(\underline{y}/z)^\tau - 1}{z/\underline{y} + (\underline{y}/z)^\tau - 1 - \tau} && < \quad 0 < \alpha
\end{aligned}$$

In addition, it can be shown that this last expression is also increasing in z on some interval $[K(\underline{y}(\ell)), +\infty)$. So in that region, $\underline{y}\bar{S}(z, \underline{y})$ is log-supermodular in $(\underline{y}, z)$.

B.4.3 Generalization: dynamic stability

In this section, I define a notion of dynamic stability of steady-states to rule out steady-states with negative assortative matching (NAM).

Definition 1. *(Dynamically stable assignment)*
A dynamically stable assignment is a pair of functions $\mathcal{A} : \ell \mapsto (z(\ell), \underline{y}(\ell))$ such that (a) $\mathcal{A}$ solves the job location problem (12) *and (b) $\mathcal{A}$ is the steady-state assignment that arises starting from a uniform assignment, and letting of jobs choose their location at Poisson rate $\mathcal{R}$, in the limit where $\mathcal{R} \to 0$.*

Definition 1 proposes a natural restriction on the set of possible equilibria that may arise. Starting from a uniform assignment of jobs to locations, the equilibrium must be attainable as jobs are slowly allowed to relocate over time. This apparently mild restriction suffices to eliminate potential coordination failures, a common source of mutliplicity in assignment problems with agglomeration economies, whose role is played here by the general equilibrium feedback of labor market tightness into employers' payoffs.[76]

[76] Exogenous differences across locations ℓ create incentives for jobs to sort, but so do endogenous differences in the vacancy contact rate $q(\ell)$. When exogenous differences in productivity ℓ are small, starting from an assignment where jobs are perfectly sorted but in reverse order relative to ℓ may still generate large enough differences in the vacancy contact rate $q(\ell)$ to sustain that assignment. Jobs' location choices would thus result in a spatial coordination failure, as aggregate output would be depressed relative to the best possible self-sustaining assignment. While examining these outcomes may be interesting per se, they are not the subject of the present paper. An alternative restriction would be to simply pick the output-maximizing self-sustaining assignment.

Proposition 9. *(Sorting)*
Under Assumption 2, conditional on the mass of entrants M_e and the value of unemployment U, there exists a unique globally stable assignment function for job quality $z(\ell)$ and a unique local cutoff function $\underline{y}(\ell)$. z and $\underline{y}$ are strictly increasing functions.

Proof. First, the limit $\mathcal{R} \to 0$ ensures that steady-state values are sufficient to characterize employers' values: all employers exit with probability one before $\mathcal{R}$ changes sufficiently to affect the values. The proof proceeds by "continuous induction" – formally, I show that the set of times $\mathcal{T}$ such that weak PAM obtains is a non-empty closed and open subset of $\mathbb{R}_+$, which then implies that it can only be $\mathbb{R}_+$. First, note that $\mathcal{T}$ is characterized by a weak inequality $\underline{y}' \geq 0$. Thus, it is a closed set.

Initialization. Consider time 0 at which employers are randomly allocated. For the fraction $\mathcal{R}dt$ of employers who can choose their location, the location FOC is equation (39), but where $q'/q = 0$. Therefore, $\underline{y}' > 0$ immediately follows at time 0, and so $0 \in \mathcal{T}$.

Recursion. Let t be the least upper bound of $\mathcal{T}$. The location FOC for employers allowed to relocate at t is (40), where by definition of $\mathcal{T}$, $z' \geq 0$ at t. It then immediately follows that $\underline{y}' \geq 0$ for a small time interval $[t, t+\varepsilon)$.

Thus, $\mathcal{T}$ is both open and closed in $\mathbb{R}_+$, and is nonempty. Thus, it is $\mathbb{R}_+$.

□

B.5 Endogenous job loss and unemployment

Intuition for the KFE (13). To provide some intuition for the KFE (13), consider a point $y > \underline{y}$. In a small time interval dt, a mass of $(\sigma^2/2 - \delta)y \cdot g(y)dy \cdot dt$ flows out from the interval $[y, y + dy)$ because of the drift. However, a mass $(\sigma^2/2 - \delta)(y - dy) \cdot g(y - dy)dy \cdot dt$ flows in from below when the drift is positive $\sigma^2/2 - \delta > 0$. In net, the change in mass is $(\sigma^2/2 - \delta)\Big[(y - dy)g(y - dy) - yg(y)\Big] \cdot dt$. Divide by dy to recover a density, and by dt to recover a change per unit of time, and take $dy \to 0$ to obtain the first term in the KFE.

B.5.1 KFE bound

Derivation of the KFE bound. First consider an intuitive version of the proof. Consider a second-order time interval $(dt)^2$. The change in log productivity is $d^2 \log z_t \approx \sigma dt N$ where N is a standard normal variable. Thus, half of the workers at the cutoff $\underline{y}$ are thrown below the cutoff $\underline{y}$ and into unemployment in an interval $(dt)^2$. Starting from $g(\underline{y})$ workers at the cutoff, only a fraction $2^{-\lfloor \frac{1}{dt} \rfloor}$ of those workers remain there after a time dt. Taking $dt \to 0$, this fraction must be zero. I now make this intution precise.

Proof. Denote $x = \log y$, and $\underline{x} = \log \underline{y}$. Omit ℓ indices for clarity. Let f be the local invariant density function. Consider the interval $[\underline{x}, \underline{x} + dx)$. The gross flows in and out of this interval between times t and $t + dt$ are:

$$\begin{aligned} \text{Inflow} &= \int_{\underline{x}+dx}^{\infty} f(z)\mathbb{P}[\underline{x} \leq z + dW_t \leq \underline{x} + dx]dz = \int_{dx}^{\infty} f(\underline{x} + y)\mathbb{P}[-y \leq dW_t \leq -y + dx]dy \\ \text{Outflow} &= \int_{\underline{x}}^{\underline{x}+dx} f(z)\mathbb{P}[z + dW_t > \underline{x} + dx \text{ or } z + dW_t < \underline{x}]dz = \int_0^{dx} f(\underline{x} + y)\Big\{\mathbb{P}[y + dW_t < 0] + \mathbb{P}[y + dW_t > dx]]\Big\}dy \end{aligned}$$

Then, denoting by Φ the cumulative distirbution function of a standard normal variable,

$$\text{Net flow}(dx, dt) = -\int_0^{dx} f(\underline{x} + y)dy + \int_0^{\infty} f(\underline{x} + y)\left\{\Phi\left(\frac{-y + dx}{\sigma\sqrt{dt}}\right) - \Phi\left(\frac{-y}{\sigma\sqrt{dt}}\right)\right\}dy$$

Then:

$$\begin{aligned} \frac{\partial f}{\partial t}(\underline{x}) &= \frac{1}{dxdt}\text{Net flow}(dx, dt) = -\frac{1}{dxdt}\int_0^{dx} f(\underline{x} + y)dy + \frac{1}{dxdt}\int_0^{\infty} f(\underline{x} + y)\left\{\Phi\left(\frac{-y + dx}{\sigma\sqrt{dt}}\right) - \Phi\left(\frac{-y}{\sigma\sqrt{dt}}\right)\right\}dy \\ &= -\frac{1}{dt}\int_0^1 f(\underline{x} + zdx)dz + \frac{1}{dt}\int_0^{\infty} f(\underline{x} + zdx)\left\{\Phi\left((1 - z)\lambda\right) - \Phi\left(-\lambda z\right)\right\}dz \end{aligned}$$

where $\lambda = \frac{dx}{\sigma\sqrt{dt}}$. Now,

$$\frac{1}{dt}\int_0^1 f(\underline{x}+zdx)dz \quad \approx_{dx\ll 1} \quad \frac{f(\underline{x})}{dt} + \frac{f'(\underline{x})dx}{2dt} + \frac{f''(\underline{x})dx^2}{6dt} + \mathcal{O}(dx^3/dt)$$

So is left to calculate:

$$\int_0^\infty f(\underline{x}+zdx)\left\{\Phi\left((1-z)\lambda\right) - \Phi\left(-\lambda z\right)\right\} dz$$

In integral form and changing variables:

$$\Phi\left((1-z)\lambda\right) - \Phi\left(-\lambda z\right) = \Phi(z\lambda) - \Phi(z\lambda - \lambda) = \int_0^\lambda \varphi(z\lambda - y)dy$$

where φ here denotes the standard normal density function. Then, after some algebra

$$\int_0^\infty f(\underline{x}+zdx)\left\{\Phi\left((1-z)\lambda\right) - \Phi\left(-\lambda z\right)\right\} dz = \frac{1}{dx}\int_{\mathbb{R}} dz\varphi(z)\int_{z\sigma\sqrt{dt}}^{z\sigma\sqrt{dt}+dx} \mathbb{1}[y\geq 0] f(\underline{x}+y)dy$$

Now, $\int_a^{a+\varepsilon} f(y)dy \approx f(a)\varepsilon + f'(a)\frac{\varepsilon^2}{2} + \frac{1}{2}f''(a)\frac{\varepsilon^3}{6} + \mathcal{O}(\varepsilon^4)$, and $\int_\delta^{\delta+\varepsilon} f(\underline{x}+y)dy \approx f(\underline{x})\varepsilon + \frac{1}{2}f'(\underline{x})\varepsilon(2\delta+\varepsilon) + \frac{1}{6}f''(\underline{x})\varepsilon[3\delta^2 + 3\delta\varepsilon + \varepsilon^2] + \ldots$. So

$$\begin{aligned}\frac{1}{dx}\int_{\mathbb{R}} dz\varphi(z)\int_{z\sigma\sqrt{dt}}^{z\sigma\sqrt{dt}+dx} f(\underline{x}+y)dy \quad &= \quad \frac{1}{dx}\int_0^\infty dz\varphi(z)\int_{z\sigma\sqrt{dt}}^{z\sigma\sqrt{dt}+dx} f(\underline{x}+y)dy \qquad (=A)\\ &+ \quad \frac{1}{dx}\int_{-\lambda}^0 dz\varphi(z)\int_0^{z\sigma\sqrt{dt}+dx} f(\underline{x}+y)dy \qquad (=B)\end{aligned}$$

Then:

$$A \quad \approx \quad \int_0^\infty dz\varphi(z)\left\{f(\underline{x}) + f'(\underline{x})dx\left(2\frac{z}{\lambda}+1\right) + \frac{1}{6}f''(\underline{x})dx^2\left[1 + 3\left(\frac{z}{\lambda}\right)^2 + 3\frac{z}{\lambda}\right]\right\}$$

Similarly,

$$B \quad \approx_{\lambda\to+\infty} \quad A_{-\infty}^0 + \frac{f(\underline{x})}{\lambda}\int_{-\infty}^0 \varphi(z)z\ dz + \mathcal{O}(\lambda^{-2})$$

and so

$$A + B \quad = \quad f(\underline{x}) + f'(\underline{x})dx + \frac{f''(\underline{x})dx^2}{6} - \frac{f(\underline{x})}{\lambda\sqrt{2\pi}} + \mathcal{O}(\lambda^{-2} + \ldots)$$

Thus,

$$\frac{\partial f}{\partial t}(\underline{x}) = -\frac{f(\underline{x})}{dt\lambda\sqrt{2\pi}} + o(1)$$

Now, $\lambda\to\infty$ but $dx\to 0$. So $\lambda dt \sim dx dt^{1/2} \to 0$. This implies:

$$\frac{\partial f}{\partial t}(\underline{x}) = -\infty$$

and thus

$$f(\underline{x}, t) = 0$$

for all times $t > 0$.

B.5.2 KFE and job losing rates in Lemma 2 and Proposition 2.

Pareto case: Proof of Lemma 2 Consider a single location ℓ and omit location subscripts ℓ. Thus, in logs $x = \log y$, the entry mass functio is $g_0(x) = g_0 e^{-\zeta(x-\underline{x})}$ where $\zeta = 1/z$. Denote g the invariant entry density, and $h(x) = g'(x)$. Then the KFE becomes

$$0 \quad = \quad \delta g(x) + s g'(x) + g_0 e^{-\zeta(x-\underline{x})}$$

The homogeneous solution is $h_H(x) = Ae^{-\kappa(x-\underline{x})}$. Varying the constant, I obtain

$$sA'(x)e^{-\kappa(x-\underline{x})} + g_0 e^{-\zeta(x-\underline{x})} = 0$$

and so

$$\begin{aligned} A(x) &= \tilde{A}_0 - \frac{g_0}{s}\int_0^{x-\underline{x}} e^{(\kappa-\zeta)t}dt \\ &= A_0 - \frac{n g_0}{s(\kappa-\zeta)} e^{(\kappa-\zeta)(x-\underline{x})} \end{aligned}$$

Therefore,

$$g'(x) = A_0 e^{-\kappa(x-\underline{x})} - \frac{g_0}{s(\kappa-\zeta)} e^{-\zeta(x-\underline{x})}$$

Given the integrability condition for g, the integration constants must cancel out, and

$$g(x) = Be^{-\kappa(x-\underline{x})} + \frac{g_0}{s\zeta(\kappa-\zeta)} e^{-\zeta(x-\underline{x})}$$

$f(\underline{x}) = 0$ then pins down B, so that

$$g(x) = \frac{g_0}{s\zeta(\kappa-\zeta)}\left[e^{-\zeta(x-\underline{x})} - e^{-\kappa(x-\underline{x})}\right] \geq 0$$

The separation flow is $sg'(\underline{x})$, which is given by

$$\varepsilon = \frac{g_0}{\zeta}$$

Normalizing g to 1 pins down g_0 which otherwise simply scales with the total mass of employed workers. This allows to compute the exit rate. Normalizing g to 1 yields $1 = \frac{g_0}{s\zeta(\kappa-\zeta)} \cdot \frac{\kappa-\zeta}{\zeta\kappa} = \frac{g_0}{s\zeta^2\kappa}$. Thus, the invariant density function is

$$f(x) = \frac{\zeta\kappa}{\kappa-\zeta}\left[e^{-\zeta(x-\underline{x})} - e^{-\kappa(x-\underline{x})}\right]$$

and the exit rate is

$$\varepsilon = s\zeta\kappa = \delta\zeta$$

To express job finding, it suffices to use the definition of workers' value of search. Under Assumption 1, they follow equation (35). The realized finding rate is thus

$$f_R(\ell) = f(\ell)\left(\frac{B}{b+v(\ell)}\right)^{1/z(\ell)} = \frac{\rho v(\ell)}{\beta(b+v(\ell))\bar{S}(z(\ell))} \equiv \Phi_R(v(\ell), 1/z(\ell))$$

Substituting in the definition of reservation wages delivers the expression for f_R in Proposition 2, with $\underline{w}_1 = b\Big(1-\beta+\beta\underline{y}_0/\rho\Big)$. The expression for the unemployment rate follows from the usual accounting equation. Under Assumption 2, re-arranging the worker's value of search yields

$$f_R(\ell) = \frac{\rho}{\beta}\frac{v(\ell)}{b+v(\ell)}\frac{1-G_0(\underline{y}(\ell)|z(\ell))}{\bar{S}(z(\ell), \underline{y}(\ell))}$$

General case: Proof of Lemma 5 The KFE now becomes

$$0 = \delta g'(x) + s g''(x) + g_0(x)$$

The homogeneous solution is the same as before. As before, I vary the constant and look for a solution $f'(x) = A(x)e^{-\kappa(x-\underline{x})}$, so that

$$A'(x) = -g_0(x)e^{-\kappa(x-\underline{x})}$$

Thus,

$$A(x) = A_0 + \int_x^{\infty} g_0(y)e^{\kappa(y-\underline{x})}dy$$

and hence

$$g'(x) = A_0 e^{-\kappa(x-\underline{x})} + e^{-\kappa(x-\underline{x})}\int_x^{\infty} g_0(y)e^{\kappa(y-\underline{x})}dy$$

Integrating once more:

$$g(x) \quad = \quad A + Be^{-\kappa(x-\underline{x})} - \int_x^{\infty} dy e^{-\kappa(y-\underline{x})}\int_y^{\infty} g_0(z)e^{\kappa(z-\underline{x})}dz = A + Be^{-\kappa(x-\underline{x})} - \frac{1}{\kappa}\int_x^{\infty} g_0(y)[e^{\kappa(y-x)} - 1]dy$$

Integrability imposes $A = 0$. B is determined by $g(\underline{x}) = 0$:

$$B = \frac{1}{\kappa}\int_{\underline{x}}^{\infty} g_0(y)[e^{\kappa(y-\underline{x})} - 1]dy$$

As before, the total mass of new jobs simply scales the invariant mass distribution. The separation flow is $sg'(\underline{x})$, where

$$g'(\underline{x}) = -\kappa B + \int_{\underline{x}}^{\infty} g_0(y)e^{\kappa(y-\underline{x})}dy = \int_{\underline{x}}^{\infty} g_0(y)dy$$

To get the separation rate, normalize g to 1. Denote by $H_0 = \int_{\underline{x}}^{\infty} g_0(y)dy$ the mass of newly created new jobs and $h_0 = g_0/H_0$ the entry density of new jobs. Using the expression for g above,

$$\frac{\kappa}{H_0} \quad = \quad \int_{\underline{x}}^{\infty} e^{-\kappa(x-\underline{x})}\int_{\underline{x}}^{x} e^{\kappa(y-\underline{x})}h_0(y)dy - \frac{1}{\kappa} + \int_{\underline{x}}^{\infty}\int_x^{\infty} h_0(y)dy = \int_{\underline{x}}^{\infty} x h_0(x)dx$$

Therefore the separate rate is

$$\frac{\delta}{\mathbb{E}^{h_0}[\log(y/\underline{y})]}$$

These arguments proove the following Lemma.

Lemma 5. *(Employment distribution)*
Denote by $g_0(y_0|z(\ell))$ the density function of successful new jobs. Then the invariant distribution g in location ℓ is

$$g(y,\ell) = B(\ell)\Big(y/\underline{y}(\ell)\Big)^{-\kappa} - \frac{1}{\kappa}\int_y^{\infty} g_0(y'|z(\ell))\Big(y'/\underline{y}(\ell)\Big)^{\kappa}\frac{dy'}{y'}$$

where $B(\ell) = \frac{1}{\kappa}\int_{\underline{y}(\ell)}^{\infty} g_0(y'|z(\ell))\Big(y'/\underline{y}(\ell)\Big)^{\kappa}\frac{dy'}{y'}$, and the job losing rate is

$$s(\ell) = \frac{\delta}{\int_{\underline{y}(\ell)}^{\infty}\left(\log\frac{y'}{\underline{y}(\ell)}\right)dy'}$$

B.6 Proof of Proposition 3

B.6.1 Pareto case

Impose Assumption 1 and consider dynamically stable steady-states. Then, PAM obtains. Denote again $\zeta = 1/z$. Because of PAM, labor market clearing in location ℓ writes

$$\theta(\ell) = -\frac{M_e f_\zeta(\zeta(\ell))\zeta'(\ell)}{u(\ell)L(\ell)f_\ell(\ell)}$$

and so

$$M_e f_\zeta(\zeta(\ell))\zeta'(\ell) = -L(\ell)u(\ell)\theta(\ell)f_\ell(\ell)$$

Using the expression of the finding rate in Proposition 2, re-express labor market tightness as a function of v, ζ:

$$\begin{aligned}\theta(\ell) &= \left[\frac{\rho}{\beta m}\frac{v(\ell)}{\big(b+v(\ell)\big)\left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)}\bar{\mathcal{S}}(\zeta(\ell))}\right]^{\frac{1}{1-\alpha}}\\ &\equiv \Theta\Big(v(\ell),\zeta(\ell)\Big)\end{aligned}$$

In what follows, it is useful to define the notation for the local unemployment rate

$$u(v(\ell),\zeta(\ell)) = \frac{\delta\zeta(\ell)}{\delta\zeta(\ell) + \Phi_R\big(v(\ell),\zeta(\ell)\big)}$$

Land market clearing writes in each location

$$r(\ell) = \omega L(\ell)\ell\Big[bu(v(\ell),\zeta(\ell)) + \big(1-u(v(\ell),\zeta(\ell))\big)(b+v(\ell))\big((1-\beta)+\beta\mathcal{E}(\zeta(\ell)\big)\Big]$$

where $\mathcal{E}(\zeta) = \underline{y}_0/\rho\frac{\kappa\zeta}{(\kappa-1)(\zeta-1)}$ is expected productivity under the invariant distribution from Lemma 2. Substituting into workers' free mobility condition $\rho U = \frac{\ell(b+v(\ell))}{r(\ell)^\omega}$, one can express population as

$$L(\ell) = U^{-\frac{1}{\omega}}\bar{L}(\ell, v(\ell), \zeta(\ell)) \tag{41}$$

with

$$\bar{L}(\ell, v(\ell),\zeta(\ell)) = \frac{1}{\omega\rho^{\frac{1}{\omega}}}\frac{\ell^{\frac{1}{\omega}-1}(b+v(\ell))^{\frac{1}{\omega}}}{bu(v(\ell),\zeta(\ell)) + (1-u(v(\ell),\zeta(\ell)))(b+v(\ell))\big((1-\beta)+\beta\mathcal{E}(\zeta(\ell)\big)}$$

Substitute back into labor market clearing:

$$K\zeta'(\ell) = -\frac{f_\ell(\ell)\tilde{L}(\ell,v(\ell),\zeta(\ell))\Theta\big(v(\ell),\zeta(\ell)\big)}{f_\zeta(\zeta(\ell))} \tag{42}$$

where $K = U^{\frac{1}{\omega}}M_e$ is a combined general equilibrium constant, and

$$\tilde{L}(\ell,v(\ell),\zeta(\ell)) = \bar{L}(\ell,v(\ell),\zeta(\ell))u(v(\ell),\zeta(\ell)) = \frac{1}{\omega\rho^{\frac{1}{\omega}}}\frac{\ell^{\frac{1}{\omega}-1}(b+v(\ell))^{\frac{1}{\omega}}}{b+\frac{\rho v(\ell)}{\delta\beta\zeta(\ell)\bar{\mathcal{S}}(\zeta(\ell))}\big((1-\beta)+\beta\mathcal{E}(\zeta(\ell)\big)}$$

Substituting into the FOC for v:

$$\begin{aligned}&\frac{v'(\ell)}{v(\ell)}\left[\alpha + \frac{v(\ell)}{b+v(\ell)}(\zeta(\ell)-1)\right] \\ = \; &\frac{1-\alpha}{\ell} - \frac{1}{K}\times\alpha\left(\frac{\bar{\mathcal{S}}'(\zeta(\ell))}{\bar{\mathcal{S}}(\zeta(\ell))} + \log\frac{B}{b+v(\ell)}\right)\frac{\tilde{L}(\ell,v(\ell),\zeta(\ell))\Theta\big(v(\ell),\zeta(\ell)\big)}{f_\zeta(\zeta(\ell))}\end{aligned} \tag{43}$$

Given K, equations (42)-(43) define a coupled system of ODEs, with two boundary conditions:

$$\zeta(\underline{\ell}) = \overline{\zeta} \qquad ; \qquad \zeta(\overline{\ell}) = \underline{\zeta}$$

Inspection of (42)-(43) indicate that the system satisfies standard regularity conditions for a unique solution to obtain if it has two initial conditions. The present system, however, has one initial and one terminal condition.

Existence of a solution to the ODE system given K. Denote $\underline{v} = v(\underline{\ell})$. Given K, inspection of (42)-(43) reveals that the system is Lipschitz continuous. Given $\underline{v}, \underline{\zeta}$ and K, there thus exists a unique solution to (42)-(43). The idea is now to study how changes in $\underline{v}$ affect $\zeta(\overline{\ell})$ in the solution to that system. Lipschitz continuity ensures that $\zeta(\overline{\ell})$ is a continuous function of $\underline{v}$. Further inspection of (42)-(43) reveals that as $v \to 0$, so do Z, V. Similarly, as $\underline{v} \to +\infty$, so do Z, V. Therefore, the same conclusion holds when $\underline{v} \to 0$ or $\underline{v} \to +\infty$. Hence, there exists at least on $\underline{v}(K)$ such that $\zeta(\overline{\ell}) = \underline{\zeta}$.

Existence of K. The equilibrium has a block-recursive structure. Free-entry alone is enough to determine K without using population adding up. Given K and thus the solution (b, ζ), population adding-up immediately determines U as per (41). Thus, it suffices to show that free-entry implies existence of K. Free-entry can be re-written

$$\begin{aligned} K \cdot c_e &= J_0 \int \left[B^{\zeta(\ell)} (b + v(\ell))^{1-\zeta(\ell)} v(\ell)^{-\alpha} \bar{\mathcal{S}}(\zeta(\ell)) \right]^{\frac{1}{1-\alpha}} \cdot \ell \cdot \bar{L}(\ell) u(\ell) \theta(\ell) d\ell \\ &= J_0' \int v(\ell) \ell \tilde{L}(\ell, v(\ell), \zeta(\ell)) d\ell \end{aligned}$$

As $K \to 0$, (42) together with the boundary conditions on ζ and an application of Rolle's theorem to $\zeta'(\ell)$ implies $\underline{v}(K)^{\frac{1}{1-\alpha}} \sim K \to 0$. As $K \to +\infty$, a similar argument implies $\underline{v}(K)^{\frac{1}{\omega}-1+\zeta_0} \sim K \to +\infty$, where $\zeta_0 \in [\underline{\zeta}, \overline{\zeta}]$. Thus, the right-hand-side integral of free-entry is of order $K^{1-\alpha}$ as $K \to 0$, and is of order $K^{\frac{1}{\zeta_0 + 1 - \omega}}$ as $K \to +\infty$. Since $\underline{\zeta} > 1$ by assumption, $\frac{1}{\zeta_0 + 1 - \omega} < 1$. Therefore, there exists at least one solution K to the free-entry condition.

Uniqueness. Now suppose that the supports of F_ℓ, F_z are small enough. This assumption allows to use a first-order approximation to the ODE system (42)-(43). In that case, to a first order,

$$K\zeta'(\ell) \approx -\frac{\tilde{L}(\ell, \underline{v}, \overline{\zeta}) \Theta(\underline{v}, \overline{\zeta})}{f_\zeta(\overline{\zeta})} = -L_0 \frac{\underline{v}^{\frac{1}{1-\alpha}}}{1 + L_1 \underline{v}} (b + \underline{v})^{\frac{1}{\omega} + \frac{\overline{\zeta}-1}{1-\alpha}} \tag{44}$$

where $L_0, L_1 > 0$ are transformations of parameters. Integrating, it implies

$$K = L_0' \frac{\underline{v}^{\frac{1}{1-\alpha}}}{1 + L_1 \underline{v}} (b + \underline{v})^{\frac{1}{\omega} + \frac{\overline{\zeta}-1}{1-\alpha}} \tag{45}$$

where $L_0' = L_0 \frac{\overline{\ell} - \underline{\ell}}{\overline{\zeta} - \underline{\zeta}}$ only depends on parameters. Similarly, free entry can be approximated to a first order by

$$K = J_0'' \frac{\underline{v}}{\underline{v} + 1/L_1} (b + \underline{v})^{\frac{1}{\omega}} \tag{46}$$

where J_0'', J_1 depend only on parameters. Substituting (46) into (45), one obtains

$$1 = L_0'' \underline{v}^{\frac{\alpha}{1-\alpha}} (b + \underline{v})^{\frac{\overline{\zeta}-1}{1-\alpha}} \tag{47}$$

where L_0'' depends only on parameters. The right-hand-side of (47) is strictly increasing in $\underline{v}$, and so (47) uniquely pins down $\underline{v}$. Then (46) uniquely pins down K. Then $\int L(\ell) F_\ell(d\ell)$ uniquely pins down U.

B.6.2 Non-Pareto case

When the entry distribution satisfies Assumption 2, all the previous arguments continue to hold. They only involve additional notation. Hence, I omit them for brevity.

B.7 Limiting economies

B.7.1 Proof of Corollary 1

This limiting economy preserve a wide support for F_z but considers the limit of a small support for F_ℓ. In that case, it is more useful to index locations by their value of search v rather than productivity ℓ.

Shrinking the support of F_ℓ implies $\ell'(v) = 0$. Thus, the FOC (38) implies

$$\alpha + \frac{v}{b+v}(\zeta(v) - 1) + \alpha v \left(\frac{\bar{\mathcal{S}}'(\zeta(v))}{\bar{\mathcal{S}}(\zeta(v))} + \log \frac{B}{b+v} \right) (-\zeta'(v)) = 0 \tag{48}$$

which defines a non-degenerate assignment $\zeta(v)$ in the limit. Given the boundary conditions, it must be that there is an interval of v's in the limit. The assignment $\zeta(v)$ implies non-vanishing dispersion in job losing and unemployment rates.

B.7.2 Vanishing search frictions

Suppose that labor market frictions become small. Should cities be expected to have similiar labor market flows? Corollary 2 below shows that cities remain different event with small labor market frictions.

Corollary 2. *(Vanishing spatial differences with vanishing search friction)*
Suppose that the conditions in Proposition 3 hold and that $\alpha > 0$. The variance of local job losing and unemployment rates remain strictly positive and bounded above zero as the matching function efficiency becomes large ($m \to +\infty$).

This result highlights that taking into account even small labor market frictions leads to substantial departures from a model that would not feature frictional unemployment. Spatial unemployment differentials survive when search frictions become small because reservation wages rise everywhere as meeting with new employers becomes easier. This adjustment offsets the direct unemployment benefits of higher contact rates.[77]

Proof. Consider the limit $m \to +\infty$. Denote $K(m) = m^{\frac{1}{1-\alpha}} M(\rho U)^{\frac{1}{\omega}}$ and $\theta_0(\ell) = \theta(\ell) m^{\frac{1}{1-\alpha}}$. I now distinguish three cases for the candidate limit as $m \to +\infty$. Impose Assumption 1 for now.

- Suppose that $v \sim 1$ remains finite.
  Then $f_R, u, \hat{L} \sim 1$ and so free-entry ensures that $K(m) \sim 1$.
- Suppose that $v \to +\infty$.
  Then $f_R, u \sim 1$. Then $\hat{L} \sim v^{\frac{1}{\omega}-1} \to +\infty$. Free-entry then implies $K(m) \sim ||v||^{\frac{1}{\omega}} \to +\infty$. Labor market clearing requires $||v||^{\frac{1}{\omega}} ||\zeta'|| \sim ||v||^{\frac{1}{\omega} + \frac{\zeta}{1-\alpha} - 1}$ and so $||\zeta'|| \to \infty$, violating the boundary conditions for ζ.
- Suppose that $v \to 0$. Then $\hat{L} \sim 1$, and from free-entry $K(m) \to 0$. From population adding up, $U \sim 1$. The definition of $K(m)$ then implies $M_e \to 0$. This cannot be an equilibrium.

Therefore, in the limit $m \to +\infty$, v remains finite and the assignement is non-degenerate. The same arguments hold under Assumption 2.

□

B.8 Planning solution

B.8.1 Optimality conditions

The planner solves chooses the number of unemployed workers $\mathcal{U}(t,\ell)$ to locate in each city ℓ at time t, the rate at which to break up existing matches $\Delta(t,y,\ell)$. The planner also chooses the consumption $c_U(t,\ell), c_E(t,y,\ell)$, $h_U(t,\ell), h_E(t,y,\ell)$ of employed and unemployed workers, as well as the consumption of the owners $C(t)$. For simplicity, I assume that unemployed workers produce $b\ell$ at home. I anticipate that the planner chooses PAM, so that it suffices to let the planner choose the matching function $\zeta(t,\ell)$ together with its slope $\xi(t,\ell)$. I assume for simplicity that the planner cannot transfer resources across time periods. Because I will focus on steady-states, this assumption is without loss of generality.

I denote by $\lambda(\ell)$ the planner's weight on individuals who live in location ℓ. Due to complementaries between housing and final good consumption in the utility function, the spatial redistribution of the final good is not neutral.

[77] See Bilal et al. (2019) and Martellini and Menzio (2018) for related results.

Only one particular set of weights implements an allocation that resembles the decentralized equilibrium, which will be the focus of this paper.[78]

The planner's objective is then

$$\begin{aligned} W &= \int_0^{\infty} dt e^{-\rho t} \int d\ell f_\ell(\ell)\lambda(\ell)\Bigg\{\mathcal{U}(t,\ell)\left(\frac{c_U(t,\ell)}{1-\omega}\right)^{1-\omega}\left(\frac{h_U(t,\ell)}{\omega}\right)^{\omega} \\ &\qquad + \int \mathcal{E}(t,y,\ell)\left(\frac{c_E(t,\ell)}{1-\omega}\right)^{1-\omega}\left(\frac{h_E(t,\ell)}{\omega}\right)^{\omega} dy\Bigg\} \\ &\quad + \int_0^{\infty} e^{-\rho t} C(t) dt \end{aligned}$$

where $\mathcal{E}$ denotes the mass distribution of employment across productivity y in location ℓ at time t. The last term is the welfare of the owners. The planner is subject to the constraints

$$\begin{aligned} \forall t,\ 1 &= \int d\ell f_\ell(\ell)\left\{\mathcal{U}(t,\ell) + \int \mathcal{E}(t,y,\ell)dy\right\} \\ \forall t,\ell,\ 1 &= \mathcal{U}(t,\ell)h_U(t,\ell) + \int \mathcal{E}(t,y,\ell)h_E(t,y,\ell)dy \\ \forall t,\ell\ 0 &= \int f_\ell(\ell)\Big\{\mathcal{U}(t,\ell)\Big(b\ell - c_U(t,\ell)\Big) + \int \mathcal{E}(t,y,\ell)\Big(y\ell - c_E(t,y,\ell)\Big)dy\Big\} - C(t) - c_e M_e(t) \\ \forall y,\ell,t,\ \frac{\partial \mathcal{E}}{\partial t}(t,y,\ell) &= L_y^*\mathcal{E}(t,y,\ell) + n(M_e(t),\xi(t,\ell),\mathcal{U}(t,\ell))g_0(y,\zeta(t,\ell)) - \Delta(t,y,\ell)\mathcal{E}(t,y,\ell) \\ \forall t,\ \zeta(t,\underline{\ell}) &= \overline{\zeta} \\ \forall t,\ \zeta(t,\overline{\ell}) &= \underline{\zeta} \\ \forall t,\ell,\ \int_{\underline{\ell}}^{\ell} \xi(t,x)dx &= 1 - F_\zeta(\zeta(t,\ell)) \\ \forall t\ n(M(t),\xi(t,\ell),\mathcal{U}(t,\ell)) &= m\Big(M(t)\xi(t,\ell)\Big)^{1-\alpha}\mathcal{U}(t,\ell)^{\alpha} \end{aligned}$$

The first constraint simply states that total population is one in the economy. The second constraint clears the land market in each location. The third constraints is the planner's aggregate resource constraint. The fourth constraint is the time-dependent KFE that encodes how the distribution of employment across productivity evolves over time. The fifth and sixth constraints are the boundary conditions for the assignment function, i.e. the location choice of jobs. The seventh constraint is simply the definition of ξ, which is the slope of the assignment function that enters into labor market tightness. The eighth constraint simply states the matching function.

The structure of the planning problem is standard. The only non-standard element is that the planner controls a full distribution of workers in each location. This distribution $\mathcal{E}$ is an infinite-dimensional object. To use standard convex optimization methods described in Luenberger (1997), some regularity conditions must be imposed on the functional space in which the distribution $\mathcal{E}$ is allowed to lie. I build on ideas developed in Moll and Nuño (2018), who propose functional spaces for such cases. There are several differences between their approach and the one in this paper. First, their results do not directly apply because of the endogenous separation margin and I must start from first principles. Second, their method in fact requires further restrictions on the functional spaces that those they outline. They propose to use square integrable functions of time and other states (section 2.1.2 p. 154). This restriction is in fact not quite sufficient for their Theorem 2 p. 168 to obtain. The reason is that Luenberger (1997)'s Theorem 1 p. 243 that they refer to also requires that the transition operator that encodes the evolution equation of the state, maps into a Banach space. Yet, there is in general no guarantee that a functional operator like a continuous-time transition operator L_y^* maps the space of square-integrable functions into a Banach space.[79] For it to map into a Banach space, the functional space in which the distribution lies must be further restricted.

It suffices to impose that the distribution $\mathcal{E}$ lies in a Sobolev-Strichartz space, which is a variant of Sobolev

[78] An alternative assumption to choosing one particular set of weights is that the planner has to provide consumption to workers with locally produced final goods.

[79] I thank Ben Moll and Galo Nuño for related discussions.

spaces:

$$\begin{aligned}H^{1,2} \quad \equiv \quad & \Big\{\mathcal{E}: \text{ for all } g \text{ among } \mathcal{E}, \text{ its first } t,y,\ell\text{-weak derivatives,} \\ & \text{and second } y,\ell\text{-weak derivatives,} \\ & \int_0^\infty e^{-\rho t}\left(\iint |\mathcal{E}(t,y,\ell)|^2 dy d\ell\right) dt < \infty\Big\}\end{aligned}$$

Sobolev-Strichartz spaces are useful precisely because infinitesimal generators such as L_y^* map Sobolev-Strichartz spaces into Lebesgue space (see Tao, 2006).

Finally, the approach I use builds on duality methods similar to Moll and Nuño (2018). However, these duality methods apply without loss of generality in my setup because the distribution endogenously satisfies the boundary condition $\mathcal{E} = 0$ at the lower point of the support. In contrast, Moll and Nuño (2018) apply duality methods for general distributions. However, when the distribution does not vanish at the edge of bounded supports, additional terms should appear in their results.

I am now ready to formulate a current-value Hamiltonian (which is equivalent to a Lagrangian):

$$\begin{aligned}H \quad = \quad & \int d\ell f_\ell(\ell)\lambda(\ell)\Big\{\mathcal{U}(t,\ell)\left(\frac{c_U(t,\ell)}{1-\omega}\right)^{1-\omega}\left(\frac{h_U(t,\ell)}{\omega}\right)^{\omega} \\ & \qquad + \int \mathcal{E}(t,y,\ell)\left(\frac{c_E(t,\ell)}{1-\omega}\right)^{1-\omega}\left(\frac{h_E(t,\ell)}{\omega}\right)^{\omega} dy\Big\} \\ - \quad & c_e M_e(t) + \int f_\ell(\ell)\Big\{\mathcal{U}(t,\ell)\Big(b\ell - c_U(t,\ell)\Big) + \int \mathcal{E}(t,y,\ell)\mathcal{U}(t,\ell)\Big(y\ell - c_E(t,y,\ell)\Big)\Big\} \\ + \quad & \iint dz d\ell f_\ell(\ell)\Big[\mathcal{E}(t,y,\ell)LS(t,y,\ell) \\ & \qquad + n(M(t),\xi(t,\ell),\mathcal{U}(t,\ell))g_0(y,\zeta(t,\ell))S(t,y,\ell) - \Delta(t,z)\mathcal{E}(t,z,\ell)S(t,y,\ell)\Big] \\ + \quad & \rho U(t)\left[1 - \int d\ell f_\ell(\ell)\left(\mathcal{U}(t,\ell) + \int \mathcal{E}(t,y,\ell)dy\right)\right] \\ + \quad & \int \Big(\partial_\ell \pi(t,\ell)\Big) F_\zeta(\zeta(t,\ell))d\ell - \int \xi(t,\ell)\pi(t,\ell) \\ + \quad & \tau(t)\left[1 - \int \xi(t,\ell)d\ell\right] \\ + \quad & \int d\ell f_\ell(\ell) r(t,\ell)\Big\{1 - \mathcal{U}(t,\ell)h_U(t,\ell) - \int \mathcal{E}(t,y,\ell)h_E(t,y,\ell)dy\Big\}\end{aligned}$$

where I have substituted out the consumption of owners using the aggregate budget constraint. I have integrated by parts the ξ constraint with multiplier A, and denoted $\pi(t,\ell) = -\int_{\ell}^{\bar{\ell}} A(t,x)dx$. I have an ad adding up constraint for total employment in each location. S is the multiplier attached to the KFE constraint, which I also integrated by parts. I have also combined the multipliers on the resource constraints, without loss of generality.

Consumption and housing. Optimality of consumption and housing choices in steady-state delivers

$$\begin{aligned}1 \quad &= \quad \frac{(1-\omega)\lambda(\ell)}{c_U(\ell)}u(c_U(\ell),h_U(\ell)) = \frac{(1-\omega)\lambda(\ell)}{c_E(y,\ell)}u(c_E(y,\ell),h_E(y,\ell)) \\ r(\ell) \quad &= \quad \frac{\omega\lambda(\ell)}{h_U(\ell)}u(c_U(\ell),h_U(\ell)) = \frac{\omega\lambda(\ell)}{h_E(y,\ell)}u(c_E(y,\ell),h_E(y,\ell))\end{aligned}$$

Re-arranging,

$$\frac{rh_i}{c_i} = \frac{\omega}{1-\omega} \qquad ; \qquad u_i = r^{-\omega} c_i (1-\omega)^{-1}$$

which then implies $r^\omega = \lambda$. Land market clearing in every location re-writes

$$\frac{\omega}{1-\omega} r(\ell) = \mathcal{U}(\ell) c_U(\ell) + \int \mathcal{E}(y,\ell) c_E(y,\ell) dy$$

where the second equality follows from the the local budget constraint. Vayring the weights $\lambda(\ell)$ thus $r(\ell) = \lambda(\ell)^\omega$, and traces out the Pareto frontier of this economy. To keep the focus on the inefficiency in the location choice of employers, I choose the specific set of weights $\lambda(\ell)$ such that the land market clearing coincides with its decentralized equilibrium counterpart when $\beta = \alpha$. Namely, I *choose* $\lambda(\ell)$ such that

$$\frac{\lambda(\ell)^\omega}{\omega} = \mathcal{U}(\ell) b\ell + \int \mathcal{E}(y,\ell) \Big[(1-\alpha)(b + v(\ell))\ell + \alpha y \ell \Big] dy \tag{49}$$

where $v(\ell)$ is defined below.

Allocation of workers. I now take FOCs w.r.t. $\mathcal{U}, \Delta$ and impose steady-state. Starting with $\mathcal{U}$:

$$\lambda(\ell) u(c_U(\ell), h_U(\ell)) + \alpha n(\ell) \int g_0 S + (b\ell - c_U(\ell)) - r(\ell) h_U - \rho U = 0$$

Using the previous FOCs to obtain that $\lambda u_U = c_U + r h_U$, and denoting $v(\ell)\ell r(\ell)^{-\omega} = \alpha n(\ell) \int g_0 S$, I obtain

$$\rho U = \frac{(b + v(\ell))\ell}{r(\ell)^\omega}$$

I guess that for now, the definition of v does not depend on r. The co-state equation for $\mathcal{E}$ is then

$$\rho S \quad = \quad u(c_E(y,\ell), h_E(y,\ell)) + LS - \rho U + (y\ell - c_E(y,\ell)) - r(\ell) I(\ell) h_E(y,\ell) - \Delta S$$

Re-arranging similarly to the $\mathcal{U}$ FOC,

$$\rho S \quad = \quad \frac{\Big(y - (b + v(\ell)) \Big) \ell}{r(\ell)^\omega} + LS - \Delta S$$

Finally, the FOC for Δ yields

$$\Delta = \begin{cases} 0 & \text{if} \quad S \geq 0 \\ +\infty & \text{if} \quad S < 0 \end{cases}$$

Therefore, $X = r(\ell)^\omega S$ solves $\rho X = (y - (b + v(\ell))\ell)\ell + LX$ in the continuation region. Hence, v is defined as $v(\ell)\ell = \alpha n(\ell) \int g_0 X$. Together, these define a pair of equations that does not directly depend on r. Thus, the guess that the definition of v does not depend on r is verified.

These multipliers correspond exactly to the shadow values of unemployed and employed workers when $\beta = \alpha$. The planner breaks up matches when the surplus S is negative, and thus the recursion for S has the same solution as in the decentralized equilibrium when replacing β with α.

Allocation of jobs. The FOC for M_e is

$$c_e \quad = \quad (1-\alpha) \frac{1}{M} \int d\ell f_\ell(\ell) \; n(\ell) \int g_0(y,\ell) S(y,\ell) dy$$

The FOCs for ξ and ζ are then

$$\begin{aligned}[\pi+\tau]\xi &= (1-\alpha)nf_\ell \int gSdz \\ 0 &= f_\ell n \cdot \left(\int \frac{\partial_\zeta g_0}{g_0} g_0 Sdy\right) + \frac{\pi'(\ell)}{\pi(\ell)+\tau}\cdot[\pi(\ell)+\tau]f_\zeta(\zeta)\end{aligned}$$

Denote $J(\ell)\equiv \tau+\pi(\ell)$ and so simplifying out f_ℓ

$$n\left(\int \frac{\partial_\zeta g_0}{g_0}\cdot g_0 Sdy\right) + \frac{J'(\ell)}{J(\ell)}\cdot\frac{f_\zeta(\zeta)}{\xi}\cdot(1-\alpha)n\int g_0 Sdy = 0$$

and hence

$$(1-\alpha)\frac{J'}{J\zeta'} = \frac{\int \frac{\partial_\zeta g_0}{g_0}\cdot g_0 Sdy}{\int g_0 Sdy}$$

Using the known solution to S, one obtains

$$\frac{\int \frac{\partial_\zeta g_0}{g_0}\cdot g_0 Sdy}{\int g_0 Sdy} = \frac{\bar{\mathcal{S}}'(\zeta)}{\bar{\mathcal{S}}(\zeta)} + \log\frac{B_0}{b+v(\ell)}$$

Finally, changing variables to $J(\ell)\equiv J(\zeta(\ell))$:

$$(1-\alpha)\frac{J'(\zeta)}{J(\zeta)} = \frac{\bar{\mathcal{S}}'(\zeta)}{\bar{\mathcal{S}}(\zeta)} + \log\frac{B_0}{b+v(\zeta)}$$

This equation corresponds to an envelope condition of the decentralized equilibrium. which coincides with the FOC/envelope condition from the competitive equilibrium.

Re-write the ξ FOC as

$$J(\ell)\xi(\ell) = (1-\alpha)q(\ell)\cdot M\xi(\ell)\cdot\int g_0 S$$

where I have used that, by definition of $q(\ell)$, $n(\ell) = q(\ell)\cdot M_e\xi(\ell)$. Thus,

$$\frac{J(\ell)}{\rho M_e(1-\alpha)} = q(\ell)\ell\left(\frac{B_0}{b+v(\ell)}\right)^{\zeta(\ell)}(b+v(\ell))\bar{\mathcal{S}}(\zeta(\ell))$$

Finally, use the definition of v to substitute q out:

$$J(\ell)^{1-\alpha} \propto \ell^{1-\alpha}v(\ell)^{-\alpha}\cdot\left(\frac{B_0}{b+v(\ell)}\right)^{\zeta(\ell)}(b+v(\ell))\bar{\mathcal{S}}(\zeta(\ell))$$

Then using the envelope condition from above, I obtain the FOC for v:

$$-\frac{v'(\ell)}{v(\ell)}\left[\alpha+\frac{v(\ell)}{b+v(\ell)}(\zeta(\ell)-1)\right] + \frac{1-\alpha}{\ell} + 0 = 0 \tag{50}$$

This FOC resembles the one in the decentralized equilibrium, except that it does not have the last term : $\left(\frac{\bar{\mathcal{S}}'(\zeta)}{\bar{\mathcal{S}}(\zeta)} + \log\frac{B_0}{b+v(\zeta)}\right)\zeta'(\ell)$. This last term is the labor market pooling externality that the planner internalizes. Finally, I can go back to the entry FOC, which re-writes:

$$c_e = (1-\alpha)\int d\ell\; q(\ell)f_\zeta(\zeta(\ell)|\zeta'(\ell)|\ell\left(\frac{B_0}{b+v(\ell)}\right)^{\zeta(\ell)}(b+v(\ell))\bar{\mathcal{S}}(\zeta(\ell))$$

which correponds to the free-entry condition when $\beta=\alpha$.

B.8.2 Proof of Proposition 4

Having laid out the planner's optimality conditions, I can now turn to the proof of Proposition 4.

Extensions of decentralized equilibrium results. Comparing the v FOC tin the decentralized equilibrium (37) and in the planning solution (50), the labor market pooling externality immediately arises. Except for this discrepancy, inspecting the the planner's optimality conditions reveal that they are identical to the decentralized equilibrium's when $\beta = \alpha$. Therefore, Propositions 1, 2 and 3 extend under the same conditions.

Efficiency. Due to the labor market pooling externality term in the v FOC tin the decentralized equilibrium (37) relative to the planning solution (50), the decentralized equilibrium is inefficient as soon as $\alpha > 0$.

Comparison of allocations. In the linearized case of small supports for F_ℓ, F_z and when $\beta = \alpha$, it is possible to compare the assignment functions. From (46) and (47), $\underline{v}$ and K are identical in both the decentralized equilibrium and the planner solution to a first order. But then from the FOC (43), $\frac{\ell v'(\ell)}{v(\ell)}$ is larger in the decentralized equilibrium due to the labor market pooling externality term. Given that $\frac{\ell \underline{w}'(\ell)}{\underline{w}(\ell)} \approx \frac{v}{v+\underline{v}} \frac{\ell v'(\ell)}{v(\ell)}$ to a first order, the comparison between reservation wages obtains.

For the comparison between assignmen functions z, it is uesful to start from (42). Re-arranging its first-order approximation delivers the first-order approximation to $v(\ell) \approx \underline{v}\big(\ell/\underline{\ell}\big)^{v_1}$, where v_1 is a constant that depends only on parameters. v_1 is higher in the decentralized equilibrium due to the labor market pooling externality. A coomin solution method in ODEs is to "bootstrap" successive approximation to derive higher orders. I follow this method in spirit and substitute back this first-order approximation into (42) and re-arrange to obtain

$$\zeta(\ell) \approx 1 + \left(1 + \frac{b}{\underline{v}}(\ell/\underline{\ell})^{-v_1}\right)\left(1 - \alpha + \alpha I_0\right)$$

where $I_0 > 0$ in the decentralized equilibrium and I_0 in the planner's solution. Using the boundary conditions, one obtains

$$\frac{z(x) - \underline{z}}{\overline{z} - \underline{z}} = \frac{x^{v_1} - 1}{\overline{x}^{v_1} - 1}$$

where $x = \frac{\ell}{\underline{\ell}}$. This functional form implies $z^{DE} < z^{SP}$ except at the boundaries.

Limit of identical locations. Consider the location FOC(48) when there is no dispersion in ℓ. It holds only if there is dispersion in v. Without the inefficiency, the last term on the left-hand-side is zero. Therefore, it implies $\frac{v}{b+v}(\zeta(v) - 1) = -\alpha < 0$ which is a contradiction. Therefore, there can be no dispersion in v in the planner's solution.

Directed search. I first briefly describe the economy with directed search. Then I show how the values of workers and employers change. Finally, I show that the location choice of employers coincides with the planner's choice.

Setup. Employers can commit to fully state-contingent contracts that promise a stream of wage payments. For simplicity, I assume without loss of generality that these contracts must be Markovian. Within each location, there can be a continuum of submarkets indexed by their contract. Workers perfectly observe each contract and each submarket and direct their search across submarkets. Once they choose a submarket, they queue and wait until they meet the employers. Meetings in each submarket are created according to the same matching function as in the random search model.

Values. The value of unemployment satisfies

$$\rho U = \frac{b\ell}{r(\ell)^\omega} + \max_{\theta \in \Theta_\ell} f(\theta) \frac{s(\ell, \theta)}{r(\ell)^\omega}$$

where without loss of generality each submarket in location ℓ is indexed by its labor market tightness θ which lies in the set Θ_ℓ. $s(\ell, \theta)$ denotes the promised value to the worker. The value of employment at wage w is

$$\rho V^E(w, \ell) = \frac{w}{r(\ell)^\omega} + L_w V^E$$

Then, $V^E - U$ solves

$$\rho\Big(V^E(w,\ell) - U(\ell)\Big) \quad = \quad \frac{w - b\ell - V(\ell)}{r(\ell)^\omega} + L_w(V^E - U)$$

where I denote

$$V(\ell) = \max_{\theta \in \Theta_\ell} f(\theta)\frac{s(\theta,\ell)}{r(\ell)^\omega}$$

the value of search in location ℓ. Denote also $v(\ell) = \frac{r(\ell)^\omega V(\ell)}{\ell}$ the value of search relative to productivity.

It is also useful to define the adjusted surplus, which satisfies

$$\rho J(\ell,y) \quad = \quad y\ell - [V(\ell) + b\ell] + L_y S$$

with boundary conditions identical to the random search case. Thus, Lemma 1 applies. The value of emloyer $\zeta = 1/z$ in location ℓ is then

$$J(\zeta,\ell) \quad = \quad \max_{\theta \in \Theta_\ell}\left\{q(\theta)\mathbb{E}_{\zeta,\ell}\Big[S(\ell,z) - s(\theta,\ell)\Big]\right\}$$

Substituting the definition of V to express tightness[80] as a function of the surplus s,

$$\begin{aligned} J(\zeta,\ell) \quad &= \quad V(\ell)^{-\frac{\alpha}{1-\alpha}} m^{\frac{1}{1-\alpha}} \max_s \left\{ s^{\frac{\alpha}{1-\alpha}} \mathbb{E}_{\zeta,\ell}\Big[S(\ell,z) - s\Big]\right\} \\ &= \quad V(\ell)^{-\frac{\alpha}{1-\alpha}} m^{\frac{1}{1-\alpha}} \max_{\hat{s} = s^{\frac{1}{1-\alpha}}} \left\{\mathbb{E}_{\zeta,\ell}[S(\ell,z)] \cdot \hat{s}^\alpha - \hat{s}\right\} \end{aligned}$$

This results in

$$\begin{aligned} s(\theta(\zeta,\ell),\ell) \quad &= \quad \alpha \mathbb{E}_{\zeta,\ell}[S(\ell,z)] \\ \theta(\zeta,\ell)^{1-\alpha} \quad &= \quad \frac{V(\ell)}{\alpha m \mathbb{E}_{\zeta,\ell}[S(\ell,z)]} \end{aligned}$$

and hence

$$J(\zeta,\ell) = \left\{\frac{(1-\alpha)^{1-\alpha}}{\alpha^\alpha} m \mathbb{E}_{\zeta,\ell}[S(\ell,z)] V(\ell)^{-\alpha}\right\}^{\frac{1}{1-\alpha}}$$

Location choice. Using Lemma 1, the value of having entering in location ℓ for employer ζ is

$$\rho J(\zeta,\ell) = \left\{\frac{(1-\alpha)^{1-\alpha}}{\alpha^\alpha} m \left(\frac{B}{b + v(\ell)}\right)^\zeta \big(b + v(\ell)\big) v(\ell)^{-\alpha} \cdot \ell^{1-\alpha} \cdot \bar{\mathcal{S}}(\zeta)\right\}^{\frac{1}{1-\alpha}}$$

which coincides with the planner's valuation.

B.9 Optimal policy

I consider five possible taxes and subsidies:

- A wage tax paid by the employer τ_w
- A profit tax τ_π
- An unemployment benefits tax τ_b
- A value added tax τ_{va}
- An employment tax $\tau_e \ell$ paid by the employer, where it is useful to define $\tau_n = \frac{\tau_e}{\tau_b \tau_w}$

[80] From the worker's indifference condition, $q(s) = V(\ell)^{-\frac{\alpha}{1-\alpha}} m^{\frac{1}{1-\alpha}} s^{\frac{\alpha}{1-\alpha}}$

Using Lemma 3, these taxes affect the decentralized equilibrium as follows.

- Effective output is $\tau_{va} y\ell$
- Unemployment benefits are $b\ell\tau_b$
- The negotiated wage is $w^* = (1-\beta)[b\tau_b + v(\ell) + \tau_e]\ell + \beta\frac{\tau_{va}\cdot z\ell}{\tau_w}$
- Employer values scale with τ_π

These taxes results in flow values for employers

$$\begin{aligned} J_0(y,\ell) &\equiv (1-\beta)\tau_\pi(\tau_{va}\cdot y - \tau_e - \tau_w\tau_b b - \tau_w v(\ell)\ell \\ &= \tau_{va}\tau_\pi(1-\beta)\left(z - \frac{\tau_e}{\tau_{va}} - \frac{\tau_w\tau_b}{\tau_{va}}\cdot b - \frac{\tau_w}{\tau_{va}}v(\ell)\right)\ell \end{aligned}$$

and workers

$$V_0 \equiv \beta\left(\frac{\tau_{va}z-\tau_e}{\tau_w}\cdot z - b\tau_b - \mathbb{E}[V^E - U]\right)\ell = \frac{\tau_{va}}{\tau_w}\cdot\beta\left(z - \frac{\tau_e}{\tau_{va}} - b\cdot\frac{\tau_w\tau_b}{\tau_{va}} - \frac{\tau_w}{\tau_{va}}\mathbb{E}[V^E-U])\right)\ell$$

The endogenous separation cutoff is

$$\underline{y} \propto \frac{\tau_e}{\tau_{va}} + \frac{\tau_w\tau_b}{\tau_{va}}\cdot b + \frac{\tau_w}{\tau_{va}}v = \frac{\tau_w}{\tau_{va}}\cdot\left(\frac{\tau_e}{\tau_w} + \tau_b b + v\right) = \frac{\tau_w\tau_b}{\tau_{va}}\cdot\left(\tau_n + b + \tilde{v}\right)$$

where $\tilde{v} = \tau_b v$. Finally, solving the worker's problem, one obtains

$$v(\ell) = \beta f(\ell)\frac{\tau_{va}}{\tau_w}\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)}\underline{y}(\ell)\bar{S}(\zeta(\ell))$$

and so

$$\tilde{v}(\ell) = \frac{\tau_{va}}{\tau_w\tau_b}\beta\cdot f(\ell)\cdot\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)}\underline{y}(\ell)\bar{S}(\zeta(\ell))$$

Denoting $T = \frac{\tau_{va}}{\tau_w\tau_b}$ one obtains Then

$$\tilde{v} = c_2 T\underline{y} - b - \tau_n$$

for a constant $c_2 > 0$, and one can use the worker's value of search to re-write

$$c_2\underline{y} - \frac{b+\tau_n}{T} = \beta c_1 f(\ell)\cdot\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)}\underline{y}(\ell)\bar{S}(\zeta(\ell))$$

which implies

$$q(\ell) \propto \left(c_2\underline{y} - \frac{b+\tau_n}{T}\right)^{-\frac{\alpha}{1-\alpha}}\cdot\beta^{\frac{\alpha}{1-\alpha}}\cdot\left[\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)}\underline{y}(\ell)\bar{S}(\zeta(\ell))\right]^{\frac{\alpha}{1-\alpha}}$$

The employers's expected value is then

$$\begin{aligned} J(\ell,\zeta)^{1-\alpha} &= \left(\frac{1-\beta}{1-\alpha}\right)^{1-\alpha}\left(\frac{\beta}{\alpha}\right)^{\alpha}(\tau_\pi(\ell)\tau_{va}(\ell))^{1-\alpha}\left[\frac{c_2\underline{y}(\ell)-b}{c_2\underline{y}(\ell)-\frac{b+\tau_n(\ell)}{T(\ell)}}\cdot\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)-\zeta}\frac{\cdot\bar{S}(\zeta(\ell))}{\bar{S}(\zeta)}\right]^{\alpha} \\ &\quad\times J_{\mathrm{SP}}(\zeta,\ell,\underline{y}(\ell))^{1-\alpha} \end{aligned}$$

where J_{SP} is the planner's shadow value of job ζ in location ℓ. To ensure that allocations in the planner's solution and the decentralized equilibrium coincide, there are three margins to correct. First, the decision to start producing together must be efficient, which can be implemented with the employment tax – the standard Hosios (1990) condition in each location. Second, the overall entry margin must be efficient, which can be implemented with the overall level of the profit tax. Third, the location choice of jobs must be efficient, which can be implemented with the spatial progressivity of the profit tax. When those three margins are corrected, it is straightforward to check that the decentralized equilibrium is efficient from the equilibrium conditions. Any transfers can be funded through non-distortionay lump-sum taxes on owners. Alternatively, a flat earnings tax (on both wages and unemployment benefits) leaves the allocation undistorted and concentrates the burden on workers.

Set $T \equiv 1$. Then there exists a τ_n that equated the separation cutoff for the planner and the decentralized equilibrium:

$$\frac{1}{\beta} \cdot \frac{c_2 \underline{y} - b - \tau_n}{\underline{y}} = \frac{1}{\alpha} \cdot \frac{c_2 \underline{y} - b}{\underline{y}} \qquad \longrightarrow \qquad \tau_n(\ell) = \frac{\alpha - \beta}{\alpha} v^{SP}(\ell)$$

Substituting back into the employer's problem,

$$J(\ell, \zeta)^{1-\alpha} \quad = \quad \left(\frac{1-\beta}{1-\alpha}\right)^{1-\alpha} \tau_\pi^{1-\alpha}(\ell) \left[\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)-\zeta} \frac{\cdot \bar{\mathcal{S}}(\zeta(\ell))}{\bar{\mathcal{S}}(\zeta)} \right]^{\alpha} \times \; J_{\text{SP}}(\zeta, \ell, \underline{y}(\ell))^{1-\alpha}$$

The efficient profit tax then satisfies

$$(1-\alpha) \frac{\tau_\pi'(\ell)}{\tau_\pi(\ell)} + \alpha \left(\frac{\mathcal{S}'(\zeta(\ell))}{\mathcal{S}(\zeta(\ell))} + \log \frac{B}{b + v^{SP}(\ell)} \right) (\zeta^{SP})'(\ell)$$

and thus $\tau_\pi'(\ell) < 0$. Given the convention that τ_π is the fraction that the employer keeps after tax, this inequality implies higher marginal profit tax rate in high ℓ locations.

C Quantitative model

C.1 Characterization

C.1.1 Values

The bargaining solution from Lemma 3 readily extends to the extended model, which allows to define values with a minimal amount of extra notation. Denote $U(p, a, h)$ the value of unemployment in location (p, a) for a worker with human capital h. With Frechet taste shocks, the continuation value from migration is

$$M(h) = \left(\int U(p, a, k)^{\nu} F_{p,a}(dp, da) \right)^{\frac{1}{\nu}}$$

where I denote $\nu = 1/\varepsilon$. Migration shares are

$$\pi(\ell, a, k) = \frac{U(\ell, a, k)^{\nu}}{M(k)^{\nu}}$$

Guess that the value of unemployment scales with k. Then the value of unemployment solves the recursion

$$(\rho + \Delta + \mu) U(p, a, k) = apr(p, a)^{-(\omega+\psi)} \cdot U_1(p, a) k + (\lambda - \varphi) k U_k + \mu M_0 k$$

where $M_0 = \left(\int U_1(p, a)^{\nu} F_{p,a}(dp, da) \right)^{\frac{1}{\nu}}$. Because there is a continuum of locations, employed workers who receive the moving opportunity always take it: there is always a location where they taste shock is high enough to make them move. The adjusted surplus then solves

$$(\rho + \Delta + \mu) S(y, k, p, a) = pr(p, a)^{-\psi} k \Big[y - U_1(p, a) \Big] + L_y S + \lambda k S_k$$

Using Lemma 1, the solution scales with k, and is

$$\begin{aligned}
(\rho + \Delta + \mu - \lambda) S(y, k, p, a) \quad &= \quad k \cdot pr(p, a)^{-\psi} U_1(p, a) \cdot \mathcal{S}\left(\frac{y}{\underline{y}(p, a)} \right) \\
(\rho + \Delta + \mu - \lambda) \frac{\underline{y}(p, a)}{\underline{y}_0} \quad &= \quad U_1(p, a)
\end{aligned}$$

where $\underline{y}_0$ is calculated using $\tilde{\rho} = (\rho+\Delta+\mu-\lambda)$ as the effective discount rate. Because workers' outside option scales with h under the guess, the separation decision is independent from k. Going back to the value of unemployment,

$$\tilde{\rho}U(p,a,k) = \frac{(b+v(p,a))ap}{r(p,a)^{\omega+\psi}}k + \mu M_0 k - \varphi k U_k$$

where

$$\tilde{\rho}v(p,a) \quad = \quad \beta f(p,a)U_1(p,a)\left(\frac{Y}{\underline{y}(p,a)}\right)^{\zeta^*(p,a)} \bar{S}(\zeta^*(p,a))$$

and the guess is verified. In addition,

$$(\tilde{\rho}+\varphi)U_0(p,a) \quad = \quad \frac{(b+v(p,a))a\ell}{r(p,a)^{\omega+\psi}} + \mu M_0$$

which can be rearranged into

$$\tilde{\rho}U_0(p,a) \quad = \quad \frac{\tilde{\rho}(b+v(\ell,a))}{\tilde{\rho}+\varphi}\cdot\frac{a\ell}{r(\ell)^{\omega+\psi}} + \mu M_0 - \frac{\varphi}{\tilde{\rho}+\varphi}\cdot\mu M_0$$

and therefore,

$$U_1(p,a) = \frac{\tilde{\rho}}{\tilde{\rho}+\varphi}(b+v(\ell,a)) - \frac{\varphi}{\tilde{\rho}+\varphi}\mu M_0$$

Under the (empirically relevant) assumption that $\mu \ll 1$, the second term is negligible, and so

$$U_1(p,a) \approx \frac{\tilde{\rho}}{\tilde{\rho}+\varphi}(b+v(p,a))$$

Going back to the joint surplus,

$$\begin{aligned}
\tilde{\rho}S(y,k,p,a) \quad &\approx \quad k\cdot pr(p,a)^{-\psi}\left[\frac{\tilde{\rho}}{\tilde{\rho}+\varphi}(b+v(p,a))\right]\cdot\mathcal{S}\left(\frac{y}{\underline{y}(p,a)}\right)\\
\tilde{\rho}\frac{\underline{y}(p,a)}{\underline{y}_0} \quad &\approx \quad \frac{\tilde{\rho}}{\tilde{\rho}+\varphi}(b+v(p,a))
\end{aligned}$$

To a first order approximation in μ, the previous results imply

$$\hat{\rho}S(y,k,p,a) = k\cdot pr(p,a)^{-\psi}(b+v(p,a))\cdot\mathcal{S}\left(\frac{y}{\underline{y}(p,a)}\right) \quad ; \quad \hat{\rho}\frac{\underline{y}(p,a)}{\underline{y}_0} = (b+v(p,a)) \tag{51}$$

$$\hat{\rho}U(p,a,k) = \frac{(b+v(p,a))ap}{r(p,a)^{\omega+\psi}}k \equiv U_0(p,a)k \quad ; \quad \pi(p,a,k) = \left(\frac{U_0(p,a)}{M_0}\right)^{\nu} \tag{52}$$

where $\hat{\rho} = \tilde{\rho}+\varphi$.

C.1.2 Human capital across locations

I now characterize the human capital distribution in each location. For now, focus on a single locationa nd omit (p,a) subscripts to facilitate exposition. The probability mass functions of rescale log human capital $h = \log k - \lambda t$ for employed and unemployed workers in a location solve:

$$\begin{aligned}
0 \quad &= \quad -sg_E(h) + f_R g_U(h) - \mu g_E(h) - \Delta g_E(h)\\
0 \quad &= \quad \varphi g_U'(h) - f_R g_U(h) + sg_E(h) - \mu g_U(h) + K(h) - \Delta g_U(h)
\end{aligned}$$

where $K(h)$ is the overall entry distribution inclusive of in-migration and newborns. This simple combination of ODEs obtains because there is no relative human capital growth while employed. This delivers the crucial simplification that separations are independent from the human capital level. Re-arranging the first equation,

$$g_E(h) = \frac{f_R}{\mu + \Delta + s} g_U(h)$$

and so, substituting back into the second equation

$$0 = \varphi g_U'(h) - \underbrace{(\mu + \Delta)\frac{\mu + \Delta + s + f_R}{\mu + \Delta + s}}_{\equiv C} g_U(h) + K(h)$$

While this ODE can be solved explicitly, computing the mean human capital is sufficient to characterize equilibrium. Multiply the KFE by e^h, integrate over h in $\mathbb{R}$ and integrate the first term by parts to obtain

$$0 = [e^h g_U(h)]_{-\infty}^{\infty} - \varphi \int_{\mathbb{R}} e^h g_U - C \int_{\mathbb{R}} e^h g_U + \int_{\mathbb{R}} e^h K$$

The first term is 0 at both extremes. Denote $k_0 = \int_{\mathbb{R}} e^h K$. To get average human capital in the location one needs to solve for total population masses in each location: $\mathcal{U}, \mathcal{E}$. They solve similar KFEs, so that

$$\mathcal{E} = \frac{f_R}{\mu + \Delta + s} \mathcal{U}$$

and so

$$(\mu + \Delta)\frac{\mu + \Delta + s + f_R}{\mu + \Delta + s} \mathcal{U} = \mathcal{K}$$

where $\mathcal{K}$ is the total mass of entrants. Hence, the unemployment *rate* is

$$u = \frac{\mathcal{U}}{\mathcal{U} + \mathcal{E}} = \frac{\mu + \Delta + s}{\mu + \Delta + s + f_R} = \frac{\mu + \Delta}{C}$$

while the *mass* of unemployed is $\mathcal{U} = u \cdot \frac{\mathcal{K}}{\mu+\Delta}$, and so population is $\mathcal{E} + \mathcal{U} = \frac{\mathcal{K}}{\mu+\Delta}$. Recall that by definition, $\int_{\mathbb{R}} g_U = \mathcal{U}$. Thus, average human capital in a location is

$$\mathbb{E}(p, a) \equiv \mathbb{E}[e^h | p, a] \equiv \bar{h}(\ell, a) = \frac{k_0}{\mathcal{U} \cdot (\varphi + C)} = (\mu + \Delta)\frac{\mathbb{E}^K[e^h]}{u \cdot (\varphi + C)} = \frac{\mu + \Delta}{\mu + \Delta + u(p, a)\varphi} \cdot \mathbb{E}^K[e^h]$$

where $\mathbb{E}^K[e^h]$ is the expected human capital of new entrants. By definition, the mass of entrants at human capital h is

$$K(h) = \mu\pi(p, a)I(h) + \Delta L(p, a)E(h)$$

where I is the economy-wide invariant distribution, and E is the entry distribution, and π are migration shares. So one obtains

$$\mathbb{E}^K[e^h] = \frac{\mu\pi}{\mu\pi + \Delta L}\mathbb{E}^I[e^h] + \frac{\Delta L}{\mu\pi + \Delta L}\mathbb{E}^E[e^h]$$

In steady-state, population density is equal to migration shares:

$$L(p, a) = \pi(p, a)$$

Therefore,

$$\mathbb{E}^K[e^h] = \frac{\mu}{\mu + \Delta}\mathbb{E}^I[e^h] + \frac{\Delta}{\mu + \Delta}\mathbb{E}^E[e^h] \equiv x_0\mathbb{E}^I + (1 - x_0)\mathbb{E}^E$$

Now,

$$\mathbb{E}^I[e^h] = \int L(p, a)F_{p,a}(dp, da) \cdot \mathbb{E}(p, a)$$

so that one obtains a linear system across locations:

$$\mathbb{E}(p,a) = \frac{z_0}{z_0+\varphi u_k}\left[x_0\mathbb{E}^I + (1-x_0)\mathbb{E}^E\right] = \frac{z_0}{z_0+\varphi u(p,a)}\left[x_0\int L(p',a')\mathbb{E}(p',a')F_{p,a}(dp',da') + (1-x_0)\mathbb{E}^E\right]$$

where $z_0 = \mu+\Delta$. Denote $X(p,a) = (1+\varphi_0 u(p,a))\mathbb{E}(p,a)$ where $\varphi_0 = \varphi/z_0$. Re-write the linear system as

$$X(p,a) = (1-x_0)\mathbb{E}^E + x_0\int \frac{L(p',a')}{1+\varphi_0 u(p',a')}\mathbb{E}(p',a')F_{p,a}(dp',da')$$

This system can be explicitly solved. Multiply by $Z(p,a) \equiv x_0 \frac{L(p,a)}{1+\varphi_0 u(p,a)} F_{p,a}(dp',da')$ and integrate to obtain

$$\int Z(p,a)F_{p,a}(dp,da) = \frac{\int Z(p,a)F_{p,a}(dp,da)}{1-\int Z(p,a)F_{p,a}(dp,da)}\cdot(1-x_0)\mathbb{E}^E$$

Then

$$Z(p,a) = \frac{(1-x_0)\mathbb{E}^E}{1-\int Z(p',a')F_{p,a}(dp',da')}$$

which finally implies

$$\mathbb{E}(p,a) = \frac{1}{1+\varphi_0 u(p,a)}\cdot\frac{(1-x_0)\mathbb{E}^E}{1-x_0\int\frac{L(p',a')}{1+\varphi_0 u(p',a')}\cdot F_{p,a}(dp',da')}$$

C.1.3 Labor market flows

Given (51), the expression for labor market flows immediately extends given an assignment $z(p,a)$ and a value of search $v(p,a)$. The only change follows from the KFE, in logs:

$$0 = \delta g'(x) + \frac{\sigma^2}{2}g''(x) - (\Delta+\mu)g(x)$$

The associated characteristic equation has only one negative (stable) root,

$$\kappa = -\frac{1}{2}\left[\frac{2\delta}{\sigma^2} + 2\sqrt{2\frac{\mu+\Delta}{\sigma^2}+\frac{\delta^2}{\sigma^4}}\right]$$

which coincides with the simple solution $\kappa_0 = \frac{2\delta}{\sigma^2}$ when $\mu+\Delta = 0$. Thus, the previous expression for the invariant distribution extends with κ instead of κ_0. In addition, the expression for the average productivity also extends. The exit rate from employment is then $\delta/z(p,a)+\mu+\Delta$. Using the flow equation for unemployment together with the steady-state migration shares, one obtains $u(p,a) = \frac{\delta/z(p,a)+\mu+\Delta}{\delta/z(p,a)+\mu+\Delta+f_R(p,a)}$.

C.1.4 Population, housing prices and composite index

Having solved for average human capital $\mathbb{E}(p,a)$ in each location, it is possible to characterize housing prices and thus the value of employers. In steady-state, total population in a location is given by migration shares:

$$L(p,a) = \left(\frac{U_0(p,a)}{M_0}\right)^{\nu} \tag{53}$$

Housing rents follow from equating total housing demand to local supply. From the Cobb-Douglas structure of the production function, employers spend a fraction ψ of output on housing. Hence, total housing demand in location

(p, a) is now

$$\begin{aligned} H_0 r(p,a)^{\eta} &= p\mathbb{E}(p,a) r(p,a)^{-\psi} \cdot L(p,a) \cdot \frac{1}{r(p,a)} \Big[\omega u(p,a) b \\ &\quad + \omega(1-u(p,a))(b+v(p,a))(1-\beta+\beta\mathbb{E}_{p,a}[y|y>\underline{y}(p,a)]) \\ &\quad + \psi(1-u(p,a))(b+v(p,a))\mathbb{E}_{p,a}[y|y>\underline{y}(p,a)]) \Big] \\ &\equiv p\mathbb{E}(p,a) r(p,a)^{-\psi-1} L(p,a) G(\zeta(p,a), v(p,a)) \end{aligned} \tag{54}$$

where η is the housing supply elasticity, and $\mathbb{E}_{p,a}[h]$ is the average local human capital. For the last equality, I anticipate that local unemployment will still be a function of v, ζ alone, and that there is PAM in equilibrium. In what follows, normalize H_0 to one without loss of generality.

After substituting equation (54) into the migration share equation, some algebra yields

$$L = M_0^{-\frac{1}{\varepsilon + \frac{\omega+\psi}{1+\eta+\psi}}} \cdot \left[\frac{a \cdot (p\mathbb{E}(p,a))^{\frac{1+\eta-\omega}{1+\eta+\psi}} \cdot (b+v(p,a))}{G(v(p,a),\zeta(p,a))^{\frac{\omega+\psi}{1+\eta+\psi}}} \right]^{\frac{1}{\varepsilon + \frac{\omega+\psi}{1+\eta+\psi}}} \tag{55}$$

After substituting equation (55) back into (54) and some algebra, and so

$$r(p,a) = M_0^{-\frac{1}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \cdot \left\{ a \cdot (p\mathbb{E}(p,a))^{1+\varepsilon} \cdot (b+v(p,a)) \cdot G(v(p,a),\zeta(p,a))^{\varepsilon} \right\}^{\frac{1}{(1+\eta+\psi)\varepsilon+\omega+\psi}}$$

Therefore the expected prefactor in the adjusted surplus is

$$p\mathbb{E}(p,a) r(p,a)^{-\psi} = M_0^{\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \cdot \frac{\left((p\mathbb{E}(p,a))^{\omega+\varepsilon(1+\eta)} a^{-\psi} \right)^{\frac{1}{(1+\eta+\psi)\varepsilon+\omega+\psi}}}{\left[(b+v(p,a)) G(v(p,a),\zeta(p,a))^{\varepsilon} \right]^{\frac{\psi}{(1+\eta+\psi)\varepsilon+\omega+\psi}}} \tag{56}$$

This equation motivates the definition of the composite index

$$\ell(p,a) = \left(p^{\omega+\varepsilon(1+\eta)} a^{-\psi} \right)^{\frac{1}{(1+\eta+\psi)\varepsilon+\omega+\psi}} \tag{57}$$

C.1.5 Location choice of employers

Using the adjusted surplus and (56), the value of opening a job in location (p, a) for employer $\zeta = 1/z$ is

$$\begin{aligned} J(\zeta,p,a)^{\frac{1}{1+\gamma}} \propto\ & \mathbb{E}(p,a)^{\frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \cdot \ell(p,a) \\ & (b+v(p,a))^{1-\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \cdot G(v(p,a),\zeta^*(p,a))^{\frac{\varepsilon\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \\ & q(p,a) \left(\frac{B}{b+v(p,a)} \right)^{\zeta} \bar{\mathcal{S}}(\zeta) \end{aligned} \tag{58}$$

where the optimal vacancy posting decision has been maximized out. Re-arranging equation (58) delivers equation (18). Using the worker's value of search to substitute out $q(p, a)$ delivers the employer's location problem

$$\begin{aligned} \max_{p,a} \quad & \tilde{\mathbb{E}}(u(v(p,a),\zeta^*(p,a))^{(1-\alpha)\frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \cdot \ell(p,a)^{1-\alpha} \\ & (b+v(p,a))^{1-(1-\alpha)\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} v(p,a)^{-\alpha} \cdot G(v(p,a),\zeta^*(p,a))^{\frac{(1-\alpha)\varepsilon\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \\ & \left(\frac{B}{b+v(p,a)} \right)^{(1-\alpha)\zeta+a\zeta^*(p,a)} \bar{\mathcal{S}}(\zeta)^{1-\alpha} \bar{\mathcal{S}}(\zeta^*(p,a))^{\alpha} \end{aligned} \tag{59}$$

where $\tilde{\mathbb{E}}(u(v(p,a),\zeta(p,a)) \equiv \mathbb{E}(p,a)$ but where the dependence on the local unemployment rate has been made explicit.

In principle, employers must take two first-order conditions for their optimal location choice: with respect to each dimension $i \in \{p, a\}$. After taking these first-order conditions and re-arranging, one obtains:

$$\partial_i v = A(v,\zeta,\ell)\partial_i \ell + B(v,\zeta,\ell)\partial_i \zeta^*$$

for some functions A, B. Now guess that ζ^* is a function of $\ell(p,a)$ only. Then one obtains for $i \in \{p,a\}$

$$\partial_i v = C(v,\zeta,\ell)\partial_i \ell$$

Combining equations, standard partial differential equation results imply that v is a function of ℓ alone. Thus, employers need only choose the unidimensional combined index $\ell(p,a)$.

With this observation at hand, the structure of equation (59) then closely mirrors its baseline model equivalent. Therefore, the assignment resuts extend under either Assumption 1 or Assumption 2 – the latter would only the expression for $\bar{\mathcal{S}}$. The FOC for the optimal location choice is then

$$\begin{aligned}
&\frac{v'(\ell)}{v(\ell)}\left\{\alpha + \frac{v(\ell)}{b+v(\ell)}\left(\zeta(\ell) - 1 - \frac{(1-\alpha)\psi}{\omega+\psi+(1+\psi)\varepsilon}\left[1+\varepsilon\frac{(B+v(x))G_v}{G}\right]\right)\right\} \\
= \quad &\frac{1-\alpha}{\ell} + \alpha\left(\frac{\bar{\mathcal{S}}'(\zeta(\ell))}{\bar{\mathcal{S}}(\zeta(\ell))} + \log\frac{B}{b+v(\ell))}\right)\zeta'(\ell) + \varepsilon\cdot\frac{(1-\alpha)\psi}{\omega+\psi+(1+\psi)}\frac{G_\zeta}{G}\zeta'(\ell) \\
&+\frac{(1-\alpha)(\omega+\varepsilon)}{\omega+\psi+\varepsilon(1+\psi)}\frac{d}{d\ell}\left(\log\frac{D}{D+\varphi u(\ell)}\right)
\end{aligned}$$

where

$$\frac{d}{d\ell}\left(\log\frac{D}{D+\varphi u(\ell)}\right) \quad = \quad \frac{\frac{\varphi u(\ell)}{D+\varphi u(\ell)}}{D+\delta\zeta(\ell)}\left\{\frac{u(\ell)f_R(\ell)}{b+v(\ell)}\frac{v'(\ell)}{v(\ell)} - \left[\delta(1-u(\ell)) + f_R(\ell)\frac{\bar{\mathcal{S}}'(\zeta(\ell))}{\bar{\mathcal{S}}(\ell)}\right]\zeta'(\ell)\right\}$$

and where

$$\begin{aligned}
G_v &= \omega b u_v + \omega(1-u)(1-\beta+\beta\mathcal{E}) - \omega(b+v)u_v(1-\beta+\beta\mathcal{E}) + \psi(1-u)\mathcal{E} - \psi u_v(b+v)\mathcal{E} \\
G_\zeta &= \omega b u_\zeta + \omega(1-u)(b+v)\beta\mathcal{E}_\zeta - \omega u_\zeta(b+v)(1-\beta+\beta\mathcal{E}) + \psi(1-u)(b+v)\mathcal{E}_\zeta - \psi u_\zeta(b+v)\mathcal{E}
\end{aligned}$$

where here $\mathcal{E}(\zeta) = \mathbb{E}_\zeta[y/\underline{y}|y \geq \underline{y}]$. It is then possible to express labor market tightness in a location ℓ:

$$\theta(\ell) = -\frac{M_e f_\zeta(\zeta(\ell)\mathcal{V}(\ell,\zeta(\ell))\zeta'(\ell)}{u(\ell)\mathcal{L}(\ell)f_\ell(\ell)} \tag{60}$$

where optimal vacancies are

$$\mathcal{V}(\ell,\zeta(\ell)) \propto J(\zeta^*(\ell),\ell)^\gamma \tag{61}$$

and the maximized value of employers is

$$\begin{aligned}
J(\zeta^*(\ell),\ell)^{\frac{1}{1+\gamma}} \quad = \quad &\mathbb{E}(u(\ell))^{\frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot\ell \\
&(b+v(\ell))^{\frac{1}{1-\alpha}-\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot v(\ell)^{-\frac{\alpha}{1-\alpha}}G(v(\ell),\zeta(\ell))^{\frac{\varepsilon\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \\
&\left(\frac{B}{b+v(\ell)}\right)^{\frac{\zeta(\ell)}{1-\alpha}}\bar{\mathcal{S}}(\zeta(\ell))^{\frac{1}{1-\alpha}}
\end{aligned} \tag{62}$$

C.1.6 Labor market clearing and population determination

After substituting equation (57) back into (55) and some algebra,

$$L(p,a) \quad \propto \quad \frac{\ell(p,a)^{\frac{1+\eta-\omega}{\omega+\varepsilon(1+\eta)}}\cdot(b+v(p,a))^{\frac{1+\eta+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}}{G(v(p,a),\zeta(p,a))^{\frac{\omega+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}}\cdot a^{\frac{1}{\omega+\varepsilon(1+\eta)}}$$

Then average population density in locations with index ℓ, $\mathcal{L}(\ell)$, is given by

$$\mathcal{L}(\ell) \propto \frac{\ell^{\frac{1+\eta-\omega}{\omega+\varepsilon(1+\eta)}}\cdot(b+v(\ell))^{\frac{1+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}}{G(v(\ell),\zeta(\ell))^{\frac{\omega+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}}\cdot\mathcal{C}(\ell)$$

where

$$\mathcal{C}(\ell) = \mathbb{E}\left[a^{\frac{1}{\omega+\varepsilon(1+\eta)}} \,\middle|\, \ell(p,a) = \ell \right]$$

By construction,

$$(\omega + \psi + \varepsilon(1+\eta+\psi)) \log \ell = (\omega + \varepsilon(1+\eta)) \log \ell - \psi \log a$$

As an example for $\mathcal{C}(\ell)$, consider the lognormal case of the estimation. Then $(\log a, \log \ell)$ is jointly lognormal, with variance matrix

$$\begin{pmatrix} \sigma_a^2 & \\ \frac{\sigma_a}{\omega+\psi+\varepsilon(1+\eta+\psi)} ((\omega + \varepsilon(1+\eta+\eta))\rho_{ap}\sigma_p - \psi\sigma_a) & \underbrace{\frac{(\omega+\varepsilon(1+\eta))^2\sigma_p^2 + \psi^2\sigma_a^2 - 2(\omega+\varepsilon(1+\eta))\psi\rho_{ap}\sigma_a\sigma_p}{(\omega+\psi+\varepsilon(1+\eta+\psi))^2}}_{\equiv \sigma_\ell^2} \end{pmatrix}$$

Using the conditional normal distributions,

$$\log a \mid \log \ell = \frac{1}{\sigma_\ell^2} \cdot \frac{(\omega+\varepsilon(1+\eta))\rho_{ap}\sigma_p\sigma_a - \psi\sigma_a^2}{\omega+\psi+\varepsilon(1+\eta+\psi)} \log \ell + N$$

where $N \sim \mathcal{N}(0, \sigma_a^2(1-\rho_{a,\ell}^2)$ is independent from $\log \ell$.[81] Therefore, the correction factor is

$$\mathcal{C}(\ell) = \mathcal{C}_0 \exp\left(\frac{\sigma_a^2}{\sigma_\ell^2} \cdot \frac{\rho_{ap}\frac{\sigma_p}{\sigma_a} - \frac{\psi}{\omega+\varepsilon(1+\eta)}}{\omega+\psi+\varepsilon(1+\eta+\psi)} \cdot \log \ell \right)$$

C.1.7 Intuition for sufficient statistic

Here I briefly discuss why the combined index $\ell(p,a)$ is a local sufficient statistic in equilibrium. Given that the direct contributions of local productivity p and amenities a are combined into the single index $\ell(p,a)$ in the location choice of jobs (18), it is natural to conjecture that this single index will be a sufficient statistic for the model's outcomes. However, one potential complication arises. Labor market clearing in each location relates the number of vacancies to the number of unemployed workers. While the volume of local vacancies is a function of $\ell(p,a)$ only as per equation (18), the number of locally unemployed workers is not. Because workers also directly care about amenities a, their location choices reflect p and a in a combination that does not align with employers'. Thus, the number of locally unemployed workers varies with $\ell(p,a)$ and with amenities a conditional on $\ell(p,a)$.

The key insight is that employers only value locations through the combined index $\ell(p,a)$ as long as labor market tightness $\theta(\ell)$ also only depends on the combined index. As illustrated by Figure 24, employers are then indifferent between all locations that have the same index $\ell(p,a)$ even if these locations have different amenities a. Jobs with the same quality z thus allocate along one-dimensional indifference curves – $\ell(p,a)$ isoquants – to ensure that labor market tightness $\theta(\ell)$ remains constant along the indifference curve. Locations with higher amenities a conditional on the local advantage index $\ell(p,a)$ have both more unemployed workers and more open jobs, but in similar proportions.[82]

C.1.8 Efficiency

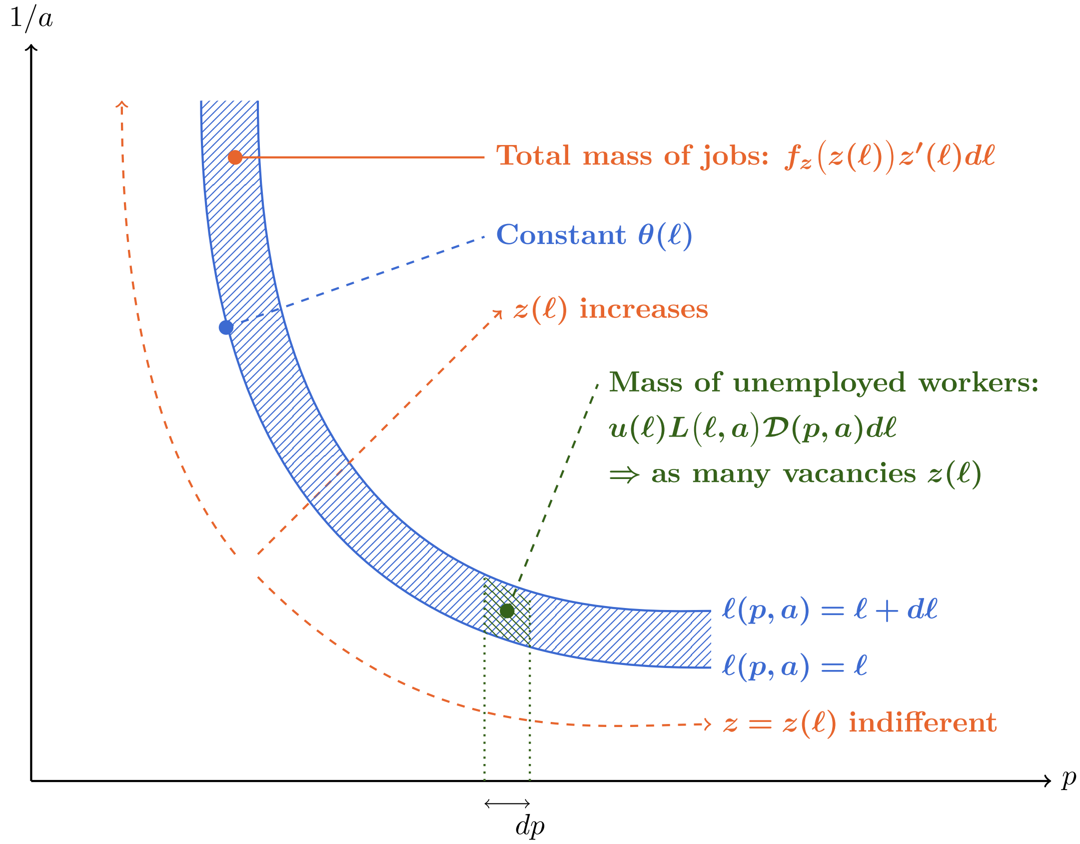
All the additional choices in the extended model are efficient. Thus, the normative results extend, with one caveat. Workers have heterogeneous human capital within a location but search in the same labor market. Therefore, low human capital workers who separate into unemployment create a negative externality on high human capital workers who are searching for a job. In general, this provides a motive for the planner to retain workers with low human capital longer on the job.

This source of inefficiency is not the focus of the paper, and thus I do not attempt to derive an optimal policy that would correct it. Rather, note that when φ and the support of F_k are small enough, there is little dispersion

[81] Note that the correlation is $\rho_{a\ell} = \frac{(\omega+\varepsilon)\rho_{ap}\sigma_p - \psi\sigma_a}{\omega+\psi+\varepsilon(1+\psi)} \Big/ \sigma_\ell$

[82] $\mathcal{D}(p,a)$ in Figure 24 is defined in the Appendix and encodes how small changes dp, da translate into small changes $d\ell$. Formally, is the determinant of the Jacobian matrix of the mapping $\ell(p,a)$.

Figure 24: The location choice of employers and workers in two-dimensional space.



between human capital levels within a location. In that case, it is possible to show that, the inefficiency is small in the sense that it is quadratic in φ, Var_k. Finally, it is possible to extend the directed search environment to the richer framework. Because human capital is not observed by employer prior to matching, the optimal contract may in principle depend nonlinearly on human capital if employers try to screen different workers. It is nonetheless possible to show that the optimal contract is still a local wage rate per unit of human capital, which makes comparisons with the random search model straightforward.

Making those arguments precise requires a substantial amount of new notation and lengthy derivations. Thus, they are omitted in the present paper, but are available upon request.

C.1.9 Welfare

The average welfare of unemployed workers in locations ℓ satisfies

$$\hat{\rho}\mathbb{E}\left[\int U(p,a,k)L(p,a)F_{p,a}(dp,da)\Big|\ell\right] \quad \approx \quad \bar{U}_0 \cdot J(M_0)U_2(\ell)\mathcal{D}(\ell)\mathbb{E}(\ell)$$

where $\bar{U}_0$ is a transformation of parameters, and

$$J(M_0) = M_0^{\frac{\omega+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \qquad ; \qquad U_2(\ell) = \ell^{\frac{\varepsilon(1+\eta-\omega)}{\omega+\varepsilon(1+\eta)}} \cdot (b+v(\ell))^{\frac{\varepsilon(1+\eta+\psi)}{\omega+\psi+\varepsilon(1+\eta+\psi)}} G(v(\ell),\zeta(\ell))^{-\frac{\varepsilon(\omega+\psi)}{\omega+\psi+\varepsilon(1+\eta+\psi)}}$$

and

$$\mathcal{D}(\ell) = \mathbb{E}[a^{\frac{\varepsilon}{\omega+\varepsilon(1+\eta)}\times\left(1+\frac{\eta\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}\right)}|\ell]$$

can be computed similarly to $\mathcal{C}(\ell)$. However, given the presence of idiosyncratic preference shocks, that act as compensating differentials, the welfare of unemployed workers of the same human capital is equalized across locations in expectation, and is simply M_0. To compute M_0, note that

$$M_0^{\nu} = \int U_0(p,a)^{\nu} F(dp,da)$$

and so

$$M_0^{\frac{1+\eta+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} = \int U_2(\ell)^{\nu} \ell^{\frac{1+\eta-\omega}{\omega+\varepsilon(1+\eta)}} \tilde{\mathcal{D}}(\ell) F_{\ell}(d\ell)$$

where

$$\tilde{\mathcal{D}}(\ell) = \mathbb{E}\Big[a^{\frac{1}{\omega+\varepsilon(1+\eta)}\left(1+\frac{\eta\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}\right)}\Big|\ell\Big]$$

The contribution of unemployed workers to total welfare is thus

$$\bar{U} = A\int J(M_0)U_2(\ell)\mathcal{D}(\ell)\mathbb{E}(\ell)u(\ell)\mathcal{L}(\ell)F_{\ell}(d\ell)$$

for some constant $A > 0$. Similarly, total welfare of employed workers is

$$\begin{aligned}
\bar{V^E} &= A\int \mathbb{E}\left[\int \mathbb{E}[V^E(\ell,a,y)]L(\ell,a)dF_{\ell,a}(\ell,a)\Big|x\right] \cdot \bar{h}(x)\cdot(1-u(x))\mathcal{L}(x)\cdot dF_x(x) \\
&= A\int J(M_0)U_2(\ell)\Big(1+\beta\bar{\mathcal{S}}(\zeta(\ell))\Big)\mathcal{D}(\ell)\mathbb{E}(\ell)(1-u(\ell))\mathcal{L}(\ell)F_{\ell}(d\ell)
\end{aligned}$$

Finally, in the decentralized equilibrium, I must take a stand on how profits from land rents and employers are redistributed. I assume that they are rebated to workers with a flat earnings subsidy. This formulation has two advantages. First, is non-distortionary. Second, the transfer is subsumed into M_0 and thus usinf the standard population adding up and free entry condition suffices to compute the equilibrium.

C.2 Alternative extensions

Here I discuss several possible extensions of the model and argue that the main theoretical predictions are robust to those alternative specifications.

Micro-foundation for local productivity. Assume for simplicity that the production function is $F(y_t,\ell,k) = y_t \cdot k \cdot \ell$. Assume also homogeneous amenities $a \equiv 1$, and a subsistence level of housing for workers stemming from Stone-Geary preferences. Denote by $K(\ell) = \mathbb{E}[k|\ell]$ the average worker productivity in location ℓ under the equilibrium population distribution. It is then straightforward to show that the job's location problem is the same as in equation (12), except that the endogeneous productivity $\ell \cdot K(\ell)$ replaces the exogeneous productivity ℓ. Thus, jobs sort in space according to a matching function $z(\ell \cdot K(\ell))$ resulting in a cutoff function $\underline{y}(\ell \cdot K(\ell))$.

In the limit when locations becomes ex-ante homogeneous $\text{Var}(\ell) \to 0$, there remain sustained differences in measured productivity $K(\ell)$ because workers systematically sort in space due to Stone-Geary preferences, as in Corollary 1. Small local differences again act as a coordinating device or sunspot. To see why, recall that when locations becomes ex-ante homogeneous $\text{Var}(\ell) \to 0$, there can be sustained differences in measured productivity $K(\ell)$ only if workers systematically sort in space, as in Corollary 1. To obtain positive sorting, I must introduce Stone-Geary preferences for housing: there is a subsistance level of housing $H_0 > 0$ that workers must rent to live in a location. Workers' location choices then follow their free-mobility condition

$$\rho U(k) = \max_{\ell} \frac{(k\ell)\Big(b+v\big(\ell K(\ell)\big) - H_0 r(\ell)}{r(\ell)^{\omega}}$$

Then, the worker's object $(k\ell)\Big(b+v\big(\ell K(\ell)\big)$ is supermodular in $\Big(k, \ell\big(b+v(\ell K(\ell))\big)\Big)$, which ensures perfect positive assortative matching between worker productivity k and the location's value of $\ell\big(b+v(\ell K(\ell))$. This last term is increasing in $K(\ell)$, and so the equilibrium sustains positive sorting between individual productivity k and average local productivity $K(\ell)$. These arguments prove the following proposition which closely mirrors Corollary 1.

Proposition 10. *(Micro-foundation for spatial differences in productivity)*
As Var(ℓ) $\to 0$, the equilibrium sorting function K converges to a strictly increasing function K_0.

Proposition 10 guarantees that even as differences in *exogenous* local productivity vanish, the endogenous sorting of heterogeneous workers sustains measured differences in endogenous local productivity $K_0(\ell)$. Small local differences act as a coordinating device or sunspot. Having established Proposition 10, Propositions 1 to 5 extend with minor modifications: the basic sorting patterns of jobs, the implications for job losing and finding rates, as well as the inefficient margins, all remain.

Skill markets. Section 1 establishes that, after city residuals, sorting of heterogeneous workers along observable skill charateristics accounts for a little over 20% of the spatial variation in local job losing rates. Suppose that workers differ in their skill x, which is exogenously distributed in the population according to a cumulative distirbution function F_x. Suppose for clarity of exposition that the production function is $F(y_t, \ell, x) = y_t \cdot x \cdot \ell$, no differences in amenities $a \equiv 1$ and free mobility. Assume that firms perfectly observe x and can choose which skill level to hire, i.e. there are segmented labor markets per skill in each location. Then Proposition 1 extends with an endogeneous profitability index $\ell(p, x) = \ell \cdot x$ that combines worker's skill and local characteristics. Conditional on skill, locations with higher ℓ attract high productivity jobs and thus have a lower job loss rate. Conditional on location, high skill workers have lower job loss rates because they match with high productivity jobs, as in the data. Propositions 1 to 5 then extend with minor modifications.

Skills also sort across space due to an endogeneous complementarity between skills and locations stemming from the job's location problem. To see this, note that jobs' value conditional on their type z and their choice ℓ, x is

$$\frac{J(z,x,\ell)}{J_0} = \left\{ \frac{b + v(\ell,x)}{v(\ell,x)^{\alpha}} \bar{S}\Big(z, \underline{y}(\ell,x)\Big)^{1-\alpha} \bar{S}\ \Big(z(\ell,x), \underline{y}(\ell,a,h)\Big)^{\alpha} \right\}^{\frac{1}{1-\alpha}} \cdot \Big(\ell x\Big)$$

The same argument as in Proposition 6 ensures that $\underline{y}, v, z$ are univariate functions of the endogenous profitability index ℓx. Worker's free mobility then writes

$$\underline{y}_0 U(x) = \max_{\ell} \frac{\ell x}{r(\ell)^{\omega}} \underline{y}(\ell x)$$

There is positive sorting between x and ℓ if $\underline{y}(\ell x)$ is log-supermodular in (x, ℓ). Now,

$$\partial_x \Big(\log \underline{y}(\ell x) \Big) = \frac{\ell x \cdot v'(\ell x)}{v(\ell x)} \cdot \frac{v(\ell x)}{b + v(\ell x)}$$

From the firm's FOC, the first term is

$$\frac{\ell x \cdot v'(\ell x)}{v(\ell x)} = \frac{1-\alpha}{\alpha + \frac{v(\ell x)}{b+v(\ell x)}(1/z(\ell x) - 1) - \alpha \left(\log \frac{B_0}{b+v(\ell x)} + \frac{\bar{S}'(z(\ell x))}{\bar{S}(z(\ell x))} \right) \frac{z'(\ell x)}{z(\ell x)^2}}$$

Thus, when α is small enough, $\partial_x \Big(\log \underline{y}(\ell x) \Big)$ is increasing in ℓ. This ensures log-supermodularity of the worker's objective and hence positive assortative matching between x and ℓ.

Industries. To include heterogeneous industries in the model, it suffices to assume that jobs in each industry j draw their quality from an industry-specific initial distribution $F_{z,j}$. This extended model generates differences in job loss rates by industry, and the mixture of those distributions directly maps into the single distribution F_z.

Agglomeration and congestion spillovers. It is straightfoward to include standard agglomeration and congestion externalities into the model, following Diamond (2016) and Fajgelbaum and Gaubert (2019). Agglomeration externalities would call for subsidies to large, high wage cities, thus going in the opposite direction as the new spatial search externality on the firm side that I highlight in this paper. Congestion externalities would call for subsidies to small, low wage cities. The net optimal policy would account for all these margins.

Trade costs. Throught the model, I maintain the assumption of a freely traded single final good, due to the lack of good data on commuting zone-level trade flows in France and in the US. In practice, some goods are more tradable than others. It is possible to include intra-national trade in this model. Assume that each job produces a differentiated variety subject to location-pair-specific trade costs $\tau(\ell, a; \ell', a')$. In each location, perfectly competitive intermediate producers aggregate varieties from all locations with a Constant Elasticity of Substitition (CES) aggregator. They produce a non-traded good, which is used as an intermediate in production by firms which use labor in production.

In this model, exogenous local differences in productivity ℓ are complemented by a general equilibrium aggregator of bilateral trade costs, and market conditions in all other locations, resulting in an effective local productivity

$P(\ell;\tau(\cdot),...)$. Propositions 1 to 5 then extend directly when replacing the exogenous productivity ℓ by the endogenous productivity P. While trade adds a layer of general equilibrium conditions, the basic sorting and inefficiency properties continue to hold along with their implications for local labor market flows and policy.

Identical jemployers and different workers. Consider the alternative model where z is attached to a worker instead of a job. Denote $\zeta = 1/z$. Suppose that all locations are ex-ante heterogeneous. Then the worker's location problem is

$$\rho U(\zeta) = \max_{\ell} \frac{(b + v(\zeta,\ell))}{r(\ell)^{\omega}}$$

where

$$v(\zeta,\ell) = \beta f(\ell) \left(\frac{B_0}{b + v(\zeta,\ell))}\right)^{\zeta} \bar{S}(\zeta)$$

Take $b \to 0$ to simplify, so that

$$v(\zeta,\ell) \propto \left[f(\ell)\bar{S}(\zeta)\right]^{\frac{1}{1+\zeta}}$$

and so

$$\zeta : \max_{\ell} \frac{1}{1+\zeta} \log f(\ell) - \omega \log r(\ell)$$

The main difference is that workers take part of their productivity cutoff with them across space. Thus, if there is sorting, there is NAM between ζ and f, and so NAM between separation rates and finding rates. Can this assignment be sustained? From jobs' free mobility,

$$J = q(\ell) \left(\frac{B_0}{b + v(\zeta(\ell),\ell))}\right)^{\zeta(\ell)} \bar{S}(\zeta(\ell)) =_{b\to 0} q(\ell) \left(\frac{B_0}{v(\zeta(\ell),\ell))}\right)^{\zeta(\ell)} \bar{S}(\zeta(\ell)$$

and so substituting back into the worker's location choice,

$$1 \quad \propto \quad f^{-\frac{\alpha}{1-\alpha}} \left(\frac{B_0}{v}\right)^{\zeta} \bar{S} \propto f^{-\frac{\alpha}{1-\alpha}} \left(\frac{B_0}{f^{\frac{1}{1+\zeta}} \bar{S}^{\frac{1}{1+\zeta}}}\right)^{\zeta} \bar{S} \propto f^{-\left[\frac{\alpha}{1-\alpha}+\frac{\zeta}{1+\zeta}\right]} B_0^{\zeta} \bar{S}(\zeta)^{\frac{1}{1+\zeta}}$$

and so

$$f(\ell) \propto \left[B_0^{\zeta(\ell)} \bar{S}(\zeta(\ell))^{\frac{1}{1+\zeta(\ell)}}\right]^{\frac{1}{\frac{\alpha}{1-\alpha}+\frac{\zeta(\ell)}{1+\zeta(\ell)}}}$$

When B_0 is above 1, then $f(\zeta)$ becomes an increasing function. It can also be non-monotonic depending on parameter values. Thus, whether the sorting can be sustained is not robust to changes in parameter values.

To see this clearly, consider a Taylor expansion when $\zeta(\ell) \gg 1$. It is not hard to see that

$$\log f(\ell) \approx_{\zeta(\ell)\to\infty} \zeta(\ell) \cdot (1-\alpha) \log B_0 + o(1)$$

and so the sign only depends on whether B_0 is larger or smaller than 1.

D Estimation

D.1 Time-dependent KFE

The first step for the estimation is to compute an explicit solution to the time-dependent KFE:

$$g_t = L_y^* g - (\Delta + \mu) g$$

where t denotes tenure at a job, and subscripts denote partial derivatives. Define

$$g(t,y) = e^{-(\Delta+\mu)t} h(t,y)$$

Then $g_t = e^{-(\Delta+\mu)t}(h_t - kh)$ so that

$$h_t = L_y^* h \tag{63}$$

The solution to this PDE is known. In logs, $x = \log y$, define

$$\Gamma(t,x) = \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{(x+\delta t)^2}{2\sigma^2 t}} \qquad ; \qquad G(t,x,y) = \Gamma(t,x-y) - e^{\frac{2\delta}{\sigma^2}(y-\underline{x})} \cdot \Gamma(t, x+y-2\underline{x})$$

Then it is straightforward to check that

$$h(t,x) = \int_{\underline{x}}^{\infty} G(t,x,y) h_0(y) dy$$

is the solution to (63) in logs, with initial distribution h_0. See Luttmer (2007) and references therein for a similar result. The details of the derivation are available upon request. Then, in logs,

$$g(t,x) = e^{-(\Delta+\mu)t} \int_{\underline{x}}^{\infty} G(t,x,y) g_0(x_0) dx_0 \tag{64}$$

is the time-dependent distribution of log productivity across employed workers in a location with log cutoff $\underline{x}$ given a starting distribution g_0.

D.2 Tenure profile of job loss

Now fix a location ℓ and omit ℓ indices for simplicity. Normalize $\underline{x} \equiv 0$ without loss of generality. The flow of workers into local unemployment is

$$\text{Endog. Sep.}(t) = \frac{\sigma^2}{2} \frac{\partial g}{\partial x}(t, \underline{x})$$

which can be calculated at all times using the explicit solution (64). Straightforward yet lengthy algebra leads to

$$\begin{aligned} & \text{Endog. Sep.}(t) \\ = \quad & e^{-(\mu+\Delta)t} \left[\frac{(s/\hat{\delta})}{\sqrt{t}} \varphi\left(\hat{\delta}\sqrt{t}\right) + \frac{(s/\hat{\delta})^2}{2} e^{\frac{(s/\hat{\delta})^2 t}{2}} \left\{ e^{st}\Phi\left(-(\hat{\delta} + s/\hat{\delta})\sqrt{t}\right) - e^{-st}\Phi\left((\hat{\delta} - s/\hat{\delta})\sqrt{t}\right) \right\} \right] \end{aligned}$$

where $s = \delta\zeta$ is the local average job losing rate (into local unemployment), and $\hat{\delta} = \delta/\sigma$.

To get the time-aggregated job losing rate in the first year, denoted $s_1(s, \hat{\delta}, D)$, integrate between 0 and 1 against g. Yet lengthier algebra leads to

$$\begin{aligned} & s_1(s, \hat{\delta}, D) \\ = \quad & \frac{(s/\hat{\delta})}{4} \left\{ \frac{e^{-(D+s-(s/\hat{\delta})^2/2)}}{D + s - (s/\hat{\delta})^2/2} (s/\hat{\delta}) + 4 \frac{\text{Erf}\left[\frac{\sqrt{2D+\hat{\delta}^2}}{\sqrt{2}}\right]}{\sqrt{2D+\hat{\delta}^2}} \right. \\ & + \frac{s/\hat{\delta}}{D + s - (s/\hat{\delta})^2/2} \left(-1 - \frac{\hat{\delta} - s/\hat{\delta}}{\sqrt{2D+\hat{\delta}^2}} \text{Erf}\left[\frac{\sqrt{2D+\hat{\delta}^2}}{\sqrt{2}}\right] + e^{-(D+s-(s/\hat{\delta})^2/2)} \text{Erf}\left[\frac{\hat{\delta} - (s/\hat{\delta})}{\sqrt{2}}\right] \right) \\ & + \frac{(s/\hat{\delta})}{D - s - (s/\hat{\delta})^2/2} \left(-1 + \frac{\hat{\delta} + s/\hat{\delta}}{\sqrt{2D+\hat{\delta}^2}} \text{Erf}\left[\frac{\sqrt{2D+\hat{\delta}^2}}{\sqrt{2}}\right] + e^{-D+s+(s/\hat{\delta})^2/2} \text{Erfc}\left[\frac{\hat{\delta} + (s/\hat{\delta})}{\sqrt{2}}\right] \right) \end{aligned}$$

where $D = \Delta + \mu$, and Erf denotes the error function (a transformation of the Gaussian cumulative function). In the limit of a small D, it is tedious but straightforward to check that s_1 is a decreasing function of $\hat{\delta}$.

D.3 Tenure profile of wages

A similar computation as for job loss first delivers average labor productivity net of human capital by tenure

$$\begin{aligned}\bar{y}(t)/\underline{y} &= \frac{s/\hat{\delta}/\sigma}{2}\Bigg(\frac{e^{A_0(s,\sigma,\hat{\delta})t}\mathrm{Erf}\big[A_1(s,\hat{\delta})\sqrt{t}\big] - \mathrm{Erf}\big[A_2(\sigma,\hat{\delta})\sqrt{t}\big]+1}{s/\hat{\delta}/\sigma - 1} \\ &\quad - e^{-A_3(\sigma,D,\hat{\delta})t}\frac{e^{A_0(s,\sigma,\hat{\delta})t}\Big(\mathrm{Erfc}\big[A_1(s,\hat{\delta})\sqrt{t}\big]-2\Big)+\mathrm{Erfc}\big[A_2(\sigma,\hat{\delta})\sqrt{t}\big]}{s/\hat{\delta}/\sigma+1-2\hat{\delta}/\sigma} \\ &\quad + \frac{e^{A_4(s,\hat{\delta},D)t}}{s/\hat{\delta}/\sigma - 1}\Bigg) \\ &\equiv Y(t,s,\sigma,\hat{\delta},D)\end{aligned}$$

where $A_0(s,\sigma,\hat{\delta}) = \frac{1}{2}\left((s/\hat{\delta})^2 - \sigma^2 - 2s + 2\sigma\hat{\delta}\right)$, $A_1(s,\hat{\delta}) = \frac{\hat{\delta}-s/\hat{\delta}}{\sqrt{2}}$, $A_2(\sigma,\hat{\delta}) = \frac{\hat{\delta}-\sigma}{\sqrt{2}}$, $A_4(s,\hat{\delta},D) = \frac{(s/\hat{\delta})^2-2s-2D}{2}$ and $A_3(\sigma,D,\hat{\delta}) = \sigma\hat{\delta} + D - \sigma^2/2$.

Average wages of continuing jobs at tenure t relative to new jobs in a given location at calendar time t_{cal} are then

$$\frac{W(t,t_{\text{cal}})}{W^{\text{new}}(t_{\text{cal}})} = \underbrace{(s/\hat{\delta}/\sigma - 1)\Big[(1-\beta) + \beta\underline{y}_0/\hat{\rho}Y(t,s,\sigma,\hat{\delta},D)\Big]}_{\equiv w(t,\hat{\delta},\sigma,s)}$$

When $\hat{\delta}, D$ are small, lengthy algebra ensures that the slope of Y with tenure falls as σ rises. When β is small enough,

$$\log w(t,\hat{\delta},\sigma,s) \approx \text{constant} + \log(s/\hat{\delta}/\sigma - 1) + \log Y(t,s,\sigma,\hat{\delta},D)$$

Assuming that β is small, this equation can be taken to the data with NLLS, using

$$\log\frac{W_{t_{\text{cal}},t_{\text{tenure}},c}}{W^{\text{new jobs}}_{t_{\text{cal}},c}} = \log w(t_{\text{tenure}},\hat{\delta},\sigma,s_c)\ , \tag{65}$$

where $W_{t_{\text{cal}},t_{\text{tenure}},c}$ denotes average wages in city c at calendar time t_{cal} and tenure t_{tenure}. $W^{\text{new jobs}}_{t_{\text{cal}},c}$ denotes average wages of new jobs. I time-aggregate (65) at the quarterly frequency.

In practice, β may not be small, so it must be estimated jointly with σ. To do so, I numerically search for the pair (β,σ) that minimizes the sum of square deviations from equations (65) and the labor share equation in the main text.

D.4 Labor share

From the bargaining solution, the labor share in location ℓ is

$$\text{Labor Share}(\ell) = \frac{(1-\beta)(b+v(\ell)+\beta\underline{y}_0/\rho H(\ell)}{E(\ell)} \tag{66}$$

where $H(\ell) = \mathbb{E}_\ell[y/\underline{y}|y\geq\underline{y}]$ is expected labor productivity in location ℓ under the invariant distribution. Using the solution to the KFE, one obtains

$$H(\ell) = \frac{\kappa\zeta(\ell)}{(\kappa-1)(\zeta(\ell)-1)} = \frac{\kappa}{(\kappa-1)(1-z(\ell))} \equiv H(z(\ell))$$

D.5 Learning parameters

Log real wages are proportional to $K_tR_t^{-\omega}$, where t is caldendar time, K_t the average knowledge of the economy and R_t average house prices. In the data, economy-wide log real wages grow by 0.0015 each quarter. In the model, $K_tR_t^{-\omega} \propto K_t^{1-\omega-\psi}$ up to a constant. Thus, $\lambda = \frac{0.0015}{1-\omega-\psi} = 0.0023$.

Then, notice that all workers who become unemployed in a given location have the same wage when they are laid off: the reservation wage. While they are unemployed, their human capital grows at rate $\lambda - \varphi$. When they find a new job, they draw a productivity from the local new job distribution, which is independent from their history. Therefore, equation (21) obtains. In empirical specifications, I follow the literature and restrict the sample to workers that held a job for at least two years before becoming unemployed. This restriction ensures that the estimates are not driven by temporary jobs.

The model abstracts from additional mechanisms that could create a correlation between new productivity draws an workers' past unemployment or employment history, as in Jarosch (2015). Thus, in practice, equation equation (21) may deliver a biased estimate of the depreciation rate of human capital. To address such concerns, I run version of equation (21) with additional controls that account flexibly for workers' past employment history. I also control for worker-level unobserved heterogeneity. For completeness, I also propose a specification where I use employed workers as a control group in a difference-in-difference specification – although this control group introduces an additional endogeneity problem. Results for the estimate of $\lambda - \varphi$ are reported in the first row of Table 7. The point estimate remains stable across specifications, although controlling for worker-level heterogeneity reduces the coefficient by half.

Table 7: Unemployment scar estimation. Dependent variable = post-unemployment log wage

	Unemployed only					DiD
	(1)	(2)	(3)	(4)	(5)	(6)
Job loss × Duration	-0.019***	-0.012***	-0.011***	-0.011***	-0.010**	-0.006*
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Job loss						-0.086*
						(0.041)
Duration						0.001
						(0.001)
Pre log wage		0.535***	0.421***	0.385***	-0.246**	-0.058
		(0.022)	(0.021)	(0.018)	(0.081)	(0.048)
Skill			0.003***	0.003***		
			(0.000)	(0.000)		
Fixed Effects						
Year	✓	✓	✓	✓	✓	✓
2-digit Industry				✓	✓	✓
City				✓	✓	✓
Worker					✓	✓
Obs.	30952	30952	30952	30952	4940	66186
R^2	0.029	0.270	0.315	0.339	0.812	0.769
W.-R^2	0.012	0.257	0.303	0.253	0.049	0.003

Standard errors in parenthesis, two-way clustered by city and 2-digit industry.

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

D.6 Lower bound of productivity draws

I propose a microfoundation of job search that allows to use data on duration since last job offer to inform Y. In the LFS, there is data on duration since last contact with the national unemployment agency (at the time called ANPE, "Agence Nationale Pour l'Emploi", now called "Pôle Emploi") and duration since last job offer. The latter is not necessarily an offer that came from the ANPE.

To leverage this data, I assume that individuals contact either the national unemployment agency, or the private sector with intensity S. Conditional on a contact, it is a contact with the ANPE with probability s, and a contact with the private sector with probability $1-s$. Conditional on contacting the ANPE, workers they receive an offer with probability ω. Conditional on receiving an offer, they accept it with probability a. Conditional on a private sector contact, they receive offers with conditional probability τ. They accept them with conditional probability a. The key is that the conditional acceptance probability a is the same. Allowing for private sector contacts is also important because only 6.58% of jobs are found through the ANPE.

Unemployment duration in sample of unemployed individuals. The rate at which an individual leaves unemployment in dt is $S(s\omega+(1-s)\tau)a\cdot dt$. Therefore, the probability that a currently unemployed individual has been unemployed for exactly n small dt periods is

$$p_n^u \propto [1-S(s\omega+(1-s)\tau)adt]^n$$

Note that a given amount of time is $T=n\cdot dt$. The expected unemployment duration in a sample of unemployed individuals is thus

$$D^U = \left\{\frac{1}{S(s\omega+(1-s)\tau)a}\cdot\frac{1}{dt}-1\right\}\cdot dt \to_{dt\to 0} \frac{1}{S(s\omega+(1-s)\tau)a}$$

So $D^U = e^{S(s\omega+(1-s)\tau)a}$.

Composition of job findings. The probability that an individual finds a job in a quarter through the ANPE is $s\omega a$, and through other channels $(1-s)\tau a$. Thefore, the probability that an employed individuals has found a job through ANPE is

$$P^{ANPE} = \frac{s\omega a}{s\omega a+(1-s)\tau a}$$

At this point, one can thus identify $x\equiv Ss\omega a$ and $y\equiv S(1-s)\tau a$: $x+y=\frac{1}{D^U}$ and $x=P^{ANPE}\times(x+y)$

ANPE contacts. The probability that a currenly unemployed individual has last contacted the ANPE n periods ago and did not find a job is thus

$$p_n \propto \Big[1-Sdt+S(1-s)(1-\tau a)dt\Big]^n$$

So the expected duration since the last ANPE contact is, similarly to before,

$$D^C = \frac{1}{S\big(1-(1-s)(1-\tau a)\big)} = \frac{1}{Ss+y}$$

which identifies Ss given x and y: $Ss=\frac{1}{D^C}-y$. Hence, $X=\omega a=\frac{x}{Ss}$ is known and $z=\frac{x}{y}=\frac{\omega}{\tau}\frac{s}{1-s}$, and so ω/τ.

ANPE offers. Similarly, the probability that a current unemployed worker has last received an offer from ANPE n periods ago is

$$q_n \propto \Big[1-Sdt+S\big((1-s)(1-\tau a)+s(1-\omega)\big)dt\Big]^n = \Big[1-S\big(1-(1-s)(1-\tau a)-s(1-\omega)\big)dt\Big]^n$$

So the expected duration since the last ANPE offer is

$$D^O = \frac{1}{S\big(1-(1-s)(1-\tau a)-s(1-\omega)\big)}$$

Re-write this as $\frac{1}{D^O}=Ss\omega+y$, which identifies $Ss\omega=\frac{1}{D^O}-y$, and therefore $a=\frac{x}{Ss\omega}$.

D.7 Local wages

Wages in location ℓ are given by

$$\overline{w}(\ell) = W_0 p\cdot\left(\frac{\Delta}{\Delta+\varphi u(\ell)}\right) r(\ell)^{-\psi}\underline{y}(\ell)\left[(1-\beta)\frac{\rho}{\underline{y}_0}+\beta H(s(\ell))\right] \tag{67}$$

where W_0 is a general equilibrium constant.

D.8 Housing elasticity

Using (54) together with the solution for average wages in a location $W(\ell)$, housing prices can be expressed as

$$r(\ell)^{1+\eta} \quad = \quad \frac{\overline{w}(\ell)}{1-\beta+\beta\underline{y}_0/\hat{\rho}E(s(\ell))} \cdot L(\ell,a)G(s(\ell)/\delta, v(\ell))$$

The right-hand-side defines r_0, and involves parameters that have been estimated or data.

D.9 Migration elasticity

To circumvent endogeneity in the OLS regression version of (23), I use changes in predicted local employment as an instrument. I break down the sample in two subperiods, and, in this section only, I use the notation Δ to refer to changes between these two periods. Specifically, I use predicted changes in local employment ΔE_c from Appendix A.2. To understand this instrument within the model, assume that there is a set $\mathcal{J} = \{1,...,J\}$ of industries. Employers in each industry draw from the same productivity distribution F_z. Locations c differ in a set of industry-specific productivities $\{p_{jc}\}_j$. Consistent with larger cross-industry flows than cross-location worker flows, suppose that there is a single labor market for all industries within a location. Suppose further that the cross-industry variance in industry productivity $\mathrm{Var}_c(p_{jc})$ is much smaller than the cross-location variance in city productivity $\mathrm{Var}_j(p_{jc})$. This assumption implies that the industrial mix is not strongly predictive of the local unemployment rate, consistent with the data. Under this assumption, the single-industry model is also a close approximation to the multi-industry model.

Now consider a set of industry-wide shocks that change p_{jc} to $p'_{jc} = p_{jc}\hat{p}_j$. Vacancy creation reacts to changes $\hat{p}_j$, so that national employment in industry j is positively correlated with $\hat{p}_j$. Similarly, employment shares $E_{jc,0}$ in the first subperiod are correlated with p_{jc}. Suppose that (1) $\hat{p}_j$ are uncorrelated with p_{jc}, (2) $\hat{p}_j$ are i.i.d. across industries. Then $\hat{p}_j$ are uncorrelated with changes in amenities Δa_c in the population-weighted distribution of cities and industries, even if $\mathbb{E}_c[p_{jc}]$ are correlated with amenities. With a large number of industries and locations, the shift share ΔE_c is thus correlated with the average change in local productivity $\mathbb{E}_c[\Delta p'_{jc}]$. In general equilibrium, employers relocate in each industry, and so ΔE_c is also correlated with $\mathbb{E}_c[\Delta\zeta_{jc}]$. If anything, this correlation makes the instrument stronger. The crucial exclusion restriction is that ΔE_c is uncorrelated with changes in amenities Δa_c. Therefore, it constitutes a valid instrument in this augmented model with small industry heterogeneity.

D.10 Productivity distribution

To estimate F_z, I first recover firm quality in each location using (22). It is easier to work with the reciprocal of firm quality z, denoted $\zeta = 1/z$. Consider locations with profitability in $(\ell - d\ell, \ell]$. Because the job losing rate is strictly decreasing in ℓ, they are are exactly those with a job loss rate in $[s(\ell), s(\ell)+ds(\ell)]$. Due to the model's sorting implications, the mass of open jobs in those locations is proportional to $f_\zeta(\zeta(\ell))d\zeta(\ell) = \delta^{-1}f_\zeta(\zeta(\ell))ds(\ell)$.

For simulations, I estimate a a Beta distribution for ζ: $f_\zeta(\zeta) \propto \left(\frac{\zeta-\underline{\zeta}}{\bar{\zeta}-\underline{\zeta}}\right)^{g_2}\left(\frac{\bar{\zeta}-\zeta}{\bar{\zeta}-\underline{\zeta}}\right)^{g_1}$ which is equivalent to a Beta distribution for z. I estimate the Beta distribution by minimizing the mean square error between the empirical density function (a histogram) and the Beta density.

D.11 Matching function elasticity

Start from

$$\theta(\ell) \quad = \quad \left(\frac{f_\zeta(\ell)|\zeta'(\ell)|}{f_\ell(\ell)\mathcal{U}(\ell)}\right) J(\ell)^\gamma$$

But

$$J(\ell) = q(\ell)\bar{J}(\ell) \propto \theta(\ell)^{-\alpha}\hat{J}(\ell)$$

where

$$\hat{J}(\ell) \quad \propto \quad \mathbb{E}(u(\ell))^{\frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \cdot \ell(b+v(\ell))^{1-\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \cdot G(v(\ell),\zeta(\ell))^{\frac{\varepsilon\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)} \bar{\mathcal{S}}(\zeta(\ell))$$

Therefore,

$$\theta(\ell)^{1+\alpha\gamma} \propto \left(\frac{f_\zeta(\ell)|\zeta'(\ell)|}{f_\ell(\ell)\mathcal{U}(\ell)}\right) \bar{J}(\ell)^\gamma$$

and so

$$f_R(\ell) \Big/ \left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)} \propto \left(\frac{f_\zeta(\ell)|\zeta'(\ell)|}{f_\ell(\ell)\mathcal{U}(\ell)}\right)^{\frac{1-\alpha}{1+\alpha\gamma}} \hat{J}(\ell)^{\frac{\gamma(1-\alpha)}{1+\alpha\gamma}}$$

Taking logs delivers

$$\log\left(\frac{f_R(\ell)}{\mathbb{P}_\ell[\text{Accept}]}\right) = \text{cste} + \frac{1-\alpha}{1+\alpha\gamma}\log\frac{f_z(z(\ell))z'(\ell)}{f_\ell(\ell)\mathcal{U}(\ell)} + \frac{(1-\alpha)\gamma}{1+\alpha\gamma}\log\hat{J}(\ell, \underline{y}(\ell), z(\ell)) \ , \tag{68}$$

where recall that $\mathcal{U}(\ell)$ denotes the number of unemployed workers in location ℓ, and $\hat{J}(\ell, \underline{y}(\ell), z(\ell))$ is now known. At this stage, both right-hand-side variables can be calculated. In the model, equation (68) can be estimated with OLS. It is not hard to add location-specific heterogeneity in the matching function efficiency or vacancy costs to the model. In that case a structural residual correlated with the right-hand-side variables arises. In contrast to the previous estimating equations, this structural residual leads to omitted variable bias in equation (68).

With OLS, α, γ are separately identified only through functional form differences between the right-hand-side variables because both are functions of the same latent variable ℓ. 2SLS also relies on functional form identification. Thus, I use the local shift-share shock and a non-linear transformation thereof as two instruments. Notice also that in the generalized model with omitted variable bias, the latter only affects the estimation of equation (68), and not the previous estimating equations. Indeed, the previous estimating equations condition on the observed job losing rate, which is enough to control for the omitted variables through local job quality $z(\ell)$.

I first-difference (68) between the two subperiods. I use as the first instruments the same shift-share shocks ΔE_c. Under the same assumptions as in section D.9, it is a valid instrument. To obtain a second instrument, I de-mean ΔE_c and use $\mathbb{1}_{\Delta E_c > 0}$. This is a nonlinear transformation of ΔE_c. Strengthening the identification assumption to conditional independence makes it a valid instrument.

D.12 Over-identification

Table 8: Cross-sectional regression of labor share onto log wages.

	(1) Data	(2) Model
log(Wage)	-0.192*	-0.114***
	(0.084)	(0.027)
Constant	0.728***	0.724***
	(0.004)	(0.003)
Obs.	348	348
R^2	0.066	0.046

Standard errors in parentheses

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Standard errors in the model regressions reported for completeness.

They are population objects and do not reflect statistical uncertainty.

Table 9: Correlation of estimated amenities with observables.

	(1)
Weather	
Sun hours	0.333*
	(0.151)
Services	
Basic public	0.051
	(0.132)
Education	0.113
	(0.107)
Health	0.223**
	(0.041)
Commercial	0.098
	(0.085)
Obs.	288
R^2	0.517

Robust S.E. in parenthesis.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.
Log amenities on log sun hours per month and log service establishments.

D.13 Validation

Table 10: Plant-level labor productivity across space.

	Log			Growth rate		
	(1) VA/N	(2) VA/N	(3) VA	(4) VA/N	(5) VA/N	(6) VA
Geography						
Job losing rate	-0.373***	-0.241	-0.140	0.009	-0.002	0.027**
	(0.079)	(0.150)	(0.143)	(0.012)	(0.016)	(0.008)
Entrant × Job losing rate	-0.433**	-0.317^{+}	-0.516***			
	(0.142)	(0.174)	(0.131)			
Skill mix controls						
Skill mix		0.287**	0.001		0.005	0.000
		(0.107)	(0.029)		(0.005)	(0.003)
Entrant × Skill mix		0.085**	0.025			
		(0.031)	(0.039)			
Fixed Effects						
Year × Entry status	✓	✓	✓	✓	✓	✓
2-digit industry × Entry status	✓	✓	✓	✓	✓	✓
Obs.	785252	383099	383099	692951	333873	333873
R^2	0.173	0.213	0.084	0.005	0.006	0.003

Standard errors in parenthesis, two-way clustered by city and 2-digit industry.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Annual frequency, 2003-2005.
Davis-Haltiwanger growth rate. Continuing plants only. Entrant defined as less than two year old.
All value added per worker regressions are employment-weighted.

Job loss at survivers and exiters. Local job losing rates can be broken down into the contribution of job loss at surviving establishments and the contribution of exiting establishments. I plot that simple accounting decomposition in Figure 25. The y-axis values of the blue circles and the green diamonds add up to the 45 degree line in orange. It shows that job loss at surviving establishments is the dominant source of geographical variation in the job losing rate, accounting for 89% of the cross-location variance.

Figure 25: Job loss from surviving and exiting establishment, by French commuting zones.

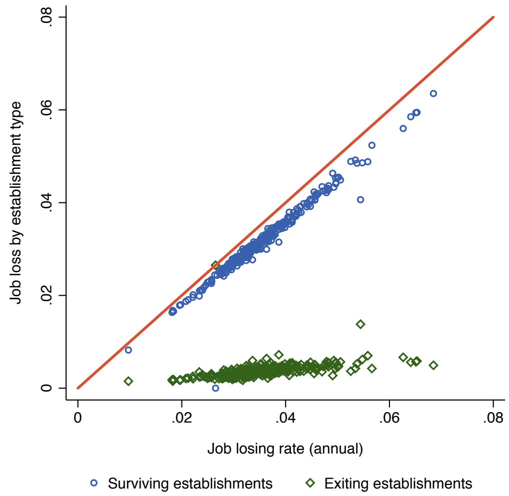


Table 11: Job losing rate at new plants within multi-plant openings.

	(1)	(2)
Local job losing rate	1.34**	0.30
	(0.42)	(0.38)
Controls and FEs		
Skill-Year	✓	✓
3-digit industry-Year	✓	✓
Firm-Year		✓
Obs.	21676	20046
R^2	0.143	0.635

S.E.s in parenthesis, two-way clustered by commuting zone and 2-digit industry.
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

E Results: robustness

Table 12: Robustness to bargaining power and searcha assumptions.

	Model					Data
Entry cost estimation	$\beta \neq \alpha$			$\beta = \alpha$		
Estimation	$\beta \neq \alpha$	$\beta \neq \alpha$	$\beta \neq \alpha$	$\beta = \alpha$	$\beta = \alpha$	
Bargaining power	$\beta \neq \alpha$	$\beta \neq \alpha$	$\beta = \alpha$	$\beta = \alpha$	$\beta = \alpha$	
Search	Random	Directed	Random	Random	Directed	
Aggregate unemployment rate	0.076	0.071	0.127	0.076	0.076	0.076
St. dev. unemployment rate	0.022	0.004	0.002	0.004	0.004	0.023
Var. log unemp. / emp.	0.079	0.002	0.002	0.012	0.001	0.112
Job losing rate	85 %	-180 %	380 %	183 %	32 %	86 %
Job finding rate	15 %	280 %	-280 %	-83 %	68%	14 %