A Dynamic Empirical Model of Frictional Spatial Job Search*

Christian Schluter[†] and Guillaume Wilemme[†]

March 2019

Work in Progress

Abstract

This paper develops a general equilibrium life-cycle model of spatial job search across heterogeneous local labour markets in the presence of search frictions. US and European labour markets exhibit very low geographic mobility. This pattern has usually been framed as resulting solely from moving costs. However, to account for the observed geographic mobility, the implied moving costs should be extremely high. Stating the problem with a search-theoretic perspective, we establish a tractable model of location choice that accounts for the spatial dimension of search frictions. The model allows disentangling the different frictions that contribute to lowering geographic mobility, with a particular emphasis on the role of age. We estimate our model structurally using French administrative individual-level job transition data. Our results suggest first that job search search frictions reduce internal migration much more than mobility costs. Second, mobility costs are more constraining for middle-aged workers than for young and senior workers.

JEL Code: J31, J61, J64, R23

Keywords: search frictions, location choice, local labor markets, life cycle

*Funding from ANR-DFG (contract ANR-15-FRAL-0007-01) is gratefully acknowledged. Address: Aix-Marseille School of Economics, 5 Boulevard Maurice Bourdet CS 50498, 13205 Marseille Cedex 01, France. Emails: `christian.schluter@univ-amu.fr, guillaume.wilemme@gmail.com`.

[†]Aix-Marseille University, CNRS, EHESS, Centrale Marseille, AMSE.

1 Introduction

Job search is both frictional and spatial: Job opportunities are distributed across space, job seekers have to decide where to locate and search, and the matching between jobs and workers is not instantaneous. Search frictions, i.e. the main forces that generate wage or utility dispersion for workers of the same productivity, have therefore important spatial aspects. In particular, the local labour markets (LLMs) that constitute the decentralised labour market exhibit empirically substantial and persistent differences in many dimensions such as unemployment rates, mean wages, and productivity. Learning about job opportunities in one's current location might also be easier than about those in distant locations. At the same time, non-market aspects (generically referred to as "amenities") might enter the location decision, while moving costs, both direct and indirect such as personal attachment to "home", constitute important barriers to mobility. As such attachments usually grow with age, and time-to-retirement and thus investment horizons are longer for younger workers, the (re)location decision needs to be embedded in a dynamic life-cycle perspective. In other words, we seek to bring together several research strands that the established literature in labour and urban economics has usually treated in isolation. To this end, we propose a new dynamic model of spatial job search over the life cycle, which is then estimated structurally using administrative individual-level transition data for France.

Labour markets in developed countries exhibit very low geographical mobility. For instance, Caliendo et al. (2017) report that yearly mobility rates in the US amount to about 3% and in Europe about 1%.[1] The standard model to rationalize this fact is the one by Kennan and Walker (2011). In their setup, moving costs are the principal barriers to mobility. These costs might be overstated, however, if the process of finding job opportunities is frictional, and information about job opportunities is location-dependent.

In this paper, we therefore introduce into the location choice problem a search-theoretic perspective, thus bringing together key strands in the usually separated literatures of internal migration and job search. Alternatively, starting from the search literature, we reinterpret job search as spatial. This spatial dimension then gives rise to new search frictions that might be labelled spatial. Taking as our notional spatial unit a local labour market (LLM),[2] these spatial search frictions arise from the differential flow of information

[1] For the sample of young workers (aged 18-27) in the US examined in Kennan and Walker (2011), the average ten-year interstate migration rate is 32%. Molloy et al. (2014) report for the period 2002-2012 an interstate migration rate of 3.3% for workers aged 20-24, 1.5% for workers aged 35-44, and rates of no more than 0.9% for older workers. For shorter movement, the within-county migration rate is 6.6% for workers aged 35-44. For France, Schmutz and Sidibé (2018) report yearly mobility rates across large regions (NUTS2) for the employed of no more than 1.5% and no more than 2.5% for smaller (NUTS3, departments) regions. The overall transition rates for all are only marginally larger.

[2] As Kennan and Walker (2011) observe: "(i)deally, locations would be defined as local labour markets; (...) even if J is the number of States, the model is computationally infeasible" (p.216). Our tractable directed spatial search paradigm allows us to overcome this challenge. The classic static, long-run,

within and across LLMs, as well as the barriers to mobility. For instance, the structurally estimated (non-spatial) search model of Postel-Vinay and Robin (2002) suggests the contribution of search frictions to wage dispersion amounts to about 50%. However, these overall costs are likely to include spatial search frictions as, for instance, moving costs and location preferences contribute to such wage dispersion. Our approach enables us to account explicitly for these factors. This joint perspective of location choice and frictional spatial job search is then further enriched by explicitly taking into account the life cycle. The worker contemplates a finite horizon problem, where e.g. productivity as well as moving costs are age-dependent, so that younger workers face different incentives compared to old workers.

Our data encompasses labour market transitions both across and within LLMs. We accommodate this as follows. Each LLM is segmented, and a worker's search within a LLM is directed, as in Menzio and Shi (2010): i.e. a worker self-selects the segment in which to search by trading off the probability of finding a new job and the expected utility gain. A worker's preference for a location is, however, subject to occassional shocks, which we model in a canonical random utility framework. The resulting model, it turns out, is very tractable, and can be estimated for a large number of locations.

This model enables us to make several contributions to the literature, both on the theoretical side and on the empirical side.

By introducing spatial search frictions into a model of search, or search frictions into a model of spatial location choice, we can identify barriers to employment mobility that relate to space (such as moving costs, the role of amenities, and spatial variations in productivities and frictional parameters) and to the life-cycle. Our model thus combines several distinct mechanisms that have usually been examined in isolation or in pairs in order to explain and quantify the forces that produce wage or utility dispersion among similar workers.[3] The model provides a rich search paradigm by nesting the canonical random utility framework into a search and matching model. Accordingly, location preferences are subject to occasional shocks. Following Lentz and Moen (2017), we explicitly account for the scale of the preference shock. This scale parameter, σ_ε, determines the randomness of search strategy: As σ_ε tends to 0, workers will tend to direct their search in the best market, while if it diverges towards infinity, search across locations becomes

modelling of LLMs is surveyed in e.g. Moretti (2011). However, Amior and Manning (2018) convincingly demonstrate that the population adjustment process in the wake of local shocks takes a very long time. Our dynamic model thus complements this modelling strategy by providing micro foundations, a shorter run perspective, a role for firms, and we account for the life-cycle.

[3]The life cycle is incorporated in a random search framework by Chéron et al. (2013), and in a directed search framework by Menzio et al. (2016). These authors do not consider spatial search frictions which reduce the incentives to move. The dynamic model of Gould (2007) focuses on career concerns (categorised as blue and white collar) and features two locations. There are no search frictions, and wages are given. In the dynamic model of Baum-Snow and Pavan (2012), the number of locations is increased to three, and search frictions are introduced. However, firms are not considered explicitly and wages are parametrically given.

increasingly more random and unrelated to labour market outcomes or amenities. This parameter that governs the directness versus randomness of location choice is estimated.

The theoretical structure makes a breakthrough in the level of heterogeneity that can be accommodated in the estimation. As a result of the directness of search, our model is very tractable, therefore allowing for a fairly straightforward analysis and estimation. First, our model accommodates individual-level heterogeneity (such as age) and firm behaviour. It thus complements recent complex modelling efforts based on random search. For instance, Schmutz and Sidibé (2018) propose an equilibrium model of spatial random search for homogeneous searchers that is challenging to solve and estimate. In Nanos and Schluter (2018) job searchers determine the optimal location by first identifying the best submarket in each location, and then picking the best location. Both contributions focus, as we do, on the spatial search frictions that considerably enrich the location choice problem, such as the spatial variation of the frictional parameters (job search and separation probabilities), the cost of posting a job vacancy, and the distribution of match-specific productivities. Second, our dynamic model can be estimated for a large number of locations. By contrast, in some leading recent papers the numerical complexity that arises from the dynamic nature of the searcher's optimisation problem has constrained the number of locations. For instance, Gould (2007) considers two locations (corresponding to a rural and an urban area), while Baum-Snow and Pavan (2012) consider three locations (small/medium/large cities). An exception is Schmutz and Sidibé (2018) that account for 100 locations.

The model is structurally estimated using the DADS, a matched employer-employee dataset from French administrative sources. We exploit individual-level transition data taking as spatial units the 21 NUTS2 regions of mainland France. We expect to increase the number of locations in the follow-up versions of the paper. In order to account carefully for the life-cycle profile of wages, productivity or moving costs, we estimate the model using indirect inference. By contrast, a method of moment approach would have required to stratify each relevant moment by age, thus requiring a large number of moments. Indirect inference in our setting thus aids in maintaining parsimoniousness.

Our results suggest that mobility costs are not the main barriers to geographic mobility, but labour markets search frictions are. We show a heterogenous impacts of mobility cost depending on age. Although mobility costs increase with age, middle-aged workers are the most constrained by mobility costs. Young workers face the lowest mobility costs, and are therefore not constrained. Senior workers face the highest mobility costs but have very low incentives to relocate. These workers would not be more geographically mobile in absence of mobility costs.

Section 2 introduces the data. Section 3 presents the model. Section 4 discusses identification and our estimation strategy. Section 5 exposes the empirical findings. Section 6 concludes.

2 Data

We estimate the model on administrative employer-employee data from France, the FH-DADS, which supplements the DADS covering employed workers with administrative data covering the unemployed. Since the DADS is well known and has been used many times,[4] we relegate details to Data Appendix C. Here we set out briefly our principal sample selection rules, which follow established DADS-practices, and data definitions, which are again further explained in the Data Appendix.

We focus on the years between 1994 and 2007, a period without recession in France. Later years are dropped because of the financial crisis. We further select individuals for whom schooling is reported (i.e. individuals born the first four days of October, January, April or July), and consider the following five groups: no degree, vocational training, high-school degree, bachelor degree, and more than a bachelor degree. Further, we consider only private sector workers in mainland France, drop outliers in wages, and consider workers only up to the age of 60, in order not to confound the analysis by post-retirement mobility (whose distinct pattern in France is documented in e.g. Gobillon and Wolff, 2011).

To bring the model to the data, we need to specify three key observed variables: age (or unit of time), the employment situation, and location. Working with the finest calendar time and geographical unit involves statistical noise. We thus aggregate over time to define observations at the yearly level. For each year observed, a worker is categorized either as unemployed or employed. If a worker has spent more than 15 days without a main job in the year, we categorize her as unemployed this year. Accordingly, unstable employment periods are categorized as unemployment. We define a yearly reference job as the main job the individuals have the 15th of December. There is job mobility between two years if the employer's establishment identifier changes. We define the empirical counterpart of w as the hourly wage. By doing so, we homogeneize jobs in terms of hours worked so that only wages differ.

For the estimation, we define location as the workplace, i.e. the establishment's location, in order to discount changes of residence unrelated to the labour market. For unemployed workers, the current location is defined as the last employer's location. At this stage, we use the administrative partition of mainland France into 21 regions.

[4] See e.g. Abowd et al. (1999), Postel-Vinay and Robin (2002), Combes et al. (2008). The DADS (Declarations Annuelles des Donnees Sociales) is based on firms' payroll reports, and information on unemployment had hitheFor instance, workers by twice The increase in mobility rates for 25 years-old workers would be twice higher than for rto to be infered. The FH version, by contrast, includes administrative information about unemployment spells.

2.1 Age patterns of mobility

Our (unbalanced panel) data set comprises about 186,000 workers, who are observed during the years 1994 to 2007. We proceed to describe our key transition data, having temporally aggregated these into annual transitions. Figure 1 depicts, as a function of the worker's age, the annual transitions on the labour market, as well as the spatial mobility of workers using different levels of spatial aggregation. In particular we consider the region (NUTS 2, with 21 units), the department (NUTS 3, with 94 units), and the commune (LAU 2, with 36,000 units) in mainland France.

For all cases considered, we have a clear life-cycle profile: From about the age of 25 all curves are rapidly decreasing and then flatten out. For young ages, the graphs are increasing, which is to be explained by composition effects, as individuals continue in education whilst other drop out of education and start work (often in unstable jobs). The measure of residential mobility depends, of course, on the extent of spatial aggregation. Mobility across fairly large (NUTS3) units in France is, however, of a similar magnitude observed in other European countries and the US. They are also consistent with the aggregate time series for France reported in Schmutz and Sidibé (2018). In particular, for our observation window the (unconditional) mean rates of residential mobility are 10.8% across communes, 4.1% across departments, and 2.3% across regions.

Turning to the joint event of changing employer and workplace department, Figure 1.A reveals that the life-cycle profile clearly mimics the unconditional employment transition rate. However, the scale of the former is 6 times smaller than of the latter. Joint changes at ages 25-30 are about 20% of the employment transition rate, and at ages 50-60 are about 14%.

3 A Life-Cycle Model of Spatial Job Search

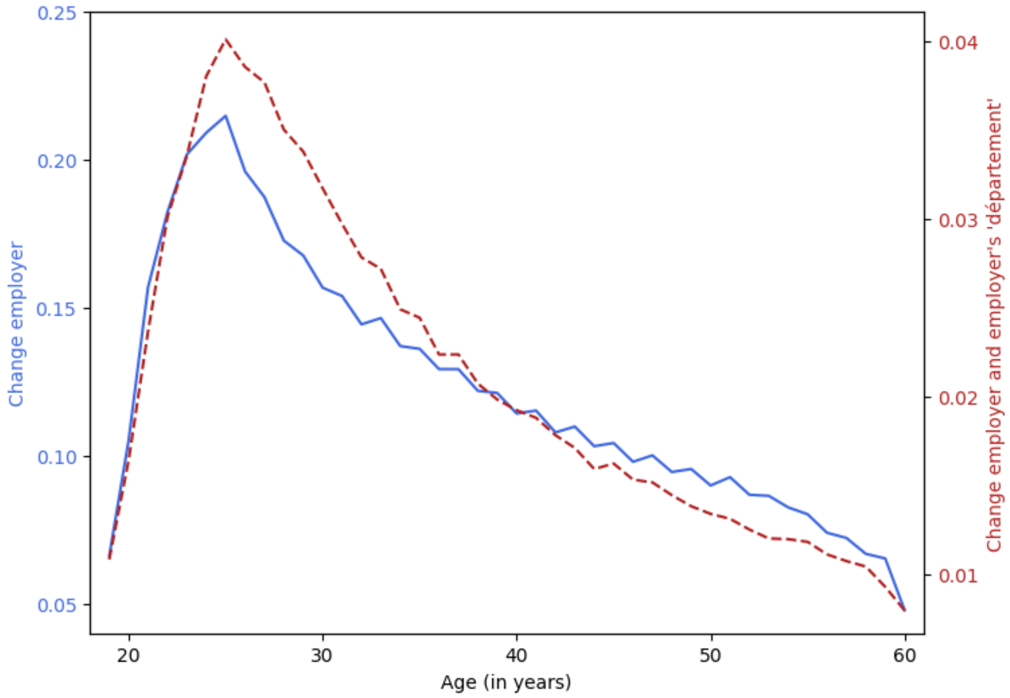
Starting from the perspective a conventional job search model, we reinterpret such search spatially. To this end, the economy is composed of L distinct and heterogeneous locations, which in our empirical application will be local labour markets.

3.1 The environment

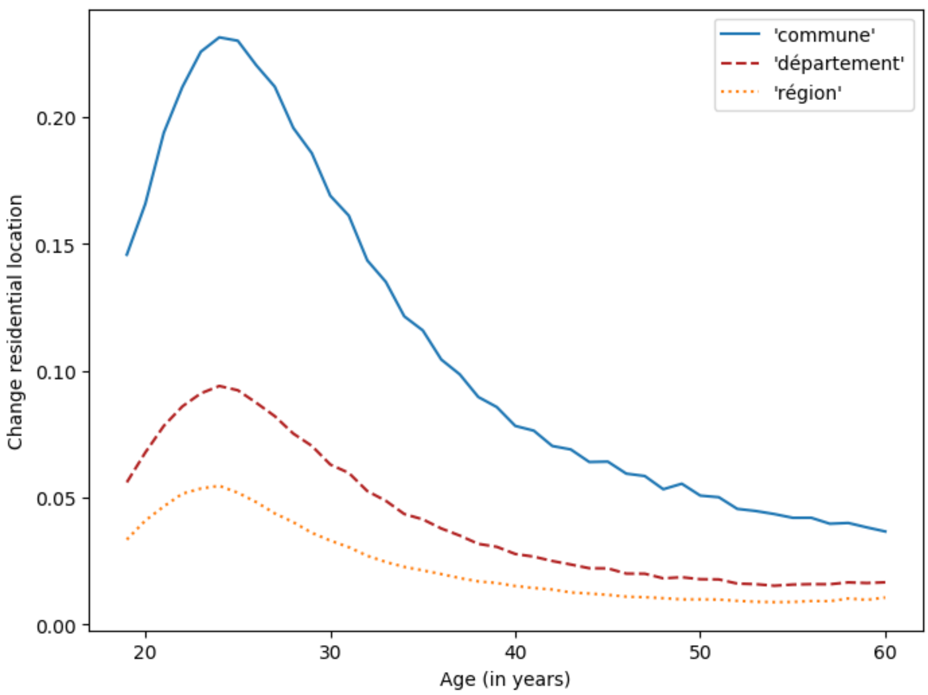
Time is discrete. Workers start their economic life at age $a = \underline{a}$ and retire at an exogenous age $a = \bar{a}$, and overlapping generations of workers are distributed across the L locations. In the cross-section, workers differ thus in terms of age, and in terms of their skill type which is denoted by x. The retirement age is the same for all workers whereas the entry age $\underline{a}$ differs according the skill type x.

Workers, both unemployed and employed, can search within and across locations. Their per-period worker utility is derived not only from income from employment (a

Figure 1: Annual transitions on the labour market and across space in France, 1994-2007



(a) Job mobility



(b) Residential mobility

Notes. Panel (b) shows the share of workers who changed their employment situation within a year (blue solid line, scale on left), and the share of annual labour market transitions that imply a change in the workplace department (red dashed line, scale on right). Panel (b) shows the share of workers who changed their residential location for different geographical scales. Spatial aggregation: Regions are NUTS 2 (with 21 units), departments are NUTS 3 (with 97 units), and communes are LAU 2 (with 36,000 units). Source: authors' calculations, FH-DADS for the years 1994-2007.

wage) or unemployment (benefits b), but also from the amenities v_l that living in location l brings. If the worker decides to search for a job, a search cost is incurred. If the worker decides to move and change location, a further moving cost has to be paid. Both costs will be specified below.

Firms have fixed locations, and produce with constant returns to scale (so that a multiple-job firm is not different from a collection of single-job positions). A firm in location l can open a job vacancy at cost μ_l per period. If matched with a type-x worker of age a, they produce the quantity $y(a, x, z)$ for the period where z denotes a match-specific component. This z remains constant throughout the lifetime of the job, and is only observed once the match is formed. Jobs are thus experience goods, meaning the productivity components are perfectly observed by the worker and the firm. In order to model the observed spatial heterogeneity of productivities, we assume that z is drawn from a distribution $\mathcal{F}_l(z)$, which depends on the new job's location l. A firm with a filled position receives as per-period profits production $y(a, x, z)$ net of the wage.

Jobs may exogenously break at the end of a period with some probability δ_l. We assume that this event is known at the beginning of the period, before searching, so that an employee has the information when searching on-the-job.[5] We further assume that firms and workers are risk-neutral and discount time with a factor $\frac{1}{1+r} < 1$. A state is then defined as a tuple containing information on age, location, worker's type and match type, (a, l, x, z).

Matching frictions and competitive search Each location can be thought of as an island where firms employ local workers. The only way for a worker to change location is to get a job offer from another location. Getting a job offer, however, is a random event that takes time and efforts.

We adopt the competitive search framework of Moen (1997) and Menzio and Shi (2010). labour markets are segmented within locations. A worker of type (a, x) located in l can look for a job in location k. She must pick a submarket among the set $\mathcal{M}_{ak}(x)$. Submarkets differ by the lifetime utility W firms promise to workers. Within each submarket, a firm with an open position in location k meets a worker of type (a, x) with probabilities governed by a constant-returns-to-scale matching function. The matching probabilities thus only depends on the ratio of vacancies to the number of job applicants on the submarket, so-called (sub)market tightness θ. A worker receives a job offer with probability $p(\theta)$ whereas a firm finds a job seeker with probability $q(\theta) \equiv \frac{p(\theta)}{\theta}$. p is increasing concave from $(0, +\infty)$ onto $(0, 1)$, and q decreasing convex from $(0, +\infty)$ onto $(1, 0)$. Free entry and exit of firms implies that a zero-profit condition is reached for each submarket at equilibrium. It will define a mapping of every promised utility W to submarket tightness

[5]This modelling is a common tool to rationalize job-to-job transitions with a job quality decrease. See Jolivet et al. (2006); Bagger et al. (2014).

θ. Each submarket is labeled by a pair (θ, W), an element of $\mathcal{M}_{ak}(x)$.

A worker is endowed with one unit of search effort. She can only choose at most one submarket per period. Picking a labour market is equivalent to choosing simultaneously a location k (where to search geographically) and a pair $(\theta, W) \in \mathcal{M}(a, k, x)$ (where to search within the location).[6] A worker can also decide optimally not to search at all this choice is labeled $k = 0$.

This simultaneous search procedure is equivalent to a two-step search: i) the worker chooses optimally a location k; ii) the worker directs her search effort toward the utility-maximizing submarket in the selected location.

Search and mobility costs All searching workers incur a search cost, but only workers who change location incur a known age-dependent mobility cost mc_{alk}, with $mc_{all} = mc_{al0} = 0$. The search cost is sunk whatever the success of search activities and does not, therefore, affect the choice of submarket (θ, W) within a location. Turning to the specification of these search costs, these have two components, $sc^s_{alk} - \varepsilon_k$. The first term sc^s_{alk} is age-dependent and non-random, which can be interpreted as a disutility of searching, or as a cost of information to apply for a job in location k from location l. For instance, a worker in location l may be less efficient in searching in location k than workers already living in location k. Workers pay nothing when they do not search, so $sc^s_{al0} = 0$.

The second component, ε_k, capturing a preference shocks regarding the location k, allows us to account for the observed heterogeneity in relocations in the data since otherwise the same (a, x) worker would always choose the same location. We then follow standard modelling practices and adopt a random utility framework (e.g. McFadden 1973; Rust 1987; Kennan and Walker 2011),[7] letting the preference shock follow a Gumbel distribution. However, we follow Lentz and Moen (2017) and consider explicitly the shape parameter of the shock distribution. Specifically, the c.d.f. of ε_k is $\exp(-\exp(-\frac{x}{\sigma_\varepsilon} - euler))$, with $\sigma_\varepsilon > 0$ and $euler$ the Euler constant. Shocks then have a zero mean and a variance equal to $\frac{\pi^2}{6}\sigma_\varepsilon^2$. The mean has no effect on the model as it only shifts expected utility by a fixed value. By contrast, the scale parameter σ_ε affects the randomness of search strategy: As σ_ε gets closer to 0, workers will tend to direct their search in the best market, as it diverges towards infinity, search across locations becomes increasingly more random and unrelated to labour market outcomes or amenities.

[6] Even if the unit of search effort were divisible, a worker would optimally allocate the entire unit to the same segment. This is not the case anymore when search has decreasing returns within labour markets (Decreuse and Zylberberg, 2011; Wilemme, 2017).

[7] In non-spatial search and matching models, the arrival rate of job offers differs between unemployed and employed workers (see for instance Burdett and Mortensen 1998; Postel-Vinay and Robin 2002; Menzio and Shi 2010). In our model, instead, the random utility framework provides microfoundations for the different job-finding rates. For instance if $sc^u_{alk} \leq sc^e_{alk}$ then employed workers are more likely not to search (to choose $k = 0$) than unemployed workers.

Complete contracts and wages Following the literature on directed search (e.g. Menzio and Shi 2011; Menzio et al. 2016), we assume that firms offer bilaterally efficient contracts to workers. It means that the search strategy maximizes the joint surplus of the worker and the firm. Bilateral efficiency arises for various specifications of the contract space. Wages are therefore undetermined. In order to exploit data on wages, we add further structure, and assume firms advertise a Nash bargaining rule parametrised by ρ: Firms commit to offer a fixed share ρ of the match surplus over the employment spell at each period.[8] When $\rho = 1$, the worker receives the full joint surplus of the match, and when $\rho = 0$, she receives the value of unemployment. This individual Nash bargaining then leads to further heterogeneity in wage outcomes, as similar workers can have different wages if the individual-level sharing rule differs.

Timing of events The timing of events between two periods is the following.

1. Production is realized. Employed workers receive the wage specified by their job contract. Unemployed workers consumes home production. Both enjoy local amenities. Firms with a filled position receive profits.
2. The separation shock is revealed. Employed workers discover if the job survives for the next period with probability $1 - \delta_l$.
3. Taste shocks for prospected locations $\boldsymbol{\varepsilon} = \{\varepsilon_k\}_{k=0,..,L}$ are revealed. Firms open vacancies after paying the cost μ_l, and workers choose a search strategy after paying the search costs.
4. Matching occurs. Workers discover whether search is successful or not. Workers who obtain a new job pay the mobility costs and then discover their match type z. Workers who do not obtain a job offer are unemployed for the next period in their current location, except if they were employed and their job survives.

3.2 Value Functions

Active submarkets The joint surplus of a firm and a (a, x)-type worker in location l and match-specific effect z is denoted by $V_{al}(x, z)$. A firm that promises lifetime utility W to workers of type (a, x) receives the expected surplus $\int V_{al}(x, z) d\mathcal{F}_l(z) - W$ when it finds an employee. Free entry of firms drives expected profits to zero. As long as the expected surplus is positive, firms open new vacancies on each submarket such that the

[8] Acemoglu and Shimer (1999) proposes this modelling as an extension of their benchmark model. An alternative approach is to assume piece-rate wages, as in e.g. Barlevy (2008) and Bagger et al. (2014) assume wages are a constant "piece-rate" fraction of productivity.

cost of vacancy is equal to the probability to match times the expected surplus:

$$\mu_l = q(\theta)\left[\int V_{al}(x,z)d\mathcal{F}_l(z) - W\right] \qquad \text{and} \quad \int V_{al}(x,z)d\mathcal{F}_l(z) \geq W. \tag{1}$$

All the pairs (θ, W) satisfying (1) define the set of active submarkets $\mathcal{M}_{al}(x)$.

Firms make a low surplus when the promised value W is high. These jobs are only profitable when the cost of finding a worker is low, or equivalently when market tightness θ is low. When the promised value W exceeds the expected surplus from matching, profits are negative profits which discourages any firm from proposing such a contract. W is therefore decreasing with θ in the set $\mathcal{M}_{al}(x)$.

The value of unemployment Denote by $U_{al}(x)$ the lifetime utility of a worker unemployed at age a in location l. She receives home production b, local amenities v_l, pays search costs $sc^u_{alk} - \varepsilon_k$ and pays moving costs mc_{alk} if she obtains a job. She obtains a new job with probability $p(\theta)$ giving the expected lifetime utility W. Otherwise she remains unemployed in the same location yielding the value $U_{a+1,l}(x)$. The prospected location k and the submarket (θ, W) are chosen after the taste shocks $\boldsymbol{\varepsilon}$ are revealed. The lifetime utility follows a recursive equation for $a < \bar{a}$,

$$\begin{aligned} U_{al}(x) &= b + v_l \\ &+ \mathbb{E}_\varepsilon \max_{\substack{k\in\{0,..,L\} \\ (\theta,W)\in\mathcal{M}_{ak}(x)}} \Big\{ - sc^u_{alk} + \varepsilon_k - p(\theta)mc_{alk} + \frac{1}{1+r}\left[p(\theta)W + (1-p(\theta))U_{a+1,l}(x)\right] \Big\}, \end{aligned} \tag{2}$$

with the termination conditions $U_{\bar{a}l}(x) = b + v_l$. $\mathbb{E}_\varepsilon$ denote the expectation operator over the $L+1$ taste shocks ε_k. By convention, we define the set of submarkets when the worker decides not to search, $\mathcal{M}_{a0}(x) = \{(0,0)\}$.

The search problem can be decomposed in a sequential search problem,

$$U_{al}(x) = b + v_l + \frac{U_{a+1,l}(x)}{1+r} + \mathbb{E}_\varepsilon \max_{k\in\{0,..,L\}} \{R^u_{alk}(x) - sc^u_{alk} + \varepsilon_k\}, \tag{3}$$

$$R^u_{alk}(x) = \max_{(\theta,W)\in\mathcal{M}_{ak}(x)} \left\{ p(\theta)\left[\frac{W - U_{a+1,l}(x)}{1+r} - mc_{alk}\right]\right\}. \tag{4}$$

In the first equation, the worker chooses one location in which she will look for a job, or does not search by choosing $k = 0$. She makes the decision based on the returns from searching within location k, R^u_{alk}. These returns are endogenous and defined by the second equation. Given a prospective location k, a worker picks the segment that provides the highest expected gains, trading-off the probability to be selected by the firm and the promised utility. Note that $R^u_{alk}(x) \geq 0$ because a worker can always choose

$\theta = 0$, and $R^u_{al0}(x) = 0$. We denote respectively $K^u_{al}(x, \boldsymbol{\varepsilon})$, $\theta^u_{alk}(x)$ and $W^u_{alk}(x)$ the three policy functions of the optimization programs (3) and (4) at age $a < \bar{a}$. The search costs are sunk even if search is unsuccessful. The choice of a submarket is therefore affected neither by the deterministic part sc^u_{alk} nor the stochastic part $\boldsymbol{\varepsilon}$.

The joint value of a match An employed worker produces $y(a, x, z)$ and benefits from amenities v_l. She also incurs search costs and mobility costs. If the job terminates with probability δ_l, the search problem is equivalent to the unemployed. If the job survives, the profile of search costs is different and the outside option if search is unsuccessful is $V_{a+1,l}(x, z)$. The recursive equation for $a < \bar{a}$ writes

$$\begin{aligned} V_{al}(x,z) &= y(a,x,z) + v_l \\ &+\delta_l.\mathbb{E}_\varepsilon \max_{\substack{k\in\{0,..,L\}\\(\theta,W)\in\mathcal{M}_{ak}(x)}} \Big\{ - sc^u_{alk} + \varepsilon_k - p(\theta)mc_{alk} + \frac{1}{1+r}\left[p(\theta)W + (1-p(\theta))U_{a+1,l}(x)\right] \Big\} \\ &+(1-\delta_l).\mathbb{E}_\varepsilon \max_{\substack{k\in\{0,..,L\}\\(\theta,W)\in\mathcal{M}_{ak}(x)}} \Big\{ - sc^e_{alk} + \varepsilon_k - p(\theta)mc_{alk} + \frac{1}{1+r}\left[p(\theta)W + (1-p(\theta))V_{a+1,l}(x,z)\right] \Big\}, \end{aligned} \tag{5}$$

with the termination conditions $V_{\bar{a}l}(x, z) = y(\bar{a}, x, z) + v_l$. Job contracts are bilaterally efficient, so the worker maximizes the joint value of a match and not her share of the surplus. The program can be decomposed in a sequential decision problem:

$$\begin{aligned} V_{al}(x,z) &= y(a,x,z) + v_l + \frac{\delta_l U_{a+1,l}(x) + (1-\delta_l)V_{a+1,l}(x,z)}{1+r} \\ &\quad + \delta_l.\mathbb{E}_\varepsilon \max_{k\in\{0,..,L\}} \{R^u_{alk}(x) - sc^u_{alk} + \varepsilon_k\} \\ &\quad + (1-\delta_l).\mathbb{E}_\varepsilon \max_{k\in\{0,..,L\}} \{R^e_{alk}(x,z) - sc^e_{alk} + \varepsilon_k\}, \end{aligned} \tag{6}$$

$$R^e_{alk}(x,z) = \max_{(\theta,W)\in\mathcal{M}_{ak}(x)} \left\{ p(\theta)\left[\frac{W - V_{a+1,l}(x,z)}{1+r} - mc_{alk}\right]\right\}. \tag{7}$$

The returns from searching within a location k depends on the job separation shock. If the job terminates, the worker behaves as if she was already unemployed. The returns form searching are $R^u_{alk}(x)$ in that case. If the job survives, the returns from searching are $R^e_{alk}(x, z) \geq 0$. We denote the choice of prospected location $K^e_{al}(x, z, \boldsymbol{\varepsilon})$. If the job does not break, the policy function is $(\theta^e_{alk}(x, z), W^e_{alk}(x, z))$. If the job survives, the policy function is $(\theta^u_{alk}(x), W^u_{alk}(x))$ as defined in the unemployed worker's search problem.

Wages Wages depend on the current state (a, l, x, z) and the bargaining rule ρ. Denote these by $w_{al}(x, z, \rho)$. The worker's outside option in the bargaining is the value of unemployment in the region. Following Nash bargaining, the firm's surplus from employing a worker in state (a,l,x,z) and bargaining rule ρ is $V_{al}^F(x,z,\rho) \equiv (1-\rho)(V_{al}(x,z)-U_{al}(x))$. It is recursively defined for $a < \bar{a}$ by

$$V_{al}^F(x,z,\rho) = y(a,x,z) - w_{al}(x,z,\rho) + (1-\delta_l).\mathbb{E}_\varepsilon \left\{1 - p(\theta^e_{alK^e_{al}(x,z,\varepsilon)}(x,z))\right\}.\frac{V^F_{a+1,l}(x,z,\rho)}{1+r}, \tag{8}$$

with the termination conditions $V_{\bar{a}l}^F(x,z,\rho) = y(\bar{a},x,z) - w_{\bar{a}l}(x,z,\rho)$. A firm receives current-period profits. The value of the match will become $V^F_{a+1,l}(x,z,\rho)$ the following period if the match does not separate with probability $(1-\delta_l).\mathbb{E}_\varepsilon \left\{1 - p(\theta^e_{alK^e_{al}(x,z,\varepsilon)}(x,z))\right\}$. The lifetime values, $V_{al}(x,z)$ and $U_{al}(x)$, solutions to the dynamic programming problem, do not depend on wages. By substituting $V_{al}^F(x,z,\rho)$ in equations (8) by the Nash bargaining solution, wages $w_{al}(x,z,\rho)$ can be easily solved for once the lifetime values are known, given a bargaining rule ρ.

To complete the model, we relate the bargaining rule to the promised utility. For a job starting at age a, the promised utility is an average of the expected value of the match and the value of unemployment weighted by ρ,

$$W = \rho \int V_{a,l}(x,z) d\mathcal{F}_l(z) + (1-\rho) U_{a,l}(x). \tag{9}$$

This equation defines $\rho_{al}(x,W)$. Note when a worker obtains a job with promised utility W starting at age a, she will enjoy the wage profile $w_{a'l}(x,z,\rho_{al}(x,W))$ for $a' > a$. As a consequence, the bargaining rule remains constant for future ages a' over the employment spell.

3.3 Equilibrium

We now define and then characterise an equilibrium.

Definition 1 *An equilibrium is defined by*

- *values of unemployment U_{al} and joint values of a match V_{al} that satisfy their recursive definition* (2) *and* (5)*;*
- *policy functions for prospected locations, K^u_{al} and K^e_{al}, solutions to problems* (3) *and* (6)*;*
- *policy functions for prospected submarket within location, $(\theta^u_{alk}, W^u_{alk})$ and $(\theta^e_{alk}, W^e_{alk})$, solutions to problems* (4) *and* (7)*;*
- *wage functions w_{al} that satisfy equation* (8)*;*
- *bargaining rules ρ_{al} that satisfy equation* (9).

An equilibrium satisfies the block-recursion property, which dramatically reduces the dimensionality of the numerical simulation:

Proposition 1 *There exists a unique equilibrium and it satisfies the block recursive property: the value functions and the policy functions does not depend on the distribution of observables over the population.*

We show that the block-recursion property holds with the nested random utility modelling of search across local labour markets.[9] A proof is given in appendix. In our case, the block-recursion property implies that value and policy functions do not depend on the dynamics of migration and mobility across regions and markets. Theoretically, we can thus save any assumptions on how generations of workers overlap. Simulating the model thus only requires the knowledge of the distribution at the starting age $\underline{a}$ for each skill level.

Solving the choice problem The equilibrium can be solved by backward induction within the period. We start by characterizing the returns to search. By substituting the job creation condition (1), the choice of a submarket is equivalent to,

$$R^u_{alk}(x) = \max_{\theta \geq 0} \left\{ p(\theta) \left[\frac{\int V_{a+1,k}(x,z) d\mathcal{F}_k(z) - U_{a+1,l}(x)}{1+r} - mc_{alk} \right] - \theta \frac{\mu_k}{1+r} \right\}, \tag{10}$$

$$R^e_{alk}(x,z) = \max_{\theta \geq 0} \left\{ p(\theta) \left[\frac{\int V_{a+1,k}(x,z) d\mathcal{F}_k(z) - V_{a+1,l}(x,z)}{1+r} - mc_{alk} \right] - \theta \frac{\mu_k}{1+r} \right\}, \tag{11}$$

If the term within brackets, meaning the expected match surplus net of the moving cost, is lower than the discounted cost of vacancy $\frac{\mu_k}{1+r}$, the submarket is not profitable and so market tightness is zero. In the other case, the first-order condition characterises the policy functions θ^u_{alk} and θ^e_{alk}. The functions $W^u_{a+1,lk}$ and $W^e_{a+1,lk}$ can be recovered with equation (1). Given the concavity of $p(.)$, the chosen market tightness θ^s_{alk} will be increasing in the term within the brackets, If the joint value of a match exceeds the value of unemployment, $V_{al}(x,z) \geq U_{al}$, then $\theta^e_{alk} \geq \theta^u_{alk}$. A worker who can keep her job for the next period searches differently from an unemployed worker. She will prefer high promised utilities with low job-finding rates, whereas an unemployed worker sacrifices more utility to obtain a job faster.

The probability to choose a specific location and the expected utility have close-form expressions thanks to the Gumbel distributions. We provide the mathematical details in Appendix A. The probability for an unemployed worker in location l to search within

[9] Block-recursion is a property of directed search models, first established by Menzio and Shi (2010). Menzio et al. (2016) extends the property to the life-cycle version.

location $k > 0$ is

$$\mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = k) = \frac{\exp(R^u_{alk}(x) - sc^u_{alk})^{\frac{1}{\sigma_\varepsilon}}}{1 + \sum_{j=1}^L \exp(R^u_{alj}(x) - sc^u_{alj})^{\frac{1}{\sigma_\varepsilon}}}, \tag{12}$$

and the probability to remain idle is

$$\mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = 0) = \frac{1}{1 + \sum_{j=1}^L \exp(R^u_{alj}(x) - sc^u_{alj})^{\frac{1}{\sigma_\varepsilon}}}. \tag{13}$$

We briefly characterise these choice probabilities. The probability to choose the prospected location $k > 0$ is decreasing in both the current costs, sc^u_{alk} and mc_{alk}. On the contrary, when the search sc^u_{alj} and mobility costs mc_{alj} of other options $j \neq k$ increase, the probability to choose option k (including $k = 0$) increases. As regards the variance of the preference shock, we note the following asymptotic properties:

$$\lim_{\sigma_\varepsilon \to 0} \mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = k) = 1 \qquad \text{if } k = \text{argmax}_{j=0,..,L}\{R^u_{alj}(x) - sc^u_{alj}\}, \tag{14}$$

$$\lim_{\sigma_\varepsilon \to \infty} \mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = k) = \frac{1}{1 + L}. \tag{15}$$

When the variance of the preference shock is low, workers direct their search to the location that maximizes the returns to search net of the deterministic search cost. As the variance increases, the deterministic part of the incentives becomes negligible compared to the magnitude of the preference shock. The search strategy gets closer to a random uniform distribution.

The expected utility from searching across locations thus equals

$$\mathbb{E}_\varepsilon \max_{k \in \{0,..,L\}} \{R^u_{alk}(x) - sc^u_{alk} + \varepsilon_k\} = \sigma_\varepsilon \log\left(1 + \sum_{k=1}^L \exp(R^u_{alk}(x) - sc^u_{alk})^{\frac{1}{\sigma_\varepsilon}}\right). \tag{16}$$

Analogous equations of (12), (13), and (16) are obtained for employed workers. The effects of costs and the variance of preference shocks on the choice probabilities are the same.

The expectation operator can now be substituted in the two recursive definitions (3) and (6) for $a < \bar{a}$,

$$U_{al}(x) = b + v_l + \frac{U_{a+1,l}(x)}{1+r} + \sigma_\varepsilon \log\left(1 + \sum_{k=1}^{L} \exp(R^u_{alk}(x) - sc^u_{alk})^{\frac{1}{\sigma_\varepsilon}}\right), \tag{17}$$

$$\begin{aligned} V_{al}(x,z) = y(a,x,z) + v_l + \frac{\delta_l U_{a+1,l}(x) + (1-\delta_l)V_{a+1,l}(x,z)}{1+r} \\ + \delta_l.\sigma_\varepsilon \log\left(1 + \sum_{k=1}^{L} \exp(R^u_{alk}(x) - sc^u_{alk})^{\frac{1}{\sigma_\varepsilon}}\right) \\ + (1-\delta_l).\sigma_\varepsilon \log\left(1 + \sum_{k=1}^{L} \exp(R^e_{alk}(x,z) - sc^e_{alk})^{\frac{1}{\sigma_\varepsilon}}\right) \end{aligned} \tag{18}$$

The transition probabilities We are now in position to state the transition probabilities across states at equilibrium.

The probabilities to transit from a state at age a to a state at age $a+1$ for $a < \bar{a}$ are enumerated in table 1. The first two rows provides the transition probabilities from the state of an unemployed worker aged x, living in l with type x. She will get a job in location k and match type z if i) she decides to look for a job within location k, ii) she matches with an employer in the optimal submarket, and iii) she draws the right match type. The formula in the first row is exactly the product the three probabilities. In the second row, the worker remains unemployed if either she picks $k = 0$ or her search activities are unsuccessful.

The last three rows define the transition probabilities from employment. An employed worker aged x, in location l, type x and match type z can behave in two different ways depending on the job separation shock. The third formula is similar to the first one, except that the worker can also remain in her job. Her job must survive with probability $1 - \delta_l$, and she then must not search at all or be unsuccessful. This probability is given in the first row. Lastly, she can become unemployed if her job breaks and she does not obtain any job offer.

Spatial search frictions Our introductory discussion of the spatial aspects of job search has led us to emphasise that the frictional process now has new spatial dimensions: spatial search frictions. We are now in a position to make this more precise.

In particular, the standard job-related search frictions are captured by the matching functions $p(\theta)$ and $q(\theta)$. These prevent workers from obtaining their preferred job, create (frictional) unemployment, and wage dispersion among ex-ante identical workers. These frictions are absent if $p(\theta) = q(\theta) = 1$.

Spatial search frictions are barriers to mobility, the location-dependent flow of information about job opportunities, and location preferences. These are captured in the model by mc, sc, and $\boldsymbol{\varepsilon}$, and are additional channels through which wage dispersion arises

Table 1: Transition probabilities across labour market states and locations between age a and $a+1$

From	To	Probability
1. Unemployment (a, l, x)	Employment $(a+1, k, x, z)$	$\mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = k) \times p(\theta^u_{alk}(x)) \times f_k(z)$
2. Unemployment (a, l, x)	Unemployment $(a+1, l, x)$	$\mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = 0)$ $+ \sum_{k=1}^L \mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = k) \times (1 - p(\theta^u_{alk}(x)))$
3. Employment (a, l, x, z)	Employment in a new job $(a+1, k, x, z')$	$\delta_l \times \mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = k) \times p(\theta^u_{alk}(x)) \times f_k(z')$ $+(1-\delta_l) \times \mathbb{P}_\varepsilon(K^e_{al}(x, z, \boldsymbol{\varepsilon}) = k) \times p(\theta^e_{alk}(x, z)) \times f_k(z')$
4. Employment (a, l, x, z)	Employment in same job $(a+1, l, x, z)$	$(1-\delta_l) \times [\mathbb{P}_\varepsilon(K^e_{al}(x, z, \boldsymbol{\varepsilon}) = 0)$ $+ \sum_{k=1}^L \mathbb{P}_\varepsilon(K^e_{al}(x, z, \boldsymbol{\varepsilon}) = k) \times (1 - p(\theta^e_{alk}(x, z)))]$
5. Employment (a, l, x, z)	Unemployment $(a+1, l, x)$	$\delta_l \times [\mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = 0)$ $+ \sum_{k=1}^L \mathbb{P}_\varepsilon(K^u_{al}(x, \boldsymbol{\varepsilon}) = k) \times (1 - p(\theta^u_{alk}(x)))]$

Notes: Conditional on the initial labour market state, it is easily verified that the transition probabilities sum up to 1 after integrating over z and z'. Spatial flows into location k (by initial labour market status) are simply obtained by summing over all source locations l, and conversely, flows out of location l are obtained by summing over all destinations k.

among ex-ante identical workers. Spatial search frictions are absent if $mc = sc = \sigma_\varepsilon^2 = 0$.

If both types of search frictions are absent in our model, all workers are employed in jobs that offers the best combination of productivity and local amenities. By contrast, if standard frictions are absent, $p(\theta) = q(\theta) = 1$, but spatial search frictions are present, the model is similar to Kennan and Walker (2011). Because of the present spatial search frictions a worker might not be able to choose the preferred location (in terms of amenities and productivity). Such a model generates residual wage dispersion across locations while there is no dispersion within locations.

3.4 Parametrisations

We measure the worker's type x by educational attainment, thus yielding five groups. As regards the location-dependent match-specific productivity distribution, we assume that $\mathcal{F}_l(z)$ are Beta distributions that we discretize on a fixed number of points. In a discrete-time model, the matching functions $p(\theta)$ and $q(\theta)$ are required to be bounded between 0 and 1. We set

$$p(\theta) = 1 - \exp(-\theta).$$

Amenities v_l are unrestricted location fixed effects, as in Kennan and Walker (2011), and in contrast to Diamond (2016) who takes the complementary approach of enumerating explicitly specific dimensions of amenities.

In order to reduce the dimensionality of the problem, we consider the following parametric productivity function:

$$\log y(a, x, z) = \alpha_1 a + \alpha_2 a^2 + \alpha_x + z \tag{19}$$

α_1 and α_2 capture the effect of age on productivity. α_x is a worker's type effect. We fix the support of the Beta distributions to be between 0 and log(5) (the latter implied by our data). This implies that a worker in the least productive match can produce 5 times more if she were employed in the most productive match. As noted by Robin (2011), the multiplicative structure (here additive in logarithm) is parsimonious as the relative worker's productivity between two jobs only depends on the relative match type: $\frac{y(a,x,z)}{y(a,x,z')} = \exp(z - z')$. We account for possible measurement errors by adding a Gaussian noise to the theoretical wage to obtain the observed wage. The error term is centered and has a variance σ_w^2.

We specify search costs, incurred by all job searchers, as

$$sc_{alk}^s = \beta_1^{sc} + \beta_2^{sc}\mathbf{1}(s = e) + \beta_3^{sc} a + \beta_4^{sc}\mathbf{1}(k \in Bord(l)) + \beta_5^{sc}\mathbf{1}(l \neq k) + \beta^{sc,k} \tag{20}$$

The dummy variables $\mathbf{1}(k \in Bord(l))$ account for geographical proximity (as in e.g. Kennan and Walker, 2011), equalling 1 if the prospected region k shares a border with region l. Searching in another location, $\mathbf{1}(l \neq k)$, might be harder. This captures the idea that information flows about job opportunities might depend on geographical distance. The costs are also affine functions of age, and depend on employment status. Finally $\beta^{sc,k}$ corresponds to a set of prospected locations fixed effects (similar to Schmutz and Sidibé, 2018).

Moving costs, by contrast, are only incurred by workers who change location, and this cost might depend on distance. We specify

$$mc_{alk} = [\beta_1^{mc} + \beta_2^{mc} a + \beta_3^{mc} \mathbf{1}(k \in Bord(l))].\mathbf{1}(l \neq k) \tag{21}$$

The spatial unit in the text is the region, which has then led us to use dummy variables $\mathbf{1}(k \in Bord(l))$ to account for geographical proximity. In the next iteration of the paper, we will use much smaller spatial units in order to capture empirically better the idea of **local labour markets**. The cost functions will then be updated to include the distance between such LLMs.

Initial conditions The dynamic program starts at different ages depending on the education degree. We look at the first year of observation for each individual by education degree to guide our choice. We define different first ages $\underline{a}$ for the five groups, respectively 18, 18, 18, 20 and 23. Each worker starts as unemployed at the first age. We assume everybody retires at age $\bar{a} = 60$. For numerical and time constraints, we do not simulate the trajectories of 186,000 individuals. We will choose a smaller number of observations and we will draw their initial conditions (degree and first location) from the empirical distribution of the first observations. More precisely we determine the distribution of geographic location for workers in the first education group at age 18, second group at age 18, etc. Then we multiply this distribution to the distribution of education group in the population.

Following the literature, we take as exogenous two parameters: the interest rate that we fix at 5% annually, and the income as unemployed. We follow Schmutz and Sidibé (2018) and fix b at 3.36 euros to match 6000 euros annually.

4 Identification and Estimation by Indirect Inference

We estimate the model by Indirect Inference (II, Gourieroux et al., 1993), since this method allows us to map parsimoniously key quantities via auxiliary models across the life cycle (without resulting in a large number of moments or parameters). At the same time, the chosen auxiliary models allow us to focus specifically on key identifying information in our data. In appendix B, we develop a theoretical argument to show how the parameters of the model are identifed.

Indirect inference consists in determining an auxiliary model that offers an easily-computable statistical description of the phenomenon studied. The criterion to minimize is a distance between i) the auxiliary model estimated on observed data, and ii) the auxiliary model estimated on simulated data from the structural model. The method of simulated moments is a particular case of indirect inference when the auxiliary model contains the moments to match.

The choice of our auxiliary models is governed by the exploitation of identifying variations in the data, and how these relate to the parameters of our structural model. Essentially, first we use transitional data such as changes in labour market status, spatial transitions, and wage changes. These data provide rich identifying information, since we have labour market transitions within locations and across locations, which thus inform about search costs, moving costs, and incentives to change employment as captured by the location-specific job quality distribution $\mathcal{F}_l$. In particular, in Table 1 we have made explicit how specific transitions relate to the key model parameters. Second, we also consider stocks which further discipline the model. For instance the population shares across locations inform on the strength of local amenities.

We procede to set out the auxilliary models more precisely. To this end, denote, for each individual i aged a, by $l_{i,a}$ the location, by $E_{i,a}$ the employment indicator (equal to 1 if the worker is employed), by $w_{i,a}$ the wage, and by x_i the measure of educational attainment. We also denote by $E_{i,a}^{new}$ the dummy variable equalling 1 if the worker just starts a new job at age a (irrespective of whether she is unemployed or employed the previous year).

The first part of the auxiliary model focuses on the joint distribution of location, employment and wages $(l_{i,a}, E_{i,a}, E_{i,a}^{new}, w_{i,a})$ conditionally on age a and education x_i, but also conditionally on the past situation $(l_{i,a-1}, E_{i,a-1}, w_{i,a-1})$. The second part of the auxiliary model describes the joint distribution of $(l_{i,a}, E_{i,a}, w_{i,a})$ conditional only on age a and education x_i.[10]

//SUGGESTION: begin// We could proceed as follows:

A) within-job wage growth with job-specific fixed effects:

$$\Delta w_{i,a} = c_1^A w_{i,a} + \left(\sum_{k=1}^{3} c_{2,k}^A a^k \right) + c_{3,x(i)}^A + c_{4,j(i,a)}^A + u_{i,a}^A, \tag{22}$$

with $\Delta w_{i,a} = w_{i,a+1} - w_{i,a}$. This regression is motivated by equation (B1). Equation

[10] If we were using a method of moments, we would have an explosive number of moments. For instance, consider a very conservative discretisation into 3 groups of age (young, middle-aged, senior) and wage (low, medium, high). By generating moments from the number of observations in each group based on $(l_{i,a}, E_{i,a}, w_{i,a}, a, x_i, l_{i,a-1}, E_{i,a-1}, w_{i,a-1})$, we would obtain $21 \times 5 \times 3 \times 3 \times 5 \times 21 \times 3 = 297,675$ empirical moments for the first part of the auxiliary model. The use of II enables us to considerably reduce the number of moments.

(B1) gives:

$$\frac{y(a,x,z) - w(a,x,z,\rho)}{y(a+1,x,z) - w(a+1,x,z,\rho)} = \frac{y(a,x,z) - b + \tilde{D}_{al}(x,z)}{y(a+1,x,z) - b + \tilde{D}_{a+1,l}(x,z)} \equiv F(a,x,z,l)$$

Importantly, $F(a,x,z,l)$ does not depend on ρ. We obtain

$$w_{al}(x,z,\rho) - w_{a+1,l}(x,z,\rho) = \\ [F(a,x,z,l) - 1] w_{al}(x,z,\rho) + y(a+1,x,z) - F(a,x,z,l) y(a,x,z)$$

Hence, the form of our regression equation. In general, we cannot avoid capturing both the job type z and the bargaining rule ρ with job fixed effects in wage regressions. However, given the structure above, we can argue that the fixed effects $c^A_{4,j}$ only captures the unobserved job type z in our case.

The distribution of effects $c^A_{4,j}$ in each location l, and at the first year of observation (to avoid selection of z over job duration) is informative about the theoretical distributions $\mathcal{F}_l$. For each individual and each job held for which we observe at least two informations, we define the variable $\tilde{z}_j = \hat{c}^A_j$ based on the previous regression. We will use $\tilde{z}_j$ in the right-hand side of other regressions as if it was the true job type z

B) Wage regression at the beginning of the job spell:

$$\ln(w_{i,a}) = \left(\sum_{k=1}^{3} c^B_{1,k} a^k \right) + c^B_{2,x(i)} + c^B_3 E_{i,a-1} + c^B_4 \tilde{z}_{j(i,a-1)} E_{i,a-1} \\ + c^B_5 \tilde{z}_{j(i,a)} + c^B_{6,l(i,a-1),l(i,a)} + u^B_{i,a}, \tag{23}$$

According to theory, first wage in the spell only depends on age a, skill x, new job type z and the bargaining rule ρ. The bargaining rule is endogenous and depend also on age, skill, previous employment situation (whether employed, job type) and the two locations. We introduce paired fixed effects on locations (we could put more structure if we want).

The identification of mobility costs crucially relies on the distribution of effects $c^B_{6,l,l'}$. If the cost to move from l to l' is large, then we should observe a large gap between $c^B_{6,l,l'}$ and $c^B_{6,l,l}$.

C) Job transitions

$$E^{new}_{i,a} = \left(\sum_{k=1}^{3} c^C_{1,k} a^k \right) + c^C_{2,x(i)} + c^C_3 E_{i,a-1} + c^C_4 \tilde{z}_{j(i,a-1)} E_{i,a-1} + c^C_{5,l(i,a-1),l(i,a)} + u^C_{i,a}, \tag{24}$$

This equation is in the same spirit of the previous one. We are interested in the probability to change jobs. The same variable impacts the probabilities. The only difference is that there is no need to control for $\tilde{z}_{j(i,a)}$.

Both mobility costs and search costs affects the odds of obtaining a new job. If these costs are large from l to l', we should observe a large gap between $c^C_{5,l,l'}$ and $c^C_{5,l,l}$.

D) Job continuation, for employed workers ($E_{i,a} = 1$) only

$$E_{i,a} = \left(\sum_{k=1}^{3} c^D_{1,k} a^k\right) + c^D_{2,x(i)} + c^D_3 \tilde{z}_{j(i,a-1)} + c^D_{4,l(i,a)} + u^D_{i,a}, \tag{25}$$

In the model, these transitions are given by row 5 of Table 1. These employment equations convey information about job losses. In particular, parameters $c^D_{4,l}$ are informative about the δ_l.

//I find this stretagy better to the previous one: 1) The regressions are all motivated by structural equations and the identification proof. 2) We have potentially less parameters despite the pair fixed effects. There are only 4 regressions in total. 3) We can predict the qualitative features by looking at the estimates: $c^B_{6,l,l'}$ and $c^C_{5,l,l'}$ for moving and search costs 4) We have a good prior of $\mathcal{F}_l$ using the distributrion of $c^A_{4,j}$ per location. //
//SUGGESTION: end//

4.1 Changes: Transitions and wage dynamics

We first decompose the conditional joint distribution of $(l_{i,a}, E_{i,a}, E^{new}_{i,a}, w_{i,a}|a, x_i, l_{i,a-1}, E_{i,a-1}, w_{i,a-1})$, and then study these decompositions using linear probability models. The distribution of interest is decomposed into :

- the distribution of employment transitions $(E^{new}_{i,a}|a, x_i, l_{i,a-1}, E_{i,a-1}, w_{i,a-1})$ and the distribution of employment continuations $(E_{i,a}|a, x_i, l_{i,a-1}, E_{i,a-1} = 1, w_{i,a-1})$;
- the distribution of spatial transitions $(l_{i,a}|a, x_i, l_{i,a-1}, E_{i,a-1}, w_{i,a-1}, E^{new}_{i,a} = 1)$;
- the distribution of wage changes $(w_{i,a}|a, x_i, l_{i,a-1}, E_{i,a-1}, w_{i,a-1}, E^{new}_{i,a} = 1, l_{i,a})$.

As $E^{new}_{i,a} = 1$ implies $E_{i,a} = 1$ by construction, one can check that the first distributions in the list captures all the five possible combinations of $(E_{i,a}, E^{new}_{i,a}, E_{i,a-1})$. We use an intuitive decomposition order. Employment change precedes location change because there is no location change without an employment change in the model and in the way data are constructed.

We then estimate the following linear probability model, grouping individuals by previous location, which describes the transitions towards new employment spells.

$$E_{i,a}^{new} = \left(\sum_{k=1}^{3} c_{1,l,k}^{A} a^{k} \right) + c_{2,l,x(i)}^{A} + c_{3,l}^{A} E_{i,a-1} + c_{4,l}^{A} \ln(w_{i,a-1}).E_{i,a-1} + u_{i,a,l}^{A},$$
$$\text{by group } l_{i,a-1} = l. \tag{26}$$

In our model, these transitions are given by rows 1 and 3 of Table 1 after summing over source and destination locations. The explanatory variables are respectively a cubic specification for age, education fixed effects, employment situation the year before and log wage the year before if the worker was employed. We do not control here for any geographical moves. However, given the movements in our data, the estimated auxiliary parameters are mainly driven by within-location determinants. Since the worker must have searched, the firm must have opened a vacancy, and the employed worker must have preferred to change employer, these auxiliary parameters provide information on location-invariant search costs β_1^{sc}, β_3^{sc} and β_4^{sc}, and the cost of vacancies. They also convey information about the job quality distribution $\mathcal{F}_l$.

Next, we estimate a linear probability model for employment status for the subset of workers previously employed, $E_{i,a-1} = 1$,

$$E_{i,a} = \left(\sum_{k=1}^{3} c_{1,l,k}^{B} a^{k} \right) + c_{2,l,x(i)}^{B} + c_{3,l}^{B} \ln(w_{i,a-1}) + u_{i,a,l}^{B}, \qquad \text{by group } l_{i,a-1} = l. \tag{27}$$

In the model, these transitions are given by row 5 of Table 1 after integrating out z. These employment equations convey information about job losses. They help identifying the exogenous job separation shocks δ_l, and the search frictions parameters that govern the probability to avoid unemployment after a job separation shock.

Next, we turn to spatial transitions. Spatial transitions require a change of employer, so will be studied on the subsample for which $E_{i,a}^{new} = 1$. We proceed in two steps. We regress first the event of a change of location $\mathbf{1}(l_{i,a} \neq l_{i,a-1})$, and then we regress the presence in location $\mathbf{1}(l_{i,a} = l)$ controlling for the presence in the region the previous year

$\mathbf{1}(l_{i,a-1} = l)$ for the same l:[11]

$$\mathbf{1}(l_{i,a} \neq l_{i,a-1}) = \left(\sum_{k=1}^{3} c^C_{1,l,k} a^k\right) + c^C_{2,l,x(i)} + c^C_{3,l} E_{i,a-1} + c^C_{4,l} \ln(w_{i,a-1}).E_{i,a-1} + u^C_{i,a,l}, \quad \text{by group } l_{i,a-1} = l; \tag{28}$$

$$\mathbf{1}(l_{i,a} = l) = \left(\sum_{k=1}^{3} c^D_{1,l,k} a^k\right) + c^D_{2,l,x(i)} + c^D_{3,l} E_{i,a-1} + c^D_{4,l} \ln(w_{i,a-1}).E_{i,a-1} + c^D_{5,l}\mathbf{1}(l_{i,a-1} = l) + u^D_{i,a,l}, \quad \text{for each } l. \tag{29}$$

The first series of equations captures the odds of leaving location l, whereas the second series captures the odds of arriving in location l. In our model, these location inflows and outflows are given by Table 1 after summing over source, respectively destination locations. The parameters help identifying the spatial search frictions β^{mc}, $\beta_2^{sc,k}$, $\beta_5^{sc,k}$ and $\beta^{sc,k}$. They also inform on location attractiveness, whether it is market-related such as the firms' type distribution, or non-market related such as amenities.

We then proceed to consider the distribution of wage changes following a job change, and estimate the equation

$$\ln(w_{i,a}) = \left(\sum_{k=1}^{3} c^E_{1,l,k} a^k\right) + c^E_{2,l,x_i} + c^E_{3,l} E_{i,a-1} + c^E_{4,l} \ln(w_{i,a-1}).E_{i,a-1} + c^E_{5,l}\mathbf{1}(l_{i,a-1} = l) + u^E_{i,a,l}, \quad \text{by group } l_{i,a} = l. \tag{30}$$

The parameters convey information the choice of submarket within a location. In the theoretical model, search costs affects the choice of a prospected location, but not the choice of a submarket. The parameters therefore inform on mobility costs β^{mc} but not search costs β^{sc}. The combination of the location change from (28) and (29), and wage changes from (30) helps to identify separately the spatial search costs from the mobility costs. The parameters are also informative on the firms's type distribution.

Lastly, we build the location-specific distributions of residuals $(u^A_{i,a,l}, u^D_{i,a,l}, u^E_{i,a,l})$ and add the mean and covariance matrices to the set of auxiliary parameters. We end up with 1,260 parameters for this part of the auxiliary model.

4.2 Levels: The distribution of workers and wages

We now complement the exploitation of transitional data by information captured in levels. Specifically, to study the distribution of $(l_{i,a}, E_{i,a}, w_{i,a}|a, x_i)$, we first decompose the joint distribution into successive conditional distributions: the distribution of locations $l(i,a)|a, x_i$, the distribution of employment status conditional on location $E_{i,a}|a, x_i, l_{i,a}$,

[11] We voluntarily adopt the semantics of statistical softwares for "by" and "for".

and the conditional distribution of wages $w_{i,a}|a, x_i, l_{i,a}, E_{i,a} = 1$. These distributions are informative about the spatial selection of workers across locations on labour incomes. Workers might remain in apparently unattractive locations because they have obtained a high-quality job, or move to or remain in a lower productivity location if the amenities are high.

The auxiliary models then consist in the following the multinomial regression model:

$$\mathbf{1}(l_{i,a} = l) = \left(\sum_{k=1}^{3} c^F_{1,l,k} a^k\right) + c^F_{2,l,x_i} + u^F_{i,a,l}, \qquad \text{for each } l. \tag{31}$$

In the model, the choice probabilities are given by equations 12 and 13. The dependent variable is the dummy equal to 1 if l is the individual's current location. We estimate a cubic polynomial in age and fixed effects of education attainment.

We then estimate a linear probability model for the employment status, grouping individuals by locations:

$$E_{i,a} = \left(\sum_{k=1}^{3} c^G_{1,l,k} a^k\right) + c^G_{2,l,x_i} + u^G_{i,a,l}, \qquad \text{by group } l_{i,a} = l. \tag{32}$$

For the subsample of employed workers, $E_{i,a} = 1$, we estimate Mincer equations by locations:

$$\ln(w_{i,a}) = \left(\sum_{k=1}^{3} c^H_{1,l,k} a^k\right) + c^H_{2,l,x_i} + u^H_{i,a,l}, \qquad \text{by group } l_{i,a} = l. \tag{33}$$

For each location l, we define the joint distribution of the residuals $(\hat{u}^F_{i,a,l}, \hat{u}^H_{i,a,l})$ for the subsample of employed workers in location l. We add the mean and the covariance matrix of the joint distribution for each location to the set of auxiliary parameters. We end up with 609 coefficients.

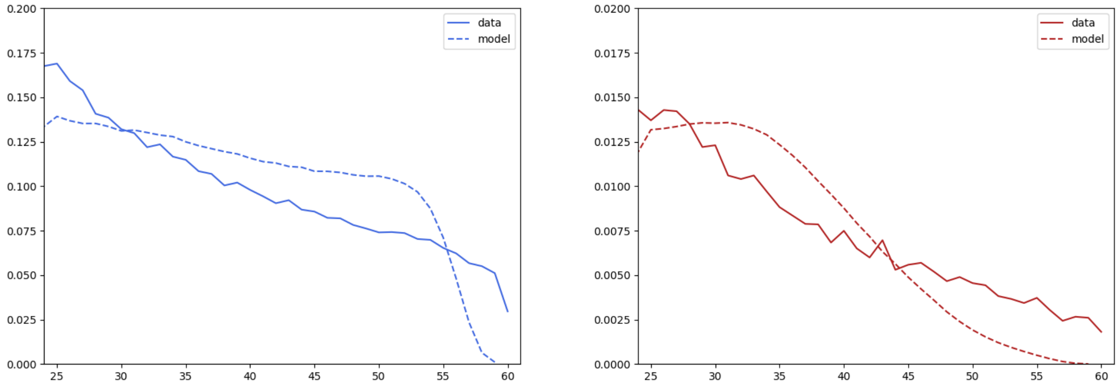
5 Empirical results

We first present how the model fits aggregate mobility rates. Then, the impact of spatial search frictions is assessed by measuring how mobility rates evolve when parameters of the spatial search frictions deviate from the estimates.

5.1 Model Fit

We measure the fit of our estimated model by comparing empirical aggregate mobility rates in the estimation sample to their expected values from the model. Aggregate mobility rates are not targets of the indirect inference estimation. This comparison exercise

Figure 2: Model fit of aggregate mobility rates



(a) Job mobility

(b) Geographic mobility

Notes. Panel (a) shows job mobility, defined as the share of workers who obtain a new job during the year, in the selected sample and in the estimated model. Panel (b) shows geographical mobility, defined as the share of workers who change job and location within a year, between the 21 continental regions.

therefore measures the ability of our model, estimated on micro-level data, to match macroeconomic variables. Figure 2 shows the fit on job mobility and geographical mobility at ages between 25 to 60 years. The model matches the trend of mobility rates over age. In the estimated model is less accurate for senior workers by predicting too few mobility. In the theory, workers close to retirement have low incentives to obtain a job if they are unemployed or to make a job-to-job transitions. In practice, however, pensions in France depend on the last periods of employment.

5.2 Mobility costs versus search costs

The average mobility costs is equivalent to a 32-euro hourly income, or 58.000 euros yearly, in the estimated model. By comparison, Kennan and Walker (2011) estimate mobility costs of 300,000 dollars for movers, and assumes non-movers have an infinite cost. We obtain much smaller mobility costs in the line of Schmutz and Sidibé (2018) who find mobility costs close to 15,000 euros. Our results thus support the idea that accounting for spatial search frictions reduce the estimates of mobility costs. We obtain larger estimates than Schmutz and Sidibé (2018) possibly because we do not use the same spatial units. They consider cities whereas we consider regions (at this stage of the paper). In our settings, a move within the same region is not accounted as a geographical move. In other words, mobility costs are higher in our case because we observe less geographic moves.

Our theoretical model defines search costs as monetary. We should be careful in interpreting these parameters in absolute terms, as they do not necessarily capture "real" search costs. We can, however, compare the spatial part of monetary search costs, β_5^{sc}, to average monetary mobility costs. This comparison is informative about the strenght of spatial search frictions. The additional cost of searching for a job in anothe region, β_5^{sc}, is estimated at 436 euros. This spatial search cost is approximately 13 times the average mobility costs. Workers are mainly constrained geographically by spatial search costs.

The spatial components of search costs may seem extremely high, but it is in the same magnitude as baseline search costs. The average search costs, conditionnally on looking for a job in the same region, is 76 euros. In addition, we estimate the dispersion parameters of random taste shocks σ_ε at 62. As individuals select the best option, taste shocks contribute in reducing realized search costs.

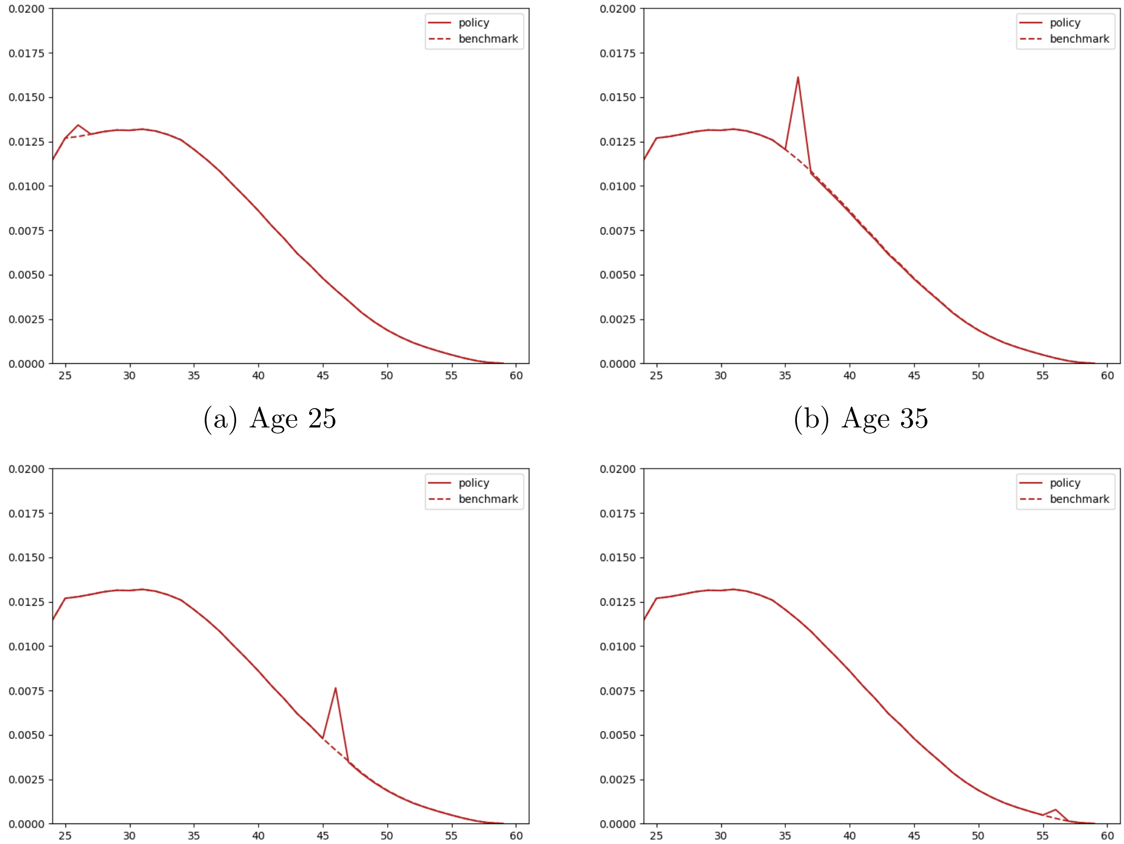
5.3 The effects of age on geographic mobility

We estimate the linear coefficient of age on mobility costs, β_2^{mc}, at 2.1. Young workers face much lower mobility costs than senior workers. For instance, 25-year-old workers have average mobility costs equal to 1200 euros annually, compared to 50-year-old workers at $96,000$ euros. This increasing patterns of mobility costs with age does not mean necessarily that young workers are less constrained by mobility costs than old workers. Workers have also different incentives to move, and age reduces the incentives for senior workers to relocate in more productive areas.

To measure how constraining mobility costs are, we compare mobility patterns in counterfactual scenarii with different mobility costs. We choose a threshold age below which workers are not affected, their trajectories are simulated from the estimated model. At the threshold age, workers unexpectedly discover a change in mobility costs, which modify their search choices and their trajectories. We repeat this procedure for different threshold ages. The block-recursion property of the equilibrium is crucial here. When a group of workers is affected by a change in parameters, like diminished costs, other workers are not affected. There is no general equilibrium externality. We can thus focus on the trajectory of this subgroup.

Figure 3 shows the mobility patterns when workers unexpectedly face zero mobility costs only for one period. Young and senior workers barely alter their search strategies compared to middle-aged workers. The reason is the following. Young workers have already low mobility costs. The policy has therefore a small impact on their behaviours. The oldest workers on the other side face high mobility costs, but they also have less incentives to relocate due to the time horizon effect. In addition, older workers are also more likely to be higher on the job ladder, i.e. working in a job of high type z. Older workers are thus more selective, they choose submarkets with a higher promised utility and lower job-finding rates. This mechanism contributes to the low impact on geographic mobility. Middle-aged workers have sufficient incentives to relocate geographically, compared to

Figure 3: Aggregate geographical mobility, after an unexpected temporary disparition of mobility costs



(a) Age 25 (b) Age 35

(c) Age 45 (d) Age 55

Notes. Each panel corresponds to change in mobility costs at different age. The policy is a drop in mobility costs to zero. This decrease occurs only at the age considered.

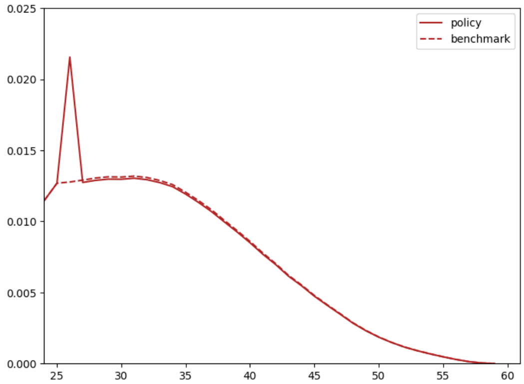
senior workers.

5.4 Mobility and dynamic patterns

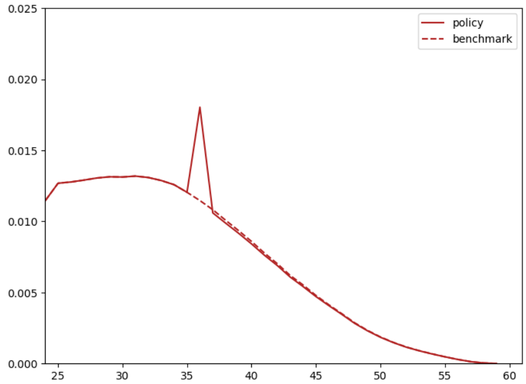
We discuss the predictions of our theoretical model in the light of a counterfactual policy. In figure 3, workers in the four cases are exposed to a subsidy to mobility costs that depends on age. Were, we assume that workers receive the same subsidy irrespective of age. We fix this subsidy arbitrarily at the average mobility costs, 32 euros.

Figure 4 emphasizes an intuitive result. As the youngest workers have the highest incentives to relocate, they react the most to a given subsidy. Our model is rich enough to have long-term impact of such a short-term policy. According to the model, the policy functions, meaning optimal search decisions and vacancy postings, are not different in the benchmark and the coounterfactual scenarii *after* the age threshold. This is due in particular to the block-recursion property. In other words, transitions rates are the

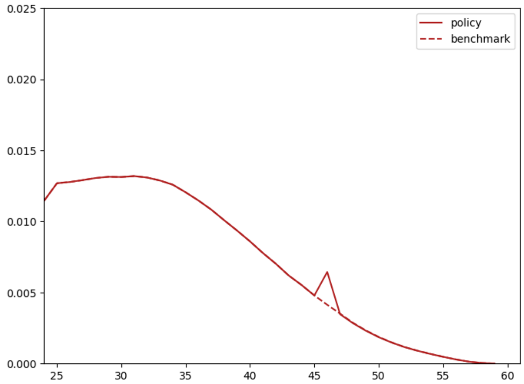
Figure 4: Aggregate geographical mobility, after an unexpected temporary change of mobility costs



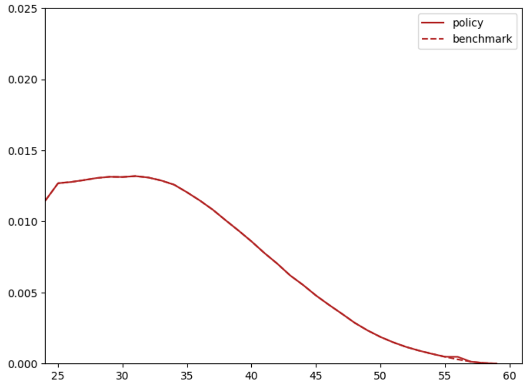
(a) Age 25



(b) Age 35



(c) Age 45



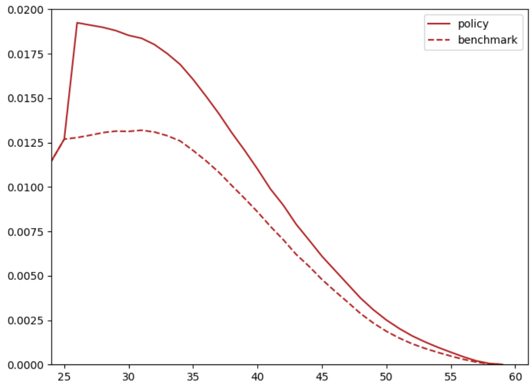
(d) Age 55

Notes. Each panel corresponds to change in mobility costs at different age. The policy is a reduce in mobility costs by 32 euros.

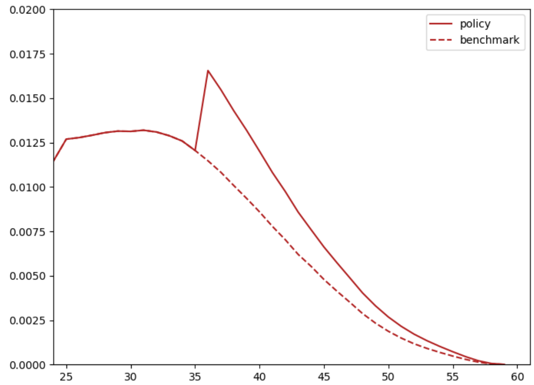
same after the age threshold. The only reason why the two curves may not superimpose comes from a distributional effect. The policy at the age considered, 25 for instance, has an impact on the distribution of workers across locations at age 26. A lower mobility following age 26 means that workers are allocated in states that involves less geographic mobility, compared to the benchmark. In our case, we only observe a small effecy on panels (a) and (b).

In the scenario of figure 4, workers can not benefit from the relocation subsidy in the subsequent periods. We consider the same policy assuming that workers are eligible at any time to this policy. Figure 5 suggests evidence of delaying behaviours. When the subsidy is temporary, workers have incentives to be less selective in searching for a job abroad to benefit from the one-chance subsidy. When the subsidy is permanent, the increase in geographic mobility is lower at the age threshold. Geographic mobility is increased in the subsequent periods.

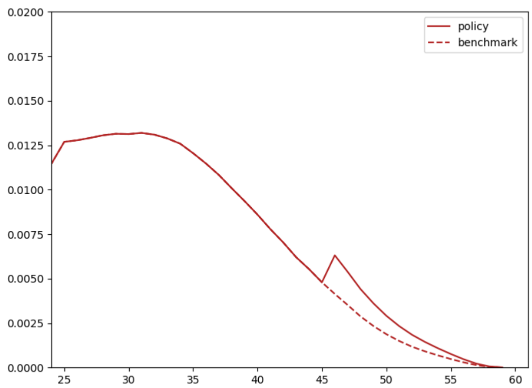
Figure 5: Aggregate geographical mobility, after an unexpected permanent change of mobility costs



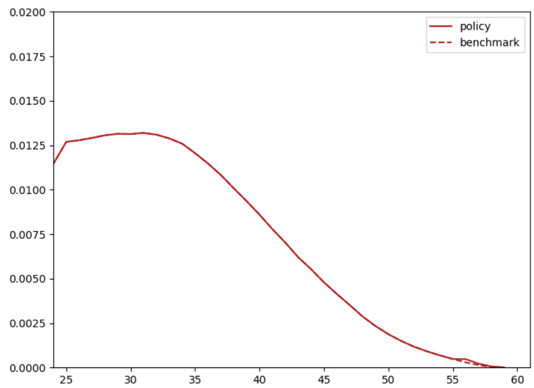
(a) Age 25



(b) Age 35



(c) Age 45



(d) Age 55

Notes. Each panel corresponds to change in mobility costs at different age. The policy is a permanent reduce in mobility costs by 32 euros.

6 Conclusion

This paper investigates how search frictions, within and between local labour markets, explain dynamic location choice decisions. We model the life-cycle location choice problem of workers in presence of spatial search frictions. Spatial search frictions constitute a barrier to mobility, distinct to mobility costs.

Age plays a key role in mobility decisions. Identical workers that differ only in age discount differently the gains from mobility as an investment. In the meantime, age captures individual attachments or inertia, in absence of household characteristics such as family size or homeownership. As age affects both gains and costs to mobility, ignoring it and only considering average mobility or search costs can mask important heterogeneities for public public efficiency. In Europe, youth unemployment indeed reaches very high levels. Our results suggest that the youngest workers are not the most constrained by mobility costs, but respond the most to public policies subsidizing mobility.

References

J. M. Abowd, F. Kramarz, and D. N. Margolis. High Wage Workers and High Wage Firms. *Econometrica*, 67(2):251–334, March 1999.

D. Acemoglu and R. Shimer. Holdups and Efficiency with Search Frictions. *International Economic Review*, 40(4):827–849, 1999.

M. Amior and A. Manning. The Persistence of Local Joblessness. *American Economic Review*, 108:1942–1970, 07 2018.

J. Bagger, F. Fontaine, F. Postel-Vinay, and J.-M. Robin. Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics. *American Economic Review*, 104(6):1551–96, June 2014.

G. Barlevy. Identification of Search Models using Record Statistics. *Review of Economic Studies*, 75(1):29–64, 2008.

N. Baum-Snow and R. Pavan. Understanding the City Size Wage Gap. *Review of Economic Studies*, 79(1):88–127, 2012.

K. Burdett and D. T. Mortensen. Wage Differentials, Employer Size, and Unemployment. *International Economic Review*, 39(2):257–73, May 1998.

M. Caliendo, S. Kunn, and R. Mahlstedt. The return to labor market mobility: An evaluation of relocation assistance for the unemployed. *Journal of Public Economics*, 148:136–151, 2017.

A. Chéron, J.-O. Hairault, and F. Langot. Life-cycle equilibrium unemployment. *Journal of Labor Economics*, 31(4):843–882, 2013.

P.-P. Combes, G. Duranton, and L. Gobillon. Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2):723–742, March 2008.

B. Decreuse and A. Zylberberg. Search Intensity, Directed Search, And The Wage Distribution. *Journal of the European Economic Association*, 9(6):1168–1186, December 2011.

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

L. Gobillon and F. Wolff. The return to labor market mobility: An evaluation of relocation assistance for the unemployed. *Urban Studies*, 48:331–347, 2011.

E. D. Gould. Cities, Workers, and Wages: A Structural Analysis of the Urban Wage Premium. *Review of Economic Studies*, 74(2):477–506, 2007.

C. Gourieroux, A. Monfort, and E. Renault. Indirect Inference. *Journal of Applied Econometrics*, 8(S):85–118, Suppl. De 1993.

G. Jolivet, F. Postel-Vinay, and J.-M. Robin. The empirical content of the job search model: Labor mobility and wage distributions in Europe and the US. *European Economic Review*, 50(4):877–907, May 2006.

J. Kennan and J. R. Walker. The Effect of Expected Income on Individual Migration Decisions. *Econometrica*, 79(1):211–251, 01 2011.

R. Lentz and E. R. Moen. Competitive or Random Search? 2017 Meeting Papers 1128, Society for Economic Dynamics, 2017.

D. McFadden. Conditional logit analysis of qualitative choice behaviour. In P. Zarembka, editor, *Frontiers in Econometrics*, pages 105–142. Academic Press New York, New York, NY, USA, 1973.

G. Menzio and S. Shi. Directed Search on the Job, Heterogeneity, and Aggregate Fluctuations. *American Economic Review*, 100(2):327–32, May 2010.

G. Menzio and S. Shi. Efficient Search on the Job and the Business Cycle. *Journal of Political Economy*, 119(3):468–510, 2011.

G. Menzio, I. Telyukova, and L. Visschers. Directed Search over the Life Cycle. *Review of Economic Dynamics*, 19:38–62, January 2016. doi: 10.1016/j.red.2015.05.002.

E. R. Moen. Competitive Search Equilibrium. *Journal of Political Economy*, 105(2): 385–411, April 1997.

R. Molloy, C. L. Smith, and A. Wozniak. Declining Migration within the US: The Role of the Labor Market. *Finance and Economics Discussion Series, Federal Reserve Board, Washington, D.C.*, 2014.

E. Moretti. Local Labor Markets. volume 4 of *Handbook of Labor Economics*, chapter 14, pages 1237–1313. 2011.

P. Nanos and C. Schluter. Search Across Local Labour Markets. *Working Paper, Aix-Marseille School of Economics and University of Sheffield*, 2018.

F. Postel-Vinay and J.-M. Robin. Equilibrium Wage Dispersion with Worker and Employer Heterogeneity. *Econometrica*, 70(6):2295–2350, November 2002.

J.-M. Robin. On the Dynamics of Unemployment and Wage Distributions. *Econometrica*, 79(5):1327–1355, 09 2011.

J. Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033, September 1987.

B. Schmutz and M. Sidibé. Frictional labor mobility. *The Review of Economic Studies*, page rdy056, 2018.

G. Wilemme. Optimal Taxation to Correct Job Mismatching. AMSE Working Papers 1723, Aix-Marseille School of Economics, Marseille, France, June 2017.

Appendices

The structure of this Appendix is as follows:

- Appendix A presents further mathematical details.
- In Appendix B, we explain how the model is identified using both wage and transition data.
- In Data Appendix C, we describe in greater detail our administrative data, and the sample selection criteria.

A Appendix: Mathematical Details

A.1 Proof of Proposition 1

We show that the equilibrium is uniquely defined recursively from the termination conditions, and that the distribution of workers across locations and jobs do not intervene. Consider the final age $a = \bar{a}$. The value functions $U_{\bar{a}l}(x)$ and $V_{\bar{a}l}(x, z)$ are uniquely defined by the termination conditions for any x, l and z. they do not depend on the distribution of workers.

Then $\theta^u_{\bar{a}-1,l,k}(x)$ and $\theta^e_{\bar{a}-1,l,k}(x, z)$ are uniquely defined as solutions to (10) and (11): if the term within brackets is positive, market tightness is an interior solution because $p(.)$ is concave; if the term is negatrive, market tightness is zero. $\theta^u_{\bar{a}-1,l,k}(x)$ and $\theta^e_{\bar{a}-1,l,k}(x, z)$ only depend on $U_{\bar{a}l}(x)$ and $V_{\bar{a},l}(x, z)$, so they do not depend on the distribution of workers.

Knowing market tightness at $\bar{a}-1$, we can recover the promised values that job seekers choose, $W^u_{\bar{a}-1,l,k}(x)$ and $W^e_{\bar{a}-1,l,k}(x, z)$, from equation (1):

$$W^u_{\bar{a}-1,l,k}(x) = \int V_{ak}(x, z) d\mathcal{F}_k(z) - \frac{\mu_k}{q(\theta^u_{\bar{a}-1,l,k}(x))}, \tag{A1}$$

$$W^e_{\bar{a}-1,l,k}(x, z) = \int V_{ak}(x, z) d\mathcal{F}_l(z) - \frac{\mu_k}{q(\theta^e_{\bar{a}-1,l,k}(x, z))}. \tag{A2}$$

When market tightness θ, the submarket is not active, and so the promised utility W does not matter. Again $W^u_{\bar{a}-1,l,k}(x)$ and $W^e_{\bar{a}-1,l,k}(x, z)$ do not depend on the distribution of workers.

The next step consists in recovering $R^u_{\bar{a}-1,l,k}(x)$ and $R^e_{\bar{a}-1,l,k}(x, z)$ using (4) and (7), and thus the location choices $K^u_{\bar{a}-1,l,k}(x, \boldsymbol{\varepsilon})$ and $K^e_{\bar{a}-1,l,k}(x, z, \boldsymbol{\varepsilon})$ from (3) and (6). These four functions are uniquely characterized.The only possible indeterminacy is when two locations provide the same utility for particular draws of $\boldsymbol{\varepsilon}$. In that case, workers are indifferent between the two options. However, we only need to define $K^u_{\bar{a}-1,l,k}(x, \boldsymbol{\varepsilon})$ and $K^e_{\bar{a}-1,l,k}(x, z, \boldsymbol{\varepsilon})$ *almost everywhere* (in probabilistic terms) on the space of $\boldsymbol{\varepsilon}$. The events of being indifferent between two options have a zero measure.

Equations (2) and (5) then can be used to recover $U_{\bar{a}-1,l}(x)$ and $V_{\bar{a}-1,l}(x, z)$. They are uniquely defined, and do not depend on the distribution of workers. We can thus repeat the same procedure, until defining the variables at the initial age.

Once the value functions are characterized, it is straightforward to uniquely define the wage functions and the bargaining rules from equation (8) and (9).

A.2 Location choice and expected utility from the Gumbel distributions

We show how to obtain equations (12), (13) and (16). We drop the notations a, l, s and x for clarity. We denote $Z_k = R_k - sc_k$. The p.d.f. of ε_k is $f_\varepsilon(x) = \frac{1}{\sigma_\varepsilon}\exp(-\frac{x}{\sigma_\varepsilon} - euler).\exp(-\exp(-\frac{x}{\sigma_\varepsilon} - euler))$. We define for $k = 0, .., L$,

$$\mathbb{P}_\varepsilon(K(\boldsymbol{\varepsilon}) = k) = \int \left[\prod_{j\neq k} \mathbb{P}_{\varepsilon_j}(\varepsilon_j < \varepsilon_k + Z_k - Z_j)\right] f_\varepsilon(\varepsilon_k) d\varepsilon_k, \tag{A3}$$

$$\mathbb{E}_\varepsilon \max_{k\in\{0,..,L\}}\{Z_k + \varepsilon_k\} = \sum_{k=0}^{L} \int (Z_k + \varepsilon_k) \left[\prod_{j\neq k} \mathbb{P}_{\varepsilon_j}(\varepsilon_j < \varepsilon_k + Z_k - Z_j)\right] f_\varepsilon(\varepsilon_k) d\varepsilon_k. \tag{A4}$$

Given $\mathbb{P}_{\varepsilon_j}(\varepsilon_j < \varepsilon_k + Z_k - Z_j) = \exp(-\exp(-\frac{\varepsilon_k + Z_k - Z_j}{\sigma_\varepsilon} - euler))$, it follows

$$\begin{aligned}\left[\prod_{j\neq k} \mathbb{P}_{\varepsilon_j}(\varepsilon_j < \varepsilon_k + Z_k - Z_j)\right] f_\varepsilon(\varepsilon_k) &= \frac{1}{\sigma_\varepsilon}\exp(-\frac{\varepsilon_k}{\sigma_\varepsilon} - euler) \\ &\times \exp\left(-\left[1 + \sum_{j\neq k}\exp(-\frac{Z_k - Z_j}{\sigma_\varepsilon})\right]\exp(-\frac{\varepsilon_k}{\sigma_\varepsilon} - euler)\right).\end{aligned} \tag{A5}$$

Denote $A_k = \frac{1}{1 + \sum_{j\neq k}\exp\left(-\frac{Z_k - Z_j}{\sigma_\varepsilon}\right)}$. We can write

$$\begin{aligned}\left[\prod_{j\neq k} \mathbb{P}_{\varepsilon_j}(\varepsilon_j < \varepsilon_k + Z_k - Z_j)\right] f_\varepsilon(\varepsilon_k) &= A_k.\frac{1}{\sigma_\varepsilon}\exp(-\frac{\varepsilon_k}{\sigma_\varepsilon} - euler - \log A_k) \\ &\times \exp\left(-\exp(-\frac{\varepsilon_k}{\sigma_\varepsilon} - euler - \log A_k)\right).\end{aligned} \tag{A6}$$

The term that follows A_k on the right-hand side is the p.d.f. of another Gumbel distribution. Its integral equals 1 and its mean is $-\sigma_\varepsilon \log A_k$. The probability and the expectation writes

$$\mathbb{P}_\varepsilon(K(\boldsymbol{\varepsilon}) = k) = A_k, \tag{A7}$$

$$\mathbb{E}_\varepsilon \max_{k\in\{0,..,L\}}\{Z_k + \varepsilon_k\} = \sum_{k=0}^{L} A_k(Z_k - \sigma_\varepsilon \log A_k). \tag{A8}$$

Using $\log A_k = \frac{Z_k}{\sigma_\varepsilon} - \log\left(\sum_{j=0}^{L}\exp(\frac{Z_j}{\sigma_\varepsilon})\right)$ and $\sum_{k=0}^{L} A_k = 1$,

$$\mathbb{E}_\varepsilon \max_{k\in\{0,..,L\}}\{Z_k + \varepsilon_k\} = \sigma_\varepsilon \log\left(\sum_{j=0}^{L}\exp(\frac{Z_j}{\sigma_\varepsilon})\right). \tag{A9}$$

$Z_0 = 0$ because there no gains nor costs in remaining idle. We therefore obtain equations (12), (13) and (16).

B Identification

In this appendix we show how the parameters of the structural model are identified using individual level data on labour market transitions and wage changes across the life cycle, and across (and within) locations.

To simplify the exposition, we first take as given the latent job type variable z, and address the identification of its distribution later by considering within-job wage growth. We also take as given the worker's observable type x. We first define some recurrent objects, namely

- the match surplus $\Omega_{al}(x,z) = V_{al}(x,z) - U_{al}(x)$;
- the expected returns to search for employed and unemployed, using equation (16),

$$S^e_{al}(x,z) = \sigma_\varepsilon \log\left(1 + \sum_{k=1}^{L} \exp(R^e_{alk}(x,z) - sc^e_{alk})^{\frac{1}{\sigma_\varepsilon}}\right)$$
$$S^u_{al}(x) = \sigma_\varepsilon \log\left(1 + \sum_{k=1}^{L} \exp(R^u_{alk}(x) - sc^u_{alk})^{\frac{1}{\sigma_\varepsilon}}\right);$$

- using equations (17) and (18), the surplus differentials are for ages $a < \bar{a}$

$$D_{al}(x,z) = \frac{1-\delta_l}{1+r}\Omega_{a+1l}(x,z) + (1-\delta_l)(S^e_{al}(x,z) - S^u_{al}(x))$$
$$\tilde{D}_{al}(x,z) = \frac{(1-\delta_l)\mathbb{E}_\varepsilon p(\theta^e_{alK^e_{al}(x,z,\varepsilon)}(x,z))}{1+r}\Omega_{a+1l}(x,z) + (1-\delta_l)(S^e_{al}(x,z) - S^u_{al}(x)),$$

  so $D_{al}(x,z)$ interprets as the extra surplus differential coming from i) the discounting of the job separation risk, ii) the differential in the returns to search. Note that $D_{\bar{a}l}(x,z) = D_{\bar{a}l}(x,z) = 0$.

We then have

$$\Omega_{al}(x,z) = y(a,x,z) - b + D_{al}(x,z) \qquad \text{for } a < \bar{a}$$
$$\Omega_{\bar{a}l}(x,z) = y(\bar{a},x,z) - b$$

With these defintions on hand, we can re-write the wage equation, using equation (8) and the Nash bargaining $V^F_{al}(x,z,\rho) = (1-\rho)\Omega_{al}(x,z)$,

$$w_{al}(x,z,\rho) = y(a,x,z) - (1-\rho)\left[y(a,x,z) - b + \tilde{D}_{al}(x,z)\right], \tag{B1}$$

for $a < \bar{a}$ with $w_{\bar{a}l}(x,z,\rho) = y(\bar{a},x,z) + (1-\rho)(y(\bar{a},x,z) - b)$.

Wages following a transition from unemployment Consider a worker who is unemployed at age $a_0 - 1$ (with $a_0 < \bar{a}$) in location l, and who finds a job at age a_0 in location k of type z. By combining equations (1), (9) and (B1), the wage in this job is for any subsequent age $a \geq a_0$,

$$y(a, x, z) - \frac{y(a, x, z) - b + \tilde{D}_{al}(x, z)}{\int [y(a_0, x, z) - b + D_{a_0 l}(x, z)] \, d\mathcal{F}_k(z)} \cdot \frac{\mu_k}{q(\theta^u_{a_0 lk}(x))}. \tag{B2}$$

Wages following a transition from employment Consider a worker who is employed at age $a_0 - 1$ (with $a_0 < \bar{a}$) in location l with job type z_0, and who finds a job of type z at age a_0 in location k. We do not observe the job destruction event as defined in the model, because workers have time to find a job before loosing the job. Taking the expectation over this event, the wage in this job is, for any subsequent age $a \geq a_0$,

$$y(a, x, z) - \frac{y(a, x, z) - b + \tilde{D}_{al}(x, z)}{\int [y(a_0, x, z) - b + D_{a_0 l}(x, z)] \, d\mathcal{F}_k(z)} \left(\delta_l \frac{\mu_k}{q(\theta^u_{a_0 lk}(x))} + (1 - \delta_l) \frac{\mu_k}{q(\theta^e_{a_0 lk}(x, z_0))} \right). \tag{B3}$$

Workers obtain as a wage the job productivity minus a share of the job creation cost.

B.1 Exploiting wage changes after transitions and life-cycle variations

We proceed by backward induction. Starting with the final transition between $\bar{a} - 1$ and $\bar{a}$, we characterize wages following a job change or a transition to employment.

Identification from wages of new jobs at age $\bar{a}$ The returns from searching $R_{\bar{a}-1lk}$ and the choice of a submarket $\theta_{\bar{a}-1lk}$ solve, using equations equations (17) and (18):

$$R^u_{\bar{a}-1lk} = \max_{\theta \geq 0} \left\{ p(\theta) \left[\frac{\bar{y}_{\bar{a}k} - b + v_k - v_l}{1 + r} - mc_{\bar{a}-1lk} \right] - \theta \frac{\mu_k}{1 + r} \right\}, \tag{B4}$$

$$R^e_{\bar{a}-1lk}(z) = \max_{\theta \geq 0} \left\{ p(\theta) \left[\frac{\bar{y}_{\bar{a}k} - y(\bar{a}, z) + v_k - v_l}{1 + r} - mc_{\bar{a}-1lk} \right] - \theta \frac{\mu_k}{1 + r} \right\}, \tag{B5}$$

where $\bar{y}_{al} = \int y(a, z) d\mathcal{F}_l(z)$ denotes the expected productivity. Obtain the first order conditions, solve for θ, and substitute into equations (B2) and (B3). In particular, the worker receives a wage

$$y(\bar{a}, z) - \frac{y(\bar{a}, z) - b}{\bar{y}_{\bar{a}k} - b} \cdot \frac{\mu_k}{q\left(p'^{-1}\left(\frac{\mu_k}{\bar{y}_{\bar{a}k} - b + v_k - v_l - (1+r) mc_{\bar{a}-1lk}}\right)\right)}, \tag{B6}$$

if the worker was unemployed at age $\bar{a}-1$; if at age $\bar{a}-1$ the worker was employed, he will now receive

$$y(\bar{a},z)-\frac{y(\bar{a},z)-b}{\bar{y}_{\bar{a}k}-b}\left(\delta_l\frac{\mu_k}{q\left(p'^{-1}\left(\frac{\mu_k}{\bar{y}_{\bar{a}k}-b+v_k-v_l-(1+r)mc_{\bar{a}-1lk}}\right)\right)}\right.$$
$$\left.+(1-\delta_l)\frac{\mu_k}{q\left(p'^{-1}\left(\frac{\mu_k}{\bar{y}_{\bar{a}k}-y(\bar{a},z_0)+v_k-v_l-(1+r)mc_{\bar{a}-1lk}}\right)\right)}\right). \tag{B7}$$

In the data, we observe the empirical counterparts of these wages.

These wage equations permits the identification of several parameters in. We assume that $R^u_{\bar{a}-1lk}>0$ or $R^e_{\bar{a}-1lk}(z)>0$, so that these transitions are observed in the data. Using wages of jobs newly obtained at age $\bar{a}$, we can the identify:
1. The job quality distributions $\mathcal{F}_l$. Having assumed that job quality z is observed, the job quality distributions $\mathcal{F}_l$ can be identified from the empirical distribution of types for the new jobs at age $\bar{a}$.
2. The productivity $y(\bar{a},z)$. consider workers who are in the same state at age $\bar{a}-1$ and who move to the same location. Within a group, workers only differ in newly drawn z. The ratio of $w-b$ thus simplifies,

$$\frac{w(z)-b}{w(z')-b}=\frac{y(\bar{a},z)-b}{y(\bar{a},z')-b}$$

Given the knowledge of b and the separability of y, $y(a,z)=f(a)e^z$, we can identify $y(\bar{a},z)$. Consequently the expectation $\bar{y}_{\bar{a}l}$ are identified by integration.
3. Vacancy costs μ_k and job destruction shocks δ_l. Consider workers who change jobs but who remains in the same location. For them, $v_k-v_l-(1+r)mc_{\bar{a}-1lk}=0$ because $k=l$. Wages can thus identify μ_l and δ_l under function inversibility conditions.
4. The parameters $v_k-v_l-(1+r)mc_{\bar{a}-1lk}$. Consider workers in the same state at age $\bar{a}-1$ except the location. By comparing the wage of identical workers moving from k to k and those from l to k, identifies the sum $v_k-v_l-(1+r)mc_{\bar{a}-1lk}$.

To summarise, we have identified all the parameters to compute $R^u_{\bar{a}-1lk}$ and $R^e_{\bar{a}-1lk}(z)$, and $\theta^u_{\bar{a}-1lk}$ and $\theta^e_{\bar{a}-1lk}(z)$. Note that we can have additional identifying information about the job destruction shock by comparing employed and unemployed workers in the same location at age $\bar{a}-1$ and unemployed in the same location at age $\bar{a}$.

Identification from observed transition frequencies between $\bar{a}-1$ and $\bar{a}$ From observable transition frequencies, we identify the choice probabilities $\mathbb{P}_\varepsilon(K^u_{\bar{a}-1l}(\boldsymbol{\varepsilon})=k)$, $\mathbb{P}_\varepsilon(K^u_{\bar{a}-1l}(\boldsymbol{\varepsilon})=0)$, $\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z,\boldsymbol{\varepsilon})=k)$ and $\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z,\boldsymbol{\varepsilon})=0)$. Given the definition of these choice proabilities in equations (12) and (12), we obtain the following relations

$$\log\left(\frac{\mathbb{P}_\varepsilon(K^u_{\bar{a}-1l}(\boldsymbol{\varepsilon})=k)}{\mathbb{P}_\varepsilon(K^u_{\bar{a}-1l}(\boldsymbol{\varepsilon})=0)}\right)=\frac{R^u_{\bar{a}-1lk}-sc^u_{\bar{a}-1lk}}{\sigma_\varepsilon} \tag{B8}$$

$$\log\left(\frac{\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z,\boldsymbol{\varepsilon})=k)}{\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z,\boldsymbol{\varepsilon})=0)}\right)=\frac{R^e_{\bar{a}-1lk}(z)-sc^e_{\bar{a}-1lk}}{\sigma_\varepsilon} \tag{B9}$$

These relations permit identification of $sc^e_{\bar{a}-1lk}$, $sc^u_{\bar{a}-1lk}$, and σ. In particular, since $sc^e_{\bar{a}-1lk}$ does not vary with z, we have

$$\frac{\log\left(\frac{\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z,\boldsymbol{\varepsilon})=k)}{\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z,\boldsymbol{\varepsilon})=0)}\right)}{\log\left(\frac{\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z',\boldsymbol{\varepsilon})=k)}{\mathbb{P}_\varepsilon(K^e_{\bar{a}-1l}(z',\boldsymbol{\varepsilon})=0)}\right)}=\frac{R^e_{\bar{a}-1lk}(z)-sc^e_{\bar{a}-1lk}}{R^e_{\bar{a}-1lk}(z')-sc^e_{\bar{a}-1lk}}$$

Since $R^u_{\bar{a}-1lk}$ and $R^e_{\bar{a}-1lk}(z)$ are already identified, this identifies $sc^e_{\bar{a}-1lk}$. Once $sc^e_{\bar{a}-1lk}$ is known, we identify σ_ε and then $sc^u_{\bar{a}-1lk}$ using (B8) and (B9).

Identification from across age comparison The preceding arguments have shown that $v_k - v_l - (1+r)mc_{\bar{a}-1lk}$ is identified using transitions between ages $\bar{a}-1$ and $\bar{a}$. We now show how we can disentangle the age-invariant amenity differentials $v_k - v_l$ from mobility costs m_{alk} by considering transitions between ages $\bar{a}-2$ and $\bar{a}-1$. For this group, the anologuous expression turns out to be $v_k - v_l - \frac{(1+r)^2}{2+r}mc_{\bar{a}-2lk}$. Hence across age comparisons identifies mobility costs, which then in turn identifies the amenity differential $v_k - v_l$.

To see how the analoguous expression arises, consider the returns from searching between $\bar{a}-2$ and $\bar{a}-1$ write:

$$R^u_{\bar{a}-2lk}=\max_{\theta\geq 0}\left\{p(\theta)\left[\frac{\bar{y}_{\bar{a}-1k}-b+\int\tilde{V}_{\bar{a}-1l}(z)d\mathcal{F}_k(z)-\tilde{U}_{\bar{a}-1l}}{1+r}+\frac{2+r}{(1+r)^2}(v_k-v_l)-mc_{\bar{a}-2lk}\right]-\theta\frac{\mu_k}{1+r}\right\},$$

$$R^e_{\bar{a}-2lk}(z)=\max_{\theta\geq 0}\left\{p(\theta)\left[\frac{\bar{y}_{\bar{a}-1k}-y(\bar{a}-1,z)+\int\tilde{V}_{\bar{a}-1l}(z)d\mathcal{F}_k(z)-\tilde{V}_{\bar{a}-1l}(z)}{1+r}+\frac{2+r}{(1+r)^2}(v_k-v_l)\right.\right.$$
$$\left.\left.-mc_{\bar{a}-2lk}\right]-\theta\frac{\mu_k}{1+r}\right\}$$

where

$$\tilde{U}_{al}=U_{al}-b-\sum_{i=0}^{\bar{a}-a}\frac{v_l}{(1+r)^i}\qquad\text{and}\qquad\tilde{V}_{al}(z)=V_{al}(z)-y(a,z)-\sum_{i=0}^{\bar{a}-a}\frac{v_l}{(1+r)^i}$$

define the values net of flow income and discounted local amenities. Using the same arguments as before, we are able to identify the productivity $y(\bar{a}-1,l)$ and the parameters

$v_k - v_l - \frac{(1+r)^2}{2+r} mc_{\bar{a}-2lk}$.

B.2 Identification from within-job wage growth

A worker may receive a high wage because i) job productivity z is high, or ii) the worker obtains a large share of the match surplus, so ρ is high. We can eliminate the role played by the sharing rule by investigating within-job wage growth. Denote w_a and w_{a+1}, the wages of a worker at ages a and $a+1$. Using the wage equation (B1), we have

$$\frac{y(a,x,z) - w_a}{y(a+1,x,z) - w_{a+1}} = \frac{y(a,x,z) - b - \tilde{D}_{al}(x,z)}{y(a+1,x,z) - b - \tilde{D}_{a+1l}(x,z)} \tag{B10}$$

so within-job wage growth does not depend on the bargaining rule, and so on the previous worker's employment situation. Under invertibility conditions, one can recover the job type z observing wages w_a and w_{a+1}, and functions $y(a,x,z)$ and $\tilde{D}_{al}(x,z)$.

C Data Appendix

C.1 Data: Additional details

The model is estimated on French linked employer-employee data, the FH-DADS. For a subset of French workers, we dispose of information about the employer and the job contract over the period 1986-2012. Observations in years prior to 2002 only correspond to workers who are born in October an even-number year and who had at least one recorded job in the period 2002-2012. Observations from 2002 correspond to workers who are born in October, even an odd-number year, or 2nd to 5th of January, or 1st to 4th of April, or 1st to 4th of July.

We focus on the years between 1994 and 2007, a period without recession in France. It is usual practice to discard the early years of the data set (as e.g. hours worked are not available before 1992).

The FH-DADS contains information about educational attainment. For individuals born the first four days of October, January (from the 2nd), April or July, we have the education degree among 8 large categories from high-school dropout to graduated. We restrict our sample to those individuals and aggregate education by number of years of education into commonly used categories. We thus end up with five groups: no degree, vocational training, high-school degree, bachelor degree, and more than a bachelor degree.

C.2 Sample selection and definition of yearly employment status

We follow standard DADS-practices in order to generate our sample. In particular, we consider only full-time workers in the private sector in mainland France. Hence civil servants are dropped, as are individuals pursuing an apprenticeship or internship.

Turning to the wage data, we drop outliers. Specifically, individuals whose net wage is abnormally low (below 0.9 gross minimum wage) or high (above 5 times the gross minimum wage) are deleted from the sample.

Our unit of time is the year (or the age of the worker). We then categorize workers as employed or unemployed as follows. The French Statistical Institute (INSEE) defines for each job a dummy variable, indicating whether the job is "main" (in french *non-annexe*) or not. This variable is based on the number of days worked, hours worked and the wage. More precisely, a job is defined as "main" if i) the net wage is more than 3 times the gross minimum wage or ii) the number of hours worked is higher than 120, the number of days worked is higher than 30, and the ratio between the two is higher than 1.5. We use this variable.

For the first year of observation, we define a worker as employed if she holds a main job the 15th of December (rather than the 31st in order to avoid artificial end-of-calender year effects that risks confounding employment-to-employment with unmployment-to-employment transitions). For the subsequent years of observation, we define a worker as employed if she holds a main job the 15th of December and if she has not been without a job more than 15 days in the year before the 15th of December. This 15-day rule is used, for instance, in Postel-Vinay and Robin (2002).

In order to deal with specific missings in the data, we employ the following rule. If the worker's employer is the same in year $t-1$ and $t+1$, we impute employment, and extrapolate the wage and hours-of-work data in the missing year in t by mean of the previous and next year wages. If the gap extends beyond one year, the spell is dropped.

C.3 Summary statistics

The sample selection rules yield a sample of 1,152,000 annualized observations of 186,000 workers that are observed during the years 1994-2007. The panel is, of course, unbalanced. Table B2 reports some summary statistics for workers in our sample.

Table B2: Summary statistics of the selected sample

	Min	Q1	Median	Mean	Q3	Max
Age at first observation (in years)	18	30	38	38.02	46	57
Mean hourly wage (EUR)	3.99	8.34	10.21	11.39	13.2	34.4
Education [%]						
No deg.				18.5		
Vocational				41.8		
High-school				17.3		
Bachelor				14.1		
¿Bachelor				8.3		
# observations per individual	4	6	6	6.19	6	15
# observations as employed	0	5	6	5.61	6	15
# job transitions per individual [%]						
0 obs				56.4		
1 obs				24.4		
2 obs				13.2		
3 obs				4.3		
>3 obs				1.8		
# spatial moves per individual [%]						
0 obs				94.3		
1 obs				4.3		
2 obs				1.2		
3 obs				0.16		
>3 obs				0.03		