The Geography of Unemployment*

Adrien Bilal
Harvard University

September 13, 2021

Abstract

Unemployment rates differ widely across local labor markets. I offer new empirical evidence that high local unemployment emerges because of elevated local job losing rates. Local employers, rather than local workers, account for most of spatial gaps in job stability. I then propose a theory in which spatial differences in job loss arise endogenously, due to the spatial sorting of heterogeneous employers across local labor markets. Labor market frictions induce productive employers to overvalue locating close to each other. The optimal policy incentivizes them to relocate to areas with high job losing rates, providing a rationale for commonly used place-based policies. I estimate the model using French administrative data. The estimated model accounts for over three fourths 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. Both real-world and optimal place-based policies can yield sizable local and aggregate welfare gains.

*Email: adrienbilal@fas.harvard.edu. Website: `https://sites.google.com/site/adrienbilal`. This paper is a revised version of the first chapter of my Ph.D. dissertation at Princeton University. 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, as well as numerous seminar and conference participants. I gratefully acknowledge financial support from the Simpson Center for Macroeconomics, the International Economics Section, and the Industrial Relations Section at Princeton University. This work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10- EQPX-17—Centre d'accès sécurisé aux données—CASD). Special thanks to 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, 5% of workers were unemployed. In southern Marseille, the unemployment rate exceeded 12%. Comparable differences arise in the United States and in most other developed countries.[1] Despite their magnitude, these spatial gaps persist over decades. 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 and 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, that are in turn tied to local employers rather than local workers. Second, I propose a theory that accounts for spatial differences in job stability through the location choice of heterogeneous employers across frictional local labor markets. Third, I estimate the framework on French administrative data. Fourth, I quantify the local and aggregate welfare gains from place-based policies in general equilibrium. I now describe each part in more detail.

In the first part of the paper, I examine how local labor market flows differ between locations. I use French matched employer-employee data to assess whether differences in unemployment rates across commuting zones reflect differences between job losing (inflow) versus job finding (outflow) rates. Differences in job losing rates emerge as the primary source of spatial unemployment differentials, accounting for 77% of the variation. In contrast, job finding rates are close to constant across locations. Using a two-way fixed effect approach, I then establish that employer-specific heterogeneity accounts for two thirds of spatial job loss differences, while worker-specific heterogeneity accounts for only one third. Similar patterns also hold in the Current Population Survey 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 composition analysis 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. Employers offer jobs that differ in initial productivity, which subsequently fluctuates due to idiosyncratic shocks. As a result, endogenous job loss arises, and initially more productive jobs are more stable. Employers value two types of location characteristics. First, they value exogenous location fundamentals such as location-specific productivity. Second, they value endogenous location characteristics such as local wages and recruiting conditions.

The distinct interaction of the location choice of employers and labor market frictions gives rise to

[1] In 2017, the unemployment rate was 5% in Boston, Massachusetts. It was 13% in Flint, Michigan. See the OECD (2005) report for more countries.

[2] Changes in the job finding rate have been found to be the dominant force in aggregate unemployment fluctuations over 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.

labor market pooling complementarities that lie at the heart of the model's implications. 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, conditional on local productivity, 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 meeting 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. Of course, when local productivity differentials are large, cross-sectional sorting patterns are more complex.

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. In locations with many unemployed workers, there are also more employers since labor is cheaper. On net, workers meet with more employers: the worker meeting rate is high. However, these employers are unproductive, and thus meetings seldom result in a viable match: the success probability of a meeting is low. When both forces closely balance, the job finding rate is flat across locations. Reduced-form evidence using administrative establishment-level productivity and vacancy data supports these implications. Labor productivity correlates negatively with job losing rates across French establishments and commuting zones. Local labor market tightness rises modestly with local unemployment.

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, resulting in a labor market pooling externality. 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 subsidize employers in high unemployment locations.

The third part of the paper develops and structurally estimates a quantitative version of the framework. There are three main additions. First, locations also differ in residential amenities that capture non-monetary compensating differentials. Second, migration frictions introduce empirically plausible migration elasticities. Third, workers differ in human capital that depreciates while they are unemployed, thereby capturing scarring effects of unemployment.

Despite its richness, the quantitative model retains the tractability of the analytical framework. As a result, it produces estimating equations that allow for transparent identification leveraging the many dimensions of the French administrative data. In particular, a recursive scheme delivers a sequence of regression equations that identify all but one of the 19 parameters without requiring to simulate from the model. Over-identification checks support estimates of key parameters by highlighting that the model can match a number of non-targeted moments. In addition, 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 75% 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. 77% stem from the job losing rate in the data, against 73% in the model. In addition, the estimated model matches the endogenous relationship between local labor market flows on the one hand, and local wages and population on the other hand. 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%.

The fourth and last part of the paper conducts two counterfactuals that explore the impact of place-based policies in general equilibrium. I start by examining the effect of the optimal policy on local and aggregate outcomes. It takes the form of a corporate tax subsidy that is more generous in high unemployment locations. The optimal policy thus offsets the labor market pooling externality, and incentivizes marginally more productive employers to relocate towards high unemployment locations. The optimal policy cuts the local unemployment rate by over 10 percentage points in cities such as Marseille. It also achieves over 20% welfare gains for their residents. Long-run scarring effects of unemployment are central to these welfare gains, accounting for three fourths of the total. The optimal policy primarily redistributes jobs across locations, but it also ameliorates aggregate outcomes. The aggregate unemployment rate falls by 0.4 percentage points and utilitarian welfare rises by 1%. Aggregate welfare gains are more modest than gains in high unemployment locations because they average over a sizeable redistribution of resources across locations. As the most productive and stable jobs leave the lowest unemployment locations, residents experience welfare losses there.

I then contrast the optimal policy with the French Enterprise Zones (EZ) program—the *"Zone Franches Urbaines."* The French EZ program was rolled out in 1996 and consisted in heavy subsidies for businesses opening jobs in high unemployment areas. Qualitatively, the French EZ policy resembles the optimal policy. As such, it should deliver positive welfare gains. Quantitatively however, the French EZ program is much smaller than the optimal policy in scale and scope. The model indicates that the French EZ program reduced unemployment in treated areas by 2 to 3 percentage points. Local welfare gains do not exceed 5%, but once more mostly reflect reduced scarring effects of unemployment. In the aggregate, the EZ program raised welfare by 0.1%. Albeit modest, the impact of the EZ program is in fact 6 times higher per dollar spent than the optimal policy, due to decreasing returns to redistribution. This comparison suggests that small-scale place-based policies are likely to be more efficient than large-scale ones in the presence of fiscal externalities or political economy constraints.

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.[4] 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. Kline and Moretti (2013) find that subsidies to high unemployment areas reduce welfare. All these papers focus on the role of the job finding rate and abstract from job losing rate differentials. In contrast, I stress that job losing rate gaps are the key empirical determinant of spatial

[4]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.

unemployment differentials.[5] As a result, a different theory is required. It brings about 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, 2021).[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 including it leads to 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, 2020, 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, similar forces push less productive jobs to inefficiently locate in places that are too productive for them. In contemporaneous work, Brancaccio et al. (2020) 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) who coined labor market pooling as a key agglomeration force, 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, 2020). 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, Slattery, 2019, Fajgelbaum and Gaubert, 2020). 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 the 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, characterizes and empirically validates a simple model of spatial unemployment differentials with endogenous job loss. Section 3 describes its normative implications and lays out quantitative extensions. Section 4 discusses identification, estimation and over-identification. Section 5 presents the estimated model's account of spatial unemployment gaps and policy counterfactuals. The last section concludes. An Appendix and Supplemental Material collect proofs and additional details.

[5] See Hall (1972) for a study of 12 U.S. cities, and Topel (1984) for an analysis across U.S. states.

[6] For structural change over time, see Diamond (2016), Giannone (2017) Caliendo et al. (2021), Glaeser et al. (2018), and Couture et al. (2019).

[7] See Jarosch (2021) and Mangin and Julien (2021) 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, in turn tied to employers rather than workers. 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.

1.1 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 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 lets me break down the analysis by city and finely disaggregated employer and worker groups to control for composition.

One drawback of the DADS is that it only enables me 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, thereby limiting concerns related to 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. I aggregate the resulting sample at the quarterly frequency. Table 10 in Supplemental Material D.1 compares aggregate statistics in this sample and in the LFS.

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

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 2.6, I use firm-level balance sheet data covering the near universe of French business for the same period, as well as establishment-level vacancy data from a large-scale survey. 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 that 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. Supplemental Material D.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.

1.2 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 12% or below 6% can be found throughout the country. The cross-sectional standard deviation is 2.5 percentage points, twice as much as 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. 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 0.91.[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.

1.3 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 out-migration reduce the number of unemployed workers. In what follows, I use standard terminology

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

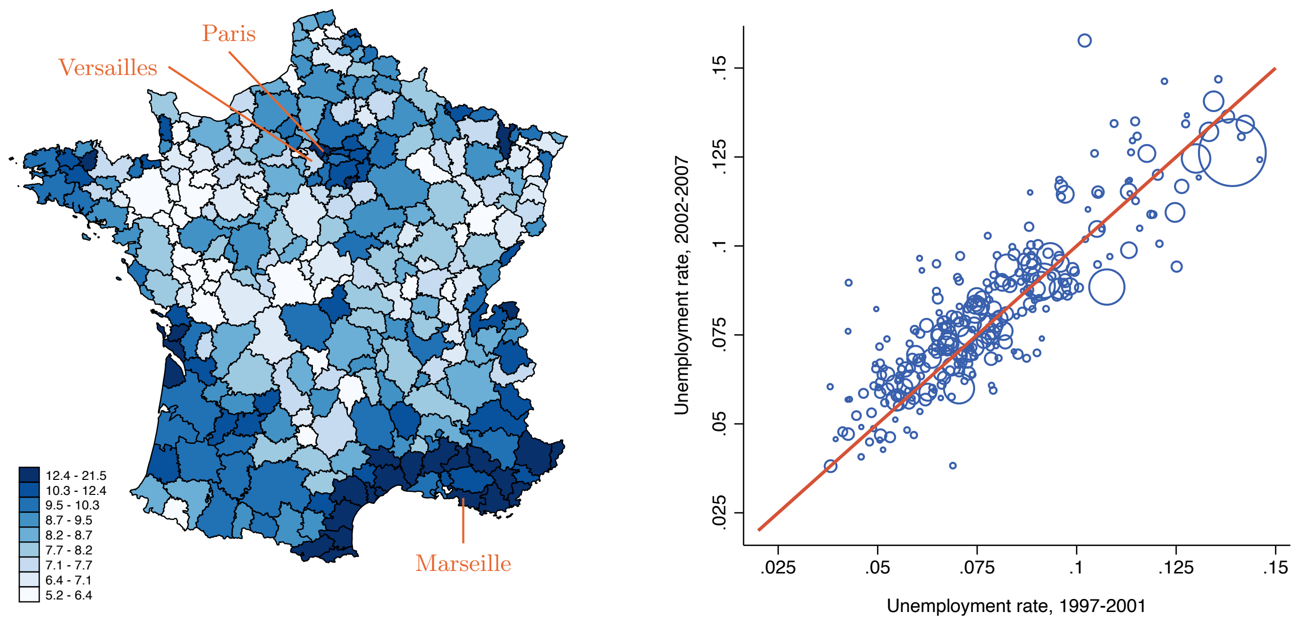
[13] I also check that using education to define skill in the CPS leaves the results unchanged.

[14] In Supplemental Material D.2, I show that controlling for economy-wide industry business cycles increases local persistence, with a conditional autocorrelation of 1.05.

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



Note: Figure 1(a) maps commuting zone unemployment rates from the DADS panel. Corsica and overseas territories omitted for exposition. Figure 1(b) plots commuting zone unemployment in two subperiods of the sample. Blue circles represent a commuting zone. Size is proportional to population.

from the literature and call 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

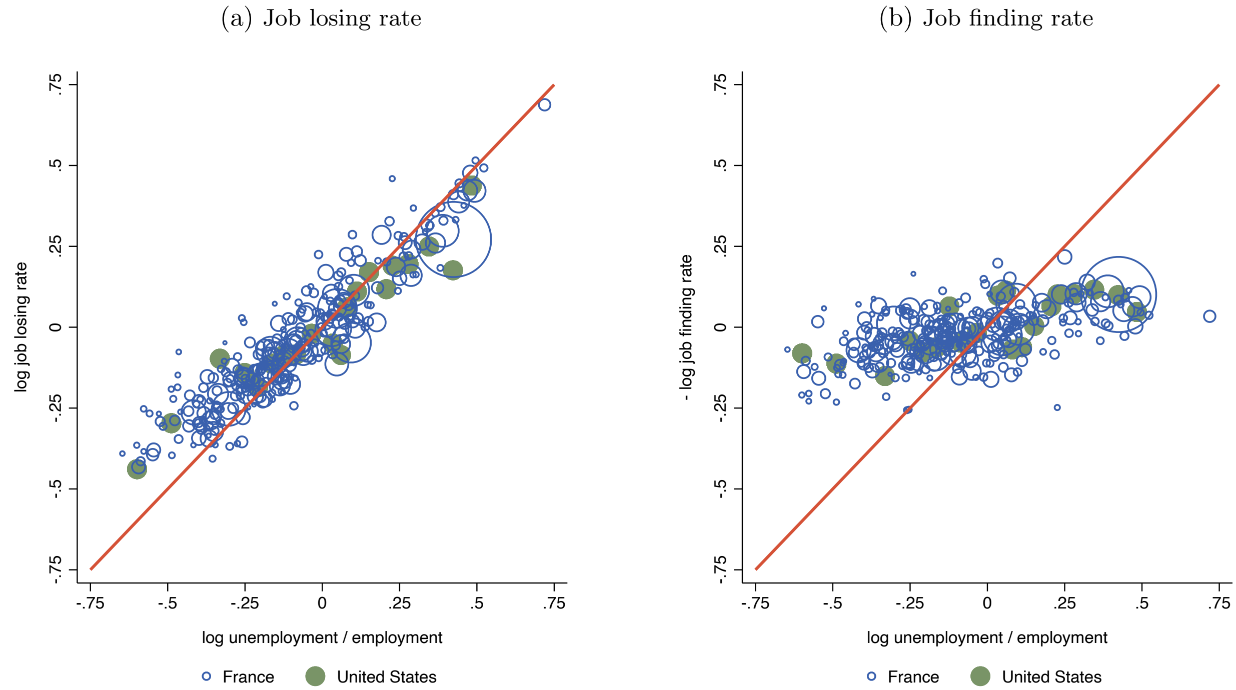
$$\log \frac{u_c}{1-u_c} = \log s_c - \log f_c. \tag{1}$$

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

To examine the respective contributions of the job losing and job finding rates to local unemployment, Figure 2(a) plots the logarithm of the job losing rate s_c against the logarithm of the unemployment-to-employment ratio $\frac{u_c}{1-u_c}$ across commuting zones, for France and the United States. The data align closely along the 45 degree line in orange for both countries, indicating that local job losing rates are the primary determinants of spatial unemployment differentials. Figure 2(b) 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.[15] Using equation (1) to run an exact variance decomposition, I find that the job losing rate accounts for 77% of the cross-sectional variation of the

[15] Similarly, Figure 14(b) in Supplemental Material D.2 reveals that the job-to-job mobility rate displays little systematic association with the local unemployment rate.

Figure 2: Local labor market flows and unemployment in France and in the United States.



Note: Figure 2(a) scatterplots the log of the job losing rate against the log of the unemployment-to-employment ratio, across commuting zones in France (DADS panel) and in the United States (CPS). Blue circles represent commuting zones in France, with size proportional to population. Green dots represent metro area groups in the US. US metro areas are grouped into 20 equally populated bins to reduce measurement error due to the smaller cross-sectional size of the CPS. 45 degree line in orange. Figure 2(b) scatterplots minus the log of the job finding rate against the log of the unemployment-to-employment ratio, across commuting zones in France and in the US.

spatial unemployment rate in France. The job finding rate accounts for the remaining 23%. In the US, the job losing rate accounts for 73% of the cross-sectional variation in spatial unemployment rate.

I establish the robustness of these results with several additional exercises. First, I verify that neither movements in and out of the labor force, migration and local transitional dynamics, nor time aggregation of quarterly probabilities into instantaneous rates, introduce a significant wedge between the left-hand-side and the right-hand-side of equation (1). Appendix A.1 derives the time aggregation correction and extends equation (1) to a three-state model with labor force participation, which can then be used for an exact variance decomposition. Table 5 in Appendix A.1 shows that the variance decomposition is robust to time-aggregating labor market flows, accounting for non-participation or using only the LFS in France.

Second, I verify that mechanical correlates of job loss such as temporary contracts, seasonality, firm exit or job reallocation can only account for a small fraction of spatial differences in job losing rates. Supplemental Material D.3 establishes that these phenomena account for 8 to 23% of spatial gaps in job loss. Thus, none of these phenomena can provide a systematic explanation for spatial differences in job losing rates.

1.4 Employer and worker composition

Differences in job losing rates across locations can arise because of two main reasons. First, workers who reside in some locations may separate into unemployment more frequently. Second, employers who open jobs in these same locations may offer jobs that are less stable.

To distinguish between these two explanations, I estimate a two-way fixed effect model in the spirit of Abowd et al. (1999):

$$\mathrm{EU}_{i,t} = \alpha_i + \beta_{J(i,t)} + \varepsilon_{i,t}, \tag{2}$$

where i indexes workers and α_i denotes a worker fixed effect. $J(i,t)$ denotes worker i's employer in quarter t, and $\beta_{J(i,t)}$ denotes an employer fixed effect. $\mathrm{EU}_{i,t}$ is an indicator variable taking the value 1 if worker i separates into unemployment in quarter t, and zero otherwise. In my main specification, I define an employer as an establishment-by-4-digit-occupation. This definition captures both spatial heterogeneity in the type of jobs across establishments within the same firm, as well as heterogeneity in the type of jobs across occupations within the same establishment. To alleviate concerns associated to limited mobility bias, I follow Bonhomme et al. (2019) and cluster worker, employers as well as commuting zones into groups before estimating (2).

After estimating (2), I retrieve the estimated group fixed effects, and average them within every commuting zone group c to obtain a sample analogue of

$$\mathrm{EU}_c = \mathbb{E}_c[\alpha_i] + \mathbb{E}_c[\beta_{J(i,t)}]. \tag{3}$$

Equation (3) breaks down the commuting zone quarterly job losing rate EU_c into an average worker component $\mathbb{E}_c[\alpha_i]$ and an average employer component $\mathbb{E}_c[\beta_{J(i,t)}]$. I use this decomposition to assess whether worker or employer composition contributes most to spatial job loss differentials.

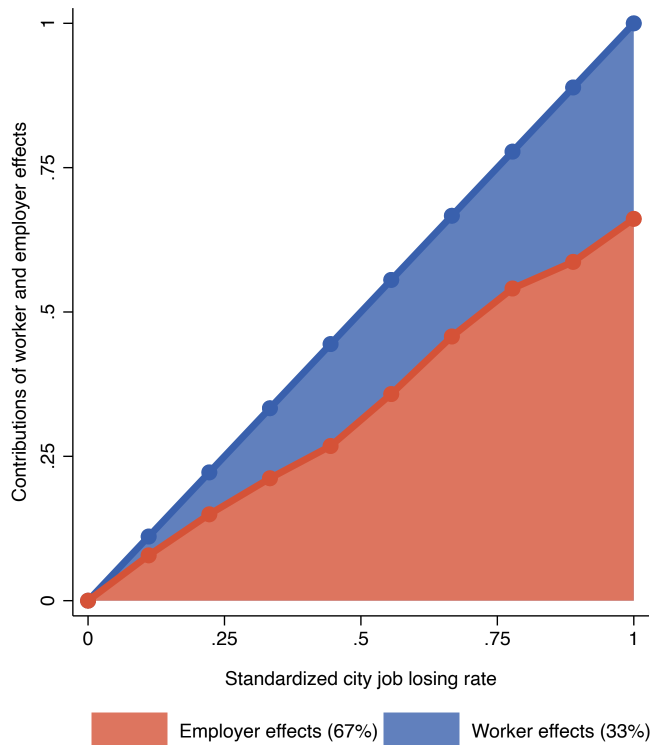
I find that systematic differences in the type of employers operating across cities are the primary reason why job losing rates differ across space. To reach this conclusion, Figure 3 plots the contributions of the average worker and employer components to the unconditional local job losing rate. The average employer effect accounts for 67% of the cross-sectional variation in job losing rates, while the average worker effect accounts only for 33%.

A number of additional exercises establish the robustness of the dominant role of employer composition. First, I assess whether my findings are related to between-industry or within-industry heterogeneity of employers. I residualize local job losing rates from industry and skill heterogeneity similarly to the specification in (2), and plot the estimated city fixed effects in Figure 12 in Appendix A.2. It reveals that the key heterogeneity driving job loss differentials arises within 3-digit industries rather than between industries—as well as within worker skills. This result implies that differences in cities' industry mixes are largely unrelated to their unemployment rate.

Second, I verify that my findings are robust to alternative econometric specifications. Table 6 in Appendix A.2 shows that results remain similar when using alternative definitions of employers, such as firms or firm-by-occupation, or when varying the number of clusters.

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 the composition of the workforce. Instead, spatial gaps in job loss primarily reflect systematic differences in the type of jobs offered by employers. 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

Figure 3: Contribution of worker and employer effects to local job losing rate in France.



Note: Orange: estimated average employer effect. Blue: estimated average worker effect. DADS panel, France. Workers, employers and commuting zones clustered into 10 population-weighted groups based on their unconditional job losing probability. X-axis represents the sum of employer and worker effects, standardized between 0 and 1.

losing rate and employer heterogeneity lie 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. Second, job loss is endogenous and tied to employers. Spatial sorting of employers leads to spatial gaps in job losing rates.

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.[16]

[16] Local productivity differences ℓ are useful to fix ideas and provide a natural ordering of locations, but are not necessary for the main mechanism. In the quantitative model, they will be microfounded through the human capital of local residents.

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 $\omega \in (0,1)$. Workers consume their income each period. They only search when unemployed.[17] 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.[18] An employer 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 their 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 $y_t \ell$: local productivity ℓ and the job's idiosyncratic productivity y_t are technological complements. Over time, the productivity of the job fluctuates according to a geometric Brownian motion

$$d \log y_t = -\delta dt + \sigma dW_t, \tag{4}$$

where $\delta > 0$ implies that 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.[19] 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 according to a Cobb-Douglas matching function $\mathcal{M}(\mathcal{U}(\ell), \mathcal{V}(\ell)) =$

[17] I do not incoporate is job-to-job mobility, for three 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 14(b) in Supplemental Material D.2 that, just like the job finding rate, the job-to-job mobility rate is only weakly correlated with the unemployment rate. Finally, adding job-to-job mobility would break the tractability of the model and make estimation and identification much more challenging.

[18] 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).

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

$m\mathcal{U}(\ell)^{\alpha}\mathcal{V}(\ell)^{1-\alpha}$. $\mathcal{U}(\ell)$ denotes the local number of unemployed workers, and $\mathcal{V}(\ell)$ denotes the local number of vacancies in that market.

Local market tightness is $\theta(\ell) = \mathcal{V}(\ell)/\mathcal{U}(\ell)$. Workers' local meeting rate is then $f(\theta(\ell)) = m\theta(\ell)^{1-\alpha}$ while the vacancy meeting rate for employers is $q(\theta(\ell)) = m\theta(\ell)^{-\alpha}$. The meeting rate may differ from the realized finding rate if some meetings 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.[20]

Wage determination. Workers and employers set wages according to generalized Nash bargaining, with worker bargaining power β. For simplicity, I assume that renegotiation occurs each instant.

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.[21]

2.2 Value functions

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.[22]

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 $V(y,\ell)$ be the value of employment at wage $w^*(y,\ell)$ in location ℓ. U and V satisfy the recursions

$$\rho U = b\ell r(\ell)^{-\omega} + f(\ell)\mathbb{E}_{\ell}\big[\max\{V(y_0,\ell) - U, 0\}\big] \tag{5}$$

$$\rho V(y,\ell) = w^*(y,\ell)r(\ell)^{-\omega} + (L_y V)(y,\ell), \tag{6}$$

where the recursion for V holds as long as the worker finds it optimal to remain in the match. The first term on the right-hand-side reflects workers' flow value when unemployed or employed. Because of Cobb-Douglas preferences and 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

[20] 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.

[21] 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.

[22] This argument is valid under a trembling-hand equilibrium refinement, since the two-sided location choice of workers and employers can otherwise result in coordinating on empty locations. I impose this trembling-hand refinement in the sequel.

on the right-hand-side of equation (5) 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 (6) reflects the expected continuation value of employment due to productivity shocks. Given the geometric Brownian motion assumption (4), the functional operator L_y is defined by

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

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

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

as long as the employer finds it optimal to keep the worker. Employers value flow profits $y\ell - w^*(y, \ell)$ as well as the contribution of future productivity changes.

Joint surplus and wage determination. Generalized Nash bargaining implies that worker-employer pairs set wages by maximizing the Nash product. Even though workers' and employers' marginal utility of a dollar differ due to housing consumption, I show in Lemma 4 in Appendix B.2 that the traditional microfoundation of generalized Nash bargaining with an alternative offers game *à la* Rubinstein (1982) continues to hold in my environment.

Lemma 4 lets me restrict attention to a single object that I call the adjusted surplus. It is defined as

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

The adjusted surplus is independent from wages because it puts each side's value on a common numeraire scale. 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:

$$r(\ell)^\omega \cdot \big(V(y, \ell) - U\big) = \beta S(y, \ell), \qquad J(y, \ell) = (1 - \beta) S(y, \ell). \tag{8}$$

In particular, both sides agree 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.2. I characterize its solution in the following Lemma.

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

$$S(y, \ell) \quad = \quad \frac{\ell \underline{y}(\ell)}{\underline{y}_0} \mathcal{S}\left(\frac{y}{\underline{y}(\ell)}\right), \quad \forall y \geq \underline{y}(\ell),$$

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

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

and $\tau, \underline{y}_0$ *are transformation of* ρ, δ, σ *given in Appendix B.2.*

Proof. See Supplemental Material E.2. □

The local endogenous separation cutoff $\underline{y}(\ell)$ increases as the worker's local value of unemployment relative to housing prices, $b + v(\ell)$, rises. $v(\ell)$ is the productivity-adjusted value of future employment opportunities to a worker.[23] The value of future employment opportunities $v(\ell)$ is 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: $S(\underline{y}(\ell), \ell) = \frac{\partial S}{\partial y}(\underline{y}(\ell), \ell) = 0$.[24]

It is also useful to define workers' reservation wage $\underline{w}(\ell)$ in each location, in efficiency units of local productivity ℓ. Using the wage equation (29) in Appendix B.2, reservation wages satisfies

$$\underline{w}(\ell) = \underline{w}_0 \underline{y}(\ell), \tag{9}$$

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 relative to local housing prices. Thus, the local reservation wage is higher.

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

$$U = \frac{\ell \underline{w}(\ell)}{\underline{w}_0 \underline{y}_0 r(\ell)^{\omega}}. \tag{10}$$

Across locations, higher housing prices compensate for 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

[23] In general equilibrium, $v(\ell)$ is itself related to the adjusted surplus, as can be seen from its expression in Lemma 1 together with the surplus-sharing rule (8).

[24] 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.

over the job's initial productivity distribution $G_0(y_0|z)$, adjusted for the vacancy meeting rate $q(\ell)$:

$$\rho\bar{J}(z,\ell) = q(\ell)(1-\beta)\int S(y_0,\ell)G_0(dy_0|z). \tag{11}$$

To facilitate 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 2.6. Nonetheless, I also provide more general distributional conditions under which my results hold in Supplemental Material E.3.

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

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

Under Assumption 1, I show in Appendix B.3.1 that the expected payoff of job z in location ℓ satisfies

$$\log\left(\left(\bar{\rho}\bar{J}(z,\ell)\right)^{\frac{z}{1-z}}\right) = \underbrace{\frac{z}{1-z}\log\bar{S}(z)}_{\substack{\text{Absolute}\\ \text{advantage}}} + \underbrace{\frac{z}{1-z}\log\ell}_{\substack{\text{Production}\\ \text{complementarities}}} + \underbrace{\frac{z}{1-z}\log q(\ell)}_{\substack{\text{Pooling}\\ \text{complementarities}}} - \underbrace{\log\underline{w}(\ell)}_{\substack{\text{Cost of}\\ \text{labor}}}, \tag{12}$$

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 (12) 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, capturing compensating differentials unrelated to labor market frictions. 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.

The third term in equation (12) 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 meeting 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 meetings do not result in a viable match, so that the vacancy filling rate and the vacancy meeting rate differ. The probability that a meeting 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 meeting 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 meeting rate $q(\ell)$ a labor market 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 (12) 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 meeting rates through labor market tightness $\theta(\ell)$. Therefore, employers face a trade-off between local vacancy meeting rates and local wages conditional on local productivity ℓ. From the pooling complementarity, more productive employers value high vacancy meeting 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 meeting 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. An employer with quality z thus solves

$$\ell^*(z) \;=\; \operatorname*{argmax}_{\ell} \;\frac{z}{1-z}\log\ell + \frac{z}{1-z}\log q(\ell) - \log\underline{w}(\ell). \tag{13}$$

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 the one in equation (13) that shares many features with standard static assignment problems.[25]

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 meeting 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. Second, well-known multiplicity issues may arise.[26]

I define an assignment 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 meetings in the labor market. $\underline{w}(\ell)$ is the equilibrium reservation wage that supports this location choice. Proposition 1 below characterizes the assignment.

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 (13) 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.

Proof. See Appendix B.3.3. □

[25] Examples thereof can be found in Sattinger (1993) and Davis and Dingel (2020). For an in-depth exposition of the underlying theory, see Topkis (1998), Villani (2003) and Galichon (2016).

[26] Gaubert (2018) also generates ex-post differences across cities when when employers' technology directly depends on city population. Grossman and Rossi-Hansberg (2012) provide an example of multiple equilibria in a spatial context with agglomeration economies and exogenous differences across locations. See Chade and Eeckhout (2019) for a similar discussion of multiplicity in a search and matching context.

Proposition 1 establishes existence of the assignment with positive assortative matching between local productivity ℓ and firm quality z: more productive jobs go to more productive locations. Restricting attention to assignments that exhibit weak positive assortative matching is only a mild restriction, for two reasons. First, positive assortative matching is the only possibility when the matching function elasticity α is not too large. Proposition 8 in Supplemental Material E.3 extends this result to more general distributional conditions for G_0. Second, any other potential steady-state assignment is dynamically unstable for any value of α, in a sense made precise in Proposition 9 in Supplemental Material E.4.

The equilibrium response of local reservation wages $\underline{w}(\ell)$ to the location choice of employers sustains the assignment. Reservation wages adjust up to the point where the marginal employer $z(\ell)$ is indifferent between locations ℓ and $\ell + d\ell$. Reservation wages reflect expected future wages conditional on starting work, which depend on equilibrium employer quality $z(\ell)$. Therefore, reservation wages rise with ℓ. However, reservation wages rise less than one-for-one relative to wages of employed workers due to discounting: unemployed workers must search for some time before finding a job and earning wages comparable to those of employed workers.

The strength of this discounting effect depends on the exogenous discount rate ρ, but also on the endogenous 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 worker meeting rate $f(\theta(\ell))$. Since the value of search $v(\ell)$ reflects both the rising expected wages conditional on work and the falling worker meeting 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.

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 (4) and is constant across locations by assumption. Therefore, any differences in local job losing rates must arise because of differences between the average starting productivity and the separation threshold. Both are related to the equilibrium assignment function $z(\ell)$ and reservation wage $\underline{w}(\ell)$.

To determine exactly 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),

$$0 = (L_y^* g)(y, \ell) + n(\ell) g_0(y, \ell), \qquad y > \underline{y}(\ell), \tag{14}$$

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.

Under the geometric Brownian motion assumption (4), it is given by

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

By construction, the density g must integrate to unity in each location: $1 = \int_{\underline{y}(\ell)}^{\infty} g(y,\ell)dy$. Because of Brownian shocks, the distribution must satisfy the additional boundary condition $g(\underline{y}(\ell),\ell) = 0$.[27] There always exists a closed-form solution to the KFE (14) with the boundary condition. Lemma 2 describes that solution under Assumption 1. The generalized solution is given in Supplemental Material E.6.

Lemma 2. *(Employment distribution)*
Let $\kappa = \frac{2\delta}{\sigma^2}$. Under Assumption 1, the solution to the KFE (14) with $g(\underline{y}(\ell),\ell) = 0$ satisfies

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

Proof. See Appendix B.4.2 □

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.4.1, I show that it is

$$s(\ell) = \frac{\sigma^2}{2}\frac{\partial g}{\partial y}(\underline{y}(\ell),\ell). \tag{15}$$

Close to the cutoff, only workers who receive a negative shock become unemployed. The number of job losers follows from the second order contribution of the mass of workers close to the cutoff $\frac{\partial g}{\partial y}(\underline{y}(\ell),\ell)$, because $g(\underline{y}(\ell),\ell) = 0$.

Expression (15) 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. The expression under the Pareto assumption is presented in the main text. Proposition 10 in Supplemental Material E.6 describes the general solution.

[27] 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. Although it is a standard mathematical result, a formal proof is provided for completeness in Supplementary Material E.5.

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

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

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

Proof. See Appendix B.4.3 □

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.[28] 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 meeting rate $f(\theta(\ell))$ and the probability that a given meeting results in a job, the success probability of a meeting $\left(Y/\underline{y}(\ell)\right)^{1/z(\ell)}$. The worker meeting rate depends positively on labor market tightness $\theta(\ell)$. More productive employers $z(\ell)$ benefit more from higher vacancy meeting rates $q(\theta(\ell))$, pushing towards a negative correlation between the worker meeting rate and $z(\ell)$ conditional on local productivity ℓ. However, the success probability of a meeting pushes in the other direction. In locations with more productive employers $z(\ell)$, meetings 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.

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

[28] 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.4. 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,

$$r(\ell) = \omega L(\ell)\Big(u(\ell)b\ell + (1-u(\ell))\overline{w}(\ell)\Big), \qquad \theta(\ell) = \frac{M_e F_z'(z(\ell))z'(\ell)}{u(\ell)L(\ell)F_\ell'(\ell)}, \tag{16}$$

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,

$$c_e = \int \bar{J}(z, \ell^*(z))dF_z(z), \qquad 1 = \int L(\ell)dF_\ell(\ell). \tag{17}$$

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 (5), (6), (8), the definitions in Lemma 1, (9), (10), (13), (14), (15), (16), and (17) hold. 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.

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.

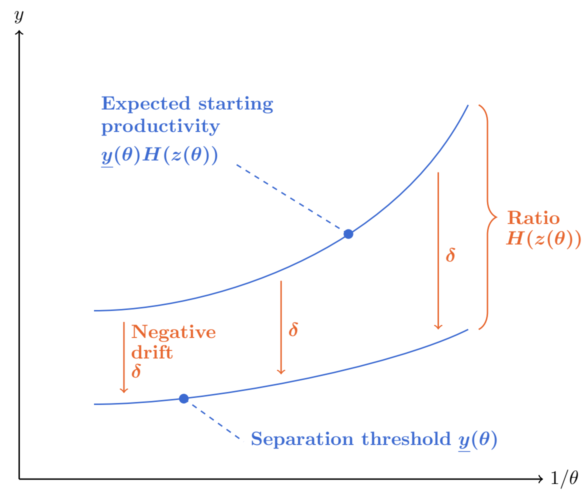
Proof. See Appendix B.5. □

With Proposition 3 at hand, it is possible to shed further light on how spatial unemployment differentials depend on the labor market pooling complementarity using a particular limiting equilibrium. Suppose that 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 ex-ante heterogeneity between locations.

Corollary 1. *(Equilibrium spatial gaps 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.

Proof. See Appendix B.6. □

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



This results highlights that the pooling complementarities suffice to sustain sorting in equilibrium, irrespectively of technological complementarities.[29] 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 Lemma 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 because free mobility (10) would equalize reservation wages across locations. With $\omega > 0$, housing prices adjust to sustain population and reservation wages differentials.

Proposition 2 and Corollary 1 conclude the description of the positive implications of the model. Before turning to the normative implications of the theory, the next section proposes reduced-form empirical evidence supporting the key mechanisms that determine job losing and job finding rates in the model.

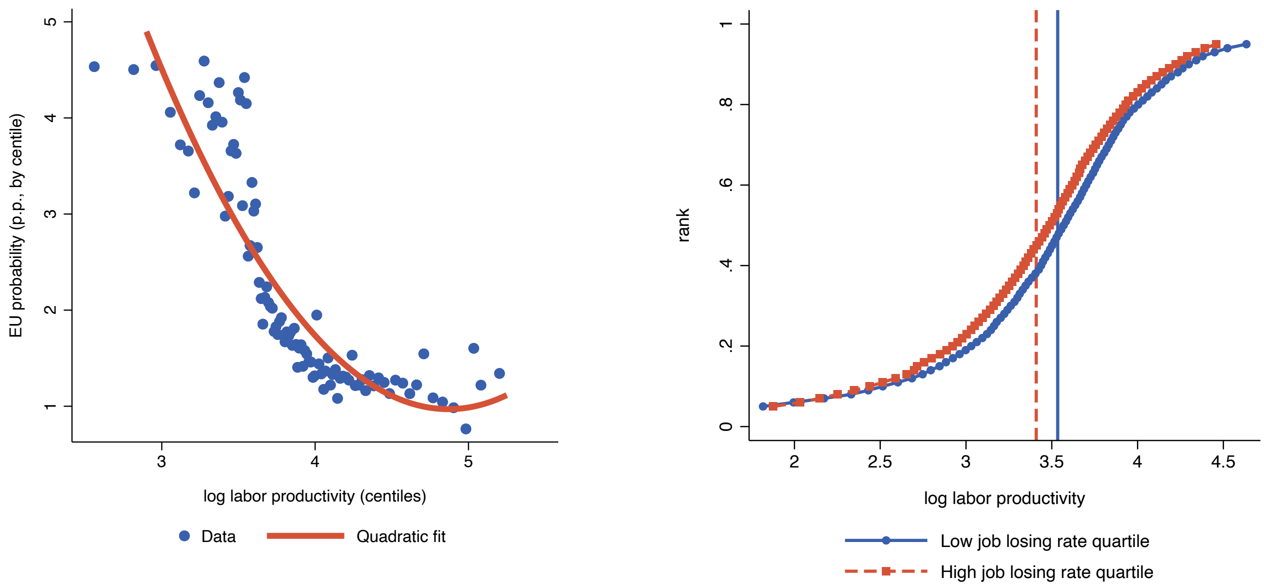
2.6 Model validation

The goal of this subsection is to provide empirical support for two crucial mechanisms. The first mechanism is the link between labor productivity and job losing rates. The second mechanism is the response of labor market tightness that determines job finding rates.

[29] 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.

Figure 5: Labor productivity and job loss in France.

(a) Job loss by employer productivity. (b) Cumulative distribution functions of labor productivity.



Note: Figure 5(a) scatterplots the employment-to-unemployment transition probabilities for workers, across centiles of their employer's labor productivity. Figure 5(b) plots empirical cumulative labor productivity distribution functions in the bottom and top quartiles of commuting zones, ranked by job losing rate. Vertical lines mark within-quartile averages of labor productivity.

Labor productivity and job losing rates. The first validation exercise emphasizes that the link between labor productivity and job losing rates in the model lines up with the data. Using the solution for the productivity distribution from Lemma 2, the model produces testable implications that tie labor productivity to job losing rates:

1. *Matches with higher labor productivity are less likely to separate into unemployment in all locations.*
2. *Average labor productivity is higher in locations with lower job losing rates.*
3. *The labor productivity distribution in low job losing rate locations first-order stochastically dominates the distribution in high job losing rate locations.*
4. *The labor productivity distribution has a Pareto tail with index $1/z(\ell)$ in each location.*
5. *The ratio of Pareto tails indices between locations is equal to the ratio of job losing rates between the same locations.*
6. *Labor productivity growth is independent from location.*
7. *Spatial labor productivity gaps at newly created jobs are larger than at older jobs.*

To test implications 1 to 7, I compute labor productivity in single-establishment firms using the firm-level balance sheet data described in Section 1.1. Figure 5(a) tests implication 1. It scatterplots job losing rates for workers across centiles of their employer's labor productivity. Consistent with persistence in the productivity process (4) that ties together the productivity of a match at a given point in time and the probability of job loss, matches at more productive employers are more stable.

Figure 5(b) tests implications 2 and 3. It 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 implication 2, average labor productivity is higher in locations with low job losing rates. Consistent with the more subtle implication 3, the cumulative distribution function of labor productivity in low job losing rate locations is always below the cumulative distribution function in high job losing rate locations: the labor productivity distribution first-order stochastically decreases with the job losing rate. Figure 13 in Appendix B.7 provides empirical support for implications 4 and 5 and therefore the Pareto assumption.

Implications 6 and 7 arise because newly created jobs are sorted across locations. Once a job is created, productivity shocks do not introduce ex-post heterogeneity that varies systematically across locations. Thus, productivity gaps should be largest for newly created jobs and vanish over time. To test implications 6 and 7, Table 7 in Appendix B.7 correlates labor productivity at entrant establishments (less than two years old), at incumbent establishments (at least two years old), and labor productivity growth with local job losing rates. All jobs are new at a firm that just entered. While it is unclear exactly what fraction of jobs are new at an incumbent firm, it is arguably less than at an entrant firm. Table 7 in Appendix B.7 shows that the data supports implications 6 and 7. Increasing the local job losing rate from the 25^{th} to 75^{th} percentile is associated with 6% lower labor productivity for incumbent establishments, but with 10% lower labor productivity for entrant establishments. In contrast, labor productivity growth rises by an economically and statistically insignificant 0.002.

Labor market tightness and job finding rates. The limiting economy of Corollary 1 is useful to interpret the role of the opposing forces entering the job finding rate as per Proposition 2, because it abstracts from the added complexity of productivity gaps ℓ across locations. 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 meeting rate and labor market tightness are positively correlated with the job losing rate in the limiting economy of Corollary 1. Therefore, but perhaps surprisingly, the labor market is tight (few vacancies per job seeker) and the worker meeting rate is small where the unemployment rate is low.

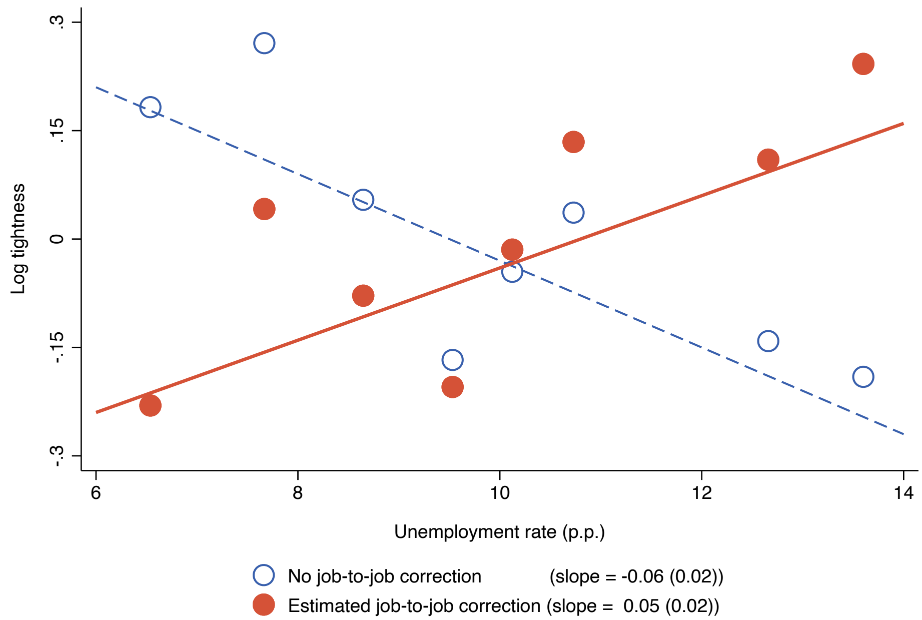
In the data, the job finding rate co-moves moderately negatively with the unemployment rate (Figure 2). Thus, the special case of Corollary 1 can rationalize the data only if two conditions are met. First, labor market tightness correlates negatively with unemployment. Second, the spatial variation in labor market tightness is moderate relative to the spatial variation in job losing rates.

Testing the implications of the limiting economy of Corollary 1 is a useful exercise, but an important caveat applies. Compensating differentials such as local productivity gaps ℓ should matter in practice for employers' location choices. In this case, the correlation between labor market tightness and unemployment rates can take any sign. The only robust implication is then that spatial variation in labor market tightness should be small relative to spatial variation in the job losing rate.

With this caveat in hand, I use establishment-level vacancy data from France to test whether labor market tightness correlates positively or negatively with local unemployment rates. To compute a measure of local labor market tightness in the data that is consistent with the model, it is important to obtain a measure of effective job seekers and correct for job-to-job search.[30] Figure 6 shows that, if anything,

[30] In the data, employed workers capture a sizable fraction of vacancies, while in the model only unemployed workers apply to vacancies. Therefore, I estimate the relative search intensity of employed workers ξ in Supplemental Material E.7. I then adjust the number of effective job seekers to compute labor market tightness in city c: $\hat{\theta}_c = \frac{\mathcal{V}_c}{\mathcal{U}_c + \xi \mathcal{E}_c}$, where $\mathcal{V}_c$ is the number of vacancies, $\mathcal{U}_c$ the number of unemployed workers, and $\mathcal{E}_c$ the number of employed workers.

Figure 6: Labor market tightness in France.



Note: Figure 6 scatterplots two measures of log labor market tightness across 8 groups of French commuting zones c. Blue circles represent the raw measure of labor market tightness without job-to-job adjustments: $\theta_c = \mathcal{V}_c / \mathcal{U}_c$, where $\mathcal{V}_c$ is the number of vacancies and $\mathcal{U}_c$ the number of unemployed workers. Orange dots represent the adjusted measure that include employed workers as effective job seekers: $\tilde{\theta}_c = \mathcal{V}_c \ (\mathcal{U}_c + \xi \mathcal{E}_c)$, where $\mathcal{V}_c$ is the number of vacancies, ξ the relative search efficiency of employed workers, and $\mathcal{E}_c$ the number of employed workers.

labor market tightness correlates positively with local unemployment rates. To reach this conclusion, Figure 6 displays a bin-scatterplot of labor market tightness across French commuting zones. Figure 6 reveals that the job-to-job correction is important to recover a positive correlation. When not adjusting the mass of job seekers for job-to-job search in the data, the correlation is negative. In both cases, the proportional variation in labor market tightness remains small in comparison to that in the job losing rate.

This section established the positive implications of the theory. It linked employer productivity, sorting, job losing and job finding rates, both in the model and in the data. The next section investigates the normative implications of the theory, before enriching the model to make it amenable to quantitative analysis.

3 Efficiency and quantitative model

This section builds on the model laid out in Section 2. I start by characterizing the efficiency properties of the economy and optimal policy. Next, I enrich the economy to make it suitable for estimation.

3.1 Efficiency and planning allocation

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 employers $z_1 = z(\ell_1) < z(\ell_2) = z_2$. Consider a marginal employer $z \in (z_1, z_2)$ contemplating entering in locations ℓ_1 or ℓ_2. If employer z enters in location ℓ_2, it is worse than the average local employer. Due to labor market frictions however, it meets as many workers as its more productive competitors. By opening in location ℓ_2, employer z exerts a negative externality on all other employers looking to hire there because it diverts workers away from them. This externality is also socially harmful, as workers are redirected towards a less productive employer, $z < z_2$. Symmetrically, the marginal employer z exterts a negative externality on other employers 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 employer, $z > z_1$. In both cases, the magnitude of the externality depends on the quality of local employers z_1 or z_2, and on the quality of the newcomer z.

On net, the marginal employer has an incentive to free-ride the favorable hiring conditions in location ℓ_2, because its vacancy meeting rate does not reflect that it is worse than average there. Wages are bargained ex-post and thus do not fully price meeting 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 heterogenenous 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. It is embedded in the labor market pooling complementarity highlighted in equation (12). Thus, I call the externality a labor market pooling externality. To make these arguments precise, I now define the planner's problem.

A utilitarian planner maximizes a possibly weighted sum of values of all individuals in the economy, taking 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 (56), Supplemental Material F.1.

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, they may not break up matches optimally. This distribution is an infinite-dimensional object. Nevertheless, a well-defined planner problem can be established building on Moll and Nuño (2018) as described in Supplemental Material F.1.1. Because it involves additional notation, I relegate the formal definition of the planning problem to Supplemental Material F.1.1 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 (56), 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* $\ell \in \{\underline{\ell}, \overline{\ell}\}$.
  - $\frac{\partial \log \underline{w}^{DE}}{\partial \ell}(\ell) > \frac{\partial \log \underline{w}^{\mathrm{SP}}}{\partial \ell}(\ell)$.
- *The planning allocation coincides with the allocation in a decentralized equilibrium in which search is directed.*

Proof. See Supplemental Material F.1.2. □

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

$$\left(\frac{\bar{J}^{DE}(z, \ell)}{\bar{J}^{SP}(z, \ell)}\right)^{\frac{1-\alpha}{\alpha}} = \frac{\bar{S}(z^{DE}(\ell))}{\bar{S}(z)} \cdot \left(Y/\underline{y}(\ell)\right)^{\frac{1}{z^{DE}(\ell)} - \frac{1}{z}} . \tag{18}$$

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 $z^{DE}(\ell)$ because of workers' spatial indifference condition. This difference between social and private values exactly encodes the labor market pooling externality, acting though the vacancy meeting 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 functions $z^{SP}(\ell)$, they internalize 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 labor market pooling externality arises because there is no price for labor market tightness in any location and matches are formed randomly. If search was fully directed, would the decentralized equilibrium be efficient? I outline an alternative setup with directed search in Supplemental Material F.1.2. 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 meeting 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 3.3. More productive employers post more vacancies than less productive ones. Thus, they meet with 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 2 in Section 5.1 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.

3.2 Optimal policy

Given that 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 Supplemental Material F.2 and show in Proposition 5 that they may 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.*[31]

Proof. See Supplemental Material F.2. □

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

[31] 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.

needs to be taxed more heavily in low productivity locations on the empirically relevant side of the Hosios (1990) condition $\beta < \alpha$. 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 calls for an optimal profit subsidy that resembles real-world place-based policies. The Empowerment Zone program in the United States and its French equivalent—the *"Zones Franches Urbaines"*—both grant large effective profit subsidies for firms opening jobs in distressed areas in the form of tax exemptions relative to a baseline tax rate. In the model, the profit subsidy corrects the labor market pooling externality that equation (18) 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.

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 3.3 below.

3.3 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 lets the model capture joint variation in population, wages, and unemployment across places. 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 migration elasticity of workers crucially affects the welfare gains from place-based policies as it governs how many move into locations that improve. 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. Those shocks stay constant until the next moving opportunity arrives.[32]

[32] 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. This normalization ensures that preferences shocks have mean 1. Supplemental Material G.4 extends standard discrete choice results to a continuum of locations.

Preferences. The flow utility function follows a standard specification and becomes $u(c, h, a, \overline{\varsigma}) = \left(\frac{c}{1-\omega}\right)^{1-\omega} \left(\frac{h}{\omega}\right)^{\omega} a\overline{\varsigma}$, where $\overline{\varsigma}$ denotes the product of all past taste shocks the worker received for locations they chose.[33]

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 affects 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 clusters of workers with low human capital. 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.[34]

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.[35] 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 $\left(ypk\right)^{\frac{1}{1+\psi}} h^{\frac{\psi}{1+\psi}}$.

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 $\frac{c_v}{1+1/\gamma} v^{1+1/\gamma}$.

[33]Formally, $\overline{\varsigma}_t = \prod_{i=1}^{N_t} \varsigma_{t_i}(\ell^*_{t_i})$, where $(t_i)_{i=1}^{N_t}$ denotes the times at which the worker received migration opportunities between calendar times 0 and t. Desmet et al. (2018) also use a multiplicative specification. Taking logs is isomorphic to standard additive discrete choice specifications, such as Caliendo et al. (2019).

[34]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.

[35]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 "traveling 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.

3.4 Characterization

The extensions preserve the basic structure of the location choice of employers. I show in Supplemental Material G that, to a first order 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 (13) becomes

$$\underset{(p,a)=\ell}{\operatorname{argmax}} \ \frac{z}{1-z}\Bigg\{ \underbrace{\log\left(p^{\mathcal{Q}}a^{-\psi\mathcal{P}}\right)}_{\substack{\text{Exogenous}\\ \text{production \& housing}\\ \text{complementarities}}} + \underbrace{\log\left(C\left(\underline{w}(\ell), z(\ell)\right)^{\psi\mathcal{P}}\right)}_{\substack{\text{Endogenous}\\ \text{housing}\\ \text{complementarity}}} + \underbrace{\log\left(\bar{k}(u(\ell))^{\mathcal{Q}}\right)}_{\substack{\text{Endogenous}\\ \text{human capital}\\ \text{complementarity}}} + \underbrace{\log q(\ell)}_{\substack{\text{Endogenous}\\ \text{pooling}\\ \text{complementarity}}} \Bigg\} - \underbrace{\log \underline{w}(\ell)}_{\substack{\text{Endogenous}\\ \text{cost of}\\ \text{labor}}}. \tag{19}$$

The constants $\mathcal{P}, \mathcal{Q}$ are combinations of parameters: $\mathcal{P} = \frac{1}{\omega+\psi+\varepsilon(1+\eta+\psi)}$ and $\mathcal{Q} = \frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}$. The average human capital $\bar{k}(u(\ell))$ in location ℓ is a decreasing function of the local unemployment rate $u(\ell)$, proportional to $\frac{\Delta}{\Delta+\varphi u(\ell)}$ up to a general equilibrium constant. Recall that $\ell = (p, a)$ now indexes productivity-amenity pairs. The function C is defined in equation (69) in Supplemental Material G.6, and is increasing in each argument.

Equation (19) highlights that several additional channels determine the location decision of employers in the extended model. First, exogenous technological complementarities still depend on productivity p. However, the housing price channel also introduces an exogenous complementarity with 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

$$\ell(p, a) \equiv p^{\mathcal{Q}}a^{-\psi\mathcal{P}}. \tag{20}$$

Second, the housing price channel also introduces an endogenous source of complementarity. It is captured by the function $C(\underline{w}(\ell), z(\ell))$ that encodes how local expenditures on housing depend on local wages. Higher wages lead workers to spend more on housing, driving up housing prices and thus employers' operation costs.

Third, equation (19) reveals the contribution of learning at the workplace for the location choice of employers. Average local human capital $\bar{k}(u(\ell))$ falls as the local unemployment rate rises and 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. Due to production complementarities with workers' human capital, high-productivity employers find it less profitable to enter a location when the local human capital mix is lower. Thus, the human capital channel microfounds part of the production complementarities that were fully exogenous in Section 2.

Fourth, equation (19) shows that labor market pooling complementarities remain unchanged and still depend only on the local vacancy meeting rate. Similarly, the expected cost of labor continues to be summarized by local reservation wages $\underline{w}(\ell)$. Because the structure of the location choice of employers in equation (19) closely resembles its more stylized version in equation (13), virtually all the analytical results from Section 2 carry through.

Proposition 6. *(Characterization of the extended model)*
To a first order when the migration rate μ and the scarring effects of unemployment φ are not too large, Propositions 1, 2, 3, 4, 5 and Corollary 1 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 the pair $\big(\ell(p,a),a\big)$: $L(p,a) \equiv L(\ell(p,a),a)$.

Proof. See Supplemental Material G. □

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 $\ell(p,a)$. I provide more details in Supplemental Material G.7. Having laid out the structure of the extended framework, I turn to the structural estimation.

4 Estimation

4.1 Identification

Despite its rich structure, the quantitative model is transparent enough to produce estimating equations for all but one of the parameters. In particular, no simulation is 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: $\rho, \Delta, \omega, \psi, \delta, \sigma, \beta, b, Y, \eta, \mu, \varepsilon, \alpha, \gamma, c_v, m, \lambda, \varphi, c_e$; together with two distributions $F_z, F_{p,a}$. While these distributions are recovered non-parametrically, I nonetheless estimate functional forms to simulate counterfactuals, adding another 7 parameters.

The 26 parameters can be divided into three groups. Parameters in the first group—$\rho, \Delta, \omega, \psi$, μ, 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 consists of distributional functional forms. The fourth group of parameters only contains the entry cost c_e, which is estimated by numerical search (Method of Simulated Moments). 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, I obtain 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 also delivers measures of aggregates such as the geographic mobility rate of workers and the average job offer acceptance probability. Finally, the DADS enable finer disaggregation of job losing rates and wages by tenure and location which is 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.

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.[36] Δ is measured by the labor force exit 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%).[37] Similarly, the expenditure share on real estate out of value added by employers ψ is equated to my estimate of 11%.[38] The remaining parameters in this first group can be normalized: $b = c_v = m = 1$.[39]

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 derived in Appendix C.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 $\hat{\delta} = \frac{\delta}{\sigma}$ given μ, Δ, and is specified in Appendix C.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

$$s_{1c} = s_1(s_c, \hat{\delta}) \tag{21}$$

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 enables to explicitly compute wage growth by tenure when β is not too large. Appendix C.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 regression similar to (21) estimates δ when β is small. When β is large, δ and β are estimated jointly. At the estimated bargaining power β however, the difference is negligible.

Bargaining power $\boldsymbol{\beta}$. Net wages relative to value added in location c is $\beta + \frac{1-\beta}{H(s_c)}$, where H only depends on δ and σ. I target aggregate net wages relative to value added to identify β by simple MDE.

[36] 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.

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

[38] 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.

[39] 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.

Learning rates λ and φ. Wage changes for workers coming out of unemployment reflect human capital losses, that grow with unemployment duration. Appendix C.5 shows that, for worker i who loses their job at time t_0 and finds a new job at time t_1 in location c, wages satisfy

$$\log W_{ict_1} = (\lambda - \varphi)(t_1 - t_0) + \Phi_c + \log W_{ict_0} + \upsilon_{ic}, \tag{22}$$

where υ_{ic} is a mean-zero random variable that reflects draws from the local new job distribution, and Φ_c 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 (22). λ is then directly obtained from aggregate real wage growth (see Appendix C.5). Thus, I recover φ.[40]

Local quality and cutoff. For the remainder of the estimation, I recover estimates of the local job quality z_c and the local productivity cutoff $\underline{y}_c$ in each city c. 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,

$$z_c = \frac{\delta}{s_c}, \qquad \underline{y}_c = \frac{b\underline{y}_0}{\hat{\rho}} \frac{\beta f_{Rc} \bar{S}(z_c)}{\hat{\rho} - \beta f_{Rc} \bar{S}(z_c)}. \tag{23}$$

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 Y. With an estimate of $\underline{y}_c$ at hand, I use data on job search behavior from the LFS to identify Y. In Supplemental Material H, I show how to use this information to recover the average success probability of a meeting. I estimate it to be 20.6%.[41] From the model, the success probability of a meeting 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 η. At this stage, it is possible to construct demand for housing in each city in the model. Appendix C.7 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.[42]

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

$$\log \pi_c = \pi_0 + \frac{1}{\varepsilon} \log \bar{U}_c + \log a \tag{24}$$

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 0 and 1 and first-difference equation (24). Then, I use local productivity shocks based on shift-share

[40] In practice, mechanisms left out from the model may generate endogeneity issues. To address those concerns, Table 8 in Appendix C.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.

[41] Faberman et al. (2017) indicate an acceptance probability of 29.6% in the United States.

[42] 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.

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 (24) 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 C.8.

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 (52) in Appendix C.6 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 (65) in Supplemental Material G.5 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 $F_{p,a}$.[43] Finally, Appendix C.9 shows that the density function of job losing rates across locations identifies f_z using (23).

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 (53) in Appendix C.10, 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 C, 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)*
To a first order when $\mu, \Delta, \delta, \beta$ 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 (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 $[\underline{z}, \overline{z}]$.

Fourth group (1 parameter). After estimating those 25 parameters, a numerical search estimates the entry cost c_e by targeting the aggregate unemployment rate.

4.2 Parameter estimates

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

[43] Altenatively, amenities could be obtained as residuals from the migration share equation (24). Because the estimation relies on observed population shares, I choose to match population rather than migration shares. In practice, they are highly correlated.

Table 1: Parameter estimates

Parameter	Interpretation	Target	Estimator	Estimate
ρ	Discount rate	Annual interest rate	MDE	0.008
Δ	Labor force exit rate	Aggregate unemployment rate	MDE	0.004
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	0.08
Y	Lower bound of init. prod.	Job acceptance probability	MDE	0.93
η	Housing elasticity	Housing prices	OLS	3.48
μ	Migration rate	Migration rate	MDE	0.002
$1/\varepsilon$	Migration elasticity	Migration shares	2SLS	4.72
α	Matching function elasticity	Local job finding rates	2SLS	0.30
γ	Vacancy cost elasticity	Local job finding rates	2SLS	1.44
φ	Learning rate	Unemployment scar	OLS	0.01
$\underline{z}$	Lowest job quality	Local job losing rates	MDE	0.03
$\overline{z}$	Highest job quality	Local job losing rates	MDE	0.10
g_1	Shape of job quality distrib.	Local job losing rates	MDE	1.36
g_2	Shape of job quality distrib.	Local job losing rates	MDE	2.19
σ_ℓ	St.d. of local productivity	Local wages	MDE	0.14
σ_a	St.d. of local amenities	Local population	MDE	0.21
$c_{\ell,a}$	Correlation prod.–amenities	Local wages and population	MDE	0.38

power $\beta = 0.08$ 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 0.28, 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 4.72, within but towards the high end of the values reported in the literature between 0.5 and 5.[44] The matching function elasticity α is 0.3, within 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 (2021).[45]

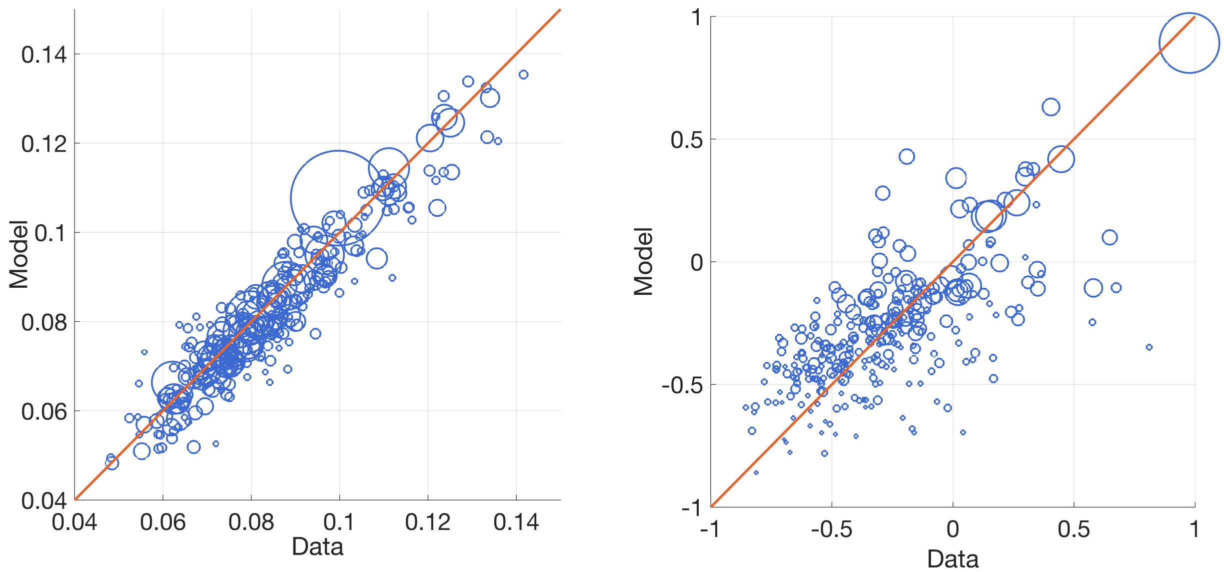
[44]The estimate for μ implies an annual migration rate of about 1%. This is lower than the overall migration rate in my sample which is about 3%. This discrepancy is due to the fact that many migrants are employed workers moving with a job at hand. However, in steady-state, the migration elasticity is the key driver of population movements, not the migration rate.

[45]Jarosch (2021) 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 (2021)'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.

Figure 7: First year job losing rates and housing prices across cities.

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

(b) Housing prices: model vs. data.



Note: Figure 7(a): annual job losing rate in first year of job, model against data. Figure 7(b): housing prices in model against data. Cities in model identified by their job losing rate. Blue circles proportional to city size.

4.3 Over-identification exercises

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

Job loss in first year. I start by discussing two exercises that lend credibility to the estimates of the productivity process δ and σ that determines job loss within and across locations. First, Figure 7(a) shows that despite relying on a single degree of freedom δ/σ to predict job losing rates in the first year as per (21), the model closely fits the full cross-sectional variation between cities. Second, I use balance sheet data to compute firm-level labor productivity growth relative to aggregate labor productivity growth, It should be close to the annualized estimate of δ, 0.4%. I obtain a relative decline of 0.5% annually.[46]

Housing prices. Dispersion in housing prices sustains spatial arbitrage. To assess how well the estimated housing supply elasticity accounts for cross-sectional dispersion in housing prices, Figure 7(b) 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.

Amenities. Unobserved local amenities allow the model to match the dispersion in city-level population. 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 establishments of various

[46]I focus on large and high labor productivity firms to minimize survival selection bias. 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 such as Bilal et al. (2019).

kinds. Table 9 in Appendix C.11 shows that more sun hours and a higher density of health or commercial services are all positively associated with higher amenities.[47] 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.

In addition to the validation exercises in Section 2.6, this section established the ability of the estimated model to speak to a number of targeted and non-targeted moments. I now turn to the main structural results of this paper: the breakdown of spatial unemployment gaps into job loss and job finding, and the welfare effects of place-based policies.

5 Spatial unemployment gaps and place-based policies

This section first shows that the model can quantitatively account for spatial unemployment differentials. Next, this section discusses the employment and welfare effects of place-based policies.

5.1 Spatial job loss differentials and unemployment

With the estimated model at hand, Figure 8 plots the model equivalent of Figure 2—the graphical decomposition of spatial unemployment gaps into job losing and job finding rates in the model. Figure 8 also reports the data for comparison purposes. Strikingly, the estimated model replicates closely the key role of job loss in the data. As employers sort across locations, job losing rate differentials account for the majority of the variation in local unemployment rates. Because of the opposing forces highlighted in Proposition 2, the job finding rate is close to flat across locations. As a result, the unemployment rate largely follows the spatial patterns of the job losing rate.

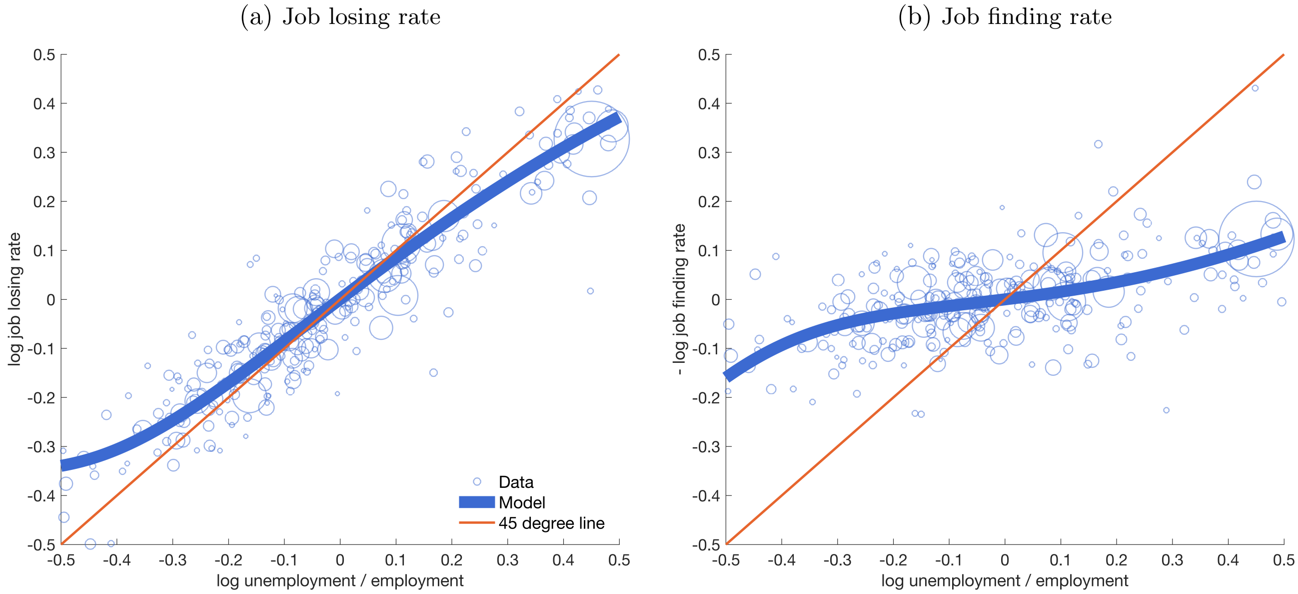
Crucially, neither the location decision of employers, nor the spatial variation in job finding rates, are constrained by the estimation. I return to this observation in more detail with Table 2 below. The spatial variation in job losing rates results from (i) the distribution f_z, that is constrained by the estimation, but also (ii) the equilibrium assignment of heterogeneous employers to locations $z(\ell)$, that is left completely free by the estimation. In addition, the estimation does not limit the spatial variation in job finding rates apart from the two coefficients in equation (53) that identify α and γ.

The job losing and job finding variance shares are therefore useful moments to assess the model's ability to speak to spatial unemployment differentials. Figure 8 establishes that the estimated model provides an empirically plausible account of the role of job loss across locations.

To assess whether the model accounts quantitatively for spatial unemployment gaps, Table 2 reports the dispersion in local unemployment rates, as well as its breakdown into job losing and job finding rates. Comparing the first and second columns of Table 2 reveals that the estimated model accounts for over 75% of the cross-sectional standard deviation of local unemployment rates. The standard deviation is 0.019 in the model against 0.025 in the data. Table 2 also highlights that the decentralized equilibrium closely replicates the contribution of job losing rates to spatial unemployment differentials. The job losing share is 73% in the estimated model against 77% in the data. Table 2 finally reports the (targeted) aggregate unemployment rate in the model and the data.

[47]For instance, doubling the number of sun hours per month raises the amenity value of a location by 12%. Doubling the density of health establishments increases amenities by 6.7%.

Figure 8: Local job losing and finding rates against unemployment-to-employment ratios in model.



Note: Figure 8(a) plots the log of the job losing rate against the log of the unemployment-to-employment ratio, across cities in the DADS data and in the estimated model. Figure 8(b) plots minus the log of the job finding rate against the log of the unemployment-to-employment ratio, across cities in the DADS data and in the estimated model.

I evaluate the role of labor market pooling complementarites in the third column of Table 2. It reports the same decomposition as the second column, but after shutting down the labor market pooling externality. The labor market pooling externality does not affect much the aggregate unemployment rate, but it is key in generating spatial differences in unemployment rates. The cross-sectional standard deviation of local unemployment rates drops to 0.003 without labor market pooling externalities, 15% of its baseline value and 12% of its value in the data.

The labor market pooling externality matters quantitatively because it magnifies the strength of labor market complementarities and thus the spatial sorting of employers. The pooling externality induces excess clustering of the most productive employers with stable jobs in the top locations. As a result, only employers that are not productive enough, with jobs that are too unstable, remain in high unemployment local labor markets. Consistent with this interpretation, the job losing rate share drops to 45%. On net, the pooling externality yields over a five-fold amplification of spatial unemployment gaps.

Finally, I assess whether the quantitative importance of the pooling externality is mitigated when fully re-estimating an efficient model without the externality. In the fourth column of Table 2, I re-estimate the entire model after imposing directed search instead of random search. The allocation from the directed search model coincides with the planner's allocation under the same parameter values as per Proposition 4. In practice, parameter estimates under the re-estimated efficient allocation are close to parameter estimates in the decentralized equilibrium.

The fourth column of Table 2 indicates that even when re-estimated, the efficient allocation falls short of replicating the empirical dispersion in local unemployment rates as well as the role of job loss in driving that variation. In the re-estimated efficient allocation, the dispersion in local unemployment rates again drops to 12% of its value in the data, and the job losing rate generates barely more than half of this variation.

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

	Data	Baseline	No pooling	Est. planner
Aggregate unemployment rate	0.097	0.097	0.093	0.097
St. dev. unemployment rate	0.025	0.019	0.003	0.003
Var. log unemp. / emp.	0.072	0.048	0.002	0.001
Job losing rate	77 %	73 %	45 %	51 %
Job finding rate	23 %	27 %	55 %	49 %

Note: All statistics are population-weighted. 'Data' column reports moments in the data. 'Baseline' column reports moments in the estimated model. 'No pooling' reports moments in the estimated model after shutting down the labor market pooling externality. 'Est. planner' reports moments for a fully re-estimated model under directed search. Its allocation coincides with the social planner's allocation.

This comparison between the baseline and the re-estimated efficient allocations confirms that the estimation does not place strong constraints on the dispersion in unemployment, job losing and job finding rates. These moments therefore provide informative over-identification restrictions to evaluate the model's ability to account for spatial unemployment gaps and discriminate between different sets of assumptions. Table 2 strongly favors the baseline version of the model that includes the labor market pooling externality, against the efficient economy that does not feature this amplification force.

Labor market flows, wages and population. In addition to matching the dispersion in labor market flows, can the model account for the co-movement between labor market flows and standard economic outcomes of interest, such as wages or population? Table 3 displays the results from OLS regressions of job losing and finding rates onto local wages and population, both in the baseline estimated model and in the data. The regressions in the data are run at the worker level to control for individual heterogeneity. Consistently, I net out human capital from wages in the model.

Wages correlate negatively with job loss both in the data and in the model, as revealed by columns (2) and (3). The coefficient is -0.12 in the data, against -0.19 in the model. This negative correlation reflects the spatial sorting of employers. Not only do productive employers offer stable jobs, they also pay high wages. Thus, cities with high wages are also those with low rates of job loss. Consistent with the moderate spatial variation in the job finding rate depicted in Figure 8(b), columns (5) and (6) in Table 3 indicate that the job finding rate is not strongly correlated with wages or population, both in the model and in the data.

Job losing and job finding rates correlate more weakly with local population conditional on local wages. Table 3 reveals that this pattern holds in the model and in the data. This weaker correlation reflects the difference in location decisions between workers and employers highlighted in Section 3.4. Employers only value local profitability encoded in the local index $\ell(p, a)$ because they do not care directly about residential amenities a. By contrast, workers enjoy residential amenities a and hence value those above and beyond $\ell(p, a)$. Thus, labor market flows and local wages are a one-to-one function of $\ell(p, a)$, but local population depends on amenities a conditional on the local index $\ell(p, a)$. Population thus exhibits additional, orthogonal dispersion relative to wages. This additional dispersion drives the coefficient on population towards zero.

Having characterized the model's ability to quantitatively account for the margins of spatial unem-

Table 3: OLS regressions of worker-level job loss and job finding probabilities.

	Job loss			Job finding		
	Data		Model	Data		Model
	(1)	(2)	(3)	(4)	(5)	(6)
Log city wage	-0.22***	-0.12***	-0.19	-0.00	0.01	0.06
	(0.02)	(0.03)		(0.01)	(0.02)	
Log city pop.	0.13***	0.04	-0.06	-0.03**	0.02	0.00
	(0.03)	(0.07)		(0.01)	(0.06)	
Fixed Effects						
Year	✓			✓		
Industry-Year		✓			✓	
Worker		✓			✓	
Obs.	2827237	2825413		405184	394678	
R^2	0.002	0.124		0.002	0.228	

Dependent variables relative to unconditional mean. Independent variables standardized to unit standard deviation. Data: 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 density by km^2. Quarterly frequency. 1997-2007. Movers only. Model: population moments.

ployment differences, I now turn to the welfare effects of place-based policies.

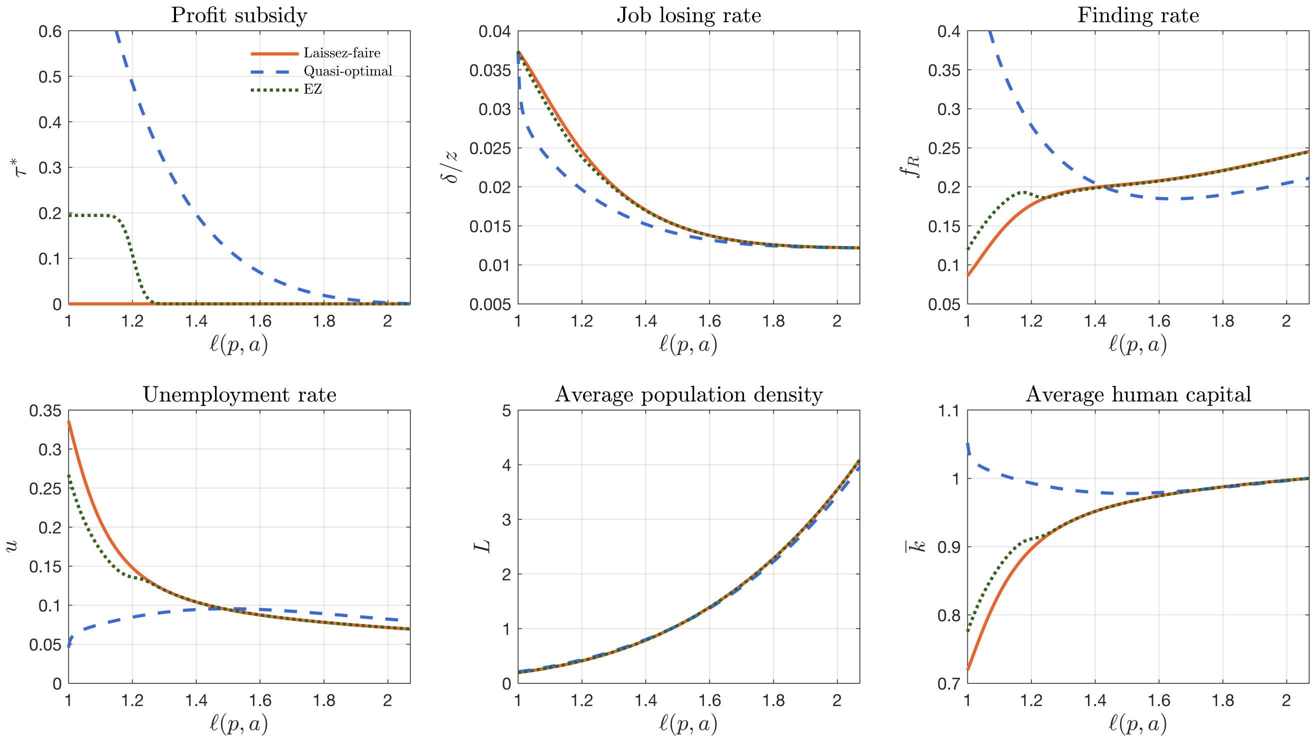
5.2 Place-based policy counterfactuals

With the estimated model at hand, this final section presents two policy counterfactuals. Both counterfactuals investigate the local and general equilibrium welfare gains from place-based policies. The first counterfactual studies the quasi-optimal policy. The second counterfactual contrasts the quasi-optimal policy with the real-world French Enterprise Zone (EZ) program in France.

As per Proposition 6, the quasi-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 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

Figure 9: Model's solution in the decentralized equilibrium, the quasi-optimal policy and the French EZ program.



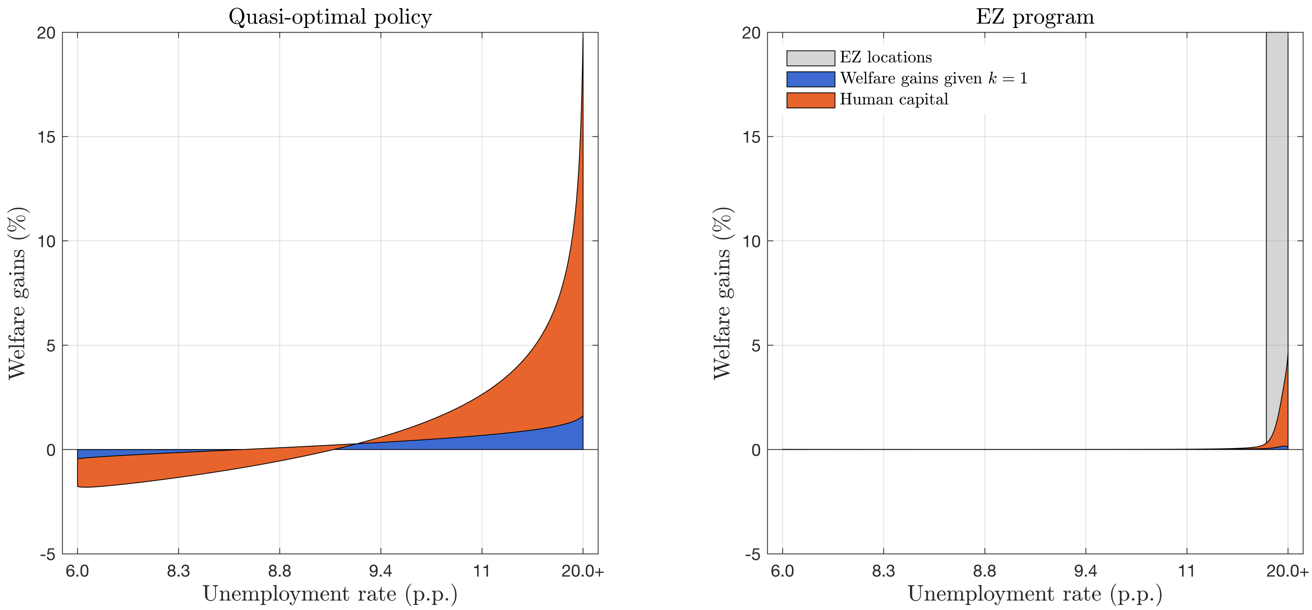
economy.[48]

Figure 9 displays the cross-sectional patterns of the equilibrium under the laissez-faire, the quasi-optimal policy and a budget-equivalent version of the French EZ program. The EZ subsidy is much smaller than the quasi-optimal one in scale and scope. However, it shares the same qualitative pattern: to incentivize high productivity employers to open jobs in high unemployment locations. Consistent with Proposition 6, the job losing rate declines as the local advantage index increases due to rising job quality $z(\ell)$ under all policy scenarios. Both policy interventions effectively relocate marginally more productive employers towards high-unemployment locations. Under the EZ program only few marginal jobs change location. The EZ program thus has minor effects on the job losing rate. By contrast, the quasi-optimal policy massively relocates productive jobs towards initially high unemployment locations, resulting in a large drop in the job losing rate.

Because of the opposing forces highlighted in Proposition 2, the job finding rate is more sensitive to policy. The EZ program succeeds in increasing the job finding rate because of a volume effect. Despite only a moderate decline in the job losing rate, the EZ program attracts more jobs to high unemployment locations. The quasi-optimal policy also strongly increases the job finding rate. Together, the reduction in job losing rates and increase in job finding rates result in large drops in local unemployment rates in initially high unemployment locations under the quasi-optimal policy, that can exceed 10 percentage points. Consequently, spatial unemployment differentials plummet.

[48]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 parameter estimates and counterfactuals.

Figure 10: Welfare gains from the quasi-optimal policy and the French EZ program.



Note: Expected welfare gains of residents who never receive the moving opportunity. Split into the average gain to a worker conditional on $k = 1$, and human capital gain defined as changes in the $W^w(\ell)$ and $\bar{k}(\ell)$ components as per equation (72) in Supplemental Material G.11. Locations ranked by their unemployment rate in the laissez-faire. X-axis reflects population-weighted quantiles. Left panel: quasi-optimal policy. Right panel: French EZ program.

Mirroring the falling unemployment rate as per (19), average human capital $\bar{k}(u(\ell))$ steeply rises across locations in the laissez-faire. The somewhat conservative estimate of φ still implies human capital gaps of over 25% between residents of the best and worst locations in the laissez-faire, due to the interaction of spatial unemployment differentials and scarring effects of job loss. Spatial gaps in human capital ameliorate under the EZ program due to the reduction in local unemployment rates in treated locations. Workers' exposure to scarring effects from unemployment falls, and they accumulate more human capital over the course of their working life. In line with the drastic reduction in unemployment under the quasi-optimal policy, human capital accumulation improves dramatically in initially high unemployment locations.

Figure 10 depicts the welfare gains for residents across 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 10 reveals that the quasi-optimal policy achieves large welfare gains in initially high unemployment locations.

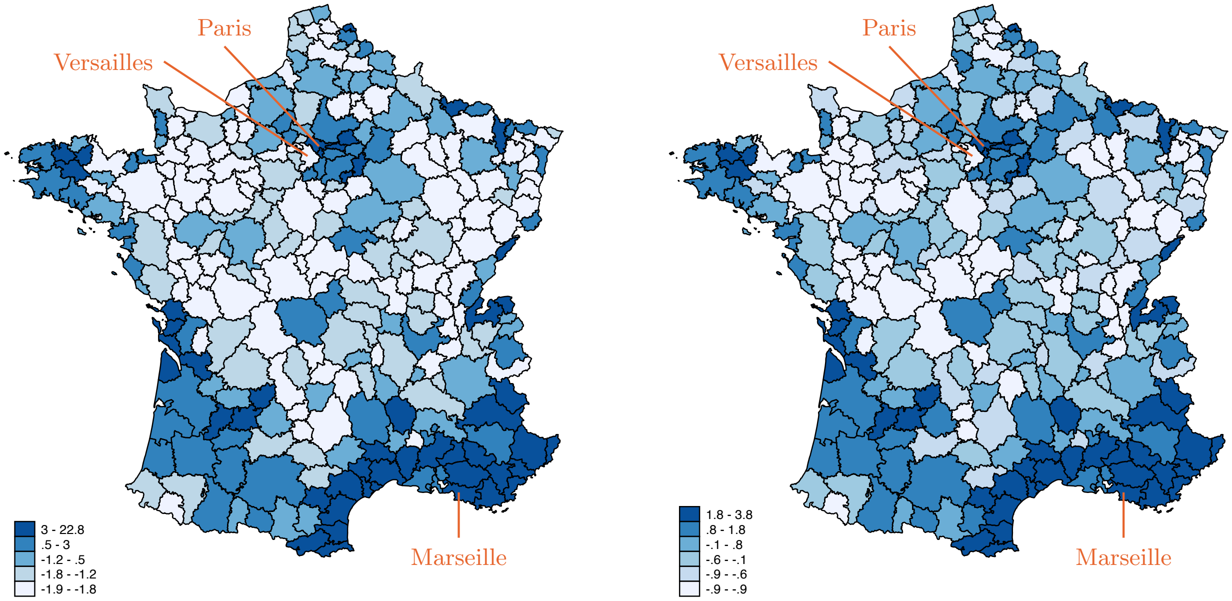
I derive an exact welfare decomposition in equation (72) in Supplemental Material G.11. This decomposition corresponds to the different colored areas in Figure 10. The blue area represents direct gains to the average resident worker, conditional on unit human capital. It is equal to the steady-state welfare gains of an unemployed worker who never received the moving opportunity and so stayed in the same location, with $k = 1$. Figure 10 shows that direct gains steadily rise with pre-policy local unemployment under both policies.[49]

The orange area displays human capital accumulation benefits. Figure 10 reveals that human capital and long-run scarring effects of unemployment are central to welfare gains from place-based policies.

[49]The blue area also corresponds to each location's contribution to the aggregate welfare gains for workers with $k = 1$.

Figure 11: Local gains from the quasi-optimal policy

(a) Welfare gains (%) (b) TFP gains (%)



Human capital benefits account for over three fourths of total welfare gains in most 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 5% and concentrated in treated, high unemployment areas.

To highlight the spatial distribution of these local welfare gains, Figure 11(a) maps the welfare 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 11(b) shows that local welfare gains are accompanied by substantial TFP improvements, as more productive employers relocate towards areas treated by the policy.

Leveraging the structure of the model, I aggregate local welfare gains and compute aggregate welfare gains from the quasi-optimal policy and the EZ program. Table 4 first highlights that the quasi-optimal policy reduces spatial unemployment differentials five-fold as in Table 2, by removing the pooling externality. This change follows a large relocation of high productivity jobs towards poorer locations. As aggregate efficiency rises, aggregate unemployment falls by half a percentage point. The quasi-optimal policy achieves just under 1% aggregate welfare gains. Second, Table 4 reveals that despite its relatively small size, the EZ program reduced spatial unemployment differentials by over 10%. While it had virtually no impact on the aggregate unemployment rate, it raised aggregate welfare by 0.1%. Most of these gains stem from better human capital accumulation in high unemployment areas.

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 9. Aggregate expenditures on the EZ policy represent redistributing 0.04% of Gross Domestic Product (GDP). Expenditures under the quasi-optimal policy are over 50 times larger. If scaling up the redistribution-efficiency ratio of the EZ policy was possible, welfare would rise by about 5% for every percent of GDP redistributed. The quasi-optimal

Table 4: Aggregate gains from place-based-policies

	Laissez-faire	Quasi-optimal	EZ program
Aggregate unemployment rate	0.097	0.093	0.096
St. dev. unemployment rate	0.019	0.003	0.016
Aggregate welfare gains (%)		0.831	0.090
Unemployed		0.150	0.005
Employed		0.039	0.002
Human capital		0.641	0.083
Resdistribution (% of GDP)		2.101	0.040

Note: Unemployed, employed and human capital gains defined as change in the W^u, W^e, W^k components in equation (74) in Supplemental Material G.11, respectively.

policy is ten times less efficient, 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.

This comparison suggests that smaller place-based policies may be more efficient than larger ones in the presence of additional frictions. Albeit a full investigation is beyond the scope of this paper, many economic forces left out of the model may generate dead-weight losses that scale with the policy intervention. Profit-shifting across establishments within firms, fiscal externalities or political economy constraints are a few examples. In this case, interventions of the scale of the EZ program are likely to be more robust than the quasi-optimal policy.

Conclusion

This paper has proposed 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 a 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 externalities 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 the 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 in 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". *Brookings Papers on Economic Activity* Spring, pp. 151–255.

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

Bilal, Adrien and Lhuillier, Hugo (2021). "Outsourcing, Inequality and Aggregate Output". *Becker Friedman Institute for Research In Economics, Working Paper 2021-05.*

Bilal, Adrien and Rossi-Hansberg, Esteban (2021). "Location as an Asset". *Forthcoming, Econometrica.*

Bilal, Adrien G., Engbom, Niklas, Mongey, Simon, and Violante, Giovanni L. (2019). "Firm and Worker Dynamics in a Frictional Labor Market". *National Bureau of Economic Research, Working Paper Series 26547.*

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

Brancaccio, Giulia, Kalouptsidi, Myrto, Papageorgiou, Theodore, and Rosaia, Nicola (2020). "Search Frictions and Efficiency in Decentralized Transportation Markets". *National Bureau of Economic Research, Working Paper Series 27300.*

Burdett, Kenneth and Mortensen, Dale T. (1998). "Wage Differentials, Employer Size, and Unemployment". *International Economic Review* 39.2, pp. 257–273.

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, Dvorkin, Maximiliano, and Parro, Fernando (2019). "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock". *Econometrica* 87.3, pp. 741–835.

Caliendo, Lorenzo, Opromolla, Luca David, Parro, Fernando, and Sforza, Alessandro (2021). "Goods and Factor Market Integration: A Quantitative Assessment of the EU Enlargement". *Forthcoming, Journal of Political Economy*.

Chade, Hector and Eeckhout, Jan (2019). "Competing Teams". *The Review of Economic Studies* 87.3, pp. 1134–1173.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, Puga, Diego, and Roux, Sébastien (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". *National Bureau of Economic Research, Working Paper Series 26142*.

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. (2020). "The comparative advantage of cities". *Journal of International Economics* 123, p. 103291.

Desmet, Klaus, Nagy, Dávid Krisztián, and Rossi-Hansberg, Esteban (2018). "The Geography of Development". *Journal of Political Economy* 126.3, pp. 903–983.

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

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". *Society for Economic Dynamics, 2018 Meeting Papers 1009*.

Engbom, Niklas and Moser, Christian (2021). "Earnings Inequality and the Minimum Wage: Evidence from Brazil". *National Bureau of Economic Research, Working Paper Series 28831*.

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

Fajgelbaum, Pablo D. and Gaubert, Cecile (2020). "Optimal Spatial Policies, Geography, and Sorting". *The Quarterly Journal of Economics* 135.2, pp. 959–1036.

Fajgelbaum, Pablo D., Morales, Eduardo, Suárez Serrato, Juan Carlos, and Zidar, Owen (Sept. 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". *Society for Economic Dynamics, 2017 Meeting Papers 190.*

Glaeser, Edward L. and Gottlieb, Joshua D. (2008). "The Economics of Place-Making Policies". *Brookings Papers on Economic Activity* Spring, pp. 155–254.

Glaeser, Edward L., Kim, Hyunjin, and Luca, Michael (2018). "Nowcasting Gentrification: Using Yelp Data to Quantify Neighborhood Change". *American Economic Association Papers and Proceedings* 108, pp. 77–82.

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. (1972). "Turnover in the Labor Force". *Brookings Papers on Economic Activity* 3.3, pp. 709–764.

— (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 (2021). "Searching for Job Security and the Consequences of Job Loss". *National Bureau of Economic Research, Working Paper Series 28481.*

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.

Krusell, Per, Mukoyama, Toshihiko, Rogerson, Richard, and Şahin, Ayşegül (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, Benoît (2021). “Efficiency in search and matching models: A generalized Hosios condition”. *Journal of Economic Theory* 193, p. 105208.

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.

Mayer, Thierry, Mayneris, Florian, and Py, Loriane (Oct. 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”. *National Bureau of Economic Research, Working Paper Series 20049*.

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

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.” *National Bureau of Economic Research, Working Paper Series 20975*.

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.

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 (Aug. 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 (2020). “Evaluating State and Local Business Incentives”. *Journal of Economic Perspectives* 34.2, pp. 90–118.

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.

Topel, Robert H. (1984). “Equilibrium Earnings, Turnover, and Unemployment: New Evidence”. *Journal of Labor Economics* 2.4, pp. 500–522.

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

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

Appendix

A Descriptive evidence

A.1 Transition rates

A.1.1 Time aggregation and three state model

Time aggregation correction. 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

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

This system admits the solution $u(t) = u_\infty + (u_0 - u_\infty)e^{-(s+f)t}$and $e(t) = e_\infty + (e_0 - e_\infty)e^{-(s+f)t}$, 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

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

Hence, the instantaneous quarterly transition rates can be recovered from time-aggregated transition probabilities from $s = \mathcal{T} \times \mathbb{P}_1[E \to U]$ and $f = \mathcal{T} \times \mathbb{P}_1[U \to E]$, where one quarter is the interval $[t, t+1)$, and the time aggregation correction factor is

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

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

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

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, $du - rn = se - fu$ and $du - rn = f_n n - s_n e$. 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,

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

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

$$\log \frac{u_c}{1 - u_c} = \log s_c - \log f_c + \log p_c + e_c, \tag{25}$$

Table 5: Variance decomposition of local unemployment-to-employment ratio.

	France		U.S.
	DADS	LFS	CPS
Direct flows: job losing and finding rates (%)	92	105	96
Job losing rate			
% direct flows	78	59	73
% total	72	62	70
% total, time-aggregated	71	55	62
Job finding rate			
% direct flows	22	41	27
% total	20	43	26
% total, time-aggregated	21	50	34
Non-participation (%)	10	-4	-4
Residual (%)	-2	-1	8

Note: Variance decomposition of log unemployment-to-employment ratio following equation (26). Direct flows represent the contributions of job losing and job finding rates.

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

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\log\frac{u_c}{1-u_c}\right] &= \overbrace{\mathbb{C}\mathrm{ov}\left[\log\frac{u_c}{1-u_c},\log s_c\right]+\mathbb{C}\mathrm{ov}\left[\log\frac{u_c}{1-u_c},-\log f_c\right]}^{\text{Direct flows}}\\ &+ \underbrace{\mathbb{C}\mathrm{ov}\left[\log\frac{u_c}{1-u_c},\log p_c\right]}_{\text{Non-participation}}+\underbrace{\mathbb{C}\mathrm{ov}\left[\log\frac{u_c}{1-u_c},e_c\right]}_{\text{Residual}}. \qquad (26)\end{aligned}$$

Time aggregation and three state model in the data. Table 5 reports the variance decomposition of local unemployment on job losing rates, job finding rates, non-participation flows and a structural residual for non-time-aggregated and time-aggregated flows.

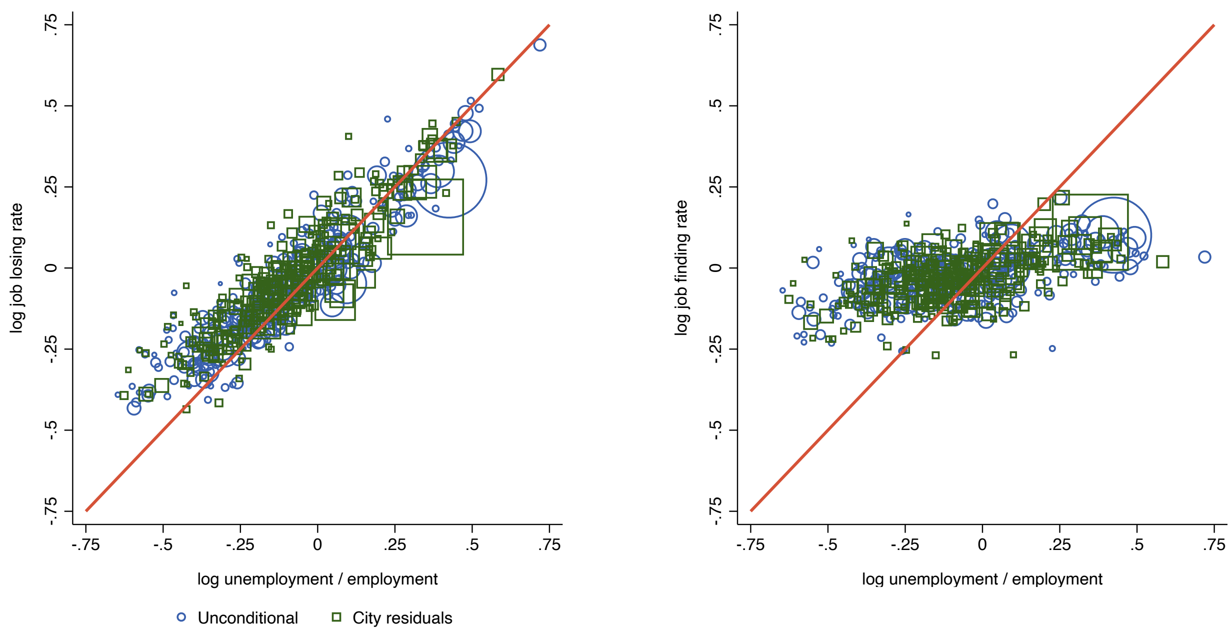
A.2 Composition

In principle, differences in the local industry mix and worker skill mix may account for some or all of the differences in the average worker and employer effects highlighted in Figure 3. To assess the role of industry and skill heterogeneity across cities, I estimate three-way fixed effect econometric models of the following form:

$$Y_{i,t} = \alpha_{C(i,t)} + \beta_{I(i,t)} + \gamma_{S(i)} + e_{i,t}, \qquad (27)$$

Figure 12: Local job losing and finding rates against unemployment-to-employment ratios in France. City fixed effects net of local industry and worker composition.

(a) Job losing rate (b) Job finding rate



Note: Scatterplots the log of the unconditional (blue circles) and residual (green squares) job losing rate and residual job finding rate against the log of the residual unemployment-to-employment ratio, across commuting zones in France (DADS panel). Residual defined as the estimated city fixed effect from the three-way fixed effect model in (27).

where C denotes a city, I denotes a 3-digit industry, S denotes a skill group, i denotes a worker identifier, and t is a quarter. α_C is a city effect, β_I an industry effect, and γ_S a skill effect. $e_{i,t}$ is a conditionally mean zero residual. $Y_{i,t}$ is an outcome of interest, that is either a job loss indicator, a job finding indicator, or an unemployment indicator. For unemployed workers, I define industry as their last industry of employment. I estimate linear probability models with 232 industry fixed effects and 300 skill fixed effects. Then, I replicate Figure 2 with the estimated city fixed effects $\hat{\alpha}_c$.

Figure 12 reveals that industry and worker composition do not contribute significantly to spatial unemployment differentials. In addition, Figure 12 shows that, even after controlling for local composition, the job losing rate remains the dominant source of spatial unemployment gaps.

Table 6 reports variance decompositions of the unconditional job losing rate across French commuting zones into employer, worker and city components with different definitions of employers.

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.

Table 6: Variance decomposition of city job losing rates on worker and employer contributions.

	Establishments			Firms		
	Raw	×Occ2	×Occ4	Raw	×Occ2	×Occ4
Variance shares (%): 10 groups						
Employer fixed effects	59	57	66	44	46	52
Worker fixed effects	41	43	34	35	36	26
City fixed effects				21	18	22
Variance shares (%): 50 groups						
Employer fixed effects	62	59	68	47	50	52
Worker fixed effects	38	41	32	33	34	24
City fixed effects				20	16	24
Variance shares (%): 300 groups						
Employer fixed effects	62	59	68	47	49	52
Worker fixed effects	38	41	32	32	34	24
City fixed effects				21	17	25

Note: Variance decomposition of average commuting zone job losing rate into worker and employer components as per equation (3). I vary the definition of an employer: an establishment (SIRET identifier, column 1), an establishment by 2-digit occupation (column 2), an establishment by 4-digit occupation (column 3), a firm (SIREN identifier, column 4), a firm by 2-digit occupation (column 5), a firm by 4-digit occupation (column 6). When using firms, equation (3) is enriched with a city fixed effect. I cluster workers and employers into groups based on their unconditional mean job losing rate. I use 10, 50 and 300 groups.

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

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

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 $V - U$ solves

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

where I denote $v(\ell)\ell = f(\ell)\mathbb{E}[V(w^*(y_0,\ell)) - U]$ 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 Adjusted surplus, wages and proof of Lemma 1

B.2.1 Adjusted surplus

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

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

which is indepedent from wages, and solves the recursion

$$\rho S(y,\ell,a) = \ell \cdot \Big(y - b - v(\ell)\Big) - L_y S \tag{28}$$

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

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

The following Lemma lets me make progress.

B.2.2 Wage determination

Lemma 3. *(Bargaining solution)*
Suppose that a worker and an employer set wages either under generalized Nash bargaining, or 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*

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

Proof. See Supplemental Material E.1. $\square$

B.2.3 Solving for the adjusted surplus

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

Lemma 4. *(Bargaining solution)*
Under either generalized Nash bargaining or alternative offers, equilibrium wages $w^*(y,\ell)$ split the adjusted surplus into constant shares:

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

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

$$w^*(y,\ell) = \Big[(1-\beta)(b+v(\ell)) + \beta y\Big]\ell. \tag{29}$$

However, all matches eventually break up. Hence, the adjusted surplus S solves an optimal stopping problem, and thus a Hamilton-Jacobi-Bellman-Variational-Inequality (HJB-VI):[50]

$$0 = \max\Big\{ \Big(y - \big(b + v(\ell)\big)\Big)\ell + (L_y S)(y,\ell) - \rho S(y,\ell)\ ,\ S(y,\ell) \Big\} \quad , \quad \forall y \geq 0. \tag{30}$$

With equation (30) at hand, Lemma 1 obtains following closely the steps and references in Luttmer (2007), with the definitions

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

For completeness, a full proof is given in Supplemental Material E.2.

B.3 Sorting

B.3.1 Proof of equations (11) and (12)

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

$$\rho J(z,\ell) = (1-\beta)q(\ell)\ell(b+v(\ell))\bar{S}(z,\underline{y}(\ell)) \quad , \quad \bar{S}(z,\underline{y}) \equiv \int \mathcal{S}\left(\frac{y_0}{\underline{y}}\right) G_0(dy_0|z). \tag{31}$$

Under Assumption 1, the integral can be explicitly computed and equation (31) becomes

$$\rho J(z,\ell) = (1-\beta)q(\ell)\ell(b+v(\ell))^{1-\frac{1}{z}}\bar{S}_0(z) \quad ; \quad \bar{S}_0(z) \equiv (\rho Y/\underline{y}_0)^{\frac{1}{z}}\frac{z}{1-z}\frac{\tau z}{\tau z+1}.$$

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

$$\rho J(z,\ell) = \frac{1-\beta}{1-\beta+\beta\underline{y}_0/\rho} q(\ell)\ell\underline{w}(\ell)^{1-\frac{1}{z}}(1-\beta+\beta\underline{y}_0/\rho)^{1/z}\bar{S}_0(z). \tag{32}$$

Define

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

Finally, substitute the definition of $\bar{S}(z)$ into (32) and raise to a power $\frac{z}{1-z}$ to deliver (12).

B.3.2 Workers' value of employment

Before turning to the proof of Propostion 1, it is useful to derive an equilibrium expression for the value of future employment opportunities to unemployed workers $v(\ell)$. In what follows, I denote $\bar{\mathcal{S}}(\zeta) = \bar{S}(1/\zeta)$ where $\zeta = 1/z$.

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

Using Lemma 1 together with the surplus-sharing rule (8), I obtain

$$\begin{aligned} v(\ell)\ell &= \rho^{-1}\beta f(\ell)\ell\big(b+v(\ell)\big)\left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)}\bar{\mathcal{S}}(\zeta(\ell)) \\ &= \rho^{-1}\beta m^{\frac{1}{\alpha}}\ell q(\ell)^{-\frac{1-\alpha}{\alpha}}\big(b+v(\ell)\big)\left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)}\bar{\mathcal{S}}(\zeta(\ell)). \end{aligned} \tag{33}$$

Therefore,

$$q(\ell)^{\frac{1-\alpha}{\alpha}} = \rho^{-1}\beta m^{\frac{1}{\alpha}} v(\ell)^{-1}\big(b+v(\ell)\big)^{1-\zeta(\ell)}B^{\zeta(\ell)}\bar{\mathcal{S}}(\zeta(\ell)). \tag{34}$$

B.3.3 Proof of Propostion 1

The proof of Proposition 1 is structured in four steps. First, re-write the location choice problem (12) into a simpler, equivalent form. Second, show that this assignment problem admits a support that is an interval. Third, show that there is positive assortative matching conditional on workers' values. Fourth, show that there is positive assortative matching when workers' values are determined in general equilibrium.

Step 1: equivalent location choice problem. To make notation lighter, denote $\zeta = 1/z$. Using again the reservation wage equation $b+v(\ell) = \frac{w(\ell)}{1-\beta+\beta \underline{y}_0/\rho}$ to replace reservation wages by $b+v$ in (12), I obtain that new jobs ζ solve the equivalent assignment problem:

$$\ell^*(\zeta) = \underset{\ell}{\operatorname{argmax}}\ [\zeta-1]\log\frac{1}{b+v(\ell)} + \log\big(\ell q(\ell)\big). \tag{35}$$

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. Using (34) to substitute for $q(\ell)$, I obtain

$$\ell^*(\zeta) = \underset{\ell}{\operatorname{argmax}}\ (\zeta-1)\log\frac{1}{b+v(\ell)} + \log\ell + \frac{\alpha}{1-\alpha}\left\{\log\frac{b+v(\ell)}{v(\ell)} + \zeta^*(\ell)\log\frac{B}{b+v(\ell)} + \log\bar{\mathcal{S}}(\zeta^*(\ell))\right\}. \tag{36}$$

Since the complementarity arises between ζ and the endogenous value of search v, it is useful to consider the inverse function $\ell(v)$ rather than $v(\ell)$, and view the problem as

$$v^*(\zeta) = \underset{\ell}{\operatorname{argmax}}\ (\zeta-1)\log\frac{1}{b+v} + \underbrace{\log\ell(v) + \frac{\alpha}{1-\alpha}\left\{\log\frac{b+v}{v} + \zeta^*(v)\log\frac{B}{b+v} + \log\bar{\mathcal{S}}(\zeta^*(v))\right\}}_{\equiv P(v),\text{ endogenous ``price'' sustaining the assignment}}. \tag{37}$$

Step 2: interval property. To use first-order conditions (FOC) to characterize the assignment in problem (37), I first show that the equilibrium results in a single interval of v's.

Now suppose for a contradiction that the function $v(\ell)$ is discontinuous, and that there is a jump at ℓ_0. Denote $v_1 = v(\ell_0^-) \neq v_2 = v(\ell_0^+)$ the lim-sup before the jump and the lim-inf after the jump. The objective function in (35) is continuous and decreasing in $v(\ell)$. Thus, the jump in $v(\ell)$ at ℓ_0 results in a jump in the objective function for a positive mass of employers ζ.[51] Thus, almost no employer would find

[51] Although $q(\ell)$ could also jump at ℓ_0 and ensure continuity of the objective function for a particular ζ_0, the term $\zeta\log\frac{1}{b+v(\ell)}$ ensures that the objective jumps for almost all ζ's.

it optimal to locate on the side of ℓ_0 that has the lowest v. Thus, locations to side of ℓ_0 that deliver the lowest v do not have any employers. But now recall that in the trembling-hand refinement, due to Inada conditions of the matching function, these locations have some (vanishing fraction) of workers. Thus, locating there for a single deviating employer would have infinite returns. This argument delivers the required contraction, and $v(\ell)$ must be a continuous function.

Step 3: sorting conditional on workers' values. Having shown that v is a continuous function, I can consider the location choice problem (37). I use standard results on monotone comparative statics—see e.g. Galichon (2016)—for problem (37), in which I temporarily consider the unknown function $v \mapsto P(v)$ as the equilibrium "price" that sustains the assignment. The supermodularity between ζ and $\log \frac{1}{b+v}$ implies that the equilibrium features a one-to-one assignment function between ζ and v. From the second-order condition, the assignment function that maps $\log \frac{1}{b+v}$ to ζ is increasing, and so the assignment function $\zeta(v)$ is decreasing: $\zeta'(v) < 0$.

Step 4: sorting in general equilibrium. In this last step, I characterize under which conditions the function $\ell \mapsto v(\ell)$ is increasing in equilibrium. Given the property $\zeta'(v) < 0$, increasing v is equivalent to $\zeta'(\ell) < 0$. The FOC in problem (37) leads to

$$(1-\alpha)\frac{v\ell'(v)}{\ell(v)} = \alpha + \frac{v}{b+v}(\zeta(v)-1) + \alpha v \underbrace{\left(\frac{\bar{\mathcal{S}}'(\zeta(v))}{\bar{\mathcal{S}}(\zeta(v))} + \log\frac{B}{b+v}\right)}_{<0} \cdot \underbrace{(-\zeta'(v))}_{>0}. \tag{38}$$

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.[52]

B.4 Endogenous job loss and unemployment

B.4.1 Proof of equation (15)

Consider a unit mass of workers who start employment. Their distribution (in logs, $x = \log y$) follows the KFE without entry:

$$g_t(t,x) = \delta g_x(t,x) + \frac{\sigma^2}{2} g_{xx}(t,x), \tag{39}$$

where subscripts denote partial derivatives. Let $M(t) = \int_{\underline{x}}^{\infty} g(t,x)dx$ denote the mass of employed workers at time t, with $M(1) = 1$. Integrating (39) over $[\underline{x}, +\infty)$ and using $g(t,\underline{x}) = 0$ for all times t leads to $M'(t) = 0 + \frac{\sigma^2}{2} g_x(t,\underline{x})$. Therefore, in steady-state, the job losing rate is $\frac{M'(0)}{M(0)} = \frac{\sigma^2}{2} g_x(\underline{x}) = \frac{\sigma^2}{2} g_y(\underline{y}) e^{\underline{y}-\underline{y}} = \frac{\sigma^2}{2} g_y(\underline{y})$.

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

B.4.2 Proof of Lemma 2

Impose Assumption 1. Consider a single location ℓ and omit location subscripts ℓ. Thus, in logs $x = \log y$, the new job distribution function is $g_0(x) = g_0 e^{-\zeta(x-\underline{x})}$ where $\zeta = 1/z$. Slightly abusing notation, denote $g(x)$ the invariant density in logs, and $h(x) = g'(x)$. Then the KFE becomes $0 = \delta g(x) + \frac{\sigma^2}{2} g'(x) + g_0 e^{-\zeta(x-\underline{x})}$.

The homogeneous solution is $h_H(x) = Ae^{-\kappa(x-\underline{x})}$. Varying the constant, I obtain $\frac{\sigma^2}{2} A'(x) e^{-\kappa(x-\underline{x})} + g_0 e^{-\zeta(x-\underline{x})} = 0$, and so $A(x) = \tilde{A}_0 - \frac{g_0}{s} \int_0^{x-\underline{x}} e^{(\kappa-\zeta)t} dt = A_0 - \frac{n g_0}{s(\kappa-\zeta)} e^{(\kappa-\zeta)(x-\underline{x})}$. 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 $g(x) = Be^{-\kappa(x-\underline{x})} + \frac{g_0}{s\zeta(\kappa-\zeta)} e^{-\zeta(x-\underline{x})}$. Finally, $f(\underline{x}) = 0$ pins down B, so that

$$g(x) = \frac{g_0}{s\zeta(\kappa - \zeta)} \left[e^{-\zeta(x-\underline{x})} - e^{-\kappa(x-\underline{x})} \right] \geq 0. \tag{40}$$

The separation flow is $\frac{\sigma^2}{2} g'(\underline{x}) = \frac{g_0}{\zeta}$. Thus, g_0 simply scales with the total mass of employed workers. Normalizing g to integrate to 1 to obtain a probability density function yields $1 = \frac{g_0}{s\zeta(\kappa-\zeta)} \cdot \frac{\kappa-\zeta}{\zeta\kappa} = \frac{g_0}{s\zeta^2\kappa}$. Substituting into (40) delivers the desired expression.

B.4.3 Proof of Proposition 2

Job losing rate. From Lemma 2 and equation (15), the job losing rate is $s\zeta\kappa = \delta\zeta = \delta/z$.

Job finding rate. To express job finding, it suffices to use the definition of workers' value of search. Under Assumption 1, they follow equation (33). 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 $\underline{w}_1 = b\Big(1 - \beta + \beta \underline{y}_0/\rho\Big)$.

Unemployment rate. The expression for the unemployment rate then follows from the two-state model as in Section 1.1.

B.5 Proof of Proposition 3

The proof of Proposition 3 is structured in four steps. First, derive the system of ODEs that determine the equilibrium. Second, show existence of solutions to this system conditional on general equilibrium aggregates. Third, show existence of general equilibrium aggregates. Fourth, show uniqueness.

Step 1: ODE system. 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 implies

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

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

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

Define also 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. (42) defines a function Z such that $\zeta'(\ell) = Z(\ell, v(\ell), \zeta(\ell))$. In addition,

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

Substituting into the FOC for v:

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

(43) defines a function V such that $v'(\ell) = V(\ell, v(\ell), \zeta(\ell))$. Given K, equations (42)-(43) define a coupled system of ODEs, with two boundary conditions: $\zeta(\underline{\ell}) = \overline{\zeta}$ and $\zeta(\overline{\ell}) = \underline{\zeta}$.

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.

Step 2: 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 one $\underline{v}(K)$ such that $\zeta(\overline{\ell}) = \underline{\zeta}$.

Step 3: 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

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

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.

Step 4: Uniqueness. Now suppose that the supports of F_ℓ, F_z are small enough. This assumption makes possible using 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 (44),

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

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

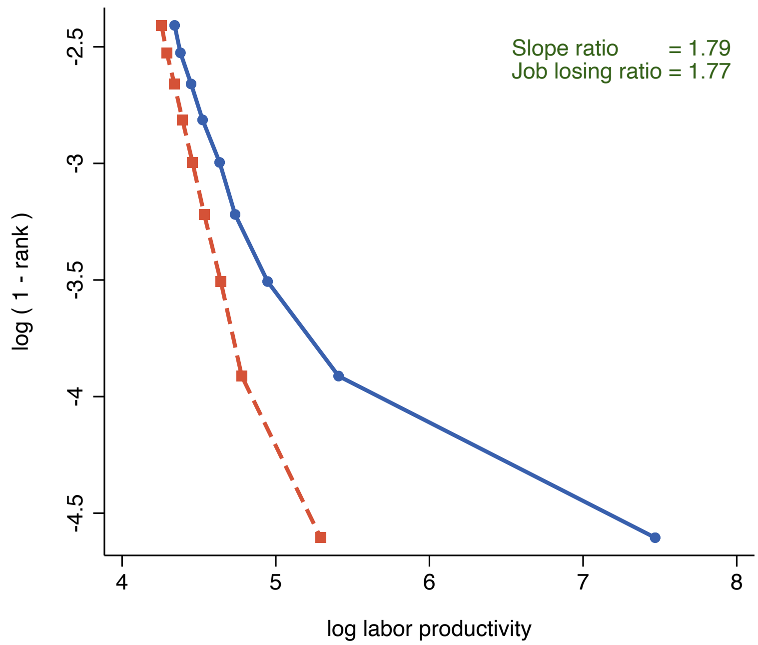
$$K = J_0'' \frac{\underline{v}}{\underline{v} + 1/L_1} (b + \underline{v})^{\frac{1}{\omega}}, \tag{46}$$

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.

Figure 13: Tail labor productivity distribution across French commuting zones.



Note: Figure 13 plots the log inverse cumulative distribution function for bottom and top quartiles of commuting zones, ranked by job losing rates

B.6 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 and $\omega > 0$, 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. If instead $\omega = 0$, the free-mobility condition (10) would equalize v across locations.

B.7 Model validation

Figure 13 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 1.79. It is close to the ratio of group averages of job losing rates, which is 1.77. Together, these results support the Pareto assumption.

Table 7 reports the results from linear regressions of labor productivity for incumbent and entrant

Table 7: Plant-level regressions.

	Level			Growth rate	
	(1) VA/N	(2) VA/N	(3) VA	(4) VA/N	(5) VA/N
Incumbent × Separation rate	-0.028**	-0.020*	-0.054***	0.001	-0.000
	(0.010)	(0.008)	(0.011)	(0.006)	(0.006)
Entrant × Separation rate	-0.050***	-0.040***	-0.064***		
	(0.011)	(0.010)	(0.013)		
Controls					
Year	✓	✓	✓	✓	✓
2-digit industry	✓	✓	✓	✓	✓
Skill mix		✓	✓		✓
Obs.	92694	92694	92694	31373	31373
R^2	0.120	0.155	0.150	0.012	0.013
W.-R^2	0.025	0.065	0.035	0.000	0.001

Note: Standard errors in parenthesis, clustered by city and 2-digit industry. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Annual frequency, 1997-2006. Entrant defined as less than two year old. Employment-unweighted regressions.

establishments on local job losing rates, as well as of annual labor productivity growth on local job losing rates.

C Estimation

C.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 $h_t = L_y^* h$. 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 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,

$$\begin{aligned} g(t, x) &= e^{-(\Delta+\mu)t} \int_{\underline{x}}^{\infty} G(t, x, y) g_0(x_0) dx_0 \\ &= \frac{\zeta}{2} e^{-\left(D+\delta\zeta-\frac{\sigma^2\zeta^2}{2}\right)t-\zeta x} \left(1 + \text{Erf}\left[\frac{x + t(\delta - \zeta\sigma^2)}{\sqrt{2t}\sigma}\right] - e^{2\left(\zeta-\frac{2\delta}{\sigma^2}\right)x} \text{Erfc}\left[\frac{x - t(\delta - \zeta\sigma^2)}{\sqrt{2t}\sigma}\right]\right) \end{aligned} \quad (49)$$

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. The second equality imposes $g_0(x_0) = \zeta e^{-\zeta(x_0-\underline{x})}$. $D = \Delta + \mu$, and Erf denotes the error function, a transformation of the Gaussian cumulative function.

C.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. Recall that the flow of workers into local unemployment is 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 (49). It can be shown that

$$\text{Endog. Sep.}(t) = e^{-(\mu+\Delta)t}\left[\frac{(s/\hat{\delta})}{\sqrt{t}}\varphi\left(\hat{\delta}\sqrt{t}\right) + \frac{(s/\hat{\delta})^2}{2}e^{\frac{(s/\hat{\delta})^2 t}{2}}\left\{e^{st}\Phi\left(-(\hat{\delta}+s/\hat{\delta})\sqrt{t}\right) - e^{-st}\Phi\left((\hat{\delta}-s/\hat{\delta})\sqrt{t}\right)\right\}\right],$$

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. I obtain

$$\begin{aligned} s_1(s,\hat{\delta},D) &= \frac{(s/\hat{\delta})}{4}\Bigg\{\frac{e^{-(D+s-(s/\hat{\delta})^2/2)}}{D+s-(s/\hat{\delta})^2/2}(s/\hat{\delta}) + 4\frac{\text{Erf}\left[\frac{\sqrt{2D+\hat{\delta}^2}}{\sqrt{2}}\right]}{\sqrt{2D+\hat{\delta}^2}} \\ &\quad + \frac{s/\hat{\delta}}{D+s-(s/\hat{\delta})^2/2}\left(-1 - \frac{\hat{\delta}-s/\hat{\delta}}{\sqrt{2D+\hat{\delta}^2}}\text{Erf}\left[\frac{\sqrt{2D+\hat{\delta}^2}}{\sqrt{2}}\right] + e^{-(D+s-(s/\hat{\delta})^2/2)}\text{Erf}\left[\frac{\hat{\delta}-(s/\hat{\delta})}{\sqrt{2}}\right]\right) \\ &\quad + \frac{(s/\hat{\delta})}{D-s-(s/\hat{\delta})^2/2}\left(-1 + \frac{\hat{\delta}+s/\hat{\delta}}{\sqrt{2D+\hat{\delta}^2}}\text{Erf}\left[\frac{\sqrt{2D+\hat{\delta}^2}}{\sqrt{2}}\right] + e^{-D+s+(s/\hat{\delta})^2/2}\text{Erfc}\left[\frac{\hat{\delta}+(s/\hat{\delta})}{\sqrt{2}}\right]\right). \end{aligned}$$

In the limit of a small D, it can be checked that s_1 is a decreasing function of $\hat{\delta}$.

C.3 Tenure profile of wages

In the model, the wage of individual i in employment spell p, location c, at calendar time τ and tenure t, is given by $w_{ipc\tau t} = w_{0c} \times (A + Be^{x_{ipct}})k_{ipc\tau t}$, with $A = 1-\beta$ and $B = \beta\frac{y_0}{\hat{\rho}}$, and w_{0c} is a location shifter. $k_{ipc\tau t}$ is the worker's human capital and is correlated with tenure. Calendar time and tenure are collinear: $\tau_{ip0} + t_{ip} = \tau$ where τ_{ip0} is the calendar time at which individual i started employment spell p. Therefore, $k_{ipc\tau t} = k_{\tau_{ip0}}e^{\lambda(\tau-\tau_{ip0})}$. Mean level wages in the economy grow at rate λ, and so $\bar{W}_\tau = \bar{W}_0 e^{\lambda\tau}$. In particular, evaluating at the starting time of the spell, $\bar{W}_{\tau_{ip0}} = \bar{W}_0 e^{\lambda\tau_{ip0}}$. Therefore, taking wages relative to aggregate wages at the starting time of the spell.

$$\hat{w}_{ipc\tau t} \equiv \frac{w_{ipc\tau t}}{\bar{W}_\tau} = \tilde{w}_{0c} \times (A + Be^{x_{ipct}}) \times \varepsilon_{\tau_{ip0}} \quad ; \quad \tilde{w}_{0c} = \frac{w_{0c}}{\bar{W}_0} \quad ; \quad \varepsilon_{\tau_{ip0}} = k_{\tau_{ip0}}e^{-\lambda\tau_{ip0}}.$$

$\varepsilon_{\tau_{ip0}}$ plays the role of a individual-spell fixed effect. Take rescaled wage growth relative to initial rescaled wages at the spell: $\omega_{ipc\tau t} \equiv \frac{\hat{w}_{ipc\tau t}}{\hat{w}_{ipc,\tau_{ip0},0}} = \frac{A+Be^{x_{ipct}}}{A+Be^{x_0}}$. Recall that x_{ipct} and x_0 are correlated by definition.

Inspection of (49) reveals that $y = x/\delta$ has a distribution that is independent from δ at all times—including time 0—conditional on $s, \hat{d}$. Anticipating a small estimate of δ, I obtain $\omega_{ipc\tau t} = \frac{\hat{w}_{ipc\tau t}}{\hat{w}_{ipc,\tau_{ip0},0}} = \frac{A+Be^{\delta y_t}}{A+Be^{\delta y_0}} \approx_{\delta\to 0} 1 + \frac{\delta B}{A+B}(y_t - y_0)$. To a first order, when δ is small relative to ρ, it can be shown that

$\frac{B}{A+B} \approx \frac{\beta}{(1-\beta)\hat{\rho}+\beta}$. Therefore, $\omega_{ipc\tau t} \approx 1 + \frac{\delta\beta}{(1-\beta)\hat{\rho}+\beta}(y_t - y_0)$. Then compute

$$\mathbb{E}[\omega_{ipc\tau t}|t, c] = 1 + \frac{\delta}{(1-\beta)\hat{\rho}+\beta}R(t) \tag{50}$$

where $R(t) = \frac{\int h_0(y_0)dy_0 \int G^y(t,y,y_0)(y-y_0)dy}{\int h_0(y_0)dy_0 \int G^y(t,y,y_0)dy}$. Mean rescaled wages in equation (50) are easily computed in the data, and can also be computed in the model at this stage of the estimation. G^y is the Green's function associated with (49) and the change of variables $y = x/\delta$. It is given by $G^y(t, y, y_0)dy = e^{-Dt}\Big(\gamma^y(t, y - y_0) - e^{2\hat{\delta}^2 y_0}\gamma^y(t, y + y_0)\Big)$ with $\gamma^y(t, z) = \frac{\hat{\delta}\exp\left(-\frac{\hat{\delta}^2}{2t}(z+t)^2\right)}{\sqrt{2\pi t}}$.

C.4 Labor share

From the bargaining solution, the labor share in location ℓ is

$$\text{Labor Share}(\ell) = \frac{(1-\beta)(b + v(\ell) + \beta\underline{y}_0/\rho H(\ell)}{E(\ell)}, \tag{51}$$

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

$$H(\ell) = \frac{\kappa\zeta(\ell)}{(\kappa-1)(\zeta(\ell)-1)} = \frac{\kappa}{(\kappa-1)(1-z(\ell))} \equiv H(z(\ell)).$$

C.5 Learning parameters

Log real wages are proportional to $K_t R_t^{-\omega}$, 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, $K_t R_t^{-\omega} \propto K_t^{1-\omega-\psi}$ 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 (22) 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 (2021). Thus, in practice, equation equation (22) may deliver a biased estimate of the depreciation rate of human capital. To address such concerns, I run version of equation (22) 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 8. The point estimate remains stable across specifications.

Table 8: Unemployment scar estimation. Dependent variable: post-unemployment spell log wage.

	Unemployed only					DiD
	(1)	(2)	(3)	(4)	(5)	(6)
Job loss × Duration	-0.02***	-0.01***	-0.01***	-0.01***	-0.01	-0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Job loss						-0.10**
						(0.04)
Duration						0.00
						(0.00)
Pre log wage		0.55***	0.44***	0.41***	-0.18*	-0.01
		(0.02)	(0.03)	(0.03)	(0.07)	(0.05)
Skill			0.00***	0.00***		
			(0.00)	(0.00)		
Fixed Effects						
Year	✓	✓	✓	✓	✓	✓
2-digit Industry				✓	✓	✓
City				✓	✓	✓
Worker					✓	✓
Obs.	35021	35021	35021	35020	6100	76700
R^2	0.027	0.287	0.331	0.353	0.802	0.775
W.-R^2	0.011	0.276	0.320	0.267	0.022	0.002

Note: Standard errors in parenthesis, two-way clustered by city and 2-digit industry. $^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

C.6 Local wages

Wages in location ℓ are given by

$$\overline{w}(\ell) = W_0 p \cdot \left(\frac{\Delta}{\Delta + \varphi u(\ell)}\right) r(\ell)^{-\psi} \underline{y}(\ell) \left[(1-\beta)\frac{\rho}{\underline{y}_0} + \beta H(\ell)\right], \tag{52}$$

where W_0 is a general equilibrium constant.

C.7 Housing elasticity

Using (64) together with the solution for average wages in a location $W(\ell)$, housing prices become

$$r(\ell)^{1+\eta} \quad = \quad \frac{\overline{w}(\ell)}{1-\beta+\beta\underline{y}_0/\hat{\rho}E(s(\ell))} \cdot L(\ell, a)G(s(\ell)/\delta, v(\ell)).$$

The right-hand-side defines r_0, and involves parameters that have been estimated or data.

C.8 Migration elasticity

To circumvent endogeneity in the OLS regression version of (24), 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 Supplemental Material D.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.

C.9 Productivity distribution

To estimate F_z, I first recover firm quality in each location using (23). 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 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.

C.10 Matching function elasticity

Start from $\theta(\ell) = \left(\frac{f_\zeta(\ell)|\zeta'(\ell)|}{f_\ell(\ell)\mathcal{U}(\ell)}\right) J(\ell)^\gamma$. Note that $J(\ell) = q(\ell)\bar{J}(\ell) \propto \theta(\ell)^{-\alpha}\hat{J}(\ell)$, where $\hat{J}(\ell) \propto \bar{k}(u(\ell))^{\mathcal{Q}} \cdot \ell(b + v(\ell))^{1-\psi\mathcal{P}} \cdot G(v(\ell), \zeta(\ell))^{\varepsilon\psi\mathcal{P}} \left(\frac{B}{b+v(\ell)}\right)^{\zeta(\ell)} \bar{\mathcal{S}}(\zeta(\ell))$. 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

$$\log\left(\frac{f_R(\ell)}{\mathbb{P}_\ell[\text{Accept}]}\right) = \text{cste} + \frac{1-\alpha}{1+\alpha\gamma}\log\frac{f_z(z(\ell))z'(\ell)}{f_\ell(\ell)\mathcal{U}(\ell)} + \frac{(1-\alpha)\gamma}{1+\alpha\gamma}\log\hat{J}(\ell, \underline{y}(\ell), z(\ell)) \ , \tag{53}$$

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 (53) 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 (53).

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 (53), 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 (53) between the two subperiods. I use as the first instruments the same shift-share shocks ΔE_c. Under the same assumptions as in section C.8, 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.

C.11 Over-identification

See Table 9.

Table 9: Correlation of estimated amenities with observables.

Sun hours	0.119*	(0.050)
Basic public services	0.072	(0.050)
Education services	-0.012	(0.039)
Health services	0.067**	(0.014)
Commercial services	0.046	(0.030)
Obs.	288	
R^2	0.457	

Heteroskedastic-robust standard errors in parenthesis. $^+$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$. Log amenities on log sun hours per month and log service establishments.

Supplemental Material

The Geography of Unemployment

Adrien Bilal

D Data and descriptive evidence

D.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.

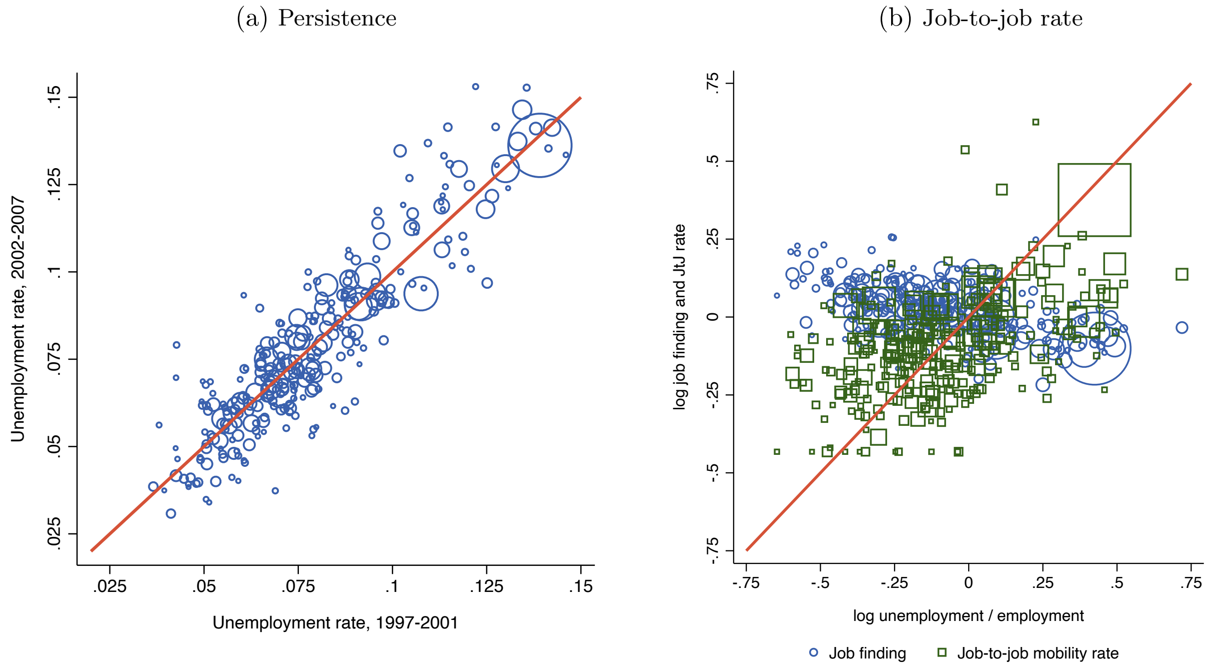
LFS. I complement the DADS panel with the LFS. I use the LFS starting in 2003 due to a large survey change in 2002. The LFS is quarterly and tracks individuals for six consecutive quarters. The LFS reports whether an individual is working, unemployed or out of the labor force. As in many surveys, the 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 probalistically impute the non-participation vs. unemployment status of individuals in the DADS panel. Table 10 shows that the DADS panel and the LFS have similar aggregate statistics.

Table 10: Summary statistics

	DADS	LFS
Unemployment rate	0.100	0.071
Implied unemp. rate from losing and finding	0.109	0.055
Participation rate	0.931	0.903
E-to-U probability	0.021	0.015
U-to-E probability	0.173	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)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: Persistence of local unemployment, and job-to-job mobility rate. France.



Note: Figure 14(a) plots commuting zone unemployment in two subperiods of the sample, after controlling for economy-wide industry cycles. Blue circles represent a commuting zone. Size is proportional to population. Figure 14(b) plots the job-to-job and job-finding rate across French commuting zones. Blue circles represent the log of the job finding rate against the log of the unemployment-to-employment ratio, across commuting zones in France (DADS panel). Green squares represent the log of the log job-to-job mobility rate across commuting zones in France (DADS panel). 45 degree line in orange. Estimating a linear regression delivers the following slopes. Job finding rate: -0.19. Job-to-job mobility rate: 0.46; excluding Paris: 0.26.

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 2-digit level. Then define skill as average occupation and age premium

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

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 and panel, which lets me identify all workers in the different establishments of the firm. For some exercises I restrict the analysis to single-establishment firms 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.

Establishment-level vacancy data. I merge the DADS data with a large-scale quarterly survey that reports vacancies at the establishment level (*"Activité et Conditions d'Emploi de la Main-d'Oeuvre—ACEMO"*).

D.2 Persistence and job-to-job rate

Figure 14(a) shows persistence in local unemployment rates after netting out country-wide industry cycles. The autocorrelation is 1.05. 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 both subperiods 0 (1997-2001) and 1 (2002-2007). Then, I construct a predicted employment change at the commuting zone level by projecting the predicted 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(a) plots $\hat{u}_{c,1}$ against $u_{c,0}$.

Figure 14(b) plots the log commuting zone-level job-to-job mobility rate, and the log job finding rate from unemployment, against the log unemployment-to-employment rate across French commuting zones.

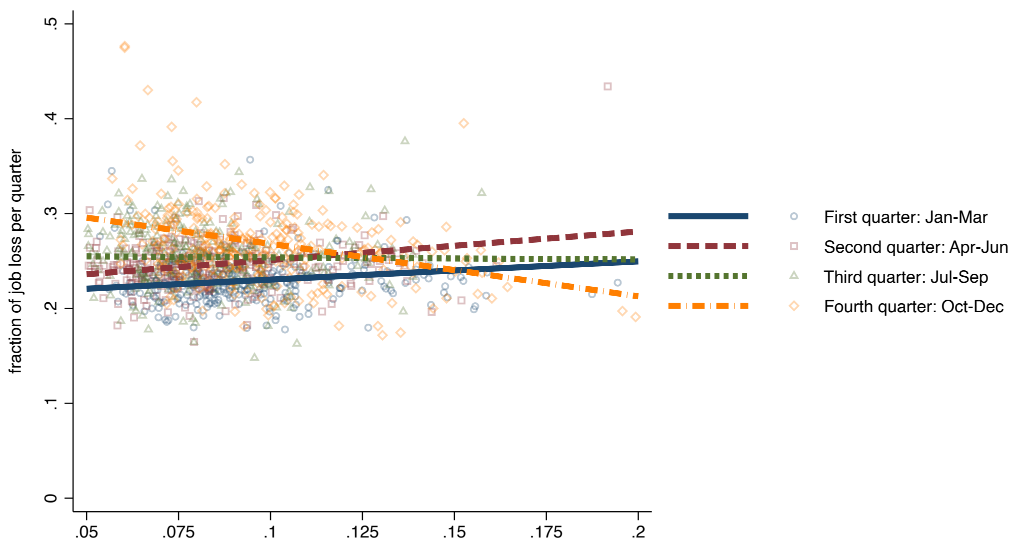
D.3 Mechanical correlates of job loss

Temporary contracts. If the proportion of workers under temporary contracts varies systematically across locations, it may mechanically lead to more job loss in locations with a high proportion of temporary contracts. I use the LFS to assess the role of temporary contracts. I first evaluate the excess probability of job loss for a worker under a temporary contract. A simple linear regression indicates that a worker under temporary contract has a 1.6 percentage point higher probability of separating into unemployment at the quarterly level—more than twice the average job losing rate.

I then run a shift-share decomposition of job loss, interacting the excess risk of job loss under temporary contracts with the share of workers under temporary contracts across locations—which varies from 18 to 23% across locations. As a result, temporary contracts account for no more than 14% of the overall differences in job loss.

Seasonality. If there are large seasonal variation in employment across locations, it may mechanically account for some of the spatial differences in job loss. Figure 15 scatterplots the fraction of job loss by quarter against the job losing rate, across French commuting zones. Figure 15 reveals that seasonality correlates somewhat with average job losing rates. Comparing the linear fits at the highest unemployment commuting zone, the fraction of job loss in the fall quarter is at most 8 percentage point lower than in other quarters. Thus, seasonality can account for no more than 8% of spatial gaps in job loss.

Figure 15: Fraction of job loss by quarter against commuting zone job losing rate, France.



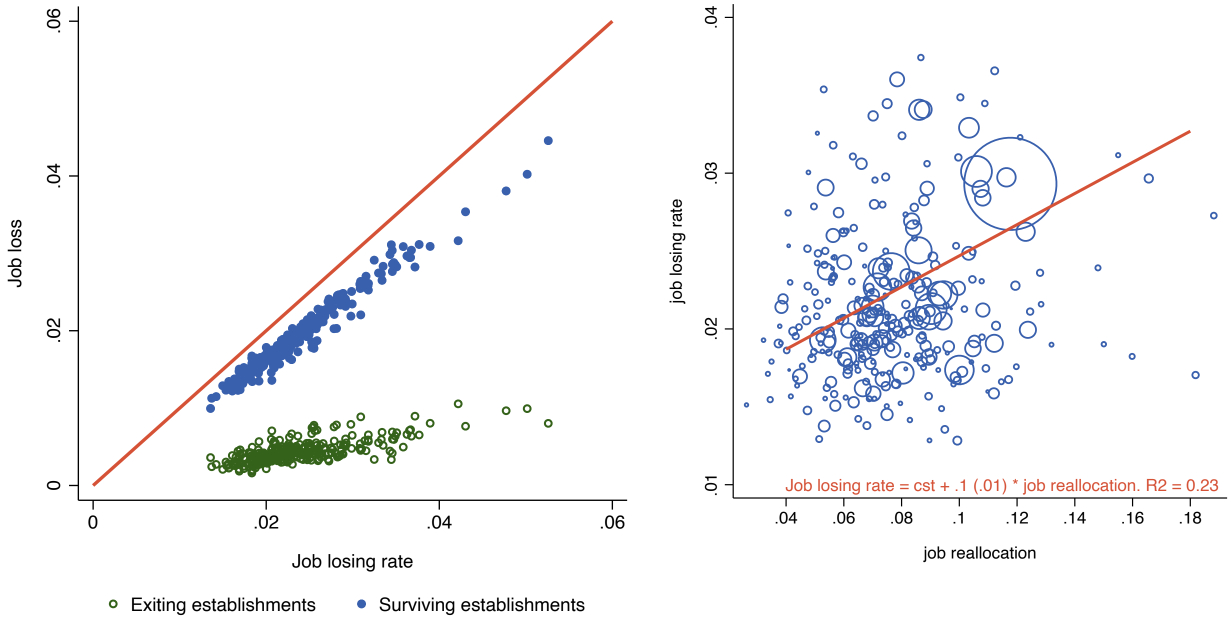
Note: Scatterplot of commuting zone-by-quarter fraction annual of job loss, by commuting zone unemployment rate. Economy-wide means of quarterly fraction of job loss adjusted using the LFS to limit reporting measurement error in the DADS.

Establishment exit and job reallocation. Figure 16(a) shows that establishment exit accounts for 11% of spatial gaps in job loss. Job loss at continuing establishments account for the remaining 89%. Thus, establishment exit is not a sizeable proximate cause for spatial gaps in job loss. Figure 16(b) shows that, even in a purely mechanical sense, job reallocation accounts for no more than 23% of spatial gaps in job loss. Therefore, job reallocation is not a sizeable proximate cause for spatial gaps in job loss.

Figure 16: Correlation between job loss, firm exit and job reallocation.

(a) Job loss at exiting and continuing establishments

(b) Job loss and job reallocation



Note: Figure 16(a): job loss from exiting establishments and continuing establishments across French commuting zones, ordered by job losing rate. Figure 16(b): commuting zone job losing rate against commuting zone job reallocation.

E Baseline model

E.1 Proof of Lemma 3

Generalized Nash bargaining. Under Nash bargaining, wages split the Nash product $w^* = \text{argmax}_w (W_0 w - W_1)^\beta (F_1 - w)^{1-\beta}$. It is straightforward to show that $\beta(F_1 - w^*) = (1-\beta)\left(w^* - \frac{W_1}{W_0}\right)$, and thus

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

where $S(w) = F(w) + \frac{W(w)}{W_0}$ is the adjusted surplus and is independent from w.

Alternating offers. To solve for wages in the alternating offers 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 Pareto frontier is exactly equal to the set of admissible wages. I now 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 $\delta^W[W_0 w^W - W_1 a] = W_0 w^F - W_1$ and $\delta^F[F_1 - w^F] = F_1 - w^W$. 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 $\beta^W = \frac{1-\delta^F}{1-\delta^F\delta^W} \in (0,1)$. Hence, $w^F = \beta^F F_1 + (1-\beta^F)\frac{W_1}{W_0}$, 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 $\frac{W(w^*)}{W_0} = \beta \cdot \Big(F(w^*) + \frac{W(w^*)}{W_0}\Big)$ and $F(w^*) = (1-\beta)\cdot \Big(F(w^*) + \frac{W(w^*)}{W_0}\Big)$. Therefore, it is enough to define an adjusted surplus $F(w) + \frac{W(w)}{W_0}$ which does not depend on wages. Rescaled values then split this adjusted surplus.

E.2 Proof of Lemma 1

The structure of the proof follows two steps. The first step uses standard results on HJB-VIs to obtain a Partial Differential Equation (PDE) formulation with boundary conditions for the adjusted surplus. The second step explicitly solves this PDE.

Step 1: from the HJB-VI to a PDE. The structure of the HJB-VI (30) has three implications. First, there exists a continuation region in which the HJB (28) 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. Pham (2009) derives the interval property and the smooth-pasting condition.

Third, the joint surplus must be smaller than the surplus of a match without any outside option, which is $\frac{y\ell}{\rho+\delta-\sigma^2/2}$. 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.

Together, the HJB (28), 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

$$\begin{aligned} \rho S(y,\ell) \quad &= \quad \Big(y - \big(b + v(\ell)\big)\Big)\ell + (L_y S)(y,\ell) \quad , \quad \forall y \geq \underline{y}(\ell) \\ &\text{s.t.} \quad S(\underline{y}(\ell), \ell) = 0 \ , \ \frac{\partial S}{\partial y}(\underline{y}(\ell), \ell) = 0 \ , \ S(y,\ell) \leq \frac{y\ell}{\rho+\delta-\sigma^2/2}. \end{aligned} \tag{54}$$

Step 2: solving the PDE (54). To lighten notation, I drop location indices ℓ and solve without loss of generality

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

First re-express the problem in logs $x = \log y$ by defining $\tilde{S}(x) = S(e^x)$. Then

$$\rho\tilde{S}(x) \quad = \quad e^x - c - \delta\tilde{S}'(x) + \frac{\sigma^2}{2}\tilde{S}''(x) \quad , \quad \forall x \geq \underline{x} \quad \text{s.t.} \ \ \tilde{S}(\underline{x}) = 0 \ , \ \tilde{S}'(\underline{x}) = 0 \ , \ \tilde{S}(x) \leq \frac{e^x}{\rho+\delta-\sigma^2/2}.$$

This problem is a second-order PDE with two boundary conditions, an unknown threshold $\underline{x}$, and a growth condition. I follow standard methods to solve the PDE given boundary conditions.

Homogeneous equation. Look for a solution $\tilde{S}(x) = e^{-\tau x}$ to $\rho\tilde{S}(x) = -\delta\tilde{S}'(x) + s\tilde{S}''(x)$ 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 $\tilde{S}_H(x) = Ae^{-\tau x}$, $A \in \mathbb{R}$.

Inhomogeneous equation. Now look for solutions $\tilde{S}(x) = Ae^{-\tau x} + Be^x - C$. Substituting into the HJB, the homogeneous term drops out and we find $B = \frac{1}{\rho+\delta-s}$ and $C = \frac{c}{\rho}$. Be^x is the value if the match 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 from the HJB. I am left with $(A, \underline{x})$ to determine, with the two boundary conditions $\tilde{S}(\underline{x}) = 0$, $\tilde{S}'(\underline{x}) = 0$. These conditions imply $Ae^{-\tau \underline{x}} + Be^{\underline{x}} = C$ and $-A\tau e^{-\tau \underline{x}} + Be^{\underline{x}} = 0$. Hence $e^{\underline{x}} = \frac{\tau}{\tau+1} \cdot \left(1 - \frac{1+\kappa}{\eta}\right) \cdot d$ and $A = \frac{B}{\tau} e^{(1+\tau)\underline{x}}$. The solution finally writes

$$\tilde{S}(x) = \frac{e^{\underline{x}}}{\rho+\delta-s} \cdot \left\{ e^{x-\underline{x}} + \tau^{-1} e^{-\tau(x-\underline{x})} \right\} - \frac{c}{\rho}.$$

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

E.3 Sorting 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)*
Let $\bar{S}(z, \underline{y})$ be the integral defined in equation (31). Assume that

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

This assumption lets me generalize the sorting results.

Proposition 8. *(Sorting 2)*
All the implications of Proposition 1 hold under Assumption 2 instead of Assumption 1.

Proof. The structure of the proof closely follows Appendix B.3.3. Steps 1 and 2 obtain analogously. The main differences to check are steps 3 and 4. First, the location choice becomes

$$\ell^*(z) = \underset{\ell}{\operatorname{argmax}} \ \log \underline{y}(\ell)\bar{S}(z, \underline{y}(\ell)) + \log(\ell q(\ell)). \tag{55}$$

Steps 3 & 4. Using the expression for the value of search,

$$q(\ell)^{\frac{1-\alpha}{\alpha}} \propto \frac{\underline{y}(\ell)\bar{S}(z,\underline{y}(\ell))}{\underline{y}(\ell) - \underline{y}_1} \quad ; \quad \underline{y}_1 \equiv b\underline{y}_0/\rho.$$

Hence

$$\ell^*(z) = \underset{\ell}{\operatorname{argmax}} \log \bar{S}(z, \underline{y}(\ell)) + \log \ell + \log \underline{y}(\ell) + \frac{\alpha}{1-\alpha} \log \frac{\underline{y}(\ell)\bar{S}(z^*(\ell), \underline{y}(\ell))}{\underline{y}(\ell) - \underline{y}_1}.$$

As before, it suffices to consider the case $\alpha \to 0$. In that case,

$$\ell^*(z) = \underset{\underline{y}}{\operatorname{argmax}} \log \left(\underline{y}\bar{S}(z,\underline{y})\right) + \log \ell(\underline{y}).$$

Because $\underline{y}\bar{S}(z,\underline{y})$ is log-supermodular in $(z, \underline{y})$, PAM between z and $\underline{y}$ obtains: $z'(\underline{y}) > 0$. Under Assumption 2, $\log \underline{y}\bar{S}(z,\underline{y})$ is increasing in z and decreasing in $\underline{y}$. Therefore, the "price" that sustains the assignment is increasing in $\underline{y}$. Hence, $\ell(\underline{y})$ is increasing. Thus, $\underline{y}(\ell)$ is increasing, and so is $z(\ell)$. $\square$

Mass point 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

$$\frac{\partial \log \underline{y}\bar{S}(z,\underline{y})}{\partial z} = \tau \frac{1 - z^{-\tau-1}\underline{y}^{1+\tau}}{\underline{y}\bar{S}(z,\underline{y})} > 0 \quad ; \quad \frac{\partial \log \underline{y}\bar{S}(z,\underline{y})}{\partial \underline{y}} = \frac{1+\tau}{\underline{y}} \frac{(\underline{y}/z)^\tau - 1}{z/\underline{y} + (\underline{y}/z)^\tau - 1 - \tau} < 0.$$

When $\underline{y}/z$ is large enough, $\frac{(\underline{y}/z)^\tau - 1}{z/\underline{y} + (\underline{y}/z)^\tau - 1 - \tau} \approx \frac{X(z)}{X(z) - \tau}$ for $X(z) = (\underline{y}/z)^\tau$ which is clearly increasing in z. Therefore, $\frac{\partial^2 \log \underline{y}\bar{S}(z,\underline{y})}{\partial \underline{y} \partial z} > 0$ on some interval $[K(\underline{y}(\ell)), +\infty)$.

E.4 Sorting 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 (13) 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.[53]

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", i.e. 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 $\mathbb{R}_+$. Define $\mathcal{T}$ to be the set of times in $\mathbb{R}_+$ such that $\underline{y}', z' \geq 0$ and such that $\partial_\ell(\ell q(\ell)) > 0$. 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 choice is given by (55), but where $q'/q = 0$. Therefore, $\underline{y}' > 0$ and $z' > 0$ immediately follows at time 0. Hence $0 \in \mathcal{T}$.

Recursion. Let t be the least upper bound of $\mathcal{T}$. The location choice for employers allowed to relocate at t is (55), where by definition of $\mathcal{T}$, $\underline{y}', z' \geq 0$ and $\partial_\ell(\ell q(\ell)) \geq 0$ at t. Monotone comparative statics and the SOC then imply that these inequalities are strict. Since only a small fraction of employers relocate every period, it then immediately follows that $\underline{y}', z' \geq 0$ and and $\partial_\ell(\ell q(\ell)) \geq 0$ 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 $\mathbb{R}_+$.

□

E.5 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.

[53] Exogenous differences across locations ℓ create incentives for jobs to sort, but so do endogenous differences in the vacancy meeting 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 meeting 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.

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:

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

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

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

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 $\int_{\delta}^{\delta+\varepsilon} f(\underline{x}+y)dy \approx f(\underline{x})\varepsilon + \frac{1}{2}f'(\underline{x})\varepsilon(2\delta+\varepsilon) + \frac{1}{6}f''(\underline{x})\varepsilon[3\delta^2 + 3\delta\varepsilon + \varepsilon^2] + \dots$. So

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

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, $B \approx_{\lambda \to +\infty} A_{-\infty}^{0} + \frac{f(\underline{x})}{\lambda}\int_{-\infty}^{0} \varphi(z)z\ dz + \mathcal{O}(\lambda^{-2})$, and so $A + B = f(\underline{x}) + f'(\underline{x})dx + \frac{f''(\underline{x})dx^2}{6} - \frac{f(\underline{x})}{\lambda\sqrt{2\pi}} + \mathcal{O}(\lambda^{-2} + \dots)$. 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 $\lambda dt \sim dxdt^{1/2} \to 0$. This implies: $\frac{\partial f}{\partial t}(\underline{x}) = -\infty$, and thus $f(\underline{x}, t) = 0$ for all times $t > 0$.

E.6 Generalization of Lemma 2 and Proposition 2 under Assumption 2

Impose Assumption 2 instead of Assumption 1. This section proves the following result.

Proposition 10. *(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 $B(\ell) = \frac{1}{\kappa}\int_{\underline{y}(\ell)}^\infty g_0(y'|z(\ell))\Big(y'/\underline{y}(\ell)\Big)^{\kappa}\frac{dy'}{y'}$. The job losing and finding rates are

$$s(\ell) = \frac{\delta}{\int_{\underline{y}(\ell)}^\infty \left(\log\frac{y'}{\underline{y}(\ell)}\right)dy'} \quad ; \quad f_R(\ell) = \frac{\rho}{\beta}\frac{v(\ell)}{b+v(\ell)}\frac{1-G_0(\underline{y}(\ell)|z(\ell))}{\bar{S}(z(\ell),\underline{y}(\ell))}.$$

The proof of Proposition 10 is structured in two main steps. First, I extend Lemma 2. Second, I extend Proposition 2.

Step 1: extending Lemma 2. Apart from the entry distribution, the KFE remains identical. The homogeneous solution is the same as in the proof of Lemma 2. Again varying the constant and looking for a solution $g'(x) = A(x)e^{-\kappa(x-\underline{x})}$, I obtain $g'(x) = A_0 e^{-\kappa(x-\underline{x})} + e^{-\kappa(x-\underline{x})}\int_x^\infty g_0(y)e^{\kappa(y-\underline{x})}dy$. Integrating once more:

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

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.

Step 2: extending Proposition 2. The separation flow is $\frac{\sigma^2}{2}g'(\underline{x})$, 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,

$$\frac{\kappa}{H_0} \;=\; \int_{\underline{x}}^\infty e^{-\kappa(x-\underline{x})}\int_{\underline{x}}^x e^{\kappa(y-\underline{x})}h_0(y)dy - \frac{1}{\kappa} + \int_{\underline{x}}^\infty\int_x^\infty h_0(y)dy = \int_{\underline{x}}^\infty xh_0(x)dx.$$

Therefore the job losing rate is $\frac{\delta}{\mathbb{E}^{h_0}[\log(y/\underline{y})]}$. Re-arranging the worker's value of search yields the expression for $f_R(\ell)$.

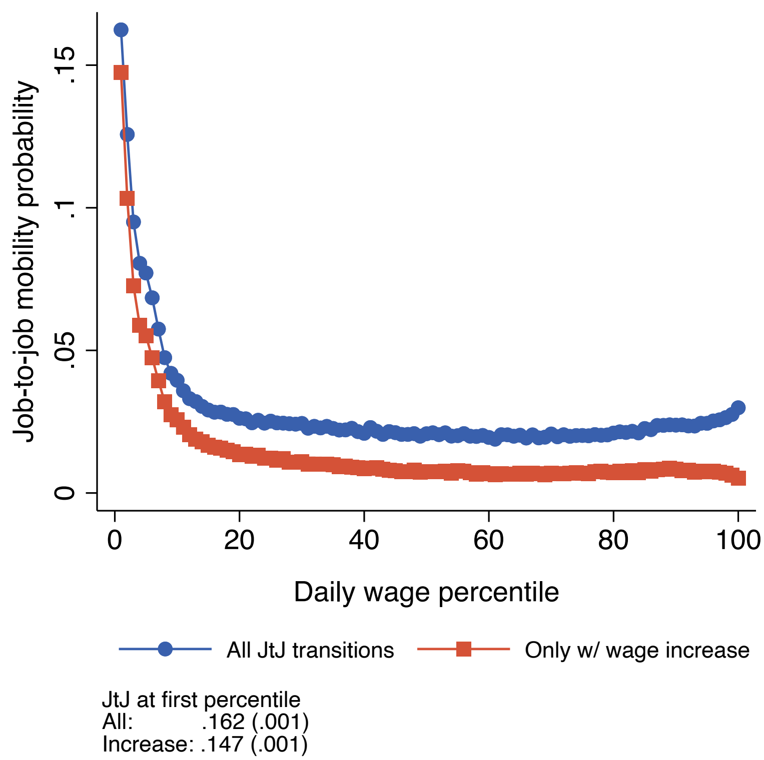
E.7 Job-to-job correction

In both constant returns models with job-to-job search and wage posting (Burdett and Mortensen, 1998, Engbom and Moser, 2021), wage bargaining (Cahuc et al., 2006), or decreasing returns models with job-to-job search with wage posting (Bilal and Lhuillier, 2021), wage bargaining (Bilal et al., 2019), the job-to-job transition rate of the lowest wage workers is equal to the job finding rate of unemployed workers. Thus, in these models, an estimate of the relative search intensity of employed workers, ξ, is

$$\xi = \frac{\text{Job-to-job transition rate of workers in first wage centile}}{\text{Job-finding rate of unemployed workers}}$$

Figure 17 displays the job-to-job transition rate of workers by wage centile. Consistent with the aforementioned models, job-to-job transition rates decline steeply with the worker's wage rank. In the first percentile, the quarterly job-to-job transition rate is precisely estimated to be 0.147 or 0.162, depending on whether job-to-job transitions involving only wage increases are used. Dividing through by the job finding rate of unemployed workers leads to an estimate of ξ of 0.92. This relatively large number is consistent with recent micro-level evidence from Faberman et al. (2017) for the U.S. The results in Figure 6 are also robust to using a more conventional value of 0.3 for ξ.

Figure 17: Job-to-job mobility rate by wage centile. France.



Note: Quarterly job-to-job transition rate of employed workers in France by wage centile. Blue circles: using all job-to-job transitions. Orange squares: using only job-to-job transitions that involve a wage increase.

F Efficiency

F.1 Planning solution

F.1.1 Optimality conditions

The planner 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 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. Only one particular set of weights implements an allocation that resembles the decentralized equilibrium, which will be the focus of this paper.[54] The planner's objective is then

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

where $\mathcal{E}$ denotes the mass distribution of employment across productivity y in location ℓ at time t. The last term

[54] 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.

is the welfare of the owners. The planner is subject to the constraints

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

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—e.g. 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.[55] 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 spaces:

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

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

Finally, the approach I use builds on duality methods similar to Moll and Nuño (2018). 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.

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

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

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

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 $\pi(t,\ell) = -\int_{\underline{\ell}}^{\overline{\ell}} A(t,x)dx$. I have an 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

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

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 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

$$
\frac{\lambda(\ell)^\omega}{\omega} = \mathcal{U}(\ell)b\ell + \int \mathcal{E}(y,\ell)\Big[(1-\alpha)(b+v(\ell))\ell + \alpha y\ell\Big]dy, \tag{56}
$$

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 $\lambda u_U = c_U + rh_U$, 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

$$\begin{aligned} [\pi + \tau]\xi \quad &= \quad (1 - \alpha)nf_\ell \int gSdz \\ 0 \quad &= \quad f_\ell n \cdot \left(\int \frac{\partial_\zeta g_0}{g_0} g_0 S dy \right) + \frac{\pi'(\ell)}{\pi(\ell) + \tau} \cdot [\pi(\ell) + \tau]f_\zeta(\zeta). \end{aligned}$$

Denote $J(\ell) \equiv \tau + \pi(\ell)$ and so simplifying out f_ℓ

$$n \left(\int \frac{\partial_\zeta g_0}{g_0} \cdot g_0 S dy \right) + \frac{J'(\ell)}{J(\ell)} \cdot \frac{f_\zeta(\zeta)}{\xi} \cdot (1 - \alpha)n \int g_0 S dy \quad = \quad 0,$$

and hence

$$(1 - \alpha)\frac{J'}{J\zeta'} \quad = \quad \frac{\int \frac{\partial_\zeta g_0}{g_0} \cdot g_0 S dy}{\int g_0 S dy}.$$

Using the known solution to S, one obtains

$$\frac{\int \frac{\partial_\zeta g_0}{g_0} \cdot g_0 S dy}{\int g_0 S dy} = \frac{\bar{S}'(\zeta)}{\bar{S}(\zeta)} + \log \frac{B_0}{b + v(\ell)}.$$

Finally, changing variables to $J(\ell) \equiv J(\zeta(\ell))$:

$$(1-\alpha)\frac{J'(\zeta)}{J(\zeta)} \quad = \quad \frac{\bar{\mathcal{S}}'(\zeta)}{\bar{\mathcal{S}}(\zeta)} + \log\frac{B_0}{b+v(\zeta)}.$$

This equation corresponds to an envelope condition of the decentralized equilibrium which coincides with the FOC 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:

$$-\frac{v'(\ell)}{v(\ell)}\left[\alpha + \frac{v(\ell)}{b+v(\ell)}(\zeta(\ell)-1)\right] + \frac{1-\alpha}{\ell} + 0 = 0. \tag{57}$$

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:

$$c_e \quad = \quad (1-\alpha)\int d\ell\; q(\ell)f_\zeta(\zeta(\ell)|\zeta'(\ell)|\ell\left(\frac{B_0}{b+v(\ell)}\right)^{\zeta(\ell)}(b+v(\ell))\bar{\mathcal{S}}(\zeta(\ell)),$$

which correponds to the free-entry condition when $\beta = \alpha$.

F.1.2 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 in the decentralized equilibrium (38) and in the planning solution (57), the labor market pooling externality immediately arises. Except for this discrepancy, inspecting 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 in the decentralized equilibrium (38) relative to the planning solution (57), 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), $\frac{\ell v'(\ell)}{v(\underline{\ell})}$ is larger in the decentralized equilibrium due to the labor market pooling externality term. Given that $\frac{\ell w'(\ell)}{\underline{w}(\ell)} \approx \frac{v}{v+\underline{v}}\frac{\ell v'(\ell)}{v(\underline{\ell})}$ 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 common 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 $I_0 = 0$ 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 $\rho V(w,\ell) = \frac{w}{r(\ell)^\omega} + L_w V$. Then, $V-U$ solves $\rho(V(w,\ell)-U(\ell)) = \frac{w-b\ell-V(\ell)}{r(\ell)^\omega} + L_w(V-U)$, 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. Finally, define the adjusted surplus, which satisfies $\rho J(\ell,y) = y\ell - [V(\ell)+b\ell] + L_y S$ with boundary conditions identical to the random search case. Thus, Lemma 1 applies. The value of employer $\zeta = 1/z$ in location ℓ is then $J(\zeta,\ell) = \max_{\theta\in\Theta_\ell}\{q(\theta)\mathbb{E}_{\zeta,\ell}[S(\ell,z) - s(\theta,\ell)]\}$. Substituting the definition of V to express tightness as a function of the surplus s,

$$J(\zeta,\ell) \quad = \quad V(\ell)^{-\frac{\alpha}{1-\alpha}} m^{\frac{1}{1-\alpha}} \max_{\hat{s}=s^{\frac{1}{1-\alpha}}} \left\{\mathbb{E}_{\zeta,\ell}[S(\ell,z)]\cdot \hat{s}^{\alpha} - \hat{s}\right\}.$$

This maximization results in $s(\theta(\zeta,\ell),\ell) = \alpha\mathbb{E}_{\zeta,\ell}[S(\ell,z)]$ and $\theta(\zeta,\ell)^{1-\alpha} = \frac{V(\ell)}{\alpha m \mathbb{E}_{\zeta,\ell}[S(\ell,z)]}$. Therefore,

$$J(\zeta,\ell) = \left\{\frac{(1-\alpha)^{1-\alpha}}{\alpha^\alpha} m\mathbb{E}_{\zeta,\ell}[S(\ell,z)]V(\ell)^{-\alpha}\right\}^{\frac{1}{1-\alpha}}.$$

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.

F.2 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 $\tau_n = \frac{\tau_e}{\tau_b\tau_w}$

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

$$J_0(y,\ell) \equiv (1-\beta)\tau_\pi(\tau_{va}\cdot y - \tau_e - \tau_w\tau_b b - \tau_w v(\ell)\ell = \tau_{va}\tau_\pi(1-\beta)\left(z - \frac{\tau_e}{\tau_{va}} - \frac{\tau_w\tau_b}{\tau_{va}}\cdot b - \frac{\tau_w}{\tau_{va}}v(\ell)\right)\ell,$$

and for workers

$$V_0 \equiv \beta\left(\frac{\tau_{va}z-\tau_e}{\tau_w}\cdot z - b\tau_b - \mathbb{E}[V-U]\right)\ell = \frac{\tau_{va}}{\tau_w}\cdot\beta\left(z - \frac{\tau_e}{\tau_{va}} - b\cdot\frac{\tau_w\tau_b}{\tau_{va}} - \frac{\tau_w}{\tau_{va}}\mathbb{E}[V-U])\right)\ell.$$

The endogenous separation cutoff is then

$$\underline{y} \propto \frac{\tau_e}{\tau_{va}} + \frac{\tau_w\tau_b}{\tau_{va}}\cdot b + \frac{\tau_w}{\tau_{va}}v = \frac{\tau_w}{\tau_{va}}\cdot\left(\frac{\tau_e}{\tau_w} + \tau_b b + v\right) = \frac{\tau_w\tau_b}{\tau_{va}}\cdot\left(\tau_n + b + \tilde{v}\right),$$

where $\tilde{v} = \tau_b v$. Finally, solving the worker's problem, one obtains

$$v(\ell) = \beta f(\ell)\frac{\tau_{va}}{\tau_w}\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)}\underline{y}(\ell)\bar{S}(\zeta(\ell)).$$

Therefore,

$$\tilde{v}(\ell) = \frac{\tau_{va}}{\tau_w\tau_b}\beta\cdot f(\ell)\cdot\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)}\underline{y}(\ell)\bar{S}(\zeta(\ell)).$$

Denoting $T = \frac{\tau_{va}}{\tau_w\tau_b}$ one obtains $\tilde{v} = c_2T\underline{y} - b - \tau_n$ for a constant $c_2 > 0$. One can then 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

$$q(\ell) \propto \left(c_2\underline{y} - \frac{b+\tau_n}{T}\right)^{-\frac{\alpha}{1-\alpha}}\cdot\beta^{\frac{\alpha}{1-\alpha}}\cdot\left[\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)}\underline{y}(\ell)\bar{S}(\zeta(\ell))\right]^{\frac{\alpha}{1-\alpha}}.$$

Finally, employers' expected value is

$$\begin{aligned} J(\ell,\zeta)^{1-\alpha} &= \left(\frac{1-\beta}{1-\alpha}\right)^{1-\alpha}\left(\frac{\beta}{\alpha}\right)^{\alpha}(\tau_\pi(\ell)\tau_{va}(\ell))^{1-\alpha}\left[\frac{c_2\underline{y}(\ell)-b}{c_2\underline{y}(\ell)-\frac{b+\tau_n(\ell)}{T(\ell)}}\cdot\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)-\zeta}\frac{\cdot\bar{S}(\zeta(\ell))}{\bar{S}(\zeta)}\right]^{\alpha} \\ &\quad\times J_{\mathrm{SP}}(\zeta,\ell,\underline{y}(\ell))^{1-\alpha}, \end{aligned}$$

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 $T = 1$. Then there exists a τ_n that equated the separation cutoff for the planner and the decentralized equilibrium: $\frac{1}{\beta}\cdot\frac{c_2\underline{y}-b-\tau_n}{\underline{y}} = \frac{1}{\alpha}\cdot\frac{c_2\underline{y}-b}{\underline{y}}$, and so

$$\tau_n(\ell) = \frac{\alpha-\beta}{\alpha}v^{SP}(\ell).$$

Substituting back into the employer's problem,

$$J(\ell,\zeta)^{1-\alpha} = \left(\frac{1-\beta}{1-\alpha}\right)^{1-\alpha}\tau_\pi^{1-\alpha}(\ell)\left[\left(\frac{B_0}{\underline{y}(\ell)}\right)^{\zeta(\ell)-\zeta}\frac{\cdot\bar{\mathcal{S}}(\zeta(\ell))}{\bar{\mathcal{S}}(\zeta)}\right]^{\alpha}\times J_{\mathrm{SP}}(\zeta,\ell,\underline{y}(\ell))^{1-\alpha}.$$

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.

G Quantitative model

G.1 Values

The bargaining solution from Lemma 4 readily extends to the extended model. Denote $U(p,a,k)$ the value of unemployment in location (p,a) for a worker with human capital k. With Frechet taste shocks and denoting $\nu = 1/\varepsilon$, the continuation value from migration and migration shares are

$$M(k) = \left(\int U(p,a,k)^\nu F_{p,a}(dp,da)\right)^{\frac{1}{\nu}} \quad ; \quad \pi(\ell,a,k) = \frac{U(\ell,a,k)^\nu}{M(k)^\nu}.$$

Guess that the value of unemployment scales with k. Then the value of unemployment solves the recursion

$$(\rho+\Delta+\mu)U(p,a,k) = apr(p,a)^{-(\omega+\psi)}\cdot U_1(p,a)k + (\lambda-\varphi)kU_k + \mu M_0 k$$

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 as 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 denoting $\tilde\rho = \rho+\Delta+\mu-\lambda$, satisfies

$$\tilde\rho S(y,k,p,a) = k\cdot pr(p,a)^{-\psi}U_1(p,a)\cdot\mathcal{S}\left(\frac{y}{\underline{y}(p,a)}\right) \quad ; \quad \tilde\rho\frac{\underline{y}(p,a)}{\underline{y}_0} \quad = \quad U_1(p,a),$$

where $\underline{y}_0$ is calculated using $\tilde\rho$ as the effective discount rate. Because workers' outside option scales with k under the guess, the separation decision is independent from k. Going back to the value of unemployment,

$$\tilde\rho U(p,a,k) = \frac{(b+v(p,a))ap}{r(p,a)^{\omega+\psi}}k + \mu M_0 k - \varphi k U_k \quad ; \quad \tilde\rho v(p,a) \quad = \quad \beta f(p,a)U_1(p,a)\left(\frac{Y}{\underline{y}(p,a)}\right)^{\zeta^*(p,a)}\bar{S}(\zeta^*(p,a)).$$

Hence, the guess is verified. I may then define U_0 such that $U(p,a,k) = U_0(p,a)k$. Then

$$(\tilde\rho+\varphi)U_0(p,a) \quad = \quad \frac{(b+v(p,a))ap}{r(p,a)^{\omega+\psi}} + \mu M_0,$$

which becomes $\tilde\rho U_0(p,a) = \frac{\tilde\rho(b+v(\ell,a))}{\tilde\rho+\varphi}\cdot\frac{a\ell}{r(\ell)^{\omega+\psi}} + \mu M_0 - \frac{\varphi}{\tilde\rho+\varphi}\cdot\mu M_0$, 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,

$$\tilde\rho S(y,k,p,a) \approx k\cdot pr(p,a)^{-\psi}\left[\frac{\tilde\rho}{\tilde\rho+\varphi}(b+v(p,a))\right]\cdot\mathcal{S}\left(\frac{y}{\underline{y}(p,a)}\right) \quad ; \quad \tilde\rho\frac{\underline{y}(p,a)}{\underline{y}_0} \approx \frac{\tilde\rho}{\tilde\rho+\varphi}(b+v(p,a)).$$

To a first order approximation in μ, the previous results imply, with $\hat{\rho} = \tilde{\rho} + \varphi$,

$$\hat{\rho}S(y,k,p,a) = k \cdot pr(p,a)^{-\psi}(b+v(p,a)) \cdot \mathcal{S}\left(\frac{y}{\underline{y}(p,a)}\right) \quad ; \quad \hat{\rho}\frac{\underline{y}(p,a)}{\underline{y}_0} = (b+v(p,a))$$
$$\hat{\rho}U(p,a,k) = \frac{(b+v(p,a))ap}{r(p,a)^{\omega+\psi}}k \equiv \hat{\rho}U_0(p,a)k \quad ; \quad \pi(p,a,k) = \left(\frac{U_0(p,a)}{M_0}\right)^{\nu}$$
$$M_0 = \left(\int U_0(p,a)^{\nu} F_{p,a}(dp,da)\right)^{\frac{1}{\nu}}. \tag{58}$$

G.2 Human capital across locations

I now characterize the human capital distribution in each location. For now, focus on a single location and omit (p,a) subscripts to facilitate exposition. The probability mass functions of rescaled log human capital $h = \log k - \lambda t$ for employed and unemployed workers in a location solve:

$$\begin{aligned} 0 &= -sg_E(h) + f_R g_U(h) - \mu g_E(h) - \Delta g_E(h) \\ 0 &= \varphi g_U'(h) - f_R g_U(h) + sg_E(h) - \mu g_U(h) + K(h) - \Delta g_U(h), \end{aligned}$$

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 feature 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

$$0 = \varphi g_U'(h) - \underbrace{(\mu+\Delta)\frac{\mu+\Delta+s+f_R}{\mu+\Delta+s}}_{\equiv C_0} g_U(h) + K(h).$$

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

$$0 = [e^h g_U(h)]_{-\infty}^{\infty} - \varphi \int_{\mathbb{R}} e^h g_U - C_0 \int_{\mathbb{R}} e^h g_U + \int_{\mathbb{R}} e^h K.$$

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 $(\mu+\Delta)\frac{\mu+\Delta+s+f_R}{\mu+\Delta+s}\mathcal{U} = \mathcal{K}$, where $\mathcal{K}$ is the total mass of entrants. Hence, the unemployment rate is $u = \frac{\mathcal{U}}{\mathcal{U}+\mathcal{E}} = \frac{\mu+\Delta+s}{\mu+\Delta+s+f_R} = \frac{\mu+\Delta}{C}$. The mass of unemployed is $\mathcal{U} = u \cdot \frac{\mathcal{K}}{\mu+\Delta}$. Population is $\mathcal{E} + \mathcal{U} = \frac{\mathcal{K}}{\mu+\Delta}$. By definition, $\int_{\mathbb{R}} g_U = \mathcal{U}$.

Average human capital in a location is

$$\bar{k}(\ell,a) = \mathbb{E}[e^h|p,a] = \frac{k_0}{\mathcal{U}\cdot(\varphi+C)} = (\mu+\Delta)\frac{\mathbb{E}^K[e^h]}{u\cdot(\varphi+C)} = \frac{\mu+\Delta}{\mu+\Delta+u(p,a)\varphi}\cdot\mathbb{E}^K[e^h],$$

where $\mathbb{E}^K[e^h]$ is the expected human capital of new entrants. By definition, the mass of entrants at rescaled 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. Hence $\mathbb{E}^K[e^h] = \frac{\mu\pi}{\mu\pi+\Delta L}\mathbb{E}^I[e^h] + \frac{\Delta L}{\mu\pi+\Delta L}\mathbb{E}^E[e^h]$. In steady-state, population density is equal to migration shares: $L(p,a) = \pi(p,a)$ Therefore,

$$\mathbb{E}^K[e^h] = \frac{\mu}{\mu+\Delta}\mathbb{E}^I[e^h] + \frac{\Delta}{\mu+\Delta}\mathbb{E}^E[e^h] \equiv x_0\mathbb{E}^I + (1-x_0)\mathbb{E}^E$$

with $x_0 = \frac{\mu}{\mu+\Delta}$. Now, $\mathbb{E}^I[e^h] = \int L(p,a)F_{p,a}(dp,da)\cdot\bar{k}(p,a)$, so that one obtains a linear system in $\bar{k}(p,a)$ across locations:

$$\bar{k}(p,a) = \frac{z_0}{z_0+\varphi u(p,a)}\left[x_0\int L(p',a')\bar{k}(p',a')F_{p,a}(dp',da') + (1-x_0)\mathbb{E}^E\right],$$

where $z_0 = \mu + \Delta$. Denote $X(p,a) = (1+\varphi_0 u(p,a))\bar{k}(p,a)$, where $\varphi_0 = \varphi/z_0$. Re-write the linear system as

$$X(p,a) = (1-x_0)\mathbb{E}^E + x_0 \int \frac{L(p',a')}{1+\varphi_0 u(p',a')}\bar{k}(p',a')F_{p,a}(dp',da').$$

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

$$\int Z(p,a)F_{p,a}(dp,da) = \frac{\int Z(p,a)F_{p,a}(dp,da)}{1-\int Z(p,a)F_{p,a}(dp,da)} \cdot (1-x_0)\mathbb{E}^E \implies Z(p,a) = \frac{(1-x_0)\mathbb{E}^E}{1-\int Z(p',a')F_{p,a}(dp',da')},$$

which finally implies

$$\bar{k}(p,a) = \frac{1}{1+\varphi_0 u(p,a)} \cdot \frac{(1-x_0)\mathbb{E}^E}{1-x_0\int \frac{L(p',a')}{1+\varphi_0 u(p',a')} \cdot F_{p,a}(dp',da')}.$$

G.3 Labor market flows

Given (58), 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 $u(p,a) = \frac{\delta/z(p,a)+\mu+\Delta}{\delta/z(p,a)+\mu+\Delta+f_R(p,a)}$.

G.4 Migration shares with a continuum of locations

This section briefly outlines how to extend standard discrete choice results with continuum of locations. For simplicity, I simply describe the case of a static discrete choice problem. Start from a discrete number of locations $i \in \{1,...,N\}$. Suppose an agent solves

$$\max_{i=1...N} u_i \varepsilon_i, \tag{59}$$

where ε_i follows a Frechet distirbution with shifter $T_i(N)$ and shape parameter ν. The shifter $T_i(N)$ may depend on N. Standard results then imply that the probability that the agent chooses i is

$$\pi_i(N) = \frac{T_i(N)u_i^\nu}{\sum_j T_j(N)u_j^\nu}, \tag{60}$$

and indirect utility is

$$V(N) = \Gamma(1-1/\nu)\left[\sum_{i=1}^N T_i(N)u_i^\nu\right]^{\frac{1}{\nu}}. \tag{61}$$

Now suppose that locations $i \in \{1,...,N\}$ lie within an interval. Without loss of generality, assume that this interval is $[0,1]$. Consider the limit in which $N \to +\infty$. For exposition, suppose that all locations $i \in \{1,...,N\}$ are equally spaced and remain in $[0,1]$ as $N \to +\infty$. Suppose also that $u_i \equiv u(i)$ and $T_i(N) \equiv T(i,N)$, where u, T are continuous functions. Postulate $T(i,N) \equiv \frac{1}{N}T(i)$. Then (60) becomes $\pi_i(N) = \frac{T(i)u(i)^\nu/N}{\frac{1}{N}\sum_{j=1}^N T(j)u(j)^\nu}$. The denominator is a standard Riemann integral approximation, and so $\frac{1}{N}\sum_{j=1}^N T(j)u(j)^\nu \to_{N+\infty} \int_0^1 T(j)u(j)^\nu dj$. Hence, to first order

$$\pi_i(N) \sim \underbrace{\frac{T(i)u(i)^\nu}{\int_0^1 T(j)u(j)^\nu dj}}_{\equiv \pi(i)} \times \frac{1}{N}. \tag{62}$$

$\pi(i)$ defines a probability density function. Just as with any continuous random variable, in the limit $N \to \infty$, the probability that any location i is chosen $\pi_i(N)$ converges to 0. However, choices admit a smooth probability density function $\pi(i)$. Similarly, indirect utility (61) becomes

$$V(N) = \Gamma(1-1/\nu)\left[\frac{1}{N}\sum_{i=1}^{N} T(i)u(i)^{\nu}\right]^{\frac{1}{\nu}} \to_{N\to\infty} \Gamma(1-1/\nu)\left[\int_0^1 T(i)u(i)^{\nu}di\right]^{\frac{1}{\nu}} \equiv V.$$

The limit of the maximization problem (59) thus always delivers a finite indirect utility V. Agents choose among an infinite number of options. Even though the shocks ε have unbounded the support, the rescaling $T_i(N) = T(i)/N$ ensures that any individual shock has a vanishing mean, keeping the maximization problem well-behaved in the limit. An alternative interpretation is that agents choose an interval $[i, i+1/N)$ that shrinks as $N \to +\infty$, and their utility function scales with the width of the interval.

G.5 Population, housing prices and composite index

Having solved for average human capital $\bar{k}(p,a)$ in each location, it is possible to characterize housing prices and thus the value of employers. From Supplemental Material G.4, total population in a location is given by migration shares:

$$L(p,a) = \left(\frac{U_0(p,a)}{M_0}\right)^{\nu} \tag{63}$$

where $\nu = 1/\varepsilon$. 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

$$\begin{aligned} H_0 r(p,a)^{\eta} \quad &= \quad p\bar{k}(p,a)r(p,a)^{-\psi}\cdot L(p,a)\cdot\frac{1}{r(p,a)}\Big[\omega u(p,a)b \\ &\qquad +\omega(1-u(p,a))(b+v(p,a))(1-\beta+\beta\mathbb{E}_{p,a}[y|y>\underline{y}(p,a)]) \\ &\qquad +\psi(1-u(p,a))(b+v(p,a))\mathbb{E}_{p,a}[y|y>\underline{y}(p,a)])\Big] \\ &\equiv \quad p\bar{k}(p,a)r(p,a)^{-\psi-1}L(p,a)G(\zeta(p,a),v(p,a)), \end{aligned} \tag{64}$$

where η is the housing supply elasticity. 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 $H_0, \mathbb{E}^E$ to one without loss of generality. After substituting equation (64) into the migration share equation (58), obtain

$$L(p,a) \quad = \quad M_0^{-\frac{1}{\varepsilon+\frac{\omega+\psi}{1+\eta+\psi}}}\cdot\left[\frac{a\cdot(p\mathbb{E}(p,a))^{\frac{1+\eta-\omega}{1+\eta+\psi}}\cdot(b+v(p,a))}{G(v(p,a),\zeta(p,a))^{\frac{\omega+\psi}{1+\eta+\psi}}}\right]^{\frac{1}{\varepsilon+\frac{\omega+\psi}{1+\eta+\psi}}}. \tag{65}$$

After substituting equation (65) back into (64), obtain

$$r(p,a) = M_0^{-\frac{1}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot\left\{a\cdot(p\bar{k}(p,a))^{1+\varepsilon}\cdot(b+v(p,a))\cdot G(v(p,a),\zeta(p,a))^{\varepsilon}\right\}^{\frac{1}{(1+\eta+\psi)\varepsilon+\omega+\psi}}.$$

Therefore the adjustment factor for the expected adjusted surplus in (58) is

$$p\bar{k}(p,a)r(p,a)^{-\psi} = M_0^{\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot\frac{\left((p\bar{k}(p,a))^{\omega+\varepsilon(1+\eta)}a^{-\psi}\right)^{\frac{1}{(1+\eta+\psi)\varepsilon+\omega+\psi}}}{\left[(b+v(p,a))G(v(p,a),\zeta(p,a))^{\varepsilon}\right]^{\frac{\psi}{(1+\eta+\psi)\varepsilon+\omega+\psi}}}. \tag{66}$$

This equation motivates the definition of the composite index

$$\ell(p,a) = \left(p^{\omega+\varepsilon(1+\eta)}a^{-\psi}\right)^{\frac{1}{(1+\eta+\psi)\varepsilon+\omega+\psi}}. \tag{67}$$

G.6 Location choice of employers

Using the adjusted surplus and (66), the value of opening a job in location (p, a) for employer $\zeta = 1/z$ is

$$J(\zeta, p, a)^{\frac{1}{1+\gamma}} \quad \propto \quad \bar{k}(p,a)^{\mathcal{Q}} \cdot \ell(p,a)(b + v(p,a))^{1-\psi\mathcal{P}} \cdot G(v(p,a), \zeta^*(p,a))^{\varepsilon\psi\mathcal{P}} q(p,a) \left(\frac{B}{b + v(p,a)}\right)^{\zeta} \bar{\mathcal{S}}(\zeta), \quad (68)$$

where the optimal vacancy posting decision has been maximized out. Re-arranging equation (68) delivers equation (19). Using the worker's value of search to substitute out $q(p, a)$ delivers the employer's location problem

$$\begin{aligned} \max_{p,a} \quad & \bar{k}(u(v(p,a), \zeta^*(p,a))^{(1-\alpha)\mathcal{Q}} \cdot \ell(p,a)^{1-\alpha}(b + v(p,a))^{1-(1-\alpha)\psi\mathcal{P}} v(p,a)^{-\alpha} \cdot G(v(p,a), \zeta^*(p,a))^{(1-\alpha)\varepsilon\psi\mathcal{P}} \\ & \cdot \left(\frac{B}{b + v(p,a)}\right)^{(1-\alpha)\zeta + a\zeta^*(p,a)} \bar{\mathcal{S}}(\zeta)^{1-\alpha} \bar{\mathcal{S}}(\zeta^*(p,a))^{\alpha}, \end{aligned} \quad (69)$$

where $\bar{k}(u(v(p,a), \zeta(p,a)) \equiv \bar{k}(p,a)$ but where the dependence on the local unemployment rate has been made explicit.

G.7 Single index property

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 $i \in \{p, a\}$ $\partial_i v = \bar{C}(v, \zeta, \ell)\partial_i \ell$ for some function $\bar{C}$. 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 $\ell(p, a)$.

Why is 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 (19), 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 (19), 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 18, 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.[56]

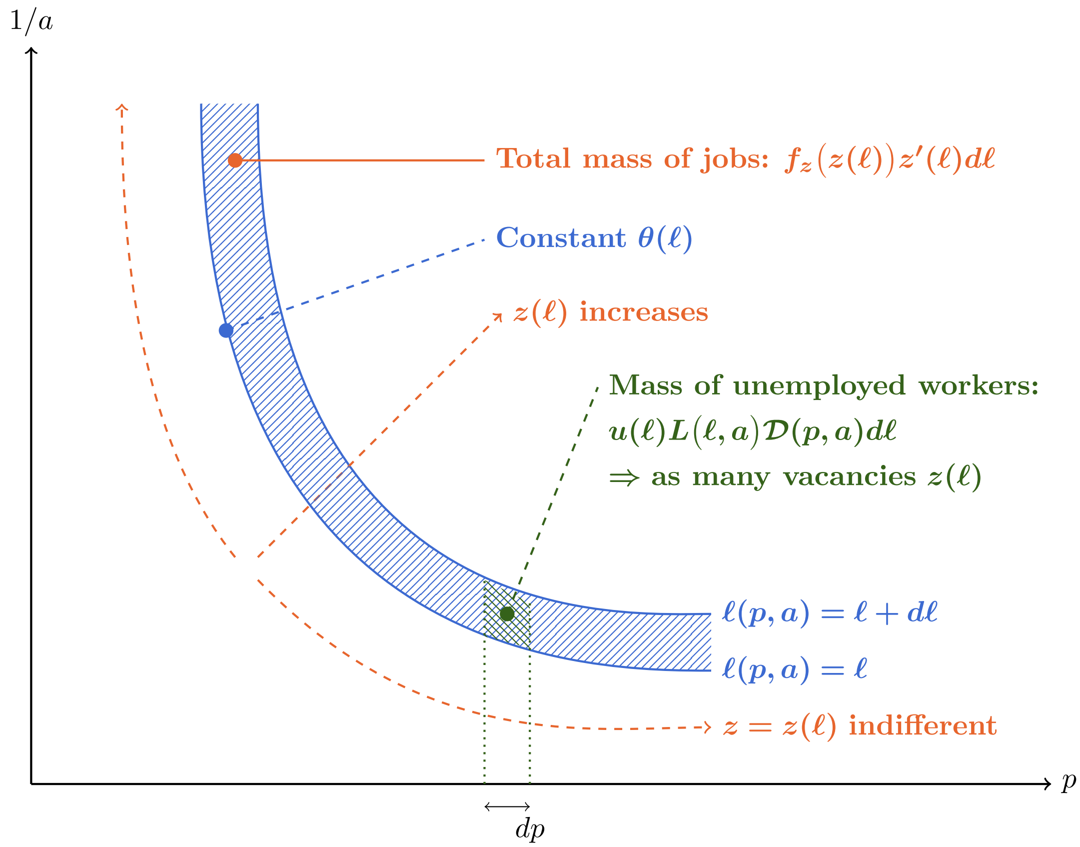
G.8 First-order condition for employers

With this observation at hand, the structure of equation (69) 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

$$\begin{aligned} & \frac{v'(\ell)}{v(\ell)} \left\{ \alpha + \frac{v(\ell)}{b + v(\ell)} \left(\zeta(\ell) - 1 - \frac{(1-\alpha)\psi}{\omega + \psi + (1+\psi)\varepsilon} \left[1 + \varepsilon \frac{(B + v(x))G_v}{G} \right] \right) \right\} \\ = \quad & \frac{1-\alpha}{\ell} + \alpha \left(\frac{\bar{\mathcal{S}}'(\zeta(\ell))}{\bar{\mathcal{S}}(\zeta(\ell))} + \log \frac{B}{b + v(\ell))} \right) \zeta'(\ell) + \varepsilon \cdot \frac{(1-\alpha)\psi}{\omega + \psi + (1+\psi)} \frac{G_\zeta}{G} \zeta'(\ell) \\ & + \frac{(1-\alpha)(\omega + \varepsilon)}{\omega + \psi + \varepsilon(1+\psi)} \frac{d}{d\ell} \left(\log \frac{D}{D + \varphi u(\ell)} \right), \end{aligned}$$

[56] $\mathcal{D}(p, a)$ in Figure 18 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 18: The location choice of employers and workers in two-dimensional space.



where

$$\frac{d}{d\ell}\left(\log\frac{D}{D+\varphi u(\ell)}\right) \;=\; \frac{\frac{\varphi u(\ell)}{D+\varphi u(\ell)}}{D+\delta\zeta(\ell)}\left\{\frac{u(\ell)f_R(\ell)}{b+v(\ell)}\frac{v'(\ell)}{v(\ell)}-\left[\delta(1-u(\ell))+f_R(\ell)\frac{\bar{S}'(\zeta(\ell))}{\bar{S}(\ell)}\right]\zeta'(\ell)\right\},$$

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:

$$\theta(\ell)=-\frac{M_e f_\zeta(\zeta(\ell)\mathcal{V}(\ell,\zeta(\ell))\zeta'(\ell)}{u(\ell)\mathcal{L}(\ell)f_\ell(\ell)}\quad;\quad\mathcal{V}(\ell,\zeta(\ell))\propto J(\zeta^*(\ell),\ell)^\gamma. \tag{70}$$

The maximized value of employers is

$$\begin{aligned}J(\zeta^*(\ell),\ell)^{\frac{1}{1+\gamma}} \;=\; & \bar{k}(u(\ell))^{\frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot\ell\\ & (b+v(\ell))^{\frac{1}{1-\alpha}-\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot v(\ell)^{-\frac{\alpha}{1-\alpha}}G(v(\ell),\zeta(\ell))^{\frac{\varepsilon\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\\ & \left(\frac{B}{b+v(\ell)}\right)^{\frac{\zeta(\ell)}{1-\alpha}}\bar{S}(\zeta(\ell))^{\frac{1}{1-\alpha}}.\end{aligned} \tag{71}$$

G.9 Labor market clearing and population determination

After substituting equation (67) back into (65),

$$L(p,a) \quad \propto \quad \frac{\ell(p,a)^{\frac{1+\eta-\omega}{\omega+\varepsilon(1+\eta)}} \cdot (b+v(p,a))^{\frac{1+\eta+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}}{G(v(p,a),\zeta(p,a))^{\frac{\omega+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}} \cdot a^{\frac{1}{\omega+\varepsilon(1+\eta)}}.$$

Then average population density in locations with index ℓ, $\mathcal{L}(\ell)$, is given by

$$\mathcal{L}(\ell) \propto \frac{\ell^{\frac{1+\eta-\omega}{\omega+\varepsilon(1+\eta)}} \cdot (b+v(\ell))^{\frac{1+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}}{G(v(\ell),\zeta(\ell))^{\frac{\omega+\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}} \cdot \mathcal{C}(\ell) \quad ; \quad \mathcal{C}(\ell) = \mathbb{E}\Big[a^{\frac{1}{\omega+\varepsilon(1+\eta)}} \Big| \ell(p,a) = \ell\Big].$$

By construction, $(\omega+\psi+\varepsilon(1+\eta+\psi))\log\ell = (\omega+\varepsilon(1+\eta))\log p - \psi\log a$. 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, $\log a \,|\, \log\ell = \frac{1}{\sigma_\ell^2}\cdot\frac{(\omega+\varepsilon(1+\eta))\rho_{ap}\sigma_p\sigma_a - \psi\sigma_a^2}{\omega+\psi+\varepsilon(1+\eta+\psi)}\log\ell + N$, where $N \sim \mathcal{N}(0, \sigma_a^2(1-\rho_{a,\ell}^2)$ is independent from $\log\ell$.[57] Therefore, the correction factor is

$$\mathcal{C}(\ell) = \mathcal{C}_0 \exp\left(\frac{\sigma_a^2}{\sigma_\ell^2} \cdot \frac{\rho_{ap}\frac{\sigma_p}{\sigma_a} - \frac{\psi}{\omega+\varepsilon(1+\eta)}}{\omega+\psi+\varepsilon(1+\eta+\psi)} \cdot \log\ell\right).$$

G.10 Efficiency

All but one of 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 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.

G.11 Welfare

Local welfare gains. To a first order approximation in μ, workers who never receive the moving opportunity face indirect utility in location ℓ equal to

$$W^u(\ell) \equiv \mathbb{E}\left[\int U(p,a,k)\frac{L(p,a)}{\mathcal{L}(\ell)}F_{p,a}(dp,da)\Big|\ell\right] \quad = \quad A_0 N_0 U_2(\ell)\mathcal{D}_0(\ell)\bar{k}(\ell),$$

where A_0 is a transformation of parameters, $\mathcal{R} \equiv \frac{1}{\omega+\varepsilon(1+\eta)}$, $N_0 = M_0^{(\omega+\psi)\mathcal{P}}$, and

$$U_2(\ell) = \ell^{\varepsilon(1+\eta-\omega)\mathcal{R}}(b+v(\ell))^{\varepsilon(1+\eta+\psi)\mathcal{P}}G(v(\ell),\zeta(\ell))^{-\varepsilon(\omega+\psi)\mathcal{P}} \qquad ; \qquad \mathcal{D}_0(\ell) = \mathbb{E}[a^{\varepsilon\mathcal{R}(1+\eta\psi\mathcal{P})}|\ell].$$

[57] Note that the correlation is $\rho_{a\ell} = \frac{(\omega+\varepsilon)\rho_{ap}\sigma_p - \psi\sigma_a}{\omega+\psi+\varepsilon(1+\psi)}\Big/\sigma_\ell$

For the average worker starting in location ℓ who never receives the moving opportunity, indirect utility is then

$$W^w(\ell) \equiv W_0^u(\ell)\big(1 + \beta\bar{\mathcal{S}}(\zeta(\ell))(1 - u(\ell))\big).$$

The following decomposition of welfare for workers who never receive the moving opportunity follows:

$$W(\ell) = W^w(\ell) \times \bar{k}(\ell). \tag{72}$$

Aggregate gains. The indirect utility of unemployed workers is equalized across locations due to preference shocks. It is equal to M_0, which is given by

$$M_0 \quad = \quad \left\{ \int U_2(\ell)^{\nu} \ell^{\frac{1+\eta-\omega}{\omega+\varepsilon(1+\eta)}} \mathcal{D}_1(\ell) F_\ell(d\ell) \right\}^{\frac{1}{(1+\eta+\psi)\mathcal{P}}}.$$

where $\mathcal{D}_1(\ell) = \mathbb{E}[a^{\mathcal{R}(1+\eta\psi\mathcal{P})}|\ell]$. The indirect utility $\overline{W}$ of the average worker is

$$\begin{aligned} \overline{W} \quad &= \quad \int \Big\{ u(\ell)\mathbb{E}[U(p,a,y)|\ell] + (1-u(\ell))\mathbb{E}[V(p,a,y)|\ell] \Big\} \bar{k}(\ell)\mathcal{L}(\ell)F_\ell(d\ell) \\ &= \quad A_1 M_0 \int \big(1 + \beta\bar{\mathcal{S}}(\zeta(\ell))\big)\bar{k}(\ell)\mathcal{L}(\ell)F_\ell(d\ell) \end{aligned} \tag{73}$$

where $A_1 > 0$ is a transformation of parameters. (73) suggests the decomposition $\overline{W} = \overline{W}^u \times \overline{W}^e \times \overline{W}^k$, where

$$\overline{W}^u = AM_0 \ \ ; \ \ \overline{W}^e = \int \big(1 + \beta\bar{\mathcal{S}}(\zeta(\ell))\big)\mathcal{L}(\ell)F_\ell(d\ell) \ \ ; \ \ \overline{W}^k = \frac{\int \big(1 + \beta\bar{\mathcal{S}}(\zeta(\ell))\big)\bar{k}(\ell)\mathcal{L}(\ell)F_\ell(d\ell)}{\int \big(1 + \beta\bar{\mathcal{S}}(\zeta(\ell))\big)\mathcal{L}(\ell)F_\ell(d\ell)}. \tag{74}$$

H Estimation: Acceptance probability

I propose a microfoundation of job search that lets me use data on duration since last job offer to inform Y. In the LFS, there is data on duration since last meeting 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 meet with either the national unemployment agency, or the private sector with intensity S. Conditional on a meeting, it is a meeting with the ANPE with probability s, and a meeting with the private sector with probability $1 - s$. Conditional on meeting 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 meeting, 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 meetings 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 $D^U = \left\{\frac{1}{S(s\omega+(1-s)\tau)a} \cdot \frac{1}{dt} - 1\right\} \cdot dt \rightarrow_{dt \to 0} \frac{1}{S(s\omega+(1-s)\tau)a}$. 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 meetings. The probability that a currenly unemployed individual has last met with 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 meeting 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 $a = \frac{x}{Ss\omega}$.