Frictional Labor Mobility[*]

Benoît Schmutz[†] Modibo Sidibé [‡]

November 15, 2017

Abstract

We build a dynamic model of migration where, in addition to usual mobility costs, workers face spatial frictions that decrease their ability to compete for distant job opportunities. We estimate the model on a matched employer-employee panel dataset describing labor market transitions within and between the 100 largest French cities. Our identification strategy is based on the premise that frictions affect the frequency of job transitions, while mobility costs impact the distribution of accepted wages. We find that: (i) controlling for spatial frictions reduces mobility cost estimates by one order of magnitude; (ii) the urban wage premium is driven by better opportunities for local job-to-job transitions in larger cities; (iii) migration dramatically reduces lifetime inequalities due to initial location; (iv) labor mobility policies based on relocation subsidies are inefficient, unlike switching from nationwide to local minimum wages.

Keywords: mobility costs, spatial frictions, migration, local labor markets

JEL Classification: J61, J64, R12, R23

[*]A previous version of this paper has been circulated under the title "Job search and migration in a system of cities" (ERID WP 181). We thank the editor as well as three anonymous referees. This paper benefited from helpful discussions with Jim Albrecht, Pat Bayer, Bruno Decreuse, Rafael Dix Carneiro, Florence Goffette-Nagot, John Kennan, Francis Kramarz, Thomas Le Barbanchon, Fabien Postel-Vinay, Sébastien Roux, John Rust and Maxime Tô, as well as seminar participants at Duke, U-Penn, Georgetown, PSE, Sciences Po, Louvain-la-Neuve, Lyon-GATE, UNC-Chapel Hill, CREST, ZEW, Polytechnique, Dauphine, Paris-Sud, and participants to the AEA, UEA and SOLE conferences, for their help, comments and discussions. Data was made available by the CASD thanks to a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-EQPX-17 - Centre d'Accès Sécurisé aux Données – CASD).

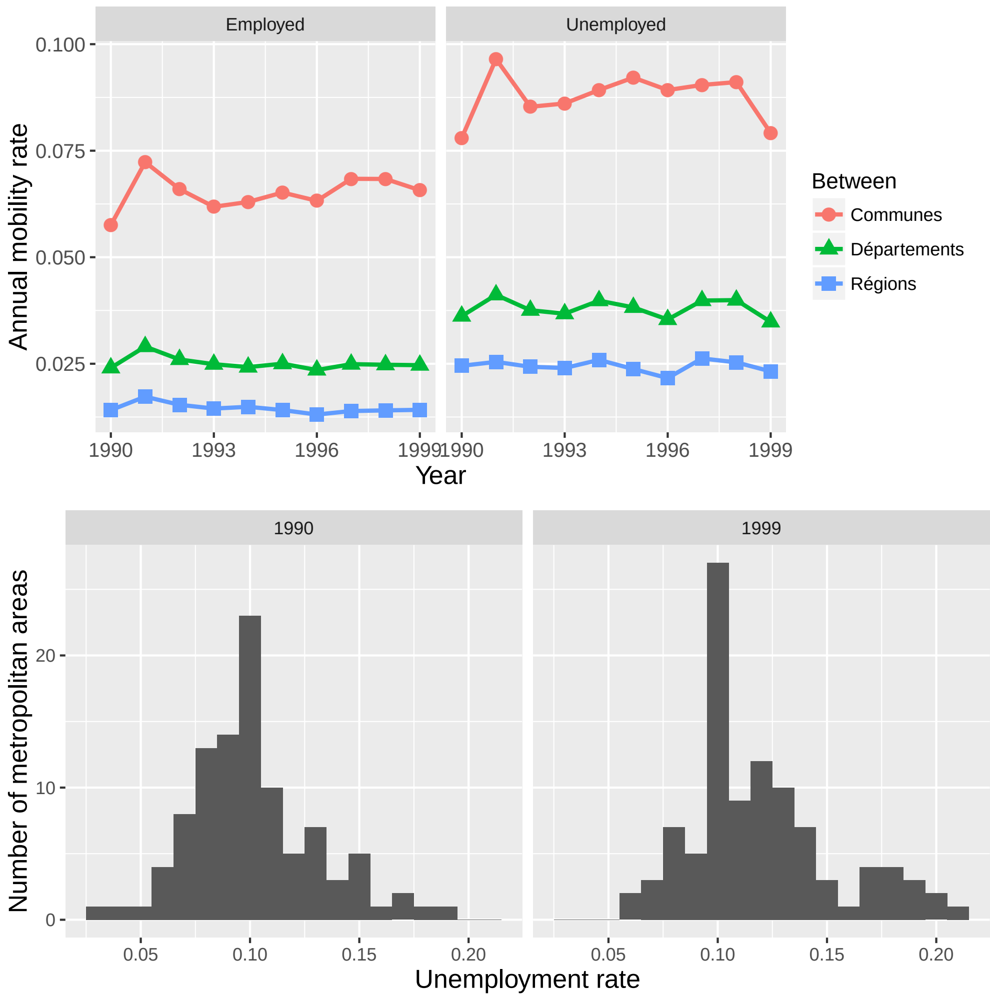
[†]Ecole Polytechnique and CREST; benoit.schmutz@polytechnique.edu

[‡]Duke University and CREST; modibo.sidibe@duke.edu

Perhaps the simplest model would be a picture of the economy as a group of islands between which information flows are costly (Phelps, 1969).

Local labor markets in developed countries are often characterized by striking economic disparities, which may persist despite substantial levels of labor mobility. As shown by Figure 1, even though France witnessed a steady 5% annual migration rate during the 1990s, the dispersion in local unemployment risk at the metropolitan area level remained almost unaffected.

Figure 1: Labor mobility and local unemployment in France in the 1990s



Notes: (i) Mobility rates: probability to have changed location in the past year, conditional on previous employment status; (ii) *communes* and *départements* are akin to US municipalities and counties, respectively; *régions* are similar to German Länder, with less autonomy; until 2016, there were 22 régions, 94 départements and over 36,000 communes in continental France; those three administrative levels form nested partitions of the French territory, unlike metropolitan areas, which are aggregates of communes and may cross département or region boundaries; (iii) Unemployment rates are computed on the 25-54 age bracket and for the 100 largest metropolitan areas in continental France, keeping a constant municipal composition based on the 2010 "Aires Urbaines" definition; (iv) Sources: Labor Force Surveys 1990-1999 and Census 1990 and 1999.

This paper aims to understand why despite similar cultural and labor market institutions, and the rapid progress of transportation and communication technologies, individuals do not take advantage of the opportunity to move into more affluent cities. The traditional solution to this puzzle is a combination of individual preferences with mobility costs. On the static side, individuals do not change regions because of a strong "home state" bias unrelated to actual economic conditions. As for mobility costs, they provide a dynamic rationale by taking on extremely high values which equalize lifetime income differences across cities. We argue that even forward-looking, profit-maximizing, location-indifferent workers may remain stuck in inauspicious locations because they face search frictions on the labor market and because objective barriers to migration (lower efficiency of job search into remote places and relocation costs) rule out off-the-job migration as an optimal behavior in spatial equilibrium. In contrast to most of the existing literature, our theory does not rest upon any kind of spatial idiosyncrasy affecting workers' utility or productivity and irrational beliefs that local economic downturns will eventually reverse.

We build on McCall's (1970) framework to model a dynamic job search model that incorporates spatial segmentation between a large number of interconnected local labor markets, or "cities". Local labor market conditions are characterized by city-specific job arrival rates and wage offer distributions, and local labor shocks are introduced through city-specific layoff rates. We consider the strategy of ex-ante identical workers, who engage in both off-the-job and on-the-job search, both within and between cities. A spatial equilibrium is achieved through the mobility of unemployed workers, who generate congestion externalities upon the non-pecuniary component of utility in each location, such that the mobility decision will be based on a cost and return analysis.

While most previous quantitative studies of migration rest upon a unidimensional conception of spatial constraints based on a black box called "mobility costs", which encompass both impediments to the mobility of workers when it takes place (actual mobility costs) and impediments to the spatial integration of the labor market (workers' ability to learn about and apply to remote vacancies), we believe that separating these mechanisms is important, as they do not take place at the same time, they do not affect the same economic outcomes nor warrant the same public policies in cases where migration is inefficiently low. In a context of high spatial frictions, relocation subsidies may not prove effective, contrary to a centralized placement agency that would increase job search efficiency across space.

In our setting, there are three barriers to migration. First, physical distance between cities reduces the efficiency of job search between cities. This dimension of "spatial frictions" determines the *centrality* of each city in the system. Second, spatial segmentation introduces heterogeneity in local nonpay component of utility, referred to as "amenity", which impacts agents' willingness to refuse a job somewhere else, even in instances where this decision appears as a sound decision from a pure labor-market standpoint. The ranking of each city according to amenities in addition to local labor market conditions generates its *attractiveness* in the system. Finally, workers face classical mobility costs, which are a lump sum that needs to be paid upon moving. As in Schwartz (1973), these costs encompass a fixed cost of losing local ties and connections, and a cost of moving from one place to another, which mostly depends on distance. Since the model is dynamic, the relative position of the city in the distribution of all possible mobility costs, which determines the level of *accessibility* of the city in the system, will also impact whether the offer was deemed acceptable in the first place.

The model is based on "mobility-compatible indifference wages" which formalize the dynamic utility trade-off between locations faced by workers. These functions of wage, which are specific to each pair of cities, are defined by the worker's indifference condition between her current state (a wage in a given city) and a potential offer in a different city. They define a complex, but intelligible relationship between wages and the model primitives. As a consequence, the model is able to cope with various wage profiles over the life cycle, including voluntary wage cuts as in Postel-Vinay & Robin (2002), which may even take place in less prosperous, yet more central, cities.

Steady state conditions on market size and unemployment level allow us to solve the model. Our estimation uses the panel version of the French matched employer-employee database *Declaration Annuelles de Données Sociales* (DADS) from 2002 to 2007, with local labor markets defined at the metropolitan area level. The identification of local labor market parameters and spatial friction parameters is based on the frequency of labor and geographical mobility whereas data on wages are used to identify mobility costs. Therefore, we can disentangle the impact of mobility costs from that of spatial frictions on the migration rate. The model is based on a partition between submarkets that can be as detailed as possible: we address the challenges raised by the high dimensionality and we allow the final level of precision to only depend on the research question. In our case, we consider that local labor markets defined at the city level provide a more accurate description of the allocation between workers and firms. Yet, the model is fractal and may apply to the analysis of spatial segmentation at the neighborhood level within a single metropolitan labor market, or even to international migration. It is also transferable to occupational segmentation.

We believe that this paper provides a complete representation of the complex dynamic trade-offs faced by workers when incorporating migration as a career decision. To do so, it relies on the existence of search frictions. Yet, we do not claim to provide a fully specified search and matching characterization of the labor market. Our current setting is not well equipped to deal jointly with individual and firm location decisions. The underlying reason is related to local matching parameters generating a finite distribution of city-specific reservation wages, which may yield discontinuous wage offer distributions. As a consequence, we assume away the location decision of firms, and the related matching problem. We view our parameters as a measure of city-level hazard rates reflecting the efficiency in the allocation process of workers across firms within and between local labor markets.

Our results consist of a set of city-specific structural parameters and a set of matrices of parameters measuring spatial constraints between each pair of cities. Our main findings are fourfold: first, we show that higher job arrival rates for employed workers are associated with higher wage dispersion; second, that cities with higher local job-finding rates are also better at sending their workers to other cities and that the ability to apply to remote vacancies plays a much larger role for employed workers, than for unemployed workers; third, that the average value of mobility costs roughly corresponds to eighteen months of work paid at minimum wage, which is one order of magnitude lower than previously reported in the literature and seems to be in large part attributable to our inclusion of spatial frictions; finally, that matching economies do exist in larger cities, which are characterized by higher job arrival rates that drive wages upwards, but do not suffer from lower cost-of-living-adjusted amenities.

We use the empirical framework provided by our estimation results to simulate the career decisions of workers over 100 local labor markets. This allows us to compute lifetime earnings based on actual realizations. While larger cities are more unequal in cross-section, we show that lifetime inequalities within cities are lower in larger cities thanks to their higher frequency of local transitions. Overall lifetime inequalities based on (initial) location are lower than cross-sectional inequalities if and only if migration is possible. Lifetime earnings are also computed under a situation where all mobility costs are paid for by the government. Our simulation shows that the resulting increase in labor mobility is very modest, while the welfare impact of this policy is twice lower than its cost. On the other hand, switching from a nationwide minimum wage to a local schedule aimed at maximizing the number of transitions between cities is costless and increases lifetime earnings by 3%.

Relationship to the literature

Our paper appeals to two strands of the literature: on the empirical side, it quantifies the determinants of migration; on the theoretical side, it uses recent advances in the search and matching literature to capture interactions between competing submarkets.

Migration Economists have long investigated the career choice of workers. Keane & Wolpin (1997) have shown that individuals make sophisticated calculations regarding work-related decisions, both in terms of pure labor market characteristics (industry, occupation, skills requirement) and location. In order to disentangle between the various underlying mechanisms, a structural approach seems natural. It was pioneered by Dahl (2002), who constructs a model of mobility and earnings over the US states and shows that higher educated individuals self-select into states with higher returns to education. However, as migration is an investment, it requires not only a static tradeoff between economic conditions, but also a comparison between expected future economic conditions (Gallin, 2004). In addition, and despite its interest and obvious links to the present paper, the classic perfect-competition approach cannot fully reconcile the equilibrium coexistence of both labor mobility and local labor market differences.

In this paper, we argue that an equilibrium model featuring search frictions can tackle this puzzle, provided it allows for spatial segmentation of the labor market. Modeling spatial segmentation in a search and matching framework is a new venue for research, despite well-documented empirical facts suggesting that the labor market may be described as an equilibrium only at a local level -in particular, because matching functions exhibit a high level of spatial instability (Manning & Petrongolo, 2017).[1] From a practical viewpoint, the absence of space in search models can be explained by the computational difficulties associated with multiple high-dimensional objects such as wage distributions.

A standard solution is to consider a very stylized definition of space, like Baum-Snow & Pavan (2012), who propose a rich model with individual ability and location-specific human capital accumulation, but only distinguish between small, medium and large cities and are therefore unable to quantify the impact of the shape and size of the spatial network on migration, nor provide policy-relevant estimates on the impact of migration on specific regions or cities.

Our paper follows on from the work of Kennan & Walker (2011), who develop and estimate a par-

[1] By restricting their estimation sample to the Paris region in their seminal work, Postel-Vinay & Robin (2002) implicitly recognize this problem.

tial equilibrium model of mobility over all US states and provide many interesting insights, including mobility costs. However, computing the model requires extra assumptions on individual information sets.[2] Moreover, the low mobility rate is rationalized by the existence of extremely high mobility costs. Finally, a focus on the state level is not fully consistent with the theory of local labor markets, which are better proxied by metropolitan areas (Moretti, 2011).

Job search and frictions between competing submarkets There is a notable effort in the recent empirical job search literature to study search patterns in competing submarkets. These papers seek to provide new dynamic micro-foundations to the old concept of dualism in the labor market. The underlying idea is that jobs are not only defined by wages, but also by a set of benefits that are only available within some submarkets. This creates potential tradeoffs, for example between a more regulated sector which offers more employment protection (in terms of unemployment risk and insurance) and a less regulated sector, which allows for more flexibility and possibly better wage paths (Postel-Vinay & Turon, 2007; Shephard, 2014; Bradley, Postel-Vinay & Turon, 2017). In doing so, these models also provide more accurate estimates of the matching parameters, which are no longer averaged across sectors.

Our main reference is Meghir, Narita & Robin (2015), who study the impact of the existence of informality in Brazil on labor market outcomes. The authors consider a very general model where workers can switch between sectors and where job arrival rates (and the number of firms in each sector) are endogenously determined by firms' optimal contracts. While our framework is less general as it leaves firms' behavior aside, it can also provide a useful complement by offering a more general definition of segmentation, where the option value of unemployment is inherited from past decisions and moves between sectors (or locations) entail switching costs.

The rest of the paper is organized as follows. In section 1, we characterize the French labor market as a spatial system; section 2 is the presentation of the model; section 3 explains our estimation strategy, the results are discussed in section 4 and section 5 describes a few experiments.

[2]For example, it is assumed that individuals have knowledge over a limited number of local wage distributions, which correspond to where they used to live. In order to learn about another location, workers need to pay a visiting cost. These assumptions may not reflect the recent increase in workers' ability to learn about other locations before a mobility (Kaplan & Schulhofer-Wohl, 2017).

1 Motivating facts

In this section, we provide descriptive evidence in favor of the modeling of the French labor market as a system of local labor markets based on metropolitan areas. These local labor markets present three salient characteristics: (i) heterogeneity in terms of economic opportunities; (ii) interconnection through workers' mobility; and (iii) stability in key economic variables. We first document the heterogeneity and the stability of the three features that will characterize a local labor market throughout the paper: population, unemployment rate and wage distribution. Then, we describe workers' mobility, both on the labor market and across space.

1.1 France as a steady state system of local urban labor markets

The functional definition of a metropolitan area brings together the notions of city and local labor market. A more precise partition of space, for instance based on municipal boundaries, would lead to a confusion between job-related motives for migration and other motives.[3] French metropolitan areas (or "aires urbaines") are continuous clusters of municipalities with a main employment center of at least 5,000 jobs and a commuter belt composed of the surrounding municipalities with at least 40% of residents working in the employment center.[4] We will focus on the 100 largest metropolitan areas in continental France, as defined by the 2010 census. Below a certain population threshold, the assumption that each of these metropolitan areas is an accurate proxy of a local labor market becomes difficult to support.[5] As shown in Figure 10 in Appendix D.1, metropolitan areas cover a large fraction of the country. More precisely, these locations make up for 65.2% of French Labor force. When considering only non-rural residents, the 100 largest metropolitan areas represent 83.4% of labor force. Paris and its 12 millions inhabitants stands out, before six other millionaire cities and eleven other metropolitan areas with more than 0.5 million inhabitants.

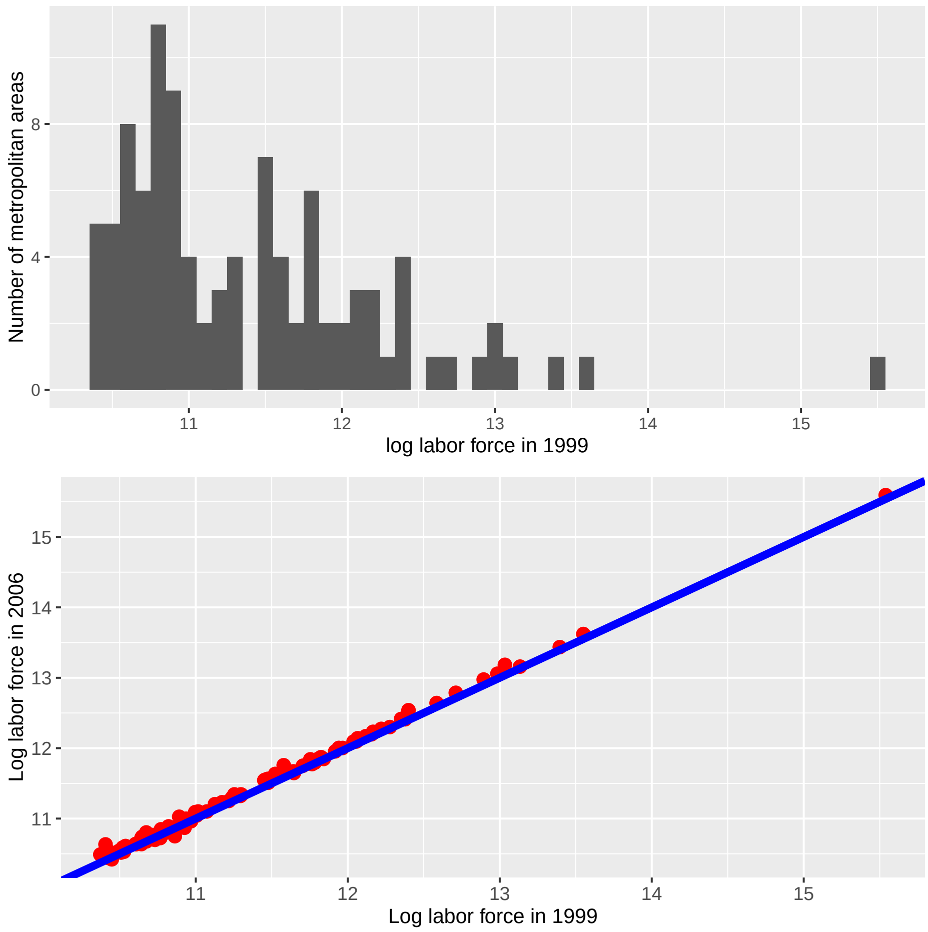
[3] According to the 2006 French Housing Survey, 16% of the households in the labor force who had been mobile in the past four years declared that the main reason for their move was job-related. However, this small proportion hides a large heterogeneity which is correlated with the scale of the migration, from 5% for the households who had stayed in the same municipality, to 12% for those who had changed municipalities while staying in the same county, to 27% for those who had changed counties while staying in the same region and to 49% for those who had changed regions.

[4] US MSAs are defined along the same lines, except the unit is generally the county and the statistical criterion is that the sum of the percentage of employed residents of the outlying county who work in the center and the percentage of the employment in the outlying county that is accounted for by workers who reside in the center must be equal to 25% or more.

[5] The smallest metropolitan area which will be isolated in our analysis is Narbonne, with 90,000 inhabitants in 2010. According to the 2010 US census, meeting the same level of precision on the US would require to distinguish between more than 360 cities.

Population Since we do not model the participation choice of workers, labor force is analogous to population. Data from 1999 and 2006 Census shows that the Paris region accounts for more than 25% of total labor force in the first 100 cities. As shown in Figure 2, local labor force is approximated by a pareto distribution (Zipf's Law) and variation in local labor force between 1999 and 2006 is negligible.[6]

Figure 2: Local labor force



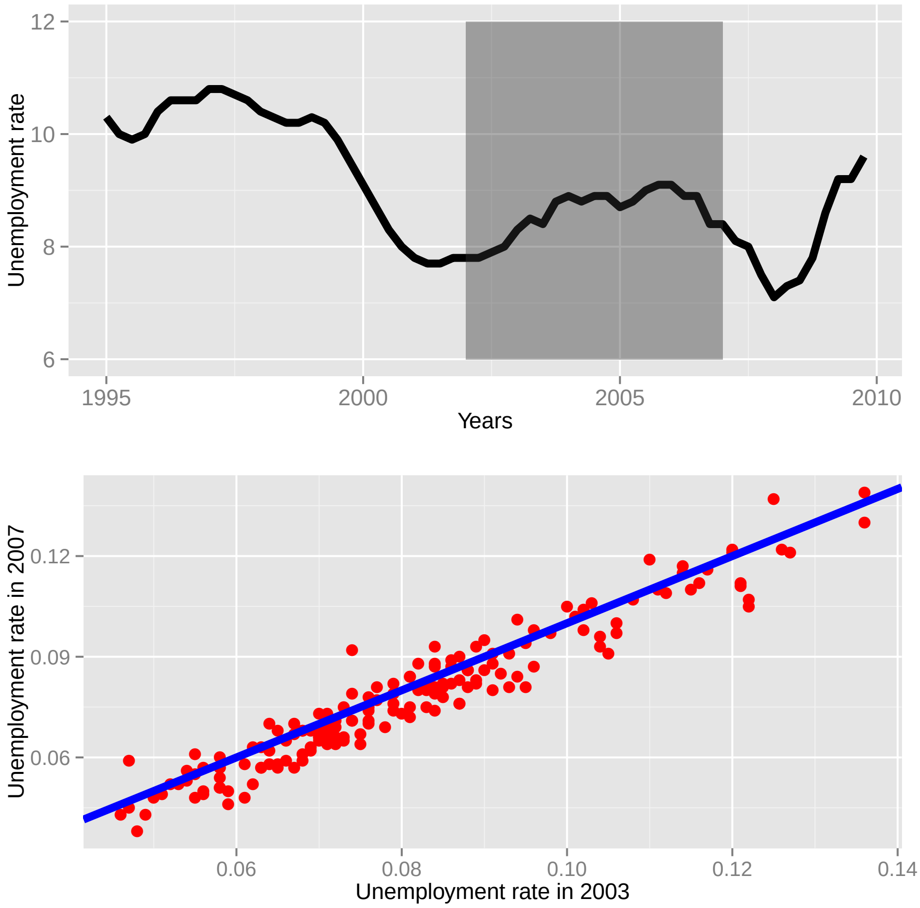
Notes: (i) Labor force is composed of unemployed and employed individuals aged between 15 and 64; (ii) the labor force in the 100 largest metropolitan areas in continental France amounts to 17.4 millions in 1999 and to 18.4 millions in 2006; (iii) The sum of the absolute values of location-by-location changes amounts to 1.1 million, i.e, 6% of total labor force in 1999. *Source: Census 1999 and 2006.*

Unemployment Figure 3 establishes that local unemployment is quite stable over time, especially over a short period of stable aggregate unemployment. Such is the case from 2002 to 2007, both in terms of range and in terms of variation of the annual moving average. During those years, which

[6]This stability is at odds with the fact that metropolitan areas face diverse net migration patterns, which are driven by the migration of nonparticipants (retired, young individuals). According to Gobillon & Wolff (2011), 31.5% of French grandparents aged 68-92 in 1992 declared that they moved out when they retired. Among them, 44.1% moved to another region. Most of these migration decisions are motivated by differences in location-specific amenities or by the desire to live closer to other family members. Obviously, family is an important determinant of migration, which we can not consider here because of data limitations.

correspond to Jacques Chirac's second term, the French economy is in an intermediate state, between a short boom in the last years of the twentieth century and the Great Recession. For this reason, we focus on this period throughout the paper.[7]

Figure 3: Aggregate and local unemployment



Notes: (i) Top graph displays the quarterly unemployment rate in France: the darker background is the period under study; bottom graph displays the yearly unemployment rate in the 168 employment areas ("zones d'emploi") intersecting with the 100 largest metropolitan areas in 2003 and 2007; (ii) Information in bottom graph is based on a combination of administrative employment and unemployment data as well as the Labor Force Survey; It is not available before 2003 and the geography of employment areas is not based on municipalities so it does not fully coincide with metropolitan areas. *Source: Série Longue Trimestrielle INSEE (top) and Taux de Chômage Localisés INSEE (bottom).*

As already illustrated in the introduction, there is a large amount of variation in local unemployment rates, with a one-to-four ratio between the high-unemployment cities and the low-unemployment ones. However, this heterogeneity is not driven by population size (see Figure 8 in Appendix C.2).

Wage distributions To compute city-specific earning distributions, we use data from the *Déclarations Annuelles des Données Sociales* (DADS). The DADS are a large collection of mandatory employer reports of the earnings of each employee of the private sector subject to French payroll taxes. The

[7]Note that this almost exactly matches the period chosen by Meghir et al. (2015) for Brazil.

DADS are the main source of data used in this paper. Table 11 in Appendix C.2 reports the main moments of the wage distributions of the nine largest cities and of nine smaller cities at various points of the distribution of city sizes. These distributions are computed over the entire 2002-2007 period. The general pattern, that has been well-documented in the literature, is that wages are higher in larger cities. The most striking example is Paris, where the average wage (36.8k€) is over 50% higher than the city-level average wage. Other large cities have similar wage premia.[8] The elasticity of wages to city size is 6.5%. This pattern is far from being a systematic rule: for example, Marseilles and Lille (rank 3 and 5) have lower wages than many large and medium cities, while Creil (rank 60) is in the opposite situation. Although the wage premium in Paris may be partly offset by the cost of living, there exist persistent wage differentials among cities with comparable size and cost of living. In addition, there is a strong positive correlation of between wage dispersion and city size, driven by the existence of high wages in large cities.

Regarding stability, Panel 3 in Table 11 in Appendix C.2 shows that wage distributions do not vary a lot between 2002 and 2007. The ratios of the three quartiles and the mean of the log-wage distributions in 2007 and 2002 are closely distributed around 1 for the whole set of metropolitan areas.

Workers' heterogeneity Apart from the size of the labor force and the unemployment rate, other dimensions, such as the skill and the sectoral composition, are also important drivers of local labor market heterogeneity and dynamics. However, we believe that, as a first-order approximation, the assumption of workers' homogeneity is not too costly when focusing on a short period, because the distribution of observable characteristics across cities remains stable (see Figure 9 in Appendix C.2).

1.2 Labor and geographical mobility

Data We now turn to the mobility patterns of jobseekers across France. To make a precise assessment regarding geographical transitions between each pair of cities, we use a specific subsample of the DADS data. Since 1976, a yearly longitudinal version of the DADS has been following all employed individuals born in October of even-numbered years. Since 2002, the panel includes all individuals born in October. Due to the methodological change introduced in 2002, and amid concerns about the stability of the business cycle, we focus on a six-year span between 2002 and 2007. The main restrictions over our sample are the following: first, to mitigate the risk of confusion between non-

[8] Our data selection procedure that excludes part-time workers and civil servants increases the wage gap between Paris and smaller locations. Using all the available payroll data in 2007, the mean wage in Paris is around 22,5k€, which is 35% higher than the national average.

participation and unemployment, we restrict our sample to males who have stayed in continental France over the period; second, we exclude individuals who are observed only once. We end up with a dataset of 384,114 individuals and 2 millions observations (see appendix C.1, for more details).

Since the DADS panel is based on firms' payroll reports, it does not contain any information on unemployment. However, it reports for each employee the duration of the job, along with the wage. We use this information to construct a potential calendar of unemployment events and, in turn, identify transitions on the labor market.[9] As in Postel-Vinay & Robin (2002), we define a *job-to-job transition* as a change of employer associated with an unemployment spell of less than 15 days and we attribute the unemployment duration to the initial job in this case. Conversely, we assume that an unemployment spell of less than 3 months between two employment spells in the same firm only reflects some unobserved specificity of the employment contract and we do not consider this sequence as unemployment.[10] Finally, we need to make an important assumption regarding the geographical transitions of unemployed individuals: we attribute all the duration of unemployment to the initial location, assuming therefore that any transition from unemployment to employment with migration is a single draw. Hence, we rule out the possibility of a sequential job search whereby individuals would first change locations before accepting a new job offer. From a theoretical viewpoint, this means that mobility has to be job-related. From a practical viewpoint, in the DADS data, the sequential job search process is observationally equivalent to the joint mobility process.

Labor market transitions Table 1 describes the 719,601 transitions of the 384,114 individuals in our sample. Over our period of study, a third of the sample has recorded no mobility. This figure is similar to the non-mobility rate of 45% reported by Postel-Vinay & Robin (2002) from 1996 to 1998. Approximately 23% of the sample records at least one job-to-job transition, and the numbers of transitions into unemployment and out of unemployment are almost identical. Average wages are almost constant over time, as shown in the last line of the table. Job-to-job transitions are accompanied by a substantial wage increase (around 7%). Transitions out of unemployment lead to a wage that is 7% lower than the wage of employed workers who do not make any transition, 25% lower than the final wage of employed workers who have experienced a job-to-job transition, and roughly equal to the initial wage of individuals who will fall into unemployment.[11]

[9] Our algorithm is available upon request.

[10] For a recent example of a similar assumption, see Bagger, Fontaine, Postel-Vinay & Robin (2014).

[11] In this table, as well as in our estimation, we assume that time starts on the first day of 2002. This left censoring is due to the fact that we do not have information about the length of unemployment for the individuals who should have entered the panel after 2002 but have started with a period of unemployment. Whereas, for employment spells, we could in

Table 1: Number and characteristics of transitions

Type of history	Number of events	Share	Initial Wage	Final Wage	Growth
	Characteristics of the spells				
No transitions while employed	126,227		26,088	28,862	5.7%
Out of unemployment	302,024		-	24,303	-
with mobility	59,605	19.8	-	24,793	-
without mobility	242,418	80.2	-	24,182	-
Job to job mobility	114,659		30,814	32,936	6.3%
with mobility	26,199	22.9	30,464	32,343	6.5%
without mobility	88,459	77.1	30,914	33,111	6.2%
Into unemployment	302,918		24,555	-	-
Full sample	719,601		27,956	28,255	6.0%
Individuals	384,114				

Notes: (i) Wages are in 2002 Euros and growth is the average log point difference between initial and final wage; (ii) Time begins on January 1st 2002. *Source: Panel DADS 2002-2007*

Geographical transitions Geographical mobility accounts for 19.8% of transitions out of unemployment and 22.9% of job-to-job transitions. As illustrated by Panel 1 in Table 12, size matters: Paris is both the most prominent destination and the city with the highest rate of transition (90.4%) with no associated mobility.[12] Panel 2 in Table 12, which compares the mobility patterns within the Lyon region (also known as "Rhône-Alpes") and between the Lyon region and Paris, illustrates that physical distance matters as well. Although Paris is the destination of a sizable share of mobile workers, geographical proximity can overcome this attractiveness, as shown for the cities of Grenoble, Saint-Etienne and Bourg-en-Bresse that are located less than 60 miles away from Lyon.

In Panels 1 and 2 of Table 2, we try to provide a more systematic picture of these phenomena by estimating gravity equations of the share of each destination city in the total number of migrations out of each origin city, on d_{jl} a measure of physical distance between city j and city l, h_{jl} a dissimilarity index based on the sectoral composition of the workforce between 35 sectors, as well as city-position fixed effects.[13]

Results show that physical distance and sectoral dissimilarity are not substitutes for one another and the correlation is negative for both migration out of unemployment and migration between two

theory use information about the year when individuals entered their current firm, we choose not to, to keep the symmetry between both kinds of initial employment status.

[12] Postel-Vinay & Robin (2002) report that 4.7% of workers from the Paris region make a geographical mobility. They conclude that this low rate allows them to discard the question of inter-regional mobility.

[13] We use the traditional Duncan index: if v is a categorical variable defined by categories k in proportions $v_j(k)$ and $v_l(k)$ in cities j and l, $h_{jl} = \sum_k |v_j(k) - v_l(k)|$. In order to construct this variable, we use the 2007 version of a firm-level census called SIRENE.

Table 2: Migration flows and wage growth: gravity estimates

Panel 1: Migration flows out of unemployment

	(1)	(2)	(3)	(4)
d_{jl}	-28.11***	-15.47***	-14.25***	-15.78***
	(5.299)	(3.506)	(3.353)	(3.681)
h_{jl}			-19.79***	-16.08***
			(4.887)	(4.648)
d_{jl}^2				13.44***
				(4.290)
h_{jl}^2				-5.381**
				(2.435)
F.E.		X	X	X
R^2	0.004	0.652	0.653	0.654

Panel 2: Migration flows between two jobs

	(1)	(2)	(3)	(4)
d_{jl}	-10.89***	-6.314***	-5.448***	-6.058***
	(2.008)	(1.433)	(1.401)	(1.554)
h_{jl}			-13.99***	-12.23***
			(2.595)	(2.690)
d_{jl}^2				5.193***
				(1.694)
h_{jl}^2				-2.830**
				(1.254)
F.E.		X	X	X
R^2	0.004	0.540	0.541	0.543

Panel 3: Average wage growth

	(1)	(2)	(3)	(4)
d_{jl}	0.0165***	0.0301***	0.0123**	0.0234***
	(0.00495)	(0.00641)	(0.00505)	(0.00690)
h_{jl}			0.00142	0.00926
			(0.00667)	(0.0109)
d_{jl}^2			-0.0133**	-0.0126**
			(0.00524)	(0.00534)
h_{jl}^2			-0.00195	-0.00285
			(0.00260)	(0.00524)
F.E.		X		X
R^2	0.005	0.120	0.009	0.123

Notes: (i) Panels 1 and 2 ($N = 9,900$): Ordinary Least Squares estimates estimates of an equation of the number of workers migrating from each origin city to each destination city, as previously unemployed (panel A) or employed (panel B), as a function of physical distance d_{jl}, sectoral dissimilarity h_{jl} and city of origin and destination fixed effects (F.E.); (ii) Panel 3 ($N = 1,932$): Ordinary Least Squares estimates of an equation of the growth of average initial wage following a job-to-job transition by city pair for all the pairs of distinct cities where a job-to-job transition is observed; (iii) Both distance and dissimilarity are normalized using their distribution on the 100×99 city matrix; (iv) Significance: * 95%, ** 99%, *** 99.9%; standard errors are clustered at the destination city level.
Source: Panel DADS 2002-2007

jobs. However, the relationship between distance and migration is convex, whereas is it concave for dissimilarity. Another noteworthy feature is that unemployed workers seem to be more sensitive to distance. Finally, city fixed effects allow us to recover more than half of the observed heterogeneity in migration flows. To account for all these features, the specification displayed in columns (4) will prove useful in the estimation of the model (see Section 3.2).

Wage dynamics between cities Wage dynamics following a job-to-job transition are characterized by several noteworthy features. First, they are not symmetric: in 55% of the cases, mean wage growth between cities is higher than mean wage growth within cities. This suggests that mobility costs often matter: if not, this rate would be 50%. On the other hand, in 23% of the cases, mean wage growth between cities is negative: in those cases, mobility costs are not large enough to counterbalance the effect of differences in local (economic) conditions.[14] Second, in many transitions originating from the same city, mean wage growth is higher towards a less remote city. For instance, mean wage growth between Lille and Paris is 13%, while it is only 6% between Lille and Toulouse. Assuming that mobility costs are mostly determined by the physical distance between two locations, such pattern cannot be fully rationalized, because Paris is four times closer to Lille than Toulouse and, as will be shown in section 4, Paris does offer many more opportunities than Toulouse. The addition of mobility costs alone cannot cope with this simple observation, unless we allow for heterogeneous local amenities (or, equivalently, local costs of living).

Panel C in Table 2 gives more general evidence that physical distance matters for wage dynamics: a one-standard deviation increase in distance is associated with a two-percentage point increase in wage growth following a job-to-job transition. The effect is concave. On the other hand, dissimilarity between cities does not seem to play an important role, suggesting that the adjustment cost related to differences in local production structures is not very high. Contrary to migration flows, wage growth is not well explained by city fixed effects.[15]

[14] In all cities but three -for which it is very close to zero, mean wage growth within the city is positive.

[15] Therefore, for parsimony, and to keep some symmetry with migration flows, the preferred specification that will be used in Section 3.2 will be the one described in column (3).

2 Job search between many local labor markets: theory

We develop a dynamic migration model where individuals can move between a set of interconnected local labor markets. We consider steady-state implications in terms of job search and migration behavior following our descriptive evidence. Our objective is to include the structural determinants of migration into a setting that allows us to quantify the respective roles of spatial frictions and mobility costs in worker's geographical mobility.

2.1 Framework

We consider a continuous time model, where infinitely lived, risk neutral agents maximize their expected steady-state discounted (at rate r) future income. The economy is organized as a system $\mathscr{J}$ of J interconnected local labor markets, or "cities", where a fixed number of M workers live and work. While the spatial position of each city j within the system is exogenous, total population m_j and unemployed population u_j, are determined by the job search process. Wage offers are drawn from a distribution $F(\cdot) \equiv \{F_j(\cdot)\}_{j\in\mathscr{J}}$ of support $[\underline{w}, \overline{w}]^J \subset (b, \infty)^J$, resulting from firms' exogenous wage posting strategy.[16] Let $\overline{F}_j(\cdot) \equiv 1 - F_j(\cdot)$. Workers do not bargain over wages. They only decide whether to accept or refuse the job offer they have received. We note $G(\cdot) \equiv \{G_j(\cdot)\}_{j\in\mathscr{J}}$ the resulting distribution of earnings observed in the economy.

Spatial segmentation and migration Cities are heterogeneous, both in terms of labor market and living conditions. Employed workers in city j face a location-specific unemployment risk characterized by the layoff probability δ_j. This probability reflects the idiosyncratic volatility of local economic conditions. When they become unemployed, workers receive uniform unemployment benefits b.[17]

All workers in city j face an indirect utility γ_j, which summarizes the difference between amenities and (housing) costs in city j. This parameter may be interpreted as the average valuation of city j among the population of workers. Amenities are an equilibrium that rationalize why perfectly mobile unemployed workers do not take advantages. The value of γ_j is separable from the level of earnings, such that the instant value of a type-i worker in city j equals $y^i + \gamma_j$, with $y^u = b$ and $y^e = w$. This specification accounts for differences in local costs of living, so that wages can still be expressed in nominal terms.

[16] The assumption of a unique maximal wage does not have any impact our estimation results.

[17] These benefits will be calibrated in monetary terms yet they also comprise a measure of the value of leisure as well as the opportunity cost of time.

Workers are ex ante identical and fully characterized by their employment status $i = e, u$, their wage level w when employed and their location $j \in \mathscr{J}$. They engage in both off-the-job and on-the-job search. Their probability of receiving a new job offer depends on their current employment status, their location, as well as on the location associated with the job offer itself. Since we do not model these rates as originating from a matching function, they do not have a fully structural interpretation, even though they may arise from the spatial distribution of jobs and the heterogeneity in local matching technologies.

Frictions reduce the efficiency of job search between cities: type-i workers living in location j receive job offers from location $l \in \mathscr{J}_j \equiv \mathscr{J} - \{j\}$ at rate $s^i_{jl}\lambda^i_l \leq \lambda^i_l$. In addition, when they finally decide to move from city j to city l, workers have to pay a lump-sum mobility cost c_{jl}. They are perfectly mobile, in the sense that anybody can always decide to pay c_{jl}, move to city l and be unemployed there. However, because of congestion externalities affecting the job finding rate for the unemployed λ^u_j and the amenity value γ_j, this type of behavior will be ruled out in equilibrium and migration will only occur in case workers have found and accepted a job.

Workers' value functions Let $(x)^+ \equiv \max\{x, 0\}$. Workers do not bargain over wages. They only decide whether to accept or refuse the job offer which they have received. The respective value functions of unemployed workers living in city j and of workers employed in city j for a wage w are recursively defined by equations 1 and 2:

$$
\begin{aligned}
rV^u_j &= b + \gamma_j + \lambda^u_j \int_{\underline{w}}^{\overline{w}} \left(V^e_j(x) - V^u_j\right)^+ dF_j(x) + \sum_{k\in\mathscr{J}_j} s^u_{jk}\lambda^u_k \int_{\underline{w}}^{\overline{w}} \left(V^e_k(x) - c_{jk} - V^u_j\right)^+ dF_k(x) \qquad (1)\\
rV^e_j(w) &= w + \gamma_j + \lambda^e_j \int_{\underline{w}}^{\overline{w}} \left(V^e_j(x) - V^e_j(w)\right)^+ dF_j(x) + \sum_{k\in\mathscr{J}_j} s^e_{jk}\lambda^e_k \int_{\underline{w}}^{\overline{w}} \left(V^e_k(x) - c_{jk} - V^e_j(w)\right)^+ dF_k(x)\\
&+ \delta_j\left[V^u_j - V^e_j(w)\right] \qquad (2)
\end{aligned}
$$

2.2 Workers' strategies

Accepting an offer in a city conveys city-specific parameters. Jobs are defined by a non-trivial combination of wage and all the structural parameters of the economy, which determines the offer's option value. By refusing an offer, workers would, in a sense, bet on their current unemployment against their future unemployment probability. A similar mechanism applies to job-to-job transitions. If workers are willing to accept a wage cut in another location, this decision is somewhat analogous to buying an unemployment insurance contract, or a path to better wage prospects. This multivariate, and dynamic trade-off allows us to define *spatial strategies*, where workers' decision to accept a job in a given city is not only driven by the offered wage and the primitives of the local labor market, but also by the employment prospects in all the other locations, which depend upon the city's specific position within the system. The sequence of cities where individuals are observed can then be rationalized as part of lifetime mobility-based careers.

Definitions In order to formalize the previous statements, we now describe the workers' strategies. These strategies are determined by the worker's location, employment status, and wage. They are defined by threshold values for wage offers. These values are deterministic and similar across individuals since we assume that workers are ex-ante identical. They consist of a set of reservation wages and a set of sequences of mobility-compatible indifference wages.

A reservation wage corresponds to the lowest wage an unemployed worker would accept in her location. Reservation wages, which are therefore location-specific, are denoted ϕ_j and verify $V_j^u \equiv V_j^e(\phi_j)$. The reservation-wage strategy is denoted $\phi \equiv \{\phi_j\}_{j\in\mathscr{J}}$. Mobility-compatible indifference wages are functions of wage which are specific to any ordered pair of locations $(j,l) \in \mathscr{J} \times \mathscr{J}_j$. These functions associate the current wage w earned in location j to a wage which would yield the same ex-post utility in location l, once the mobility cost c_{jl} is taken into account. They are denoted $q_{jl}(\cdot)$ and verify $V_j^e(w) \equiv V_l^e(q_{jl}(w)) - c_{jl}$. The migration strategy is denoted $q(\cdot) \equiv \{q_{jl}(\cdot)\}_{(j,l)\in\mathscr{J}\times\mathscr{J}_j}$. The definition of $q_{jl}(\cdot)$ extends to unemployed workers in city j who receive a job offer in city l: we have $V_j^u \equiv V_l^e(q_{jl}(\phi_j) - c_{jl}$. Finally, let $\chi_{jl}(w)$ denote another indifference wage, verifying $V_j^e(w) \equiv V_l^e(\chi_{jl}(w))$. This indifference wage equalizes the utility levels between two individuals located in cities j and l. We shall therefore refer to it as the "ex-ante" indifference wage, unlike the "ex-post" indifference wage $q_{jl}(w)$, which equalizes the utility level between one worker located in city j and the same worker after a move into city l. By definition, ex-ante indifference wages have a stationary property, whereby $\chi_{lk}(\chi_{jl}(w)) = \chi_{jk}(w)$. As will be made clear later, the introduction of $\chi_{jl}(w)$ is important to understand the role of mobility costs in the dynamics of the model.

Proposition 1 *OPTIMAL STRATEGIES*

- *Let* $\zeta_{jl} = \frac{r+\delta_l}{r+\delta_j}$. *The reservation wage for unemployed workers in city j and the mobility-compatible indifference wage in city l for a worker employed in city j at wage w are defined as follows:*

$$\phi_j = b + \left(\lambda_j^u - \lambda_j^e\right) \int_{\phi_j}^{\overline{w}} \Xi_j(x)dx + \sum_{k \in \mathscr{J}_j} \left(s_{jk}^u \lambda_k^u - s_{jk}^e \lambda_k^e\right) \left(\int_{q_{jk}(\phi_j)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(\phi_j))c_{jk} \right) \tag{3}$$

$$\begin{aligned} q_{jl}(w) &= \zeta_{jl} w + \left(\zeta_{jl}\gamma_j - \gamma_l\right) + (r+\delta_l)c_{jl} + \left(\zeta_{jl}\delta_j V_j^u - \delta_l V_l^u\right) + \zeta_{jl}\lambda_j^e \int_w^{\overline{w}} \Xi_j(x)dx - \lambda_l^e \int_{q_{jl}(w)}^{\overline{w}} \Xi_l(x)dx \quad (4) \\ &+ \zeta_{jl} \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \left(\int_{q_{jk}(w)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(w))c_{jk} \right) - \sum_{k \in \mathscr{J}_l} s_{lk}^e \lambda_k^e \left(\int_{q_{lk}(q_{jl}((w))}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{lk}(q_{jl}(w)))c_{lk} \right) \end{aligned}$$

with:

$$V_j^u = \frac{1}{r}\left[b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\overline{w}} \Xi_j(x)dx + \sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \left(\int_{q_{jk}(\phi_j)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(\phi_j))c_{jk} \right) \right] \tag{5}$$

$$\Xi_j(x) = \frac{\overline{F}_j(x)}{r + \delta_j + \lambda_j^e \overline{F}_j(x) + \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(x))} \tag{6}$$

- *Equations 3 and 4 define a system of J^2 contractions and admit a unique fixed point.*
- *The optimal strategy when unemployed in city j is:*
  1. *accept any offer φ in city j strictly greater than the reservation wage ϕ_j*
  2. *accept any offer φ in city $l \neq j$ strictly greater than $q_{jl}(\phi_j)$.*

  The optimal strategy when employed in city j at wage w is:
  1. *accept any offer φ in city j strictly greater than the present wage w*
  2. *accept any offer φ in city $l \neq j$ strictly greater than $q_{jl}(w)$.*

Proof In appendix A.1, we derive equations 3 and 4 using the definitions of ϕ_j and $q_{jl}(\cdot)$ and integration by parts. Then, in appendix A.2, we demonstrate the existence and uniqueness of the solution through an application of the *Banach fixed-point theorem.*

Interpretation The interpretation of Equation 3 is straightforward: the difference in the instantaneous values of unemployment and employment ($\phi_j - b$) reflects an opportunity cost, which must be perfectly compensated for by the difference in the option values of unemployment and employment. Those are composed of two elements: the expected wages through local job search and the expected wages through mobile job search, net of mobility costs.[18]

The interpretation of Equation 4 is similar. The difference in the instant values of employed workers in location l and location j is $[q_{jl}(w) + \gamma_l] - \zeta_{jl}[w + \gamma_j]$. The term $[\zeta_{jl}\gamma_j - \gamma_l]$ is a measure of

[18] Note that the classical result whereby reservation wages are not binding stands true here, because agents are homogeneous and workers are allowed to transition into unemployment within the same city at no cost. Therefore, no firm will ever post a wage that is never accepted by a worker.

the relative attractiveness of city j and city l in terms of amenities. The third term states that for job offers to attract jobseekers from distant locations, they have to overcome mobility costs. As for the difference in the option values of employment in city j and employment in city l, it is threefold. The first part is independent of the wage level and given by the difference in the value of unemployment, weighted by unemployment risk δ_j or δ_l. The second part is the difference in the expected wage following a local job-to-job transition and the third part is the difference in the expected wages that will be found through mobile job search, net of mobility costs.

This last term introduces the relative centrality and accessibility of city j and city l. Centrality stems from the comparison of the strength of spatial frictions between the two locations j and l and the rest of the world: a worker living in city j who receives an offer from city l must take into account the respective spatial frictions from city j and from city l to any tier location k that she may face in the future, in order to maximize her future job-offer rate. As for accessibility, it stems from the difference in the expected costs associated with mobile on-the-job search from city j and from city l: an individual living in city j who receives an offer from city l must take into account the respective mobility cost from city j and from city l to any tier location k that she may face in the future, in order to minimze the cost associated with the next move. Note that both the relative centrality and the relative accessibility measures depend on the current wage level w: cities may be more or less central and accessible depending on where workers stand in the earning distribution. Finally, note the dynamic feedback effect whereby mobility costs impact the wage that will be accepted in the new city, which in turn impacts *future* wage growth prospects in this new city.

2.3 Steady-state

We use steady-state conditions on labor market flows along with spatial equilibrium conditions to solve our model.

Spatial equilibrium Workers in city j are free to move into city l upon paying a mobility cost c_{jl} and becoming unemployed. Given the reservation wage strategy, this type of migration out of the labor market will mostly be an option for unemployed workers. However, the inflow of unemployed workers into an attractive location will generate congestion externalities which will negatively impact local amenities. In equilibrium, local amenities γ_j adjust such that no individual agent has an incentive to move without a job offer, and leads to the following definition:

Definition CONGESTION — the vector of city amenities $\Gamma = \{\gamma_j\}_{j \in \mathscr{J}}$ satisfies the set of constraints 7:

$$V_j^u \geq \max_{k \in \mathscr{J}_j} \{V_k^u - c_{jk}\} \tag{7}$$

As already explained in section 1, a cross-sectional description of the labor market as a system of cities is characterized by a set of city-specific populations and unemployment rates. If these multidimensional variables are constant, the economy can be described as a steady-state. We now describe the theoretical counterparts to these components.

Steady state distribution of unemployment rates At each point in time, the number of unemployed workers in a city j is constant. A measure $u_j \lambda_j^u \overline{F}_j(\phi_j)$ of workers leave unemployment in city j by taking a job in city j, whereas others, of measure $u_j \sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))$, take a job in another city $k \neq j$. These two outflows are perfectly compensated for by a measure $(m_j - u_j)\delta_j$ of workers who were previously employed in city j but have just lost their job. This equilibrium condition leads to the following definition:

Definition STEADY STATE UNEMPLOYMENT — the distribution of unemployment rates is given by $\mathscr{U} = \left\{\frac{u_j}{m_j}\right\}_{j \in \mathscr{J}}$, where:

$$\frac{u_j}{m_j} = \frac{\delta_j}{\delta_j + \lambda_j^u \overline{F}_j(\phi_j) + \sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))} \tag{8}$$

Steady state distribution of city populations Similarly, at each point in time, population flows out of a city equal population inflows. For each city j, outflows are composed of employed and unemployed workers in city j who find and accept another job in any city $k \neq j$; conversely, inflows are composed by employed and unemployed workers in any city $k \neq j$ who find and accept a job in city j. The equality between population inflow and outflow defines the following equation:

$$\begin{aligned}(m_j - u_j) \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x) + u_j \sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j)) \equiv \\ \lambda_j^e \sum_{k \in \mathscr{J}_j} s_{kj}^e (m_k - u_k) \int_{\underline{w}}^{\overline{w}} \overline{F}_j(q_{kj}(x)) dG_k(x) + \lambda_j^u \sum_{k \in \mathscr{J}_j} s_{kj}^u u_k \overline{F}_j(q_{kj}(\phi_k))\end{aligned} \tag{9}$$

Let $\mathscr{A}$ the matrix of typical element $\{\mathscr{A}_{jl}\}_{(j,l)\in\mathscr{J}^2}$ defined by:

$$
\begin{aligned}
\mathscr{A}_{jj} &= \frac{\left[\lambda_j^u \overline{F}_j(\phi_j) + \sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))\right] \times \left[\sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x)\right] + \delta_j \left[\sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))\right]}{\delta_j + \lambda_j^u \overline{F}_j(\phi_j) + \sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))} \\
\mathscr{A}_{jl} &= -\frac{\left[\lambda_l^u \overline{F}_l(\phi_l) + \sum_{k\in\mathscr{J}_l} s_{lk}^u \lambda_k^u \overline{F}_k(q_{lk}(\phi_l))\right] \times \left[s_{lj}^e \lambda_j^e \int_{\underline{w}}^{\overline{w}} \overline{F}_j(q_{lj}(x)) dG_l(x)\right] + \delta_l s_{lj}^u \left[\lambda_j^u \overline{F}_j(q_{lj}(\phi_l))\right]}{\delta_l + \lambda_l^u \overline{F}_l(\phi_l) + \sum_{k\in\mathscr{J}_l} s_{lk}^u \lambda_k^u \overline{F}_k(q_{lk}(\phi_l))} \qquad \text{if } j \neq l
\end{aligned}
$$

where off-diagonal elements equal the fraction of the population in the city in column who migrates into the city in row at any point in time, and diagonal elements equal the fraction of the population in the city in question who moves out at any point in time. Plugging Equation 8 into Equation 9, we recover a closed form solution for the system, written as:

$$
\mathscr{A}\mathscr{M} = 0 \tag{10}
$$

where $\mathscr{M}$ is the vector of city sizes $\{m_j\}_{j\in\mathscr{J}}$. This yields the following definition:

Definition STEADY STATE POPULATION — The distribution of city sizes is the positive vector $\mathscr{M} \equiv \{m_j\}_{j\in\mathscr{J}} \in \ker\mathscr{A}$ s.t. $\sum_{j\in\mathscr{J}} m_j = M$.

Note that Equation 9 defines a relationship between m_j and all the other city sizes in $\mathscr{M}$, whereas it is not the case for u_j, which is determined by a single linear relationship to m_j. The flow of workers into unemployment in city j is only composed of workers previously located in city j, whereas in Equation 9, the population in city j is also determined by the flow of workers who come from everywhere else and have found a job in city j.

Summary At steady state, this economy is characterized by a set of structural parameters and a wage offer distribution such that:

1. The reservation wage strategy ϕ in Equation 3 describes the job acceptation behavior of immobile unemployed workers.

2. The mobility strategy between two locations $q(\cdot)$ is defined by the indifference wage described in Equation 4.

3. The set of local amenities Γ satisfies the market clearing constraints described by equation 7.

4. The set of unemployment rates $\mathscr{U}$ is given by Equation 8.

5. The set of city populations $\mathscr{M}$ is solution to the linear system 10.

3 Estimation

The model is estimated by simulated method of moments. The estimator minimizes the distance between a set of empirical moments and their theoretical counterparts, which are constructed by solving the steady-state conditions of the model.

In Appendix B, we present a full set of solutions to solve the indifference wages and the functional equations. We take advantage of the structure of the model and use an embedded algorithm that allows us to recover a piecewise approximation of all indifference wages, and wage offer distributions. We present here the identification strategy as well as the moments used for estimation.

3.1 Identification

Let $\theta = \{\lambda_j^i, \delta_j, s_{jl}^i, c_{jl}\}_{(i,j,l)\in\{e,u\}\times\mathscr{J}\times\mathscr{J}_j}$ be the set of parameters to be estimated. The main challenge consists of identifying separately the matching rates $\{\lambda_j^i, \delta_j\}_{(i,j)\in\{e,u\}\times\mathscr{J}}$ from spatial and mobility frictions parameters $\{s_{jl}^i c_{jk}\}_{(j,k)\in\{e,u\}\times\mathscr{J}\times\mathscr{J}_j}$ and more specifically to this study, spatial frictions from mobility costs. While the parameters cannot be sequentially identified, we argue that a matching between a vector of simulated and empirical moment is likely to pin down separately the parameters of the model. Our main insight is that while spatial friction parameters are identified off of transition rates between pairs of cities, mobility costs are identified off of data on wages. In the rest of this section, we present this intuition more formally.

Lemma 1 *The indifference wages are identified from transition wages.*

Lemma 1, whose proof is trivial, is key to the rest of identification. It is an extension of the identification of minimum wage result in Flinn & Heckman (1982). Let $\mathbb{Q}_{jl}$ be the set of wages accepted following transitions off unemployment between a pair of cities (j,l). The minimum of $\mathbb{Q}_{jl}$ is a super-consistent estimator of $q_{jl}(\underline{w})$. The same logic applies to any $q_{jl}(w)$.[19] A corollary of Lemma 1 is that reservation wages $\{\phi_j\}_{j\in\mathscr{J}}$ are identified as well. The next result deals with the wage offer distribution.

Lemma 2 *Assume observed wages are i.i.d. draws from a stationary distribution, then there is a unique mapping between accepted wage distribution $\{G_j(\cdot)\}_{j\in\mathscr{J}}$ and wage offer distribution $\{F_j(\cdot)\}_{j\in\mathscr{J}}$.*

In standard partial search model, the identification of the wage offer distribution is based on a trun-

[19] In practice, the data requirement may be stringent even with large matched dataset like ours. One can alleviate this problem by grouping wages within a small bandwidth.

cation rule at the reservation wage, usually $g(x) = f(x|x > \phi)$.[20] This strategy is not feasible in our setting for two reasons. First, The existence of multiple cities complicates tremendously this derivation, as one has to condition not only on local reservation wage, but also on all the indifference wages of all other locations. Second, the presence of on-the-job search induces non-trivialities between offered and accepted wage. As a consequence, we opt against this strategy. Instead, we assume that exogenous wage offer distributions are defined as the solution to a steady-state constraint on local earning distributions (see Appendix A.3). While used mostly in equilibrium models, the steady-state relationship between observed and offered wages is a powerful tool to uncover their dependence. Equipped with lemmas 1 and 2, we can now study the identification of matching rates, spatial frictions and mobility cost.

Proposition 2 *The matching parameters* $\{\lambda_j^e, \lambda_j^u, \delta_j\}_{j\in\mathcal{J}}$ *are identified.*

Lemma 1 ensures that indifferences wages $q_{jl}(\cdot)$ are identified, while lemma 2 provides the wage offer distributions $F_j(\cdot)$. Following Magnac & Thesmar (2002), we can construct moment conditions to identify the model. Within city transition rates are natural choices. Transitions from unemployment to employment identify λ^u, since both $F(\cdot)$ and ϕ are known from lemmas 1 and 2. The same reasoning applies to the on-the-job search rate λ^e. Finally, job destruction rates δ are identified off of transitions into unemployment.

Proposition 3 *The moving cost and spatial frictions parameters are separately identified.*

The identification of spatial frictions can be established along the same lines as the other matching parameters. Using transition rates between cities for unemployed and employed workers, we can construct moments to identify the matrices s_{jl}^u and s_{jl}^e given that $q_{jl}(\cdot)$ The identification of the mobility cost parameters c_{jl} follows almost immediately: at this stage, they are the only unknown variables in the nonlinear equation defining the indifference condition evaluated at the minimum wage.

3.2 Parameterization

In addition to J city-specific wage offer distributions, the model is based on a set of parameters θ such that $|\theta| = 30,000$ with $J = 100$. In practice, estimating all these parameters would be too computationally demanding and would require to drastically restrict $\mathcal{J}$. We take an alternative path and we posit

[20] This raises some issues as multiple wage offer distribution could rationalize the accepted wage distribution. Flinn & Heckman (1982) provide a set of recoverability conditions for identification.

and estimate two parsimonious parametric models, inspired by the results in Table 2:

$$s^i_{jl} = \frac{\exp\left(s^i_{j0} + s^i_{0l} + s^i_1 d_{jl} + s^i_2 d^2_{jl} + s^i_3 h_{jl} + s^i_4 h^2_{jl}\right)}{1+\exp\left(s^i_{j0} + s^i_{0l} + s^i_1 d_{jl} + s^i_2 d^2_{jl} + s^i_3 h_{jl} + s^i_4 h^2_{jl}\right)} \tag{11}$$

$$c_{jl} = c_0 + c_1 d_{jl} + c_2 h_{jl} + c_3 d^2_{jl} + c_4 h^2_{jl} \tag{12}$$

where the city-pair specific variables are the ones used in Table 2, s^i_{j0} is a sending-city fixed effect and s^i_{0l} is a receiving-city fixed effect.

The model rests upon the premise that spatial friction parameters take on values in $[0,1]$, whereas the range of values for mobility costs is unrestricted. Given the lack of existing literature on the explicit structure of spatial frictions, we use a logistic function in Equation 11 because of its analytical properties.[21] Equation 11 is akin to a standard gravity equation: the fixed effects measure the relative openness of the local labor markets: either the ability of each city to dispatch its jobseekers to jobs located elsewhere (s_{j0}) or to fill its vacancies with workers coming from other locations (s_{0l}), and the other parameters account for the effect of distance between two locations.[22]

Physical distance is arguably the most important characteristic and both equations 11 and 12 rely on it. As for sectoral dissimilarity, it is used in order to reflect human capital specificity, which may be of particular importance to rationalize, for instance, job-to-job mobility rates between highly specialized but distant cities, between which workers will be likely to move because they will a good fit for local production needs (Bryan & Morten, 2017). We let returns to these two measures of distance vary by considering a second-order polynomial. Note that, in order to ensure continuity at the reservation wage, we assume that moving costs do not vary with labor market status, unlike spatial frictions.[23] Also, our estimates of mobility costs will depend on the pair of cities involved, but not on the direction of the move.[24]

Finally, in order to reduce the computational burden and ensure the smoothness of the density functions, we assume that $F(\cdot)$ follows a parametric distribution:

$$\hat{F}_j(x) = \text{betacdf}\left(\frac{x-b}{\overline{w}-b}, \alpha_j, \beta_j\right) \tag{13}$$

[21] See Zenou (2009) for a theoretical approach in terms of endogenous search intensity.
[22] See Head & Mayer (2014) for the current state of the art about gravity equations.
[23] This assumption may not be fully innocuous if unemployed jobseekers have access to some specific segments of the housing market, such as public housing.
[24] This symmetry assumption could easily be relaxed, for instance by including an indicator variable on whether the destination city is larger or smaller than the departure city, as in Kennan & Walker (2011). However, as shown by Levy (2010) on US data, this may not be empirically relevant.

where betacdf$(\cdot, \alpha_j, \beta_j)$ is the cdf of a beta distribution with shape parameters α_j and β_j. As argued by Meghir et al. (2015), this distribution is very flexible, without sacrificing parsimony. Under the specifications detailed in equations 11, 12 and 13, the number of parameters to be estimated amounts to 913.

3.3 Moments

Table 3 summarizes our choice of theoretical and empirical moments. In addition to the raw transitions between employment and unemployment, we use city-specific populations and unemployment rates to identify λ^u and δ.

Given the parameterization of s^i_{jl}, the model is over-identified: in particular, the $2J(J-1)$ transition rates at the city-pair level that would be required to identify each parameter s^i_{jl} are no longer needed. In order to identify the fixed-effect components, we use the $2(J-1)$ total transitions rates into and out of any given city. On the other hand, the identification of the parameters related to the distance and the dissimilarity between two cities does still require transition rates at the city-pair level. Given that Equation 11 only specifies four parameters for each labor market status, we drastically restrict the set of city pairs, down to a subset $\mathscr{T}_1 \subset \mathscr{J} \times \mathscr{J}_j$, with $|\mathscr{T}_1| = 48$, which we use in the estimation.[25]

While spatial friction parameters are identified off of transition rates between pairs of cities, mobility costs are identified from data on wages. Given there only are five parameters to estimate, we select a subset of city pairs $\mathscr{T}_2 \subset \mathscr{J} \times \mathscr{J}_j$ such that $|\mathscr{T}_2| = 12$.[26]

Finally, when all the parameters described in Table 3 have been estimated, we can recover the amenity parameters γ_j through an embedded algorithm that aims to satisfy the set of constraints described in Equation 7 (see Appendix B.4 for details).

[25]In practice, we use the off-the-job and job-to-job transitions rates from the urban areas ranked fourth to eleventh (Toulouse, Lille, Bordeaux, Nice, Nantes, Strasbourg, Grenoble and Rennes) to the urban areas ranked fifteenth, nineteenth to twenty-second and twenty-fifth (Montpellier, Clermont-Ferrand, Nancy, Orléans, Caen and Dijon). This selection is designed to include locations that are widely scattered across the French territory (see Figure 11 in appendix D.1 for details).

[26]In practice, we use the average accepted wages following a job-to-job transition between the cities ranked second to fifth (Lyon, Marseille, Toulouse and Lille). This subset has to be more restrictive than $\mathscr{T}_1$ because, while very low transitions rates convey reliable information since they are drawn from large initial populations, they do not allow to compute accurate measures of average accepted wages. Note that for homogeneity concerns, we do not include Paris, because its size is too large compared to the other cities.

Table 3: Moments and Identification

Empirical moments	Theoretical moments
Unemployment rate in city $j \in \mathscr{J}$	u_j / m_j
Labor force in city $j \in \mathscr{J}$	m_j
Transition rate ee within city $j \in \mathscr{J}$	$\lambda_j^e \int_{\underline{w}}^{\overline{w}} \overline{F}_j(x) dG_j(x)$
Earning distribution in city $j \in \mathscr{J}$	$G_j(\cdot)$
Transition rate ue out of city $j \in \mathscr{J}$	$\sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\underline{w}))$
Transition rate ue into city $l \in \mathscr{J}$	$\lambda_l^u \sum_{k \in \mathscr{J}_l} s_{kl}^u \overline{F}_l(q_{kl}(\underline{w}))$
Transition rate ee out of city $j \in \mathscr{J}$	$\sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x)$
Transition rate ee into city $l \in \mathscr{J}$	$\lambda_l^e \sum_{k \in \mathscr{J}_l} s_{kl}^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{kl}(x)) dG_k(x)$
Transition rate ue from city j to city l, $(j,l) \in \mathscr{T}_1$	$s_{jl}^u \lambda_l^u \overline{F}_l(q_{jl}(\underline{w}))$
Transition rate ee from city j to city l, $(j,l) \in \mathscr{T}_1$	$s_{jl}^e \lambda_l^e \int_{\underline{w}}^{\overline{w}} \overline{F}_l(q_{jl}(x)) dG_j(x)$
Accepted wage ee between city j and city l, $(j,l) \in \mathscr{T}_2$	$q_{jl}(\cdot)$

Notes: For details on the construction of the empirical moments, see Appendix C.3.

3.4 Fit

We parametrize unemployment benefit $b = €6{,}000$ annually (an approximation of the minimum guaranteed income, which amounts to about half of the minimum wage) and r = 2%. The minimum wage is set to 10,874 €annually. The model is optimized using sequentially derivative free (Nelder-Mead, BoByQa, and Subplex) and Quasi-Newton optimization techniques (BFGS). Integrals are evaluated numerically using trapezoidal integration rule. Standard deviations are obtained by bootstrap (we use a total of 500 replications).

Figures 12 and 13 in Appendix D show that the model allows us to reproduce well many features of the data. The model predicts almost perfectly city-level job arrival rate for unemployed workers. In addition, the distribution of city sizes is well replicated by the solution of a linear system of equations (Figure 12). The mobility rates, which are not directly targeted by our estimation, are also very well predicted, especially the migration rates for the unemployed (Figure 13). Finally, Figure 14 shows that our model is quite able to replicate average wage growth following job-to-job transitions, while this moment is not directly targeted by our estimation either. Our standard errors suggest that all our parameters are estimated very precisely. Over the 913 parameters, only 42 are non-significant at 5%. The bulk of these parameters are job destruction rates.

4 Results

In this section, we first present our city-specific parameter estimates and their correlations. Then, we discuss the quantitative implications of our estimations of spatial frictions and mobility costs. Finally, we provide a decomposition exercise to quantify the respective impact of off and on-the job search, both within and between cities, on the city size wage gap.

4.1 A dataset of city-specific parameters

Table 4 provides summary statistics of the city-specific matching and amenity parameters and the two transition rates out of each city.[27] All the values are given in yearly terms. The estimated values of λ_j^u, which range from 0.37 to 1.8, show substantial heterogeneity across cities, suggesting city average unemployment durations from 7 months to 2 years and 6 months in the absence of migration. The median value of 0.95 confirms the low transition rate of the French economy as documented by Jolivet, Postel-Vinay & Robin (2006). Local job arrival rates for employed workers are low in comparison, but encompass more heterogeneity than λ_j^u. In addition, this difference is partly compensated for by the relative weight of mobile job-to-job transitions, which is far more important: the median value of λ_j^e and $e_j e_l$ is very similar, while it is four times lower for $u_j e_l$ than for λ_j^u (see section 4.2

Table 4: City-specific parameters: summary statistics

	λ_j^u	λ_j^e	δ_j	γ_j	$u_j e_l$	$e_j e_l$
Min	0.37	0.015	0.017	-0.2	0.11	0.0095
1st De.	0.61	0.018	0.11	-0.075	0.18	0.016
1st Qu.	0.77	0.022	0.12	-0.041	0.21	0.02
Median	0.95	0.027	0.13	-0.0064	0.26	0.026
Mean	0.95	0.029	0.13	0.00015	0.28	0.028
Sd	0.26	0.0096	0.027	0.066	0.092	0.013
3rd Qu.	1.1	0.034	0.14	0.047	0.33	0.032
9th De.	1.3	0.041	0.16	0.079	0.42	0.04
Max	1.8	0.073	0.19	0.16	0.53	0.085

Notes: (i) Yearly values of the matching parameters; (ii) $u_j e_l$ is the out-migration rate for unemployed workers and $e_j e_l$ is the out-migration rate for employed workers. (iii) Each distribution is evaluated on 99 cities; (iv) The numeraire of γ_j is the maximum wage, roughly equal to 96,000 in 2002 €.

[27] Those are respectively given by $u_j e_l = \sum_{k \in \mathcal{J}_j} u_j e_k$ and $e_j e_l = \sum_{k \in \mathcal{J}_j} e_j e_k$, where $u_j e_k \equiv s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\underline{w}))$ and $e_j e_k \equiv s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x)$.

below). Overall, there still is considerable heterogeneity across cities in terms of job arrival rates, whereas involuntary job separation rates δ_j are much less dispersed.

The relationships between the parameters and the local distributions of earnings, captured by the average and the coefficient of variation, are revealing. The correlation matrix, displayed in Table 5, shows a positive co-movement ($\rho = 0.22$) between on-the-job search rate and wage dispersion at the city level, which is in line with the insights of the wage posting theory, as outlined by Burdett & Mortensen (1998). Second, a positive correlation between the average level of earnings and the likelihood of receiving job offers is observed, unsurprisingly, for the job arrival rate of employed workers ($\rho = 0.60$), but also for the job arrival rate of unemployed workers ($\rho = 0.38$) as well. Third, cities with high local job finding opportunities are also well integrated in the urban system, both for unemployed workers ($\rho = 0.59$) or for employed workers ($\rho = 0.29$). Finally, there is a large negative correlation between the average level of earnings and the value of local amenities ($\rho = -0.33$), which is in line with the compensation mechanism that has been put forward in the spatial equilibrium literature (Roback, 1982). However, only 9% of the differences in local wage levels are explained by differences in local amenities.

Table 5: Correlations between estimates and labor market primitives

	λ_j^u	λ_j^e	δ_j	γ_j	$u_j e_l$	$e_j e_l$	$\mathbb{E}(w)$
λ_j^e	0.16						
δ_j	0.10	0.31**					
γ_j	-0.85**	-0.06	-0.34**				
$u_j e_l$	0.59**	-0.30**	-0.15	-0.70**			
$e_j e_l$	-0.10	0.29**	-0.39**	0.49**	-0.24*		
$\mathbb{E}(w)$	0.38**	0.60**	0.13	-0.33**	0.07	0.13	0.10
cv	0.12	0.22*	0.06	0.01	-0.04	0.24*	-0.04

Notes: (i) Wage distributions are evaluated over the six-year span 2002-2007 (ii) ** and * denote respectively significance at the 99% and the 95% confidence levels; (ii) $\mathbb{E}(w)$ is the average wage in each city and cv is the coefficient of variation of wages in each city; (iii) Each correlation is evaluated on 99 cities.

4.2 Spatial frictions and mobility costs

We now quantify spatial constraints using our parameter estimates. We first discuss the relative magnitude of local job arrival rates and mobile job arrival rates, for a given worker. Our estimation results allow us to quantify by which factor the local job arrival rates would need to be raised to make up for a hypothetical situation where workers would no longer be able to migrate to other cities.

Table 6 shows the distribution of the ratios to local job finding rate to the migration rates for both employed and unemployed workers. The heterogeneity documents the diversity of cities with respect to how much workers rely on migration to receive new job offers. For employed workers, local

Table 6: Relative magnitude of local and mobile job arrival rates

	Unemployed workers (1)	Employed workers (2)
Min	0.092	0.53
1st De.	0.21	0.71
1st Qu.	0.24	0.78
Median	0.29	0.89
Mean	0.3	1
Sd	0.086	0.64
3rd Qu.	0.35	1
9th De.	0.42	1.4
Max	0.64	5.3

Notes: (i) Ratio of between to within job arrival rates: it is equal to $\sum_{k\in\mathcal{J}_j} u_j e_k/\lambda_j^u$ in column (1) and to $\sum_{k\in\mathcal{J}_j} e_j e_k/\lambda_k^e$ in column (2); (ii) The distribution is evaluated on 99 cities.

job finding rate would need to be raised by at least 53%, and in half of the cities, by more than 89% (column (2)) to compensate for the absence of migration. On the other hand, unemployed workers are less reliant on mobile job search: in almost all cities, the local job finding rate would only need to be raised by half to make up for the impossibility to search in other cities (column (1)). This difference between the situation of unemployed and employed workers, which could not be observed from raw transition data (as described in Table 1, mobile transitions make up for one fifth of all transitions in the data, regardless of initial employment status), is economically meaningful and may not warrant the same policies.

Reassessing the magnitude of mobility costs According to our results, average mobility costs between cities amount to between €13,700 and €16,900. These figures are twenty times lower than the mobility cost found by Kennan & Walker (2011), who estimate a value of $312,000 for the average mover, which yields a negative average cost of $80,768 for realized moves. There is a number of reasons that might explain these differences. Inter-state migration in the U.S. is in-between domestic migration and international migration within the E.U. However, we do believe the introduction of spatial frictions to be the main driver of this stark difference. In order to give support to this intuition, we re-estimated a simplified version of the model where all spatial friction parameters were set equal to one. The resulting estimates for mobility costs ranged between €90,448 and €100,847.

Our estimates for Equation 12 indicate that the mobility cost function is given by $\hat{c}_{jl} = 12,401(2,361) + 137.56(7.05)d_{jl} + 22.06(18.1)h_{jl} + 7.95(18.1)d_{jl}^2 - 39.7(38.4)h_{jl}^2$, with 100 km as the unit for distance (which yields a maximum value for distance equal to about 12.4). This function is positive and increasing for all possible values of distance, which means that we do not find any evidence of negative mobility costs (or relocation subsidies) in the French labor market, at least at the city-pair level. As for the impact of sectoral dissimilarity, it is much lower in magnitude (the maximum value of the dissimilarity index is 0.48 in the data) and not statistically significant, in line with the empirical evidence presented in Section 1.2.

The value of $\hat{c}_{jl}$ amounts to an (unweighted) average value of €15,461, which is approximately equal to 1.5 times the annual minimum wage, or, more significantly, to the first quartile of the annual wage distribution. While much lower than previously reported in the literature, mobility cost may still prevent some workers at the bottom of the wage distribution from taking advantage of distant job opportunities, as will be discussed in section 5.2. The fixed component of the mobility cost, equal to €12,401, accounts for more than 80% of the average, which may strike as high. However, one has to bear in mind that it may include both relocation costs, transaction costs on the housing market and psychic costs related to the loss of local network connections.

In addition, as shown in Table 7, distance still generates some variation in mobility costs, even if the average values are quite close to each other. The dispersion of mobility costs involving a move to Paris, which is the most central city among the cities with more than 500K inhabitants, is 14% lower than the corresponding value for Lyon (the second city, which is also fairly central), 16% for Lille (the

Table 7: Distribution of the mobility costs involving all cities or some of the largest cities

	All	Paris	Lyon	Lille	Nice
Min	13,739	14,253	14,609	14,253	14,251
1st Qu.	15,170	15,427	15,430	15,347	15,520
Median	15,533	15,700	15,678	15,675	15,936
Mean	15,515	15,662	15,638	15,650	15,850
Sd	516	365	425	437	548
3rd Qu.	15,879	15,947	15,844	15,997	16,187
Max	16,933	16,341	16,729	16,427	16,910

Notes: (i) There are 9,702 possible moves between 99 cities; however, given the symmetry assumption $c_{jl} = c_{lj}$, only half of these moves are needed to compute these distributions; (ii) Costs are given in €2002.

fifth city, on the Belgian border) and 33% for Nice (the seventh city, in the far south-eastern corner of the country). Comparing the cumulative distributions, one may also note, for instance, that 75% of the moves involving Lyon are cheaper than half of the moves involving Nice.

4.3 Decompositions

Having documented the differences between cities in terms of local factors (internal labor market conditions) and spatial factors (integration to the city network), we now provide a decomposition analysis of the two main proxies for cross-sectional economic well-being at the city level: average wage and unemployment rate.

Local unemployment In our framework, the unemployment rate is determined by the local layoff rate, the local job arrival rate and a compound job arrival rate accruing from all other cities. In order to quantify their respective impact on unemployment, Table 8 provides the regression estimates of a linear equation of local unemployment rate as a combination of these three factors. Results show that accounting for job search between cities leads to a 23% increase in the explanatory power of the model and to a 30% decrease in the estimate associated with the local job arrival rate (columns (1) to (2)). When the three factors are included, they explain almost 80% of the heterogeneity in local unemployment rates (column (3)).

Table 8: Decomposition of the local unemployment rate

	(1)	(2)	(3)
λ_j^u	-0.017**	-0.011**	-0.013**
	(0.001)	(0.002)	(0.001)
$u_j e_l$		-0.029**	-0.023**
		(0.004)	(0.004)
δ_j			0.066**
			(0.010)
R^2	0.57	0.70	0.79

Notes: (i) Ordinary-least squares estimates of equations of the local unemployment rate as a function of the estimated parameters of equation 8; (ii) ** denotes significance at the 99% level; (iii) A constant term is included in each estimation; (iv) The regression is performed over the 99 first cities.

The city size wage premium There is broad evidence that wages are higher in larger cities. In the presence of imperfect information, employers may retain high wage-setting power even in the presence of many competing firms, because, among other reasons, of search frictions, or mobility costs. The ability of workers to search for new jobs while employed compels firms to increase wages. This argument is at the core of thickness theories of the city-size wage premium, whereby workers earn more in larger cities also because they get more offers. While quantitative evidence on this phenomenon is scarce and mostly focused on spatial bargaining mechanisms, our framework allows us to tackle with this question.

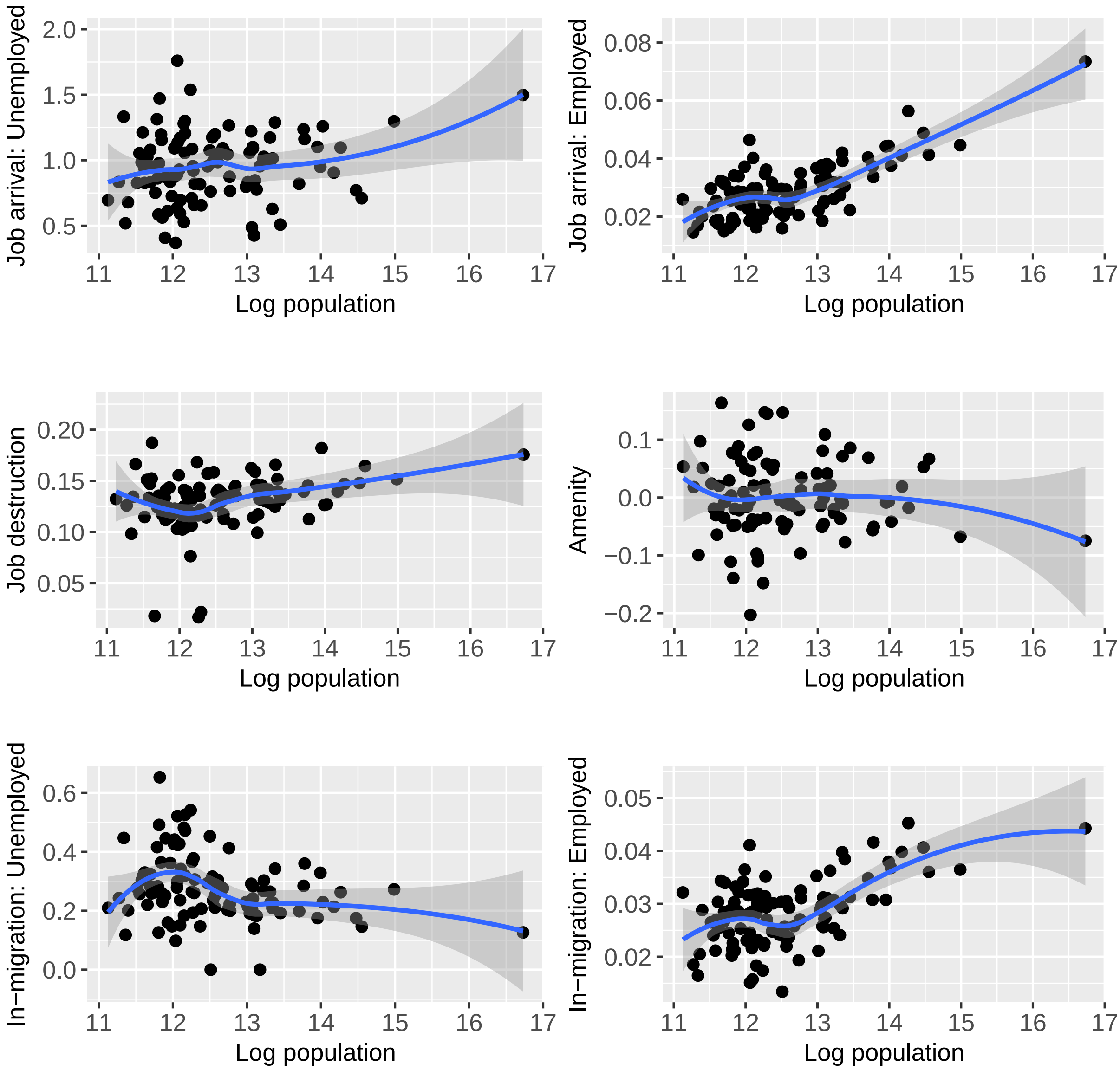
As already shown in Table 5, average wages are positively correlated with job-finding rates, uncorrelated with layoff rates and negatively correlated with amenities. Figure 4 completes this picture by displaying the distribution of the matching, layoff and amenity parameters according to city size.

Local job arrival rates, both for unemployed and employed workers, are larger in larger cities, unlike layoff rates and amenity parameters, which are not correlated with city size. This is suggestive evidence that workers in big cities benefit from high job finding rates possibly originating from matching economies which are not fully offset by congestion externalities or capitalized into local price levels. This is particularly true of employed workers: for unemployed workers, the positive correlation between the job arrival rate and city size is driven by one point, the city of Paris, which plays a very specific role in the French economy. The same pattern may be observed for the distribution of migration rates into each city: the bottom two graphs in Figure 4 show that unemployed workers do not tend to move more often into larger cities, whereas higher job arrival rates from larger cities attract previously employed workers into them.

Given larger cities are characterized by both higher local and spatial job-to-job transition rates, one may wonder which of the two features, if any, has more impact on the local level of wages. In our controlled environment, the expected wage in city j, denoted $\mathbb{E}(w|j) \equiv \int_{\underline{w}}^{\overline{w}} x g_j(x) dx$, may be expressed as a function of labor market primitives using equation 31. This yields:

$$\begin{aligned}\mathbb{E}w|j \quad &= \underbrace{\int_{\underline{w}}^{\overline{w}} x \urcorner_j(x) \lambda_j^u u_j dx}_{\text{Local UE}} + \underbrace{\int_{\underline{w}}^{\overline{w}} x \urcorner_j(x) \lambda_j^e (m_j - u_j) G_j(x) dx}_{\text{Local OTJ}} \\ &+ \underbrace{\int_{\underline{w}}^{\overline{w}} x \urcorner_j(x) \lambda_j^u \sum_{k \in \mathscr{J}_j} s_{kj}^u \psi_{kj}(x) u_k dx}_{\text{Migration UE}} + \underbrace{\int_{\underline{w}}^{\overline{w}} x \urcorner_j(x) \lambda_j^e \sum_{k \in \mathscr{J}_j} s_{kj}^e (m_k - u_k) G_k\left(q_{kj}^{-1}(x)\right) dx}_{\text{Migration OTJ}}\end{aligned} \tag{14}$$

Figure 4: Job-finding and layoff rates, local amenities and city size

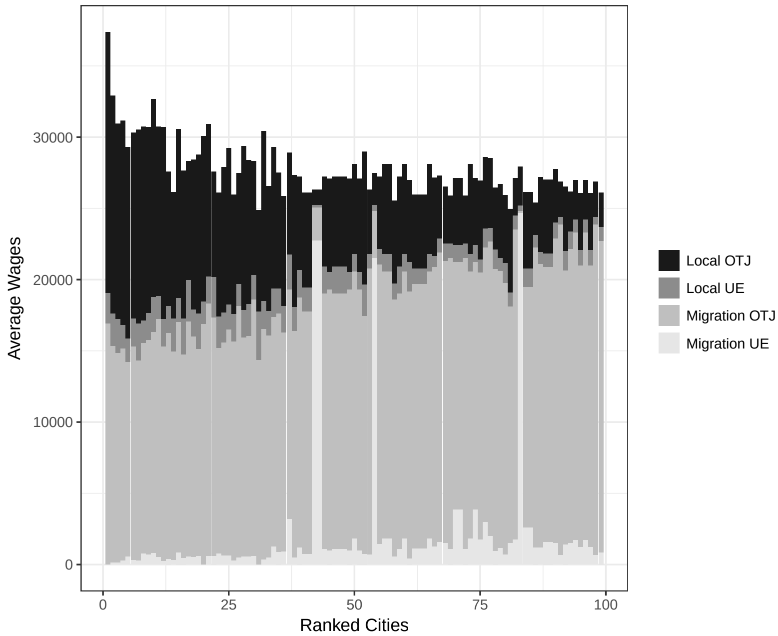


Notes: (i) Estimated values of the structural parameters $(\lambda^u, \lambda^e, \delta, \gamma)$ for the first 99 cities; (ii) See Table 4 for more details.

where $\psi_{kj}(w) = \mathbb{1}_{w>q_{kj}(\phi_k)}$ is a dummy variable indicating whether unemployed jobseekers in city k are willing to accept the job paid at wage w in city j and $\daleth_j(w) := f_j(w) \Big/ [(m_j - u_j)(\delta_j + \lambda^e_j \overline{F}_j(w) + \sum_{k\in\mathscr{J}_j} s^e_{jk} \lambda^e_k \overline{F}_k(q_{jk}(w)))]$. Local UE and Local OTJ denote respectively the contribution of local search by unemployed and employed workers, while Migration UE and Migration OTJ correspond to the contributions from search between cities. Note that the density $f_j(w)$ enters the four components in a similar fashion, so that the decomposition will allow us to assess the impact of higher job arrival rates on the level of wages, net of differences in local wage offer distributions.

Results are given for each city in Figure 5. Three main features stand out: first, the urban wage premium is entirely driven by local job-to-job transitions; second, transitions out of unemployment play almost no role, apart from a handful of exceptions, in wage disparities; third, job-to-job transitions between cities act as a mitigating factor of the impact of their local counterparts. Observed high wages in smaller cities are almost entirely due to incoming on-the-job search transitions, but those workers do not experience any wage increase after their mobility. On the contrary, in larger cities, the contribution of migration OTJ is lower because workers accept lower initial wages, relying on subsequent local wage growth opportunities.

Figure 5: Decomposition of the urban wage premium



Notes: (i) Decomposition of the wage level between the four components described in Equation 14;
(ii) The x-axis is the rank of the city according to population size.

To summarize, our results suggest that city-size wage premium results from higher frequencies of on-the-job search in larger cities. The fact that smaller cities offer initial high wage for incoming on-the-job workers, as well as the parameters of the wage offer distribution, provide evidence against the hypothesis of right-shifted wage offer distributions in larger cities.

5 Experiments

In this section, we take advantage of our empirical setting to describe several experiments based on simulating forward a system of cities where a population of 100,000 agents live for 40 years.[28]

5.1 Birthplace, migration and inequalities

Do individuals born in larger cities achieve more lifetime earnings? To what extent can migration help individuals initially located in less affluent cities achieve higher lifetime utilities? A definite answer to these questions is crucial to our understanding of persistent spatial inequalities. We use our simulation results to assess the contribution of space (initial location and mobility) to lifetime earnings.

Spatial inequalities are evaluated from the elasticity of average local income and local income dispersion with respect to city size. In cross-section, earnings are both higher and more dispersed in larger cities. Our estimates of the elasticity of average income to city size is equal to 6.5%, and the elasticity of the inter-quartile ratio to city size is equal to 6.3%.[29] Accounting for differences in local unemployment risk does not change this conclusion, because city-specific unemployment rates are not correlated with city size (see Figure 8 in Appendix C.2).

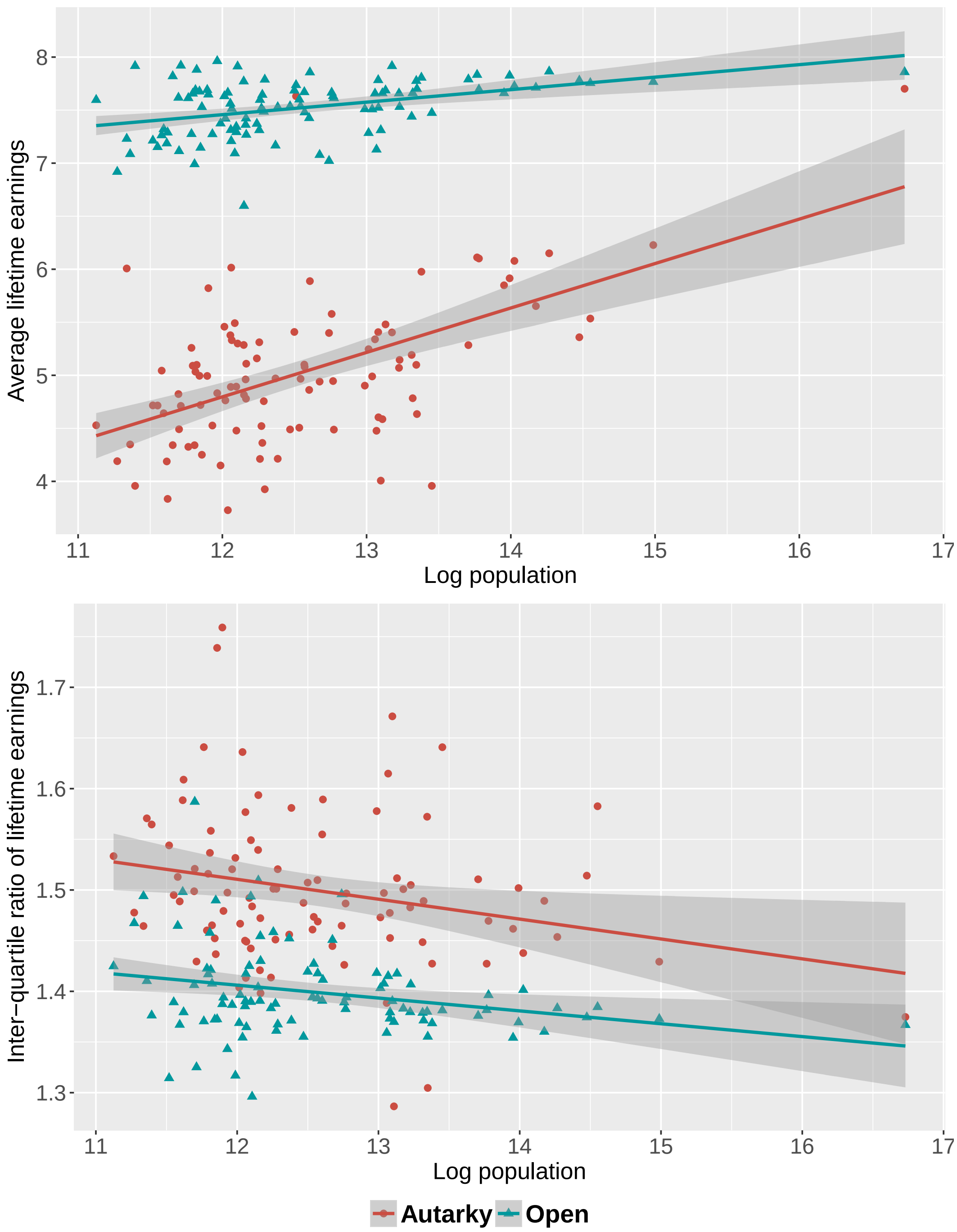
We now turn to lifetime inequalities. On-the-job search is the main source of earning dispersion in our setting, and given the strong correlation between city size and on-the-job arrival rates, one could expect higher level of inequalities in the long-run. However, our simulation results show the opposite. Within-city dispersion of lifetime earnings is not higher in larger cities: the elasticity of the inter-quartile ratio to city size is even slightly negative. The higher frequency of transitions in larger cities acts as an equalization device for lifetime values. The result is similar to findings by Flinn (2002), who shows that higher labor market transition rates achieve a more equitable welfare in the long run.

In addition, the elasticity of average lifetime earnings to city size decreases to 2.2%. Arguably, this result shows that migration allows some workers to insure themselves against bad initial location draws. To confirm this intuition, we perform a second simulation where the economy consists of a set of autarkic local labor markets ($\forall (i,j,l) \in \{e,u\} \times \mathcal{J} \times \mathcal{J}_j, s^i_{jl} = 0$) and workers cannot move from their initial location. Results are summarized in Figure 6. Under autarky, the elasticity of citywide mean lifetime income rises up to 10.9%: more dynamic labor markets in larger cities exacerbate initial in-

[28]This simulation rests upon the assumption that the steady state observed between 2002 and 2007 will last over the agents' lifetime.

[29]These estimates are in line with the literature. For instance, the raw wage elasticity to city size is equal to 6.8% in Bosquet & Overman's (2016) British data. After controlling for worker sorting, these estimates decrease by half (Combes, Duranton & Gobillon, 2008).

Figure 6: Lifetime earnings and city size



Notes: (i) Simulation results for two economies, one characterized by the entire set of estimated parameters: "Open" and one characterized by the same set of parameters to the exclusion of spatial friction parameters, which are all set to zero: "Autarky". (ii) Average lifetime earnings are divided by 10^5. They are equal to the mean observed value of discounted lifetime income, by city. Inter-quartile ratios of lifetime earnings correspond to the ratio of the 75^{th} percentile to the 25^{th} percentile of the lifetime earning distribution evaluated for each city.

equalities. On the other hand, the possibility of migration plays little role on within-city inequalities. The mechanism highlighted above is still much in play, which shows that the tendency of bigger cities to be more open to the city network is of second order for within-city inequalities, compared to differences in local search opportunities.

Interestingly, Paris is almost unaffected by autarky, thanks to the strength of its internal labor market. From an aggregate viewpoint, under local autarky, the economy as a whole is more unequal, with an inter-quartile ratio equal to 1.7 against 1.4 in the open economy and other measures of inequality follow the same pattern. Workers are also poorer, with an average value of lifetime income that decreases by 15% (770,000 euros against 650,000). Losses are relatively bigger in small or medium cities. To sum up, inequalities between cities are lower across workers' lifetime than in cross-section. Within cities lifetime inequalities are lower in large cities while the opposite stands true for cross-sectional inequalities. Migration decreases lifetime inequality between cities but not within cities.

5.2 Inefficiency of relocation subsidies

Given the lifetime gains from migration exhibited in the previous section, one might want to encourage labor mobility through relocation subsidies. There is not a lot of empirical evidence that these policies are effective. In reduced-form analysis, this type of policy may only be evaluated through natural experiments. A recent and isolated example is Caliendo, Künn & Mahlstedt's (2017) study of a German policy targeted towards unemployed workers, which shows some evidence of a positive impact on subsequent wages and job stability. However, if this policy incentivizes workers to accept offers in depressed areas, the long-run impact may prove less positive. In addition, competition effects with other workers (employed, or unemployed but not targeted) may also be large. Finally, the cost-effectiveness of the program is not discussed.

Thanks to our framework, our evaluation of the impact of this type of policy does not suffer from the same caveats. We perform a third simulation where mobility costs are set to zero. Given our estimates of mobility costs, this amounts to a €15,000 average subsidy, which is comparable to the $10,000 subsidy considered by Kennan & Walker (2011). The yearly mobility rate increases from 6.98% to 7.12%. This very modest increase comes in sharp contrast with Kennan & Walker's (2011) finding, whereby under such subsidy, the interstate migration would rise from 2.9% to 4.9%. As already discussed, and considering our different settings, such a gap is not surprising.

Another result of the simulation is that the entire increase in mobility is due to unemployed workers, for whom the mobility rate rises from 20.50% to 22.57%, while mobility actually decreases from 4.73% to 4.55% for employed workers: for employed workers, mobility costs seldom are a deterrent to migration, and they face more competition from more mobile unemployed workers. Finally, the policy is not cost-effective: it increases lifetime earnings by 1%, but its cost is, on average, equal to 2% of lifetime earnings. Targeting the policy towards unemployed workers only does not change this conclusion, because employed workers suffer from it even more, and the welfare implications are even less positive.

5.3 Local Minimum Wages

Minimum wages policies are prominent around the world. In Western Europe, most countries have a unique national wage, while in the U.S., there may exist a state-specific or even city-specific minimum wage. The impact and the desirability of minimum wage laws have been debated for decades. Despite a voluminous literature on the subject, there is very little consensus on the effect on unemployment and the wage distribution. From a theoretical perspective, the effect of minimum wage is only ambiguous in the presence of frictions or multiple equilibria. As such, our frictional model may be able to offer some answers, although the scope for analyzing optimal minimum wages is limited because of potentially large general equilibrium effect.

The indiscriminate effect of a single minimum wage on heterogenous locations may be the most serious flaw. As local labor markets convey different job opportunities and amenities, a national minimum wage introduces large distortions, destroying jobs that workers from distressed local labor markets would have accepted. This argument was already the core hurdle in the first attempt to establish a minimum wage in the U.S. because of the differences between local labor markets, notably between the North and the South.[30]

As a consequence, we define a concept of optimal minimum wage based on maximal labor mobility. That is, the minimum wage is binding when the indifference wage between a pair of locations is lower than national minimum wage. Let $\underline{w}_j$ the minimum wage in city j satisfying the following constraint:[31]

$$\underline{w}_j = \min_{k \in \mathscr{J}_j} q_{kj} \left(\underline{w}_k \right) \tag{15}$$

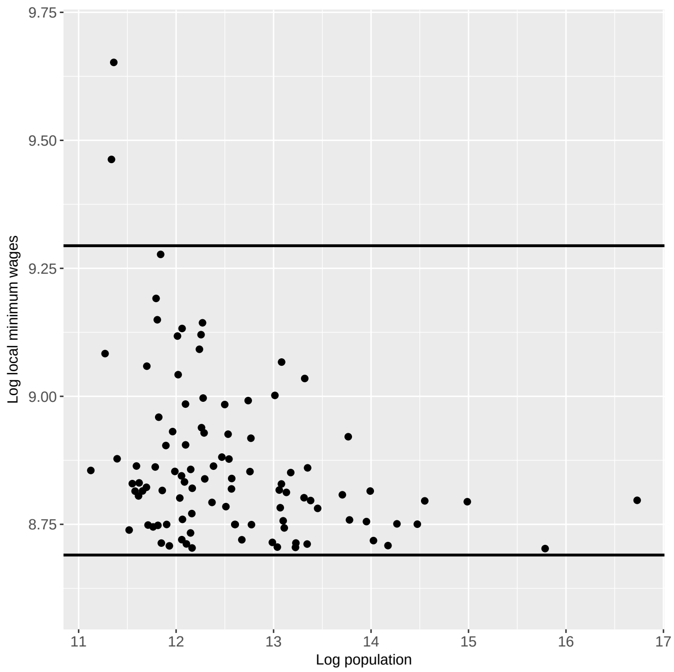
Figure 7 reports the estimated local minimum wages. Interestingly, almost all local minimum wages

[30] A more thorough discussion is provided by Flinn (2010).

[31] The presence of amenities ensures that the constraint $\underline{w}_j > b$ will always hold.

are lower than the current national minimum wage, and a large proportion of them are close to the value of unemployment benefits. This may indicate that minimum wage in France is too high.

Figure 7: Local Minimum Wages



Notes: (i) Local minimimum wages as given by equation 15; (ii) The upper horizontal line represents annual national minimum wage, while the lower stands for unemployment benefit.

To ascertain the quantitative effect of such a policy, we perform a fourth simulation where the minimum wage is defined as in equation 15. Our results show that total welfare increases by 3.1%. While the notion of local minimum wages suggests a competition between cities that may produce winners and losers, we report gains that are almost universal: average lifetime earnings increases in all locations but two. These losing cities correspond to the locations that end up with a local minimum wage higher than the previous national minimum wage. Finally, national unemployment rate decreases by 3.5%. These figures may represent a lower bound of the effect of local minimum wages since firms are not allowed to post lower wages in response.

Conclusion

In this paper, we propose a job search model to quantify the impact of mobility costs and spatial frictions on workers' migration. Our setting provides a rationale for a setting where workers can both be forward-looking profit-maximizers and remain stuck in unfavorable locations. From a computational standpoint, in contrast to the reference work by Kennan & Walker (2011), we show that the random search technology allows us to consider the full state space of a discrete choice model at the city level. The main take-away of our estimation results is that mobility costs are lower one order of magnitude when we take into account the frictional dimension of job-related migration. This result potentially has numerous public policy implications. In particular, policies based on relocation subsidies are likely to be inefficient.

Our results also shed new light on the determinants of the city size wage premium. Although individual wages are disconnected from productivity in our setup, the existence of search frictions allows us to reproduce both the upwards shift and the greater variability of the earning distributions, without resorting to the main mechanisms that have been put forward in the economic geography literature, human capital accumulation and production externalities. Markovian dynamics of on-the-job search between labor markets of unequal size are strong enough to generate such spatial pattern.

Notwithstanding, our model has several important limitations. First, it cannot be used to analyze the sorting of workers across cities, which has been shown to be a major driver of spatial wage differences (Combes et al., 2008). Second, in the spirit of Cahuc, Postel-Vinay & Robin (2006), one might want to incorporate the fact that cities vary in the number and size of firms and so that some locations provide workers with more opportunities for wage bargaining than others: in order to fully understand the contribution of location to lifetime inequalities, this dimension cannot be overlooked. More generally, we leave largely unexplored the firms' side of the dynamic location model. Whereas a mere extension à-la Meghir et al. (2015) would not convey much interest without an explicit theory of location choice, agglomeration economies and wages, we believe such explicit theory to be a promising venue for future research.

References

Bagger, J., Fontaine, F., Postel-Vinay, F., & Robin, J.-M. (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, *104*(6), 1551–96.

Baum-Snow, N. & Pavan, R. (2012). Understanding the city size wage gap. *Review of Economic Studies*, *79*(1), 88–127.

Bosquet, C. & Overman, H. G. (2016). Why Does Birthplace Matter So Much? Sorting, Learning and Geography. SERC Discussion Papers 0190, Spatial Economics Research Centre, LSE.

Bradley, J., Postel-Vinay, F., & Turon, H. (2017). Public sector wage policy and labor market equilibrium: A structural model. *Journal of the European Economic Association*, jvw026.

Bryan, G. & Morten, M. (2017). The aggregate productivity effects of internal migration: Evidence from indonesia. Working Paper 23540, National Bureau of Economic Research.

Burdett, K. & Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, *39*(2), 257–73.

Cahuc, P., Postel-Vinay, F., & Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, *74*(2), 323–364.

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

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

Dahl, G. B. (2002). Mobility and the return to education: Testing a roy model with multiple markets. *Econometrica*, *70*(6), pp. 2367–2420.

Flinn, C. & Heckman, J. (1982). New methods for analyzing structural models of labor force dynamics. *Journal of Econometrics*, *18*(1), 115–168.

Flinn, C. J. (2002). Labour market structure and inequality: A comparison of italy and the u.s. *Review of Economic Studies*, *69*(3), 611–45.

Flinn, C. J. (2010). *The Minimum Wage and Labor Market Outcomes*. MIT Press.

Gallin, J. H. (2004). Net migration and state labor market dynamics. *Journal of Labor Economics*, *22*(1), 1–22.

Gobillon, L. & Wolff, F.-C. (2011). Housing and location choices of retiring households: Evidence from france. *Urban Studies*, *48*(2), 331–347.

Hale, J. K. (1993). *Theory of functional differential equations*. Springer–Verlag, Berlin–Heidelberg–New York.

Head, K. & Mayer, T. (2014). Gravity equations: Workhorse,toolkit, and cookbook. In *Handbook of International Economics*, volume 4 chapter Chapter 3, (pp. 131–195). Elsevier.

Jolivet, G., Postel-Vinay, F., & Robin, J.-M. (2006). 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.

Kaplan, G. & Schulhofer-Wohl, S. (2017). Understanding the long-run decline in interstate migration. *International Economic Review*, *58*(1), 57–94.

Keane, M. P. & Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, *105*(3), 473–522.

Kennan, J. & Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, *79*(1), 211–251.

Levy, M. (2010). Scale-free human migration and the geography of social networks. *Physica A: Statistical Mechanics and its Applications*, *389*(21), 4913 – 4917.

Magnac, T. & Thesmar, D. (2002). Identifying dynamic discrete decision processes. *Econometrica*, *70*(2), 801–816.

Manning, A. & Petrongolo, B. (2017). How local are labor markets? evidence from a spatial job search model. *American Economic Review*, *107*(10), 2877–2907.

McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, *84*(1), 113–126.

Meghir, C., Narita, R., & Robin, J.-M. (2015). Wages and informality in developing countries. *American Economic Review*, *105*(4), 1509–1546.

Moretti, E. (2011). Local labor markets. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics*, volume 4, Part B chapter Chapter 14, (pp. 1237 – 1313). Elsevier.

Phelps, E. S. (1969). The new microeconomics in inflation and employment theory. *American Economic Review*, *59*(2), 147–60.

Postel-Vinay, F. & Robin, J.-M. (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, *70*(6), 2295–2350.

Postel-Vinay, F. & Turon, H. (2007). The public pay gap in britain: Small differences that (don't?) matter. *Economic Journal*, *117*(523), 1460–1503.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy*, *90*(6), 1257–78.

Schwartz, A. (1973). Interpreting the Effect of Distance on Migration. *Journal of Political Economy*, *81*(5), 1153–69.

Shephard, A. (2014). Equilibrium search and tax credit reform. Working Papers 1336, Princeton University, Department of Economics, Center for Economic Policy Studies.

Zenou, Y. (2009). Search in cities. *European Economic Review*, *53*(6), 607–624.

A Theory: proofs and discussions

A.1 Expressions

Reservation wages ϕ_j and indifference wages $q_{jl}(w)$ and $\chi_{jl}(w)$ verify:

$$V_j^u \equiv V_j^e(\phi_j) \tag{16}$$

$$V_j^e(w) \equiv V_l^e(\chi_{jl}(w)) \tag{17}$$

$$V_j^e(w) \equiv V_l^e(q_{jl}(w)) - c_{jl} \tag{18}$$

Equations 1 and 2 can be rewritten as:

$$rV_j^u = b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\overline{w}} \left(V_j^e(x) - V_j^u\right) dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \int_{q_{jk}(\phi_j)}^{\overline{w}} \left(V_k^e(x) - c_{jk} - V_j^u\right) dF_k(x) \tag{19}$$

$$\begin{aligned} rV_j^e(w) &= w + \gamma_j + \lambda_j^e \int_{w}^{\overline{w}} \left(V_j^e(x) - V_j^e(w)\right) dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} \left(V_k^e(x) - c_{jk} - V_j^e(w)\right) dF_k(x) \\ &+ \delta_j \left[V_j^u - V_j^e(w)\right] \end{aligned} \tag{20}$$

After integration by parts of equations 19 and 20, we get:

$$V_j^u = \frac{1}{r}\left[b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\overline{w}} \Xi_j(x)dx + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \left(\int_{q_{jk}(\phi_j)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(\phi_j)c_{jk}\right)\right] \tag{21}$$

$$V_j^e(w) = \frac{1}{r+\delta_j}\left[w + \gamma_j + \delta_j V_j^u + \lambda_j^e \int_{w}^{\overline{w}} \Xi_j(x)dx + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \left(\int_{q_{jk}(w)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(w)c_{jk}\right)\right] \tag{22}$$

where:

$$\Xi_j(x) \equiv \overline{F}_j(x) dV_j^e(x) = \frac{\overline{F}_j(x)}{r + \delta_j + \lambda_j^e \overline{F}_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(x))} \tag{23}$$

Finally, using Equations 16 and 18, we find that ϕ_j and $q_{jl}(w)$ are given by Equations 3 and 4.

A.2 Existence and uniqueness of indifference wages

From Equation 4, we derive the following proposition:

Proposition 4 *Let's denote by $\mathcal{W} = [\underline{w}, \overline{w}]$ the support of the wage distribution. $\mathcal{W}$ is a closed subset of a* Banach space. *The set of functions $q_{jl}(\cdot)$ defines a contraction. In addition, they have a unique fixed point.*

Proof We show that q_{jl} can be written in a differential form, and dq_{jl} can be bounded. These two implications allows us to show that the contraction mapping theorem applies. The derivative of $q_{jl}(\cdot)$ is given by:

$$dq_{jl}(w) = \frac{dV_l^e(q_{jl}(x))}{dV_j^e(x)} = \frac{r+\delta_l+\lambda_l^e\overline{F}_l(q_{jl}(w))+\sum_{k\in\mathscr{J}_l} s_{lk}^e\lambda_k^e\overline{F}_k(q_{lk}(q_{jl}(w)))}{r+\delta_j+\lambda_j^e\overline{F}_j(w)+\sum_{k\in\mathscr{J}_j} s_{jk}^e\lambda_k^e\overline{F}_k(q_{jk}(w))} \tag{24}$$

Consider the associated Ordinary Differential Equation (ODE) problem with $q_{jl}(w_0) = q_{jl}^0$. The integral equation version of equation 24 is given as:

$$q_{jl}(w) = q_{jl}^0 + \int_{\underline{w}}^{w} h_{jl}(x, q_j(x))dx, \tag{25}$$

where $q_j(x) \equiv \{q_{jk}(x)\}_{k\in\mathscr{J}_j}$, and $h_{jl}(x, q_j(x)) = dq_{jl}(x, q_j(x))$. Since $dV_j^e(\cdot) > 0$ and $dV_l^e(\cdot) > 0$, we have $dq_{jl}(\cdot) > 0$; moreover, given that all the structural matching parameters (s^i, λ^i, δ) are positive and the interest rate r is strictly positive, $dq_{jl}(\cdot)$ can be bounded. Therefore, it is easy to see that $dh_{jl}(\cdot) = d^2q_{jl}(\cdot)$ is also bounded. As a consequence, $dq_{jl}(x, q_j(x))$ is *Lipschitz continuous*. The *Banach fixed-point theorem* states that $q_{jl}(\cdot)$ has a unique solution. In addition, the Lipschitz continuity of $dq_{jl}(x, q_j(x))$ ensures that the solution that does not depend on the initial condition.

A.3 Uniqueness of wage offer distribution

A unique wage offer distribution is recovered through a steady-state condition on the observed wage distribution. Outflows from city j are given by all the jobs in city j with a wage lower than w that are either destroyed or left by workers who found a better match. If it is located in city j, such match will correspond to a wage higher than w. However, if it is located in any city $k \neq j$, this match will only need to correspond to a wage higher than $q_{jk}(x)$, where $x < w$ is the wage previously earned in city j. The measure of this flow, which stems from the fact that we consider several separate markets, requires an integration over the distribution of observed wages in city j. Inflows to city j are first composed of previously unemployed workers who find and accept a job in city j with a wage lower than w. These workers may come from city j or from any city $k \neq j$. However, they will only accept such a job if w is higher than their reservation wage ϕ_j or than the mobility-compatible indifference wage of their reservation wage $q_{kj}(\phi_k)$. The second element of inflows is made of workers who were previously employed in any city $k \neq j$ at a wage x lower than the maximum wage such that moving to city j would yield a utility of $V_j^e(w)$ (we denote this wage $q_{jk}^{-1}(w)$) and find a job at a wage between $q_{kj}(x)$ and w. Because of the existence of mobility costs, $q_{jk}^{-1}(w) \neq q_{kj}(w)$.

This is all summarized in Equation 26:

$$
\begin{aligned}
(m_j - u_j)\Big[G_j(w)\big(\delta_j + \lambda_j^e \overline{F}_j(w)\big) + \sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{w} \overline{F}_k(q_{jk}(x))dG_j(x)\Big] \equiv \\
\lambda_j^u\Big[\psi_{jj}(w)u_j\big(F_j(w) - F_j(\phi_j)\big) + \sum_{k\in\mathscr{J}_j} s_{kj}^u \psi_{kj}(w)u_k\big(F_j(w) - F_j(q_{kj}(\phi_k))\big)\Big] \\
+\lambda_j^e \sum_{k\in\mathscr{J}_j} s_{kj}^e (m_k - u_k)\int_{\underline{w}}^{q_{kj}^{-1}(w)} [F_j(w) - F_j(q_{kj}(x))]dG_k(x)
\end{aligned}
\tag{26}
$$

where $\psi_{kj}(w) = \mathbb{1}_{w>q_{kj}(\phi_k)}$ is a dummy variable indicating whether unemployed jobseekers in city k are willing to accept the job paid at wage w in city j. Similarly, the integral in the last term gives the measure of job offers in city j that are associated with a wage lower than w yet high enough to attract employed workers from any city $k \neq j$ and it is nil if $q_{kj}^{-1}(w) < \underline{w}$. These restrictions mean that very low values of w will not attract many jobseekers. We can differentiate Equation 26 with respect to w. This yields the following linear system of functional differential equations:

$$
f_j(w) = \frac{g_j(w)(m_j - u_j)\Big[\delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(w))\Big]}{\lambda_j^u\Big(\psi_{jj}(w)u_j + \sum_{k\in\mathscr{J}_j} s_{kj}^u \psi_{kj}(w)u_k\Big) + \lambda_j^e\Big((m_j - u_j)G_j(w) + \sum_{k\in\mathscr{J}_j} s_{kj}^e (m_k - u_k)G_k(q_{kj}^{-1}(w))\Big)}
\tag{27}
$$

In equilibrium, the instant measure of match creations associated with a job paid at wage w and located in city j equals its counterpart of match destructions.

The system of differential equations $\mathfrak{f} : \mathbb{R}^J \to (0,1)^J$ has a unique fixed point. Existence stems from a direct application of *Schauder fixed-point theorem*. Regarding uniqueness, first note that since each $f_j(\cdot)$ is a probability density function, it is absolutely continuous and its nonparametric kernel estimate is Lipschitz continuous; then, by contradiction, it is easy to show that two candidate solutions $h^0(\cdot)$ and $h^1(\cdot)$ cannot at the same time solve the differential equation, define a contraction, and be Lipschitzian (under the initial condition given by a minimum wage policy whereby $F_j(\underline{w}) = 0$). For more details, see Theorem 2.3 in Hale (1993).

B Algorithm and numerical solutions

B.1 Algorithm

Let $g(\cdot) \equiv \{g_j(\cdot)\}_{j\in\mathscr{J}}$. The set of theoretical moments $m(\theta)$ is simulated thanks to an iterative algorithm, which can be summarized as follows:

1. Given data on wage, evaluate $G(\cdot)$ and $g(\cdot)$
2. Set an initial guess for θ and $F(\cdot)$
3. Given θ and $F(\cdot)$, solve Equation 4 to recover indifference wages $q(\cdot)$
4. Solve Equation 10 to recover equilibrium population $\mathscr{M}$
5. Solve Equation 27 to update the distribution of job offers $F(\cdot)$
6. Solve Equation 7 to update the distribution of local amenities Γ
7. Update θ using the maximum of $\mathscr{L}(\theta)$.
8. Repeat steps 3 to 7 until convergence.

B.2 Indifference wages

The model raises several numerical challenges, in particular in steps 3 and 5. In step 3, $q(\cdot)$ defines a system of $J^2 - J$ equations, to be solved $\dim(\mathbb{W})$ times, where $\mathbb{W}$ is a grid of wages. Because the derivative of $q_{jl}(\cdot)$ can be easily recovered, we use a Newton-based method. Remember that for any $w^* > w$ and a relatively small $h = w^* - w$, Newton's formula yields:

$$q_{jl}(w^*) = q_{jl}(w) + h\,dq_{jl}(w) \tag{28}$$

Indifference wages can then be recovered using a sequential process, on a grid of wages w_i. We initialize $q_{jl}(\underline{w})$ using mobile transition rates across locations. That is, for a guess of spatial frictions parameters s^u_{jl}, use moments to compute $F_l(q_{jl}(\underline{w}))$, and the full sequence of $q_{jl}(\cdot)$ can be computed using the derivative given in equation 24.

However, while the sequential nature of the algorithm uncovers $q_{jl}(w)$, it does not allow us to find $q_{lk}(q_{jl}(w))$. As a consequence, we embed an inner loop step where we recover the full sequence of

$q_{jl}(\cdot)$ under the assumption that the derivative may be written using equation 29:

$$dq^m_{jl}(w) = \frac{r + \delta_l + \lambda^e_l \overline{F}_l(q_{jl}(w)) + \sum_{k \in \mathscr{J}_l} s^e_{lk} \lambda^e_k \overline{F}_k(q_{jk}(w))}{r + \delta_j + \lambda^e_j \overline{F}_j(w) + \sum_{k \in \mathscr{J}_j} s^e_{jk} \lambda^e_k \overline{F}_k(q_{jk}(w))} \tag{29}$$

and each level of indifference wages may be written using equation 30:

$$q^m_{jl}(w^*) = q_{jl}(w) + h dq^m_{jl}(w) \tag{30}$$

The full solution method is therefore as follows:

1. Evaluate $F_j(q_{lj}(\underline{w}))$ for all $(j,l) \in \mathscr{J} \times \mathscr{J}_j$.
2. Calculate $q_{lj}(\underline{w})$ for all $(j,l) \in \mathscr{J} \times \mathscr{J}_j$.
3. Recover $q^m_{jl}(\cdot)$ using the modified derivative $dq^m_{jl}(\cdot)$ under Newton Formulas.
4. Recover $q_{jl}(\cdot)$ using $q^m_{jl}(\cdot)$ to evaluate $q_{lk}(q_{jl}(w))$.
5. Repeat step 4 until $q_{jl}(\cdot)$ converges.

B.3 Wage distributions

Once the indifference wages are recovered, we can turn to the evaluation of the wage distributions (step 5 in the general algorithm). There are two difficulties when solving for the system defined by Equation 27. First, for any system of three cities or more, the system can only be solved numerically.[32] Second, the system is composed of functional equations, which standard differential solvers are not designed to handle. Our solution is twofold. First, as explained in Section 3.2, we assume that $F(\cdot)$ can be proxied by a beta distribution. Then, since our empirical counterparts are based on real wages, we treat the empirical cdf $G(\cdot)$ as unknown and we estimate the set of parameters $\alpha \equiv \{\alpha_j\}_{j\in\mathscr{J}}$ and $\beta \equiv \{\beta_j\}_{j\in\mathscr{J}}$ which minimize the distance between the empirical cdf $G(\cdot)$ and its theoretical counterpart. This theoretical counterpart is given as the solution to the following functional equation, derived from Equation 27:

$$g_j(w) = f_j(w) \times \frac{\lambda^u_j\Big(\psi_{jj}(w)u_j + \sum_{k\in\mathscr{J}_j} s^u_{kj}\psi_{kj}(w)u_k\Big) + \lambda^e_j\Big((m_j - u_j)G_j(w) + \sum_{k\in\mathscr{J}_j} s^e_{kj}(m_k - u_k)G_k(q^{-1}_{kj}(w))\Big)}{(m_j - u_j)\Big[\delta_j + \lambda^e_j\overline{F}_j(w) + \sum_{k\in\mathscr{J}_j} s^e_{jk}\lambda^e_k\overline{F}_k(q_{jk}(w))\Big]} \tag{31}$$

[32] Two-sector models, such as the one presented in Meghir et al. (2015), yield systems of two ordinary differential equations. These systems can be rewritten in a way such that they still admit a closed-form solution.

The original algorithm is modified to take into account the estimation of α and β. At step 2, we set an initial guess (α^0, β^0). At step 5, we need a solution $G(\cdot)$ to Equation 31 in order to update (α, β). We develop a simple iterative process based on Euler's approach. That is, given an initial $G_j(\underline{w})$, and a guess $G^0 \sim \beta(\alpha^0, \beta^0)$, update $G(\cdot)$ using equation 26. The full solution is therefore:

1. Set the step size h and use Euler's method to approximate the sequence of $G_j(\cdot)$.
2. Derive estimate for $G_l(q_{jl}(w))$ for all $j \in \mathscr{J}$.
3. Use estimates of $G_l(q_{jl}(w))$ to solve the functional differential equation 31.
4. Repeat steps 5.3 to 5.4 until convergence.

Once a solution for $G_j(\cdot)$ is recovered, we update α and β by minimizing the distance between $G(\cdot)$ and $\hat{G}(\cdot)$ over the space of beta distributions.

B.4 Local amenities

Step 6 is completed using an embedded ranking algorithm which proceeds as follows:

1. Set the initial guess $\forall j \in \mathscr{J} : \gamma_j^0 = 0$
2. Order the corresponding values V_j^{u0} and let $j^0 = \text{argmin}_{j \in \mathscr{J}} V_j^{u0}$
3. Set $\forall k \in \mathscr{J}_{j^0} : \gamma_k^1 = V_k^{u0} - (V_{j^0}^{u0} + c_{j^0 k})$ and update $V_k^{u1} = V_k^{u0} + \gamma_k^1$
4. Repeat steps 2 and 3 until convergence.

C Data

C.1 Data selection

The initial sample is composed of 12,379,415 observations of 3,265,759 workers when we restrict the dataset to observations related to the main jobs of individuals who are always located in continental France over the 2002-2007 period. We then drop all female workers; all workers who were less than 15 years old or more than 58 years old at some point, to avoid confusion about participation decision; all workers who at some point were working in the public sector, as apprentice, from home or part time. We drop all individuals who at some point had a reported wage below 900 euros per month or above 8,000 euros: the first case is considered as measurement error (some income is not reported, or workers should be registered as part time workers combining different jobs, since over 75% of these "full-time" employment spells are associated with less than the legal 35-hour workweek), whereas the second case extends the support of wage distributions too dramatically for less than 1% of all workers.[33] Individuals who lived at some point in a non-metropolitan area are also dismissed. Finally, for computational reasons, we drop all individuals observed only once. The selection process is summarized in Table 9 and we end up with the dataset described in Table 10.

Table 9: Selection process

Criteria	Individuals	Observations
Woman	1,498,656	5,545,284
Observed only once	1,142,409	1,142,409
Apprenticeship	108,076	473,772
Home worker	15,954	73,504
Part time	1,353,081	5,639,525
Public sector	1,194,902	4,703,876
Non-metropolitan area	496,703	1,592,928
Full time below minimum wage	372,259	1,620,337
Above 8,000 euros per month	28,890	150,823
Age lower than 15 or higher than 58	214,676	687,101

Notes: (i) A non-metropolitan area is outside one of the 770 2010 metropolitan areas; (ii) Source: Panel DADS 2002-2007

[33] Part time status is part of automated declaration by firms, which is based either on legal work duration or sectoral collective bargaining agreement. When not filled, the variable is imputed using daily hours of work.

Table 10: Structure of the dataset

Year	Number of indiv.	Number of obs.	Number of obs. by metro Min	Mean	Median	Max	Number of indiv. by metro Min	Mean	Median	Max
2002	324,091	352,751	343	28,214	5,430	81,648	365	30,822	5,863	89,283
2003	312,043	329,977	315	27,407	5,208	79,046	327	29,112	5,455	84,040
2004	322,039	339,709	330	28,210	5,400	81,416	341	29,876	5,623	86,301
2005	323,732	343,635	341	28,585	5,618	81,998	349	30,629	5,919	88,021
2006	329,814	352,091	333	28,889	5,770	83,264	344	31,007	6,139	89,469
2007	326,554	355,542	328	28,645	5,750	82,523	340	28,645	6,180	90,423
Total	384,114	2,073,705	2,066	182,788	35,496	527,537	451	36,430	7,492	104,810

Notes: (i) Metros are here the clusters of municipalities forming the 99 largest metropolitan areas in 2010; (ii) Source: Panel DADS 2002-2007

C.2 Descriptive statistics

Table 11: Local wage distributions

Panel 1: the nine largest cities

	City 1 Paris	City 2 Lyon	City 3 Mars.	City 4 Toul.	City 5 Lille	City 6 Bord.	City 7 Nice	City 8 Nantes	City 9 Stras.
$\mathbb{E}$	36,855	32,230	30,884	31,296	29,534	30,060	31,410	30,277	31,097
σ	24,511	21,898	21,299	21,457	20,936	20,699	22,260	20,368	20,460
Q_1	22,558	21,083	19,955	20,154	19,279	19,934	19,997	20,526	21,142
Q_2	31,770	27,357	26,239	26,651	24,634	25,445	26,127	25,838	27,153
Q_3	46,746	38,442	37,066	37,875	34,724	35,238	37,69	35,282	36,591

Panel 2: nine other cities

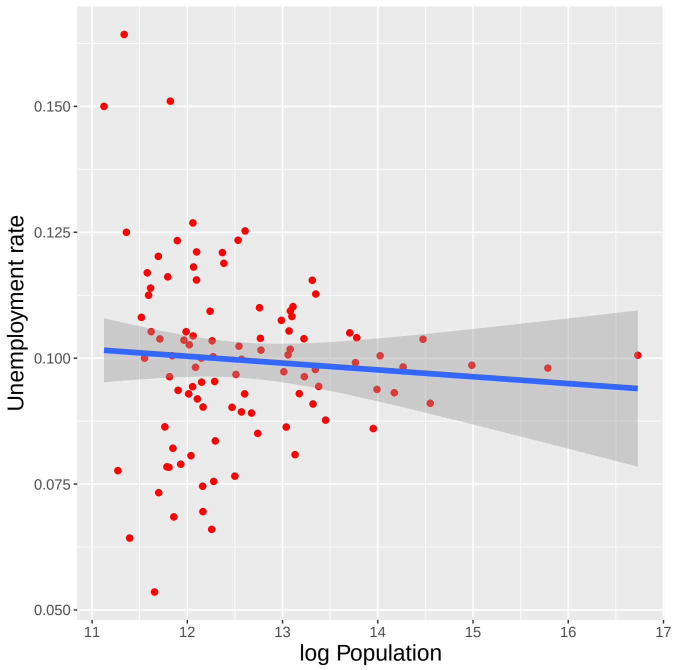
	City 10 Gren.	City 20 Nancy	City 30 Brest	City 40 Nîmes	City 50 Valence	City 60 Châlon	City 70 Creil	City 80 Charl.	City 90 Albi
$\mathbb{E}$	32,479	29,510	28,549	26,163	28,172	29,119	30,105	26,714	27,454
σ	20,980	19,843	20,073	18,283	19,672	19,104	20,712	16,991	19,710
Q_1	21,694	20,072	19,318	18,090	19,336	20,401	19,956	19,887	18,661
Q_2	28,660	25,290	23,973	22,506	23,869	25,115	25,687	23,886	23,015
Q_3	39,026	34,526	32,571	29,463	32,220	33,335	35,779	30,206	31,089

Panel 3: Stability of the wage distributions

	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}
$\tilde{Q}_1^{07/02}$	1.005	1.006	1.007	1.008	1.008	1.009	1.009	1.010	1.011
$\tilde{Q}_2^{07/02}$	1.005	1.006	1.007	1.007	1.008	1.009	1.009	1.010	1.011
$\tilde{\mathbb{E}}^{07/02}$	1.005	1.007	1.007	1.007	1.008	1.009	1.009	1.010	1.011
$\tilde{Q}_3^{07/02}$	1.003	1.006	1.006	1.007	1.008	1.009	1.009	1.011	1.013

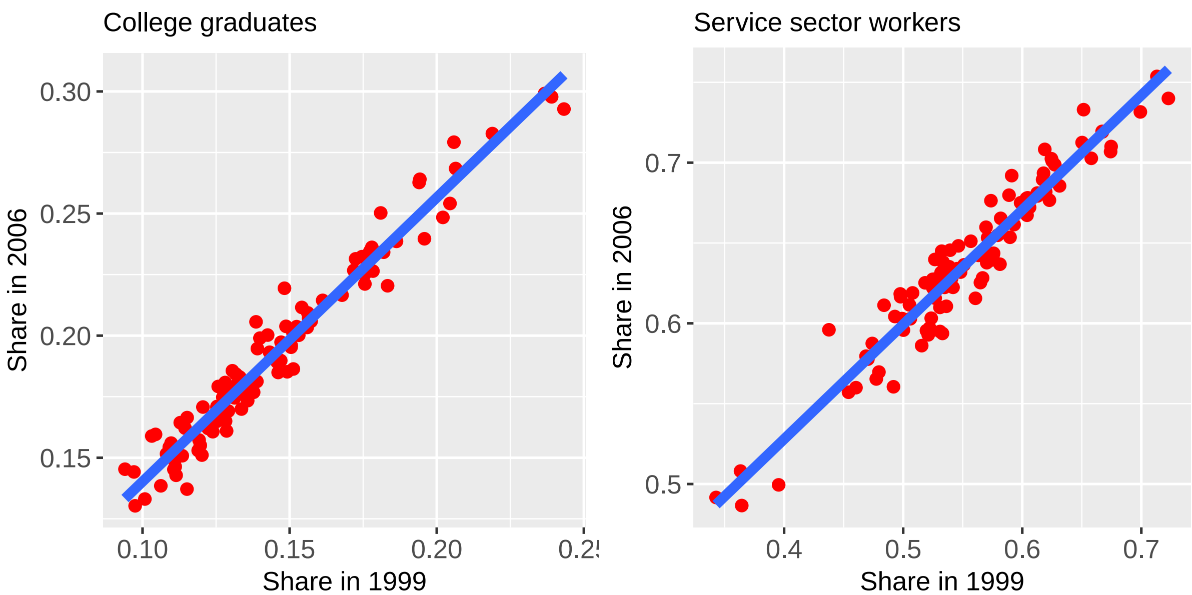
Notes: i) $\mathbb{E}$ and σ are respectively average and standard deviation of wages (in €2002), while Q_1, Q_2, and Q_3 are respectively the 1st quartile, the median and the third quartile. ii) Panels 1 and 2: Distributions are evaluated over the six-year span 2002-2007; (iii) Mars. is Marseilles, Toul. is Toulouse, Bord. is Bordeaux, Stras. is Strasbourg, Gren. is Grenoble and Charlev. is Charleville; (iv) Panel 3: $\tilde{Q}_1^{07/02}$, $\tilde{Q}_2^{07/02}$, $\tilde{\mathbb{E}}^{07/02}$ and $\tilde{Q}_3^{07/02}$ are the ratios of the moments defined in Panels 1 and 2 for the city-specific log-wage distributions in 2007 and in 2002 and P_x is the x th percentile of the distribution of these ratios for the first 100 cities *Source: Panel DADS 2002-2007*; for details on the sample, see Section 1.2 and Appendix C.1.

Figure 8: Local unemployment by city size



Notes: Unemployment rate is constructed using the transition data from the DADS Panel. It corresponds to the empirical moment used in the estimation and described in Appendix C.3. (ii) The estimate of the slope of the regression line is equal to −0.001 with a standard error of 0.002. *Source: Panel DADS.*

Figure 9: Heterogeneity and stability in skill and sectoral composition



Notes: (i) Shares are computed on the 25-54 age bracket for the population of men (left) and the population of men workers (right) and for the 100 largest metropolitan areas in continental France, keeping a constant municipal composition based on the 2010 "Aires Urbaines" definition; (ii) The respective equations of the least squares line are $\hat{C}_{06} = 1.16 \times C_{99} + 0.02$ (left) and $\hat{S}_{06} = 0.71 \times S_{99} + 0.24$ (right). *Source: INSEE, Census 1999 and 2006*

Table 12: Size, distance and migration flows: examples

Panel 1: Paris and the largest cities

		Destination					
Origin		Paris	Lyon	Marseille	Toulouse	Lille	Rest of France
Paris	UE	90.704	0.693	0.519	0.478	0.349	7.257
	EE	92.096	0.880	0.554	0.416	0.411	5.643
Lyon	UE	4.384	81.804	0.792	0.285	0.238	12.497
	EE	6.930	80.890	1.148	0.349	0.492	10.191
Marseille	UE	4.299	1.283	82.112	0.589	0.150	11.567
	EE	7.548	2.157	75.200	0.522	0.417	14.157
Toulouse	UE	4.555	0.581	0.533	82.765	0.242	11.323
	EE	5.162	0.667	0.632	83.778	0.140	9.621
Lille	UE	4.708	0.506	0.287	0.246	78.278	15.973
	EE	5.543	0.720	0.251	0.376	77.231	15.878
Rest of	UE	0.041	0.012	0.005	0.007	0.008	-
France	EE	0.033	0.016	0.008	0.010	0.012	-

Panel 2: Paris and the Lyon region

		Destination					
Origin		Lyon	Grenoble	St-Etienne	Valence	Bourg	Paris
Lyon	UE	81.804	1.523	1.146	0.277	0.584	4.384
	EE	80.890	1.517	0.964	0.349	0.328	6.930
Grenoble	UE	4.685	81.664	0.269	0.458	0.000	3.312
	EE	11.905	72.247	0.074	1.637	0.000	4.092
St-Etienne	UE	6.434	0.402	81.144	0.089	0.000	2.904
	EE	8.313	0.372	82.382	0.372	0.000	1.365
Valence	UE	2.860	1.049	0.286	73.117	0.000	3.337
	EE	4.290	6.271	0.660	63.366	0.330	3.300
Bourg	UE	9.091	0.455	0.227	0.455	73.182	0.682
	EE	13.333	0.000	0.000	0.833	62.500	1.667

Notes: (i) UE stands for transition out of unemployment and EE stands for job-to-job transition; (ii) Reading panel 1: among the transitions out of unemployment originating from the city of Lyon, 81.8% led to a job in Lyon, 4.4% led to a job in Paris and 0.8% led to a job in Marseille; reading panel 2: among the transitions out of unemployment that started in the city of Valence, 84.1% led to a job in Valence, 1.6% led to a job in Lyon and 2.2% led to a job in Paris. *Source: Panel DADS 2002-2007*

C.3 Empirical moments used in the first column in Table 3

Unemployment rate in city j: ratio of the number of individuals who should be in the panel in city j on January 1st 2002 but are unobserved (henceforth, assumed unemployed) to the sum of this number and the number of individuals observed in city j on January 1st 2002

Population in city j: number of individuals observed in the panel between 2002 and 2007 in city j

Transition rate ee within city j: ratio of the number of job-to-job transitions within city j observed over the period, to the potentially-employed population in city j (population as defined above multiplied by one minus the unemployment rate as defined above)

Earning distribution in city j: quantiles in city j on a grid of 17 wages over the period

Transition rate ue (resp., ee) out of city j: ratio of the number of transitions out of unemployment (resp., the number of job-to-job transitions) out of city j observed over the period, to the potentially-unemployed (resp., potentially-employed) population in city j (population as defined above multiplied by the unemployment rate as defined above)

Transition rate ue (resp., ee) into city l: ratio of the number of transitions out of unemployment (resp., the number of job-to-job transitions) into city l observed over the period, to the potentially-unemployed (resp., potentially-employed) population in all cities $k \neq l$

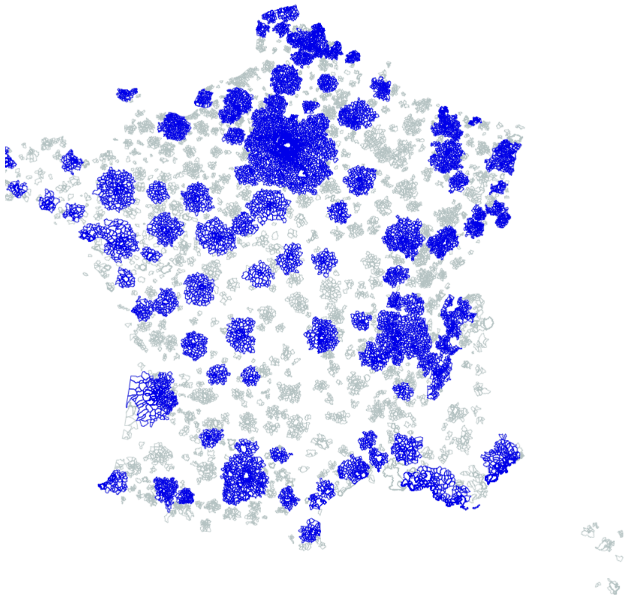
Transition rate ue (resp., ee) from city j **to city** l: ratio of the number of transitions out of unemployment (resp., the number of job-to-job transitions) from city j to city l observed over the period, to the potentially-unemployed (resp., potentially-employed) population in city j

Accepted wages ee from city j **to city** l: average wage following a job-to-job transition from city j to city l observed over the period.

D Figures

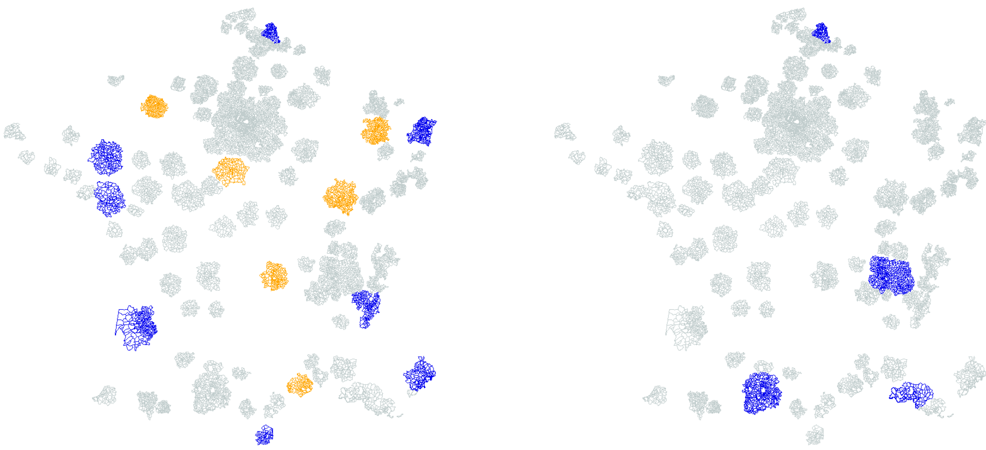
D.1 Maps

Figure 10: The French urban archipelago



Notes: the spatial unit is the municipality. There are more than 700 metropolitan areas according to the 2010 definition. In dark, the border of the municipalities that constitute the largest 100 metropolitan areas. In light, the border of all the other municipalities within a metropolitan area. *Source: INSEE, Census 2007*

Figure 11: The metropolitan areas in subset $\mathscr{T}_1$ (left) and subset $\mathscr{T}_2$ (right)



Notes: (i) In light: the 100 largest metropolitan areas; (ii) Subset $\mathscr{T}_1$ is used to identify the effect of physical distance and dissimilarity on spatial frictions based on pair-specific out-of-unemployment and job-to-job transition rates, with origin cities in blue and destination cities in orange; subset $\mathscr{T}_2$ is used to identify the effect of physical distance on moving costs based on pair-specific average accepted wages after a job-to-job transition with mobility.

D.2 Fit of the model

Figure 12: Fit: local transition rates, unemployment rate, city size, and average wages

Notes: (i) Comparison of the empirical moments and the theoretical moments, as predicted by the model and evaluated (see Table 3 and appendix C.3 for details); (ii) Log population is equal to $\log(6.5 \times 10^7 m_j)$; (iii) Job arrival rate for unemployed is $\lambda_j^u \overline{F}_j(\phi_j)/5$, Job arrival rate for employed is $\lambda_j^e \int \overline{F}_j(x) dG_j(x)/5$, job destruction rate is $\delta_j/5$ and average wage is $\int x g_j(x) dx$.

Figure 13: Fit: migration rates

Notes: (i) Comparison of the empirical moments and the theoretical moments, as predicted by the model and evaluated (see Table 3 and appendix C.3 for details); (ii) Out-migration for unemployed is $\sum_{k\in\mathcal{J}_j} s^u_{jk}\lambda^u_k\overline{F}_k(q_{jk}(\phi_j))$, out-migration for employed is $\sum_{k\in\mathcal{J}_j} s^e_{jk}\lambda^e_k\int\overline{F}_k(q_{jk}(x))dG_j(x)$, in-migration for unemployed is $\sum s^u_{kj}\lambda^u_j\overline{F}_j(q_{kj}(\phi_k))$ and in-migration for employed is $\sum s^e_{kj}\lambda^e_j\int\overline{F}_j(q_{kj}(x))dG_k(x)$.

Figure 14: Fit: wage growth

Notes: (i) Comparison of the empirical moments and the theoretical moments, as predicted by the model and evaluated (see Table 3 and appendix C.3 for details); (ii) Wage growth Δ_{jl}^{ee} is the growth of average wage along a job-to-job transition; the data corresponds to what is used in Table 2; as for the theoretical moment, it is given by equation 32.

$$\Delta_{jl}^{ee} = \frac{\int\int_{q_{jl}(w)}^{\overline{w}} x dF_l(x) dG_j(w) - \int\int_{\underline{w}}^{q_{jl}^{-1}(w)} x dG_j(x) dF_l(w)}{\int\int_{\underline{w}}^{q_{jl}^{-1}(w)} x dG_j(x) dF_l(w)} \tag{32}$$