Supplemental Material

For Online Publication Only

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 11 shows that the DADS panel and the LFS have similar aggregate statistics.

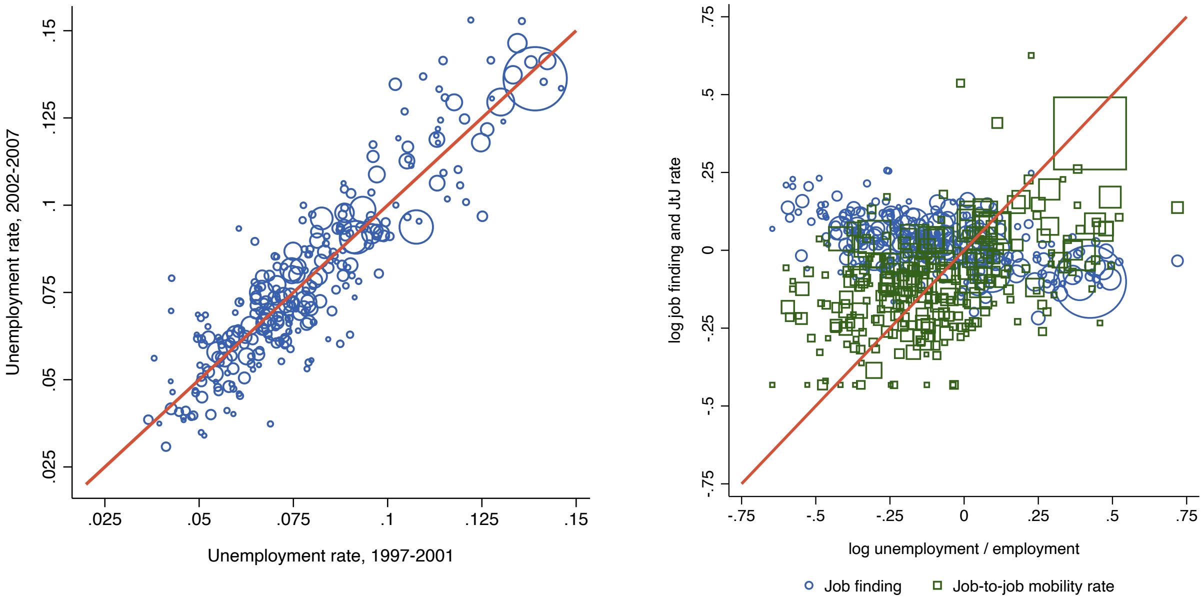
Table 11: Summary statistics

	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

Figure 17: Persistence of local unemployment, and job-to-job mobility rate. France.

(a) Persistence (b) Job-to-job rate



Note: Figure 17(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 17(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-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.

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.

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 17(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 17(a) plots $\hat{u}_{c,1}$ against $u_{c,0}$.

Figure 17(b) plots the log commuting zone-level job-to-job mobility rate, and the log job finding rate from unemployment, against the log unemployment-employment rate across French commuting zones.

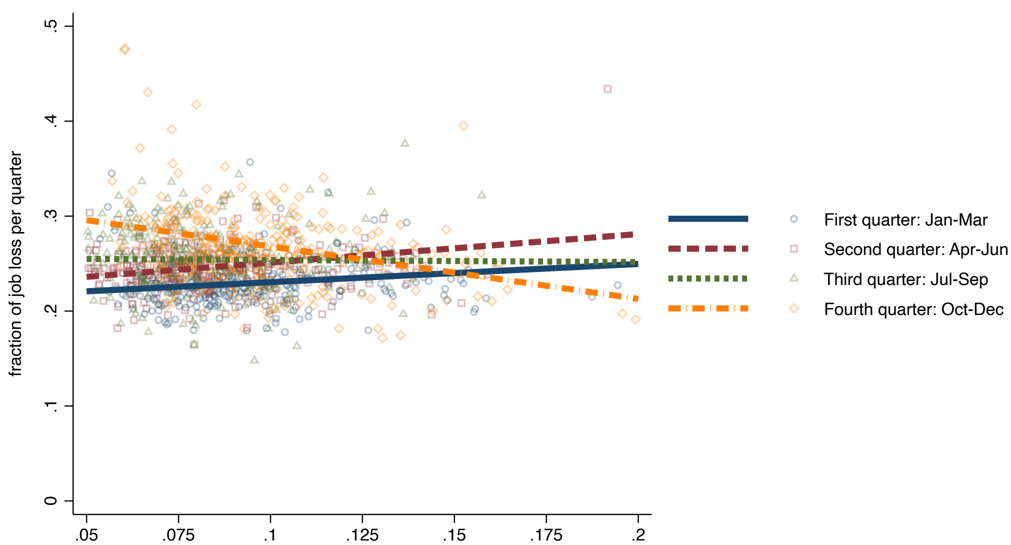
D.3 Mechanical correlates of job loss

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 18 scatterplots the fraction of job loss by quarter against the job losing rate, across French commuting zones. Figure 18 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 18: Fraction of job loss by quarter against commuting zone job losing rate, France.



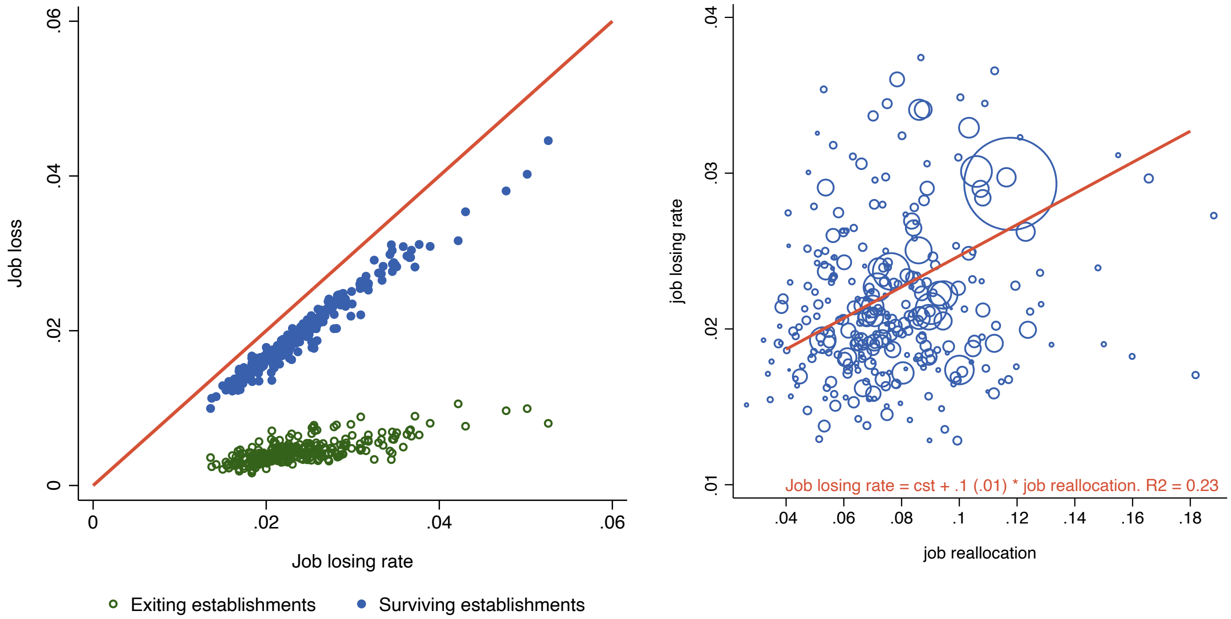
Note: Scatterplot of commuting zone-by-quarter fraction annual of job loss, by commuting zone unemployment rate. Economy-wide means of quarterly fraction of job loss adjusted using the LFS to limit reporting measurement error in the DADS.

Establishment exit and job reallocation. Figure 19(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 19(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 19: 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 19(a): job loss from exiting establishments and continuing establishments across French commuting zones, ordered by job losing rate. Figure 19(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 threshold productivity $\underline{y}(\ell)$ below which the match breaks up. Then, at that threshold, the surplus must be zero: $S(\underline{y}(\ell), \ell) = 0$. This condition is sometimes called the value-matching condition.

Second, because the threshold is chosen optimally, a first-order-condition with respect to the threshold 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 threshold $\underline{y}(\ell)$, which I summarize as

$$\begin{aligned} \rho S(y,\ell) &= \Big(y - \big(b + v(\ell)\big)\Big)\ell + (L_y S)(y,\ell) \quad , \quad \forall y \geq \underline{y}(\ell) \\ &\text{s.t.} \quad S(\underline{y}(\ell),\ell) = 0 \ , \ \frac{\partial S}{\partial y}(\underline{y}(\ell),\ell) = 0 \ , \ S(y,\ell) \leq \frac{y\ell}{\rho+\delta-\sigma^2/2}. \end{aligned} \tag{53}$$

Step 2: solving the PDE (53). To lighten notation, I drop location indices ℓ and solve without loss of generality

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

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

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

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 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 threshold $\underline{y}$ are thrown below the threshold $\underline{y}$ and into unemployment in an interval $(dt)^2$. Starting from $g(\underline{y})$ workers at the threshold, only a fraction $2^{-\lfloor \frac{1}{dt} \rfloor}$ of those workers remain there after a time dt. Taking $dt \to 0$, this fraction must be zero. I now make this intution precise.

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

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

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

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

Then:

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

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

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

So is left to calculate: $\int_0^\infty f(\underline{x}+zdx)\left\{\Phi\left((1-z)\lambda\right) - \Phi\left(-\lambda z\right)\right\}dz$. In integral form and changing variables: $\Phi\left((1-z)\lambda\right) - \Phi\left(-\lambda z\right) = \Phi(z\lambda) - \Phi(z\lambda - \lambda) = \int_0^\lambda \varphi(z\lambda - y)dy$, where φ here denotes the standard normal density function. Then, after some algebra

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

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

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

Then:

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

Similarly, $B \approx_{\lambda\to+\infty} A^0_{-\infty} + \frac{f(\underline{x})}{\lambda}\int_{-\infty}^0 \varphi(z)z\ dz + \mathcal{O}(\lambda^{-2})$, 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.4 Generalization of Lemma 2 and Proposition 1

Instead of Assumption 1, assume that new jobs are created with probability density function $\tilde{g}_0(y)$. The probability density function of successful new jobs is $g_0(y|\ell) \equiv \frac{\tilde{g}_0(y)}{1-\tilde{G}_0(\underline{y}(\ell))}$ for $y \geq \underline{y}(\ell)$. Define $\bar{S}(z,\ell) = \int_0^\infty \mathcal{S}(y/\underline{y}(\ell))\tilde{g}_0(y)dy$. This section proves the following result.

Proposition 8. *(Employment distribution)*
Denote by $g_0(y_0|\ell)$ the density function of successful new jobs. Then the invariant distribution g in location ℓ is

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

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

$$s(\ell) = \frac{\delta}{\int_{\underline{y}(\ell)}^\infty\left(\log\frac{y'}{\underline{y}(\ell)}\right)g_0(y'|\ell)dy'} \quad ; \quad f_R(\ell) = \frac{\rho}{\beta}\frac{v(\ell)}{b+v(\ell)}\frac{1-\tilde{G}_0(\underline{y}(\ell))}{\bar{S}(z(\ell),\underline{y}(\ell))}.$$

The proof of Proposition 8 is structured in two main steps. First, I extend Lemma 2. Second, I extend Proposition 1.

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 $g'(x) = A_0e^{-\kappa(x-\underline{x})} + e^{-\kappa(x-\underline{x})}\int_x^\infty g_0(y)e^{\kappa(y-\underline{x})}dy$. Integrating once more:

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

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 measure of new jobs simply scales the invariant mass distribution.

Step 2: extending Proposition 1. 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 measure of newly created new jobs and $h_0 = g_0/H_0$ the entry density of new jobs. Using the expression for g above,

$$\frac{\kappa}{H_0} = \int_{\underline{x}}^\infty e^{-\kappa(x-\underline{x})}\int_{\underline{x}}^x e^{\kappa(y-\underline{x})}h_0(y)dy - \frac{1}{\kappa} + \int_{\underline{x}}^\infty\int_x^\infty g_0(y)dy = \int_{\underline{x}}^\infty xg_0(x)dx.$$

Therefore the job losing rate is $\frac{\delta}{\mathbb{E}^{g_0}[\log(y/\underline{y})]}$. Re-arranging the worker's value of search yields the expression for $f_R(\ell)$.

E.5 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 (32). 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 9. *(Sorting 2)*
All the implications of Proposition 2 hold under Assumption 2 instead of Assumption 1.

Proof. The structure of the proof closely follows Appendix B.4.4. Steps 1 and 2 obtain analogously. The main differences to check are steps 3 and 4. First, the location choice becomes

$$\ell^*(z) = \underset{\ell}{\operatorname{argmax}} \;\log \underline{y}(\ell)\bar{S}(z,\underline{y}(\ell)) + \log(\ell q(\ell)). \tag{54}$$

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 $z(\ell)$. □

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.6 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 (15) 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.[59]

Proposition 10. *(Sorting)*
Under Assumption 2, conditional on the measure 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 thresholdf 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 (54), 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 (54), 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.7 Job-to-job correction

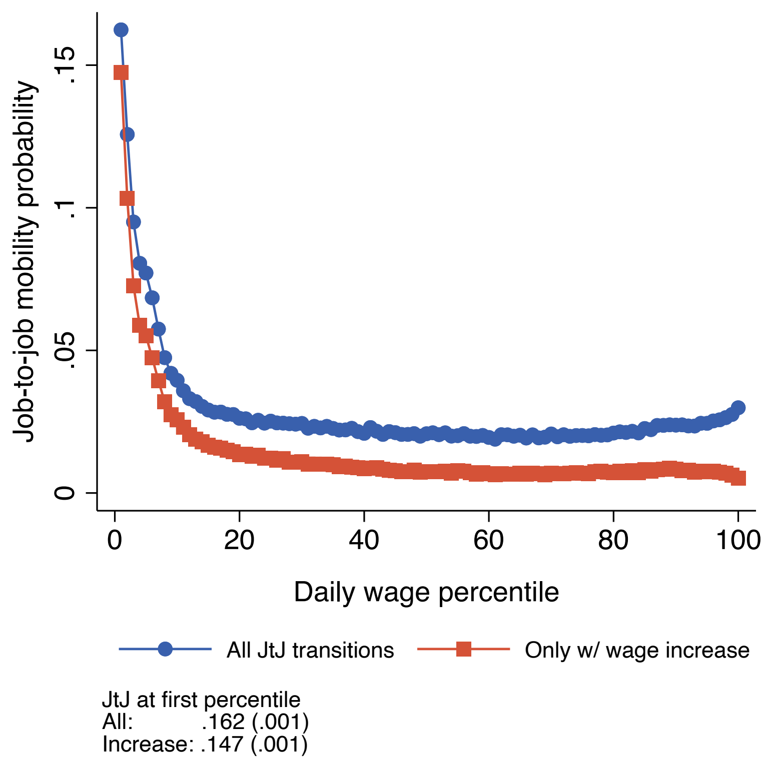
In both constant returns models with job-to-job search and wage posting (Burdett and Mortensen, 1998, Engbom and Moser, 2021), wage bargaining (Cahuc et al., 2006), or decreasing returns models with job-to-job search with wage posting (Bilal and Lhuillier, 2021), wage bargaining (Bilal et al., 2022), 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 20 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 4(c) are also robust to using a more conventional value of 0.3 for ξ.

[59]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.

Figure 20: 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.[60] The planner's objective is then

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

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

[60] 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.[61] 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:

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

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.

[61]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 $\pi(t,\ell) = -\int_\ell^{\bar{\ell}} A(t,x)dx$. 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

$$\frac{\lambda(\ell)^\omega}{\omega} = \mathcal{U}(\ell)b\ell + \int \mathcal{E}(y,\ell)\Big[(1-\alpha)(b+v(\ell))\ell + \alpha y\ell\Big]dy, \tag{55}$$

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:

$$-\frac{v'(\ell)}{v(\ell)}\left[\alpha + \frac{v(\ell)}{b+v(\ell)}(\zeta(\ell)-1)\right] + \frac{1-\alpha}{\ell} + 0 = 0. \tag{56}$$

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 (39) and in the planning solution (56), 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 2, 1 and 3 extend under the same conditions.

Efficiency. Due to the labor market pooling externality term in the v FOC in the decentralized equilibrium (39) relative to the planning solution (56), 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 (45) and (46), $\underline{v}$ and K are identical in both the decentralized equilibrium and the planner solution to a first order. But then from the FOC (42), $\frac{\ell v'(\ell)}{v(\underline{\ell})}$ 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 (41). 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 (41) 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(47) 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 threshold 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 $\tilde{v} = c_2 T\underline{y} - b - \tau_n$ 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 threshold 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,

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

The efficient profit tax then satisfies

$$0=(1-\alpha)\frac{\tau_\pi'(\ell)}{\tau_\pi(\ell)}+\alpha\left(\frac{\bar{\mathcal{S}}'(\zeta(\ell))}{\bar{\mathcal{S}}(\zeta(\ell))}+\log\frac{B}{b+v^{SP}(\ell)}\right)(\zeta^{SP})'(\ell),$$

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.

Under the assumptions of Proposition 3, I can write:

$$\frac{\tau_\pi'(\ell)}{\tau_\pi(\ell)}\approx-\frac{\alpha r}{1-\alpha}\frac{(s^{SP})'(\ell)}{s^{SP}(\ell)}\quad,\quad r\equiv-\left(\frac{\zeta_0^{SP}\bar{\mathcal{S}}'(\zeta_0^{SP})}{\bar{\mathcal{S}}(\zeta_0^{SP})}+\zeta_0^{SP}\log\frac{B}{b+v_0^{SP}}\right). \tag{57}$$

In addition, $\frac{\zeta_0^{SP}\bar{\mathcal{S}}'(\zeta_0^{SP})}{\bar{\mathcal{S}}(\zeta_0^{SP})}=-\frac{\zeta_0^{SP}}{\zeta_0^{SP}-1}-\frac{\zeta_0^{SP}}{\tau+\zeta_0^{SP}}$. When α is small enough, $\zeta^{SP}=\zeta^{DE}+\mathcal{O}(\alpha)$. To first order, I can thus use the equilibrium allocation in (57). Then integrating (57) over ℓ delivers the formula in Proposition 5.

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

$$(\rho+\Delta+\mu)U(p,a,k)=apr(p,a)^{-(\omega+\psi)}\cdot U_1(p,a)k+(\lambda-\varphi)kU_k+\mu M_0k$$

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

$$(\rho+\Delta+\mu)S(y,k,p,a)=pr(p,a)^{-\psi}k\Big[y-U_1(p,a)\Big]+L_yS+\lambda kS_k$$

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,

$$\tilde\rho U(p,a,k)=\frac{(b+v(p,a))ap}{r(p,a)^{\omega+\psi}}k+\mu M_0k-\varphi kU_k\quad;\quad\tilde\rho v(p,a)\quad=\quad\beta f(p,a)U_1(p,a)\left(\frac{Y}{\underline{y}(p,a)}\right)^{\zeta^*(p,a)}\bar{\mathcal{S}}(\zeta^*(p,a)).$$

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 $\hat{\rho} = \tilde{\rho}+\varphi$,

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

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 measure 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 measure 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 measure 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

$$H_0 r(p,a)^\eta \quad = \quad p\bar{k}(p,a)r(p,a)^{-\psi-1}L(p,a)G(\zeta(p,a), v(p,a)),$$

where η is the housing supply elasticity, and

$$\begin{aligned} G(\zeta, v) \quad &= \quad \omega u(v,\zeta)b + \omega(1-u(v,\zeta))(b+v)(1-\beta+\beta\mathbb{E}_{p,a}[y|y>\underline{y}(v)]) \\ &+ \quad \psi(1-u(v,\zeta))(b+v)\mathbb{E}_{p,a}[y|y>\underline{y}(v)]) \end{aligned} \tag{64}$$

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

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

where the optimal vacancy posting decision has been maximized out, and G is defined in equation (64). Re-arranging equation (68) delivers equation (19), with $\mathcal{P} = \frac{1}{\omega+\psi+\varepsilon(1+\eta+\psi)}$, $\mathcal{Q} = \frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}$, and

$$C(\underline{w},z) = (b+v(\underline{w}))^{-1} G(v(\underline{w}),1/z)^{\varepsilon} \tag{69}$$

Using the worker's value of search to substitute out $q(p,a)$ delivers the employer's location problem

$$\begin{aligned}\max_{p,a} \quad & \bar{k}(u(v(p,a),\zeta^*(p,a))^{(1-\alpha)\mathcal{Q}} \cdot \ell(p,a)^{1-\alpha}(b+v(p,a))^{1-(1-\alpha)\psi\mathcal{P}} v(p,a)^{-\alpha} \cdot G(v(p,a),\zeta^*(p,a))^{(1-\alpha)\varepsilon\psi\mathcal{P}} \\ & \cdot \left(\frac{B}{b+v(p,a)}\right)^{(1-\alpha)\zeta+a\zeta^*(p,a)} \bar{\mathcal{S}}(\zeta)^{1-\alpha}\bar{\mathcal{S}}(\zeta^*(p,a))^{\alpha},\end{aligned} \tag{70}$$

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 $\ell(p,a)$.

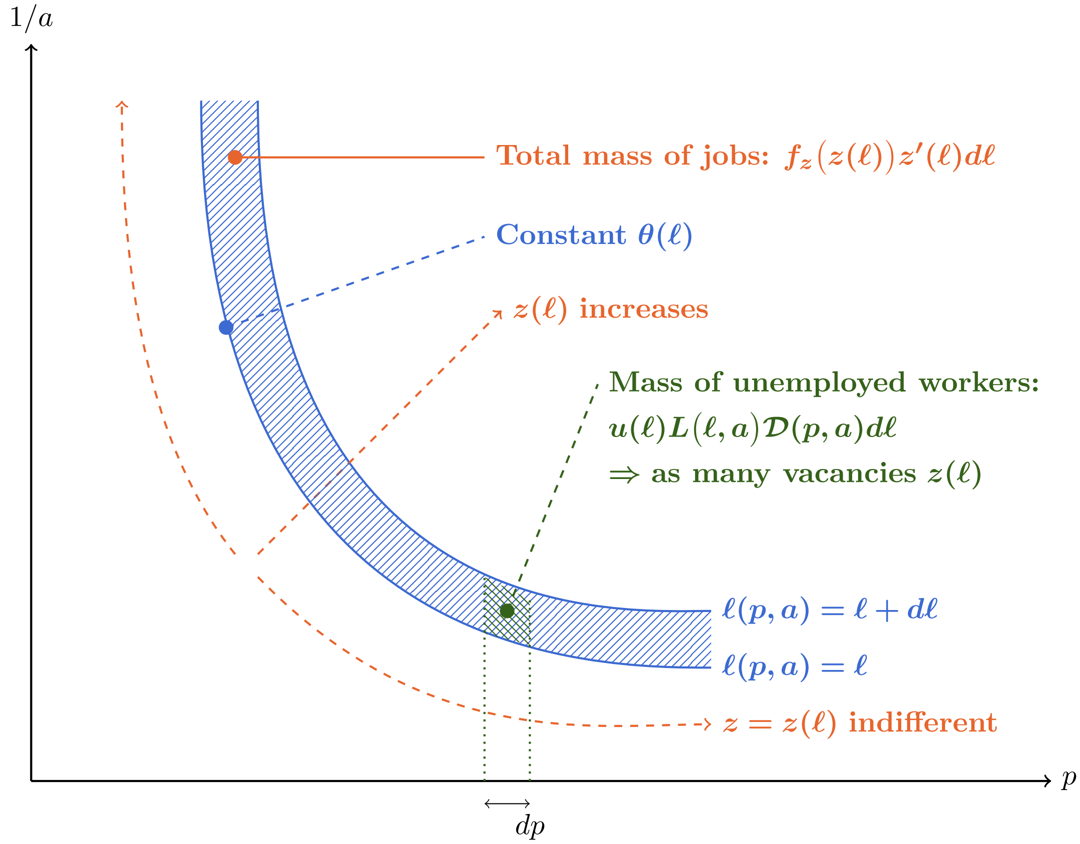
The key insight is that employers only value locations through the combined index $\ell(p,a)$ as long as labor market tightness $\theta(\ell)$ also only depends on the combined index. As illustrated by Figure 21, employers are then indifferent between all locations that have the same index $\ell(p,a)$ 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.[62]

G.8 First-order condition for employers

With this observation at hand, the structure of equation (70) 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

[62] $\mathcal{D}(p,a)$ in Figure 21 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 21: The location choice of employers and workers in two-dimensional space.



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

where

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

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:

$$
\theta(\ell)=-\frac{M_e f_\zeta(\zeta(\ell)\mathcal{V}(\ell,\zeta(\ell))\zeta'(\ell)}{u(\ell)\mathcal{L}(\ell)f_\ell(\ell)} \quad ; \quad \mathcal{V}(\ell,\zeta(\ell))\propto J(\zeta^*(\ell),\ell)^\gamma. \tag{71}
$$

The maximized value of employers is

$$\begin{aligned} J(\zeta^*(\ell),\ell)^{\frac{1}{1+\gamma}} &= \bar{k}(u(\ell))^{\frac{\omega+\varepsilon(1+\eta)}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot \ell \\ &\quad (b+v(\ell))^{\frac{1}{1-\alpha}-\frac{\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}}\cdot v(\ell)^{-\frac{\alpha}{1-\alpha}} G(v(\ell),\zeta(\ell))^{\frac{\varepsilon\psi}{\omega+\psi+\varepsilon(1+\eta+\psi)}} \\ &\quad \left(\frac{B}{b+v(\ell)}\right)^{\frac{\zeta(\ell)}{1-\alpha}} \bar{\mathcal{S}}(\zeta(\ell))^{\frac{1}{1-\alpha}}. \end{aligned} \tag{72}$$

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 $\log\ell$.[63] 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.

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

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

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

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:

$$W(\ell) = W^w(\ell) \times \bar{k}(\ell). \tag{73}$$

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

$$\begin{aligned}\overline{W} \quad &= \quad \int\Big\{u(\ell)\mathbb{E}[U(p,a,y)|\ell] + (1-u(\ell))\mathbb{E}[V(p,a,y)|\ell]\Big\}\bar{k}(\ell)\mathcal{L}(\ell)F_\ell(d\ell)\\ &= \quad A_1 M_0 \int \big(1+\beta\bar{\mathcal{S}}(\zeta(\ell))\big)\bar{k}(\ell)\mathcal{L}(\ell)F_\ell(d\ell) \end{aligned} \tag{74}$$

where $A_1 > 0$ is a transformation of parameters. (74) suggests the decomposition $\overline{W} = \overline{W}^u \times \overline{W}^e \times \overline{W}^k$, where

$$\overline{W}^u = AM_0 \;\; ; \;\; \overline{W}^e = \int \big(1+\beta\bar{\mathcal{S}}(\zeta(\ell))\big)\mathcal{L}(\ell)F_\ell(d\ell) \;\; ; \;\; \overline{W}^k = \frac{\int\big(1+\beta\bar{\mathcal{S}}(\zeta(\ell))\big)\bar{k}(\ell)\mathcal{L}(\ell)F_\ell(d\ell)}{\int\big(1+\beta\bar{\mathcal{S}}(\zeta(\ell))\big)\mathcal{L}(\ell)F_\ell(d\ell)}. \tag{75}$$

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 $D^U = \left\{\frac{1}{S(s\omega+(1-s)\tau)a}\cdot\frac{1}{dt}-1\right\}\cdot dt \to_{dt\to 0} \frac{1}{S(s\omega+(1-s)\tau)a}$. So $D^U = e^{S(s\omega+(1-s)\tau)a}$.

Composition of job findings. The probability that an individual finds a job in a quarter through the ANPE is $s\omega a$, and through other channels $(1-s)\tau a$. Thefore, the probability that an employed individuals has found a job through ANPE is $P^{ANPE} = \frac{s\omega a}{s\omega a+(1-s)\tau a}$. At this point, one can thus identify $x \equiv Ss\omega a$ and $y \equiv S(1-s)\tau a$: $x + y = \frac{1}{D^U}$ and $x = P^{ANPE} \times (x + y)$.

ANPE 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 $a = \frac{x}{Ss\omega}$.