Job Search and Migration in a System of Cities*

Benoît Schmutz† Modibo Sidibé ‡

November 11, 2013

Abstract

We build an equilibrium job search model, where agents engage in both off-the-job and on-the-job search over a set of cities, to quantify the impact of matching and spatial frictions on the job search process. The mobility problem of agents introduces the concept of "mobility-compatible indifference wages", based on a dynamic utility trade-off between locations. As a consequence, the model can cope with various wage dynamics, including voluntary wage cuts, as part of forward-looking spatial strategies. We estimate the model by simulated method of moments on the 200 largest French cities. Our results allow us to characterize each local labor market by a set of location-specific matching parameters, a job offer distribution and to quantify the level of spatial frictions between each pair of markets. Our most notable findings include a robust positive correlation between on-the-job arrival rates and wage dispersion at the city level, which provides novel empirical evidence supporting the wage posting framework. Inference on the determinants of the matching parameters shows that most of the variance of job arrival rates can be explained by a few basic city characteristics. Finally, we run a counterfactual experiment to assess the number of cities that minimizes aggregate unemployment, keeping city's location and relative size fixed. Our findings suggest that the reshuffling of the urban population into the first 28 cities would reduce the French unemployment rate by a factor of three.

Keywords: local labor market; frictions; on-the-job search; migration

JEL Classification: J2; J3; J6

---

*We thank Jim Albrecht, Pat Bayer, Bruno Decreuse, Thomas Le Barbanchon and John Rust, as well as participants to the Second Joint Paris Search and Matching Workshop, for helpful comments and discussions. This project was initiated when both authors were junior members at CREST and pursued when Schmutz was visiting Georgetown University and Sidibé was a post-doctoral fellow at Duke University. The generosity of these three institutions is gratefully acknowledged. Data was made available through the Centre d'Accès Sécurisé à Distance.

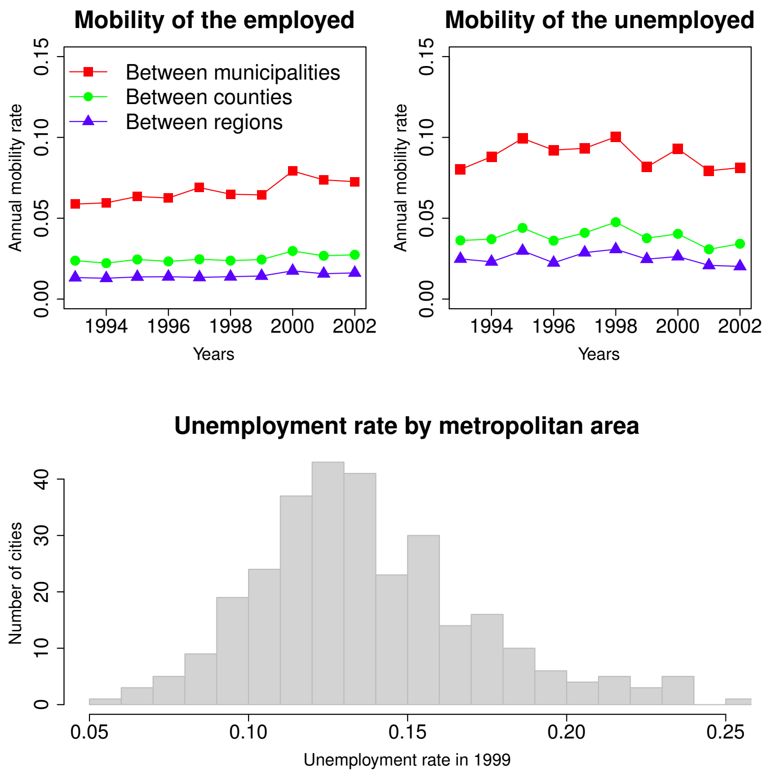
†Howard University; benoit.schmutz@howard.edu

‡Duke University; modibo.sidibe@duke.edu

Introduction

Local labor markets in developed countries are often characterized by striking and persistent disparities in key economic outcomes.[1] As shown by Figure 1 for France, this situation is compatible with steady labor flows across space. Such observation challenges the traditional explanations offered by competitive migration models and spaceless job search models, as both imply a theory of steady-state with regional convergence.[2]

Figure 1: Geographical mobility and local unemployment in France



Notes: (i) Mobility rates: probability to have changed location in the past year, conditional on previous employment status; (ii) Counties stand for the French "départements"; there are 22 regions, 96 départements and over 36,000 municipalities in France; each of these three levels form a partition of the French territory, unlike metropolitan areas, which are a non-nested combination of these three levels; (iii) Each year, $N \approx 23,000$ for the employed population and $N \approx 2,500$ for the unemployed population; (iv) Unemployment rates are computed for the 300 largest metropolitan areas in continental France defined in the 1999 Census; (v) Source: Census 1999 and Labor Force Surveys 1993-2002.

[1] See Moretti (2012) for an overview of the US; for instance, the unemployment rate at the MSA level ranges from 2.1% in Bismarck (ND) to 31.8% in Yuma (AZ)(BLS, 2013).

[2] See Elhorst (2003) for an overview of the (mostly empirical) literature on the determinants of regional differences in unemployment rates. Following Harris & Todaro (1970), competitive models have first sought to explain rural/urban migration patterns in developing economies (see Lucas (1993) for an overview). While both mean wages and unemployment risk are taken into account in the Harris-Todaro framework, it is assumed that unemployment is confined to one area (cities). The equilibrium is reached when the expected urban wage, adjusted for unemployment risk, is equal to the marginal product of an agricultural worker. By definition, these models define equilibrium as a situation where migration stops. While dynamic search models can, in theory, yield a definition of equilibrium that does not preclude migration, they fail to account for equilibrium regional heterogeneity. For example, while Mortensen & Pissarides (1999) and Jolivet, Postel-Vinay & Robin (2006) show that the search framework can explain a substantial part of the unemployment differential between respectively Europe and the US and among European countries, their model would still generate similar unemployment rate if they were to introduce individual mobility. That is, mobility would lead to a pattern of convergence in unemployment rates as in Phelps (1969)'s framework.

The goal of this paper is to fill this gap. We develop a job search model featuring both matching and spatial frictions to study the impact of these imperfections on individual mobility decisions and the resulting distribution of economic opportunities. The labor market is modeled as a system of interconnected local labor markets, or "cities". Each city is characterized by its own structural job arrival rates, layoff probability as well as a specific wage offer distribution. We consider the optimal strategy of ex-ante identical agents, who can engage in both off-the-job and on-the-job search within and between cities. After receiving a job offer from a different city, individuals are allowed to move. In addition, we introduce a difference in the efficiency of job search between locations and within locations in order to capture potential spatial frictions. Equilibrium conditions on market size, unemployment level and wage distributions allow us to recover the underlying structure of this economy. Our estimation results consist of a set of location-specific parameters and can be used to conduct an empirical investigation of the determinants of the matching functions.

In recent years, structural estimations of equilibrium job search models have proven very useful to study various features of the labor market.[3] However, job search models rest upon a rather unified conception of the labor market, where segmentation, if any, is based on sectors or qualifications. In particular, they do not account for spatial heterogeneity, even though several well-documented empirical facts suggest that the labor market may be described as an equilibrium only at a local level.[4] From a practical viewpoint, the absence of space in search models can be explained by computational difficulties. Indeed, solving for search models with multiple labor markets requires to handle high-dimensional objects such as wage distributions over numerous locations. One way to overcome this issue is to consider a very stylized definition of space. This is the path taken by Baum-Snow & Pavan (2012), who consider a frictionless model which includes several appealing characteristics such as individual ability and location-specific human capital accumulation, but have to resort to a ternary partition of space between small, mid-sized and large cities. In addition, the model does not handle the selection into mobility which is crucial in both job mobility and

[3]The original job search literature emerges as an attempt to capture the existence of frictional unemployment. Interestingly, Phelps (1969)'s island parable is, at least metaphorically, related to this paper. The major breakthrough, due to Burdett & Mortensen (1998), allows to generate ex-post wages differential from ex-ante identical agents, and provides an intuitive way to evaluate the individual unemployment probability as well as the wage offer distribution without solving the value functions. Bontemps, Robin & Van den Berg (1999) take into estimation this model to study an economy segmented by sector where firms are heterogenous in productivity. Postel-Vinay & Robin (2002) develop a model of bargaining that allows to decompose the wage dispersion. Dey & Flinn (2005) develop a framework for the provision of health insurance. Cahuc, Postel-Vinay & Robin (2006) and Flinn (2006) estimate search models with an explicit Nash-bargaining between workers and firms. Lentz (2010) uses endogenous search intensity to generate sorting at the firm level. Current topics include non-stationary models (Menzio & Shi, 2010; Robin, 2011; Moscarini & Postel-Vinay, 2013).

[4]As shown by Burda & Profit (1996) on Czech Republic or by Patacchini & Zenou (2007) and Manning & Petrongolo (2011) on the UK, matching functions exhibit a high level of spatial instability. In addition, the interregional mobility of labor and labor market outcomes are clear determinants of each other (see, e.g., Topel (1986), Pissarides & Wadsworth (1989) and Blanchard & Katz (1992)). Postel-Vinay & Robin (2002), who restrict their estimation sample to the Paris region, implicitly recognize this problem.

migration literature.[5] Our paper follows on from the path-breaking work of Kennan & Walker (2011), who develop and estimate a partial equilibrium model of mobility over the 50 US states and provide many interesting insights with respect to the mobility decision of agents, including mobility costs. However, their model relies on several simplifications. For example, it is assumed that individuals have knowledge over a limited number of wage distributions, which correspond to where they have lived before. This simplification does not reflect the recent increase in workers' ability to learn about other locations before a mobility (Kaplan & Schulhofer-Wohl, 2012). Moreover, the low mobility rate is rationalized by the existence of extremely high mobility costs, whereas the existence of spatial frictions provides a credible alternative explanation. Finally, a focus on the state level is not fully consistent with the theory of local labor markets, according to which they are better proxied by metropolitan areas (Moretti, 2011). In this paper, we consider the most precise partition of space for which enough data is available, between the 200 largest French metropolitan areas.[6]

The key innovation of our model is to allow for location-specific job arrival and destruction rates. As a consequence, we are able to provide a definition of the "mobility-compatible indifference wages", based on a dynamic utility trade-off between locations. These functions of wage, which are specific to each ordered pair of cities, are defined by the agent's indifference condition between her current state (a given wage in a given city) and another possible job in a different city. They admit a closed form which can be decomposed into three terms: the current wage, the difference in the value of on-the-job search between the two locations, and a constant term representing the difference in the value of unemployment between the locations, weighted by local unemployment risk. As a consequence, our model is able to cope with a wide variety of wage profiles over the life-cycle, including voluntary wage cuts as in Postel-Vinay & Robin (2002). These indifference wages are strictly increasing in wages and can be used, in combination with the observed earning distributions, to recover the underlying wage offer distributions through a system of first-order non-homogenous functional differential equations.

We estimate this model using the panel version of the French matched employer-employee database *Déclaration Annuelles de Données Sociales* (DADS) from 2002 to 2007. Our results provide a dataset of local labor market primitives: lay-off rates, job arrival rates for the unemployed, job arrival rates for the employed, as well as parameters measuring the efficiency of out-of-zone job search, for each of these two groups of

[5] Although the empirical probabilities of transition are weighted by the conditional probability $f(w^*|w^* > w)$, the value functions, which determine the optimal reservation value, do not include the wage distribution. This feature would make their problem similar to the one we address in this paper.

[6] The remaining metropolitan areas are pooled together with non-metropolitan areas into a single residual location. According to the US Census 2012, in order to achieve this level of precision on US data, one would have to consider a partition of space between 800 Metropolitan or Micropolitan Statistical Areas.

jobseekers, and for each pair of locations. Among other findings, we show that most of the wage variation is explained by on-the-job search in large cities whereas it is mostly explained by off-the-job search in smaller cities. We also use this set of parameter estimates as outcome variables to assess the determinants of the structural features of a labor market, using census and other administrative variables as covariates in a least-squares approach. A parsimonious linear combination of seven variables (number of firms, population density, share of people below thirty, share of males, share of people without qualifications, share of blue-collar jobs and share of manufacturing jobs) accounts for 90% of the variation in the job arrival rate for unemployed agents, against 31% for on-the-job arrival rates and 18% for job separation rates. Finally, we run a counterfactual experiment where our results are used to find the number of cities that minimizes aggregate unemployment, keeping city location and city relative size fixed. The two competing forces are that larger cities make up for more dynamic markets but the distance between them generates large spatial frictions. We find that the unemployment rate is minimal when the urban population is reshuffled into the first 28 cities.

Additional literature

The career choice of agents has long been investigated by economists. Works by Keane & Wolpin (1997) and Neal (1999) 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 choice. To understand this kind of mechanisms, Dahl (2002) proposes a model of mobility and earnings over the 50 US states and shows that higher educated individuals self-select into states with higher returns to education; however, the problem is not dynamic since migration is only one-shot in his model. Gallin (2004) shows that since migration is an investment, it cannot be only explained by cross-sectional variation in economic conditions, but also requires a comparison between expected future economic conditions. Finally, despite its interest and obvious links to the present paper, the classic perfect-competition approach faces a low-mobility puzzle which questions its ability to adequately model processes of job search in space.

Our model sheds new light on the determinants of the city size wage premium. In the frictionless economic geography literature, most of the debate has focused on the possible determinants of the gradient of wage growth across cities or between cities and rural areas (Gould, 2007). Using French data, Combes, Duranton, Gobillon, Puga & Roux (2012) show that agglomeration effects make firms more productive in large cities and tend to both dilate and shift upwards wage distributions. Using Spanish data, De la Roca & Puga (2012) reach similar conclusions on the worker's side. The authors argue that the city size wage premium does not reflect initial sorting by ability, but is rather the result of a more efficient learning process in larger cities. Although wages are disconnected from productivity in our setup, the existence of search frictions allows us to reproduce both the upwards shift and the dilation of the earning distributions, without resorting, neither to human capital accumulation, nor to production externalities. The optimal strategy of an agent consists of accepting any wage higher than her reservation wage, and working her way up to the top of the wage distribution by on-the-job search. As posited by Burdett & Mortensen (1998), these simple Markovian dynamics within labor markets of unequal size are strong enough to generate such spatial pattern.

In frictional markets, the impact of spatial constraints on labor market outcomes has already been studied extensively (Zenou, 2009b). However, the bulk of this literature focuses on intra-urban issues, namely spatial mismatch. Moreover, it is mainly theoretical or at most based on calibrations.[7] We intend to complement the existing literature regarding the following two aspects. First, we intend to study regional differences in economic opportunities, and especially, to disentangle the impact of location-specific matching frictions from the efficiency of the job search process between cities. Second, we are able to come up with estimates of the underlying structural characteristics of each local labor market and therefore, to have a detailed discussion about spatial constraints on the job search process. This ability to estimate our model comes at a cost: we do not explicitly model the housing market.[8] Throughout the paper, spatial frictions remain difficult to interpret in the absence of a separate housing market. Modelling the housing market would constitute an interesting extension, but it would require additional information on agents than the contents of French administrative data. In an extension, we consider two additional dimensions of spatial heterogeneity: location-specific amenities, to be understood as population-average valuations of each city net of local costs of living, and moving costs. Both dimensions will help capture some of the features related to the housing market.

Finally, closer to this paper in terms of method, there is the recent effort in the empirical job search literature to look at search patterns in competing submarkets. These papers seek to provide new dynamic

[7] See, among others, Coulson, Laing & Wang (2001), Brueckner & Zenou (2003) or Wasmer & Zenou (2006). More recently, Rupert & Wasmer (2012) have incorporated endogenous mobility decisions into a job search model with an explicit housing market. However, the location process they consider still takes place within a city and the job-finding rate is not location specific: therefore, the impact of location on job opportunities takes place through commuting costs only.

[8] Head & Lloyd-Ellis (2012) construct and calibrate an equilibrium job search model with heterogeneous locations, endogenous construction, and search frictions in the markets for both labor and housing. However, while their framework is very rich in many respects, the authors resort to a binary partition of space, between high-wage and low-wage US cities, which would not fit our purpose as well. Moreover, and despite this simple partition, the identification and estimation of their model would be very challenging. Similarly, Karahan & Rhee (2013) construct a rich directed search model with collateral constraints to look at the impact of the recent housing bust in the US on the spatial reallocation of labor; however, when it comes to calibrating their model, they also consider a binary partition of space where cities are classified according to the extent of the decline in house prices they have experienced.

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 available only in some submarkets. As a consequence, the papers study potential tradeoffs between dual submarket: 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. In doing so, these models also provide more accurate estimates of the matching parameters, which are no longer averaged over sectors. Postel-Vinay & Turon (2007) study the public/private pay gap in Britain and detect a positive wage premium in favor of the public sector both in instantaneous and in dynamic terms. However, the paper is only empirical, in that its focus is on the modelling of individual heterogeneity rather than on the explanation of economic mechanisms. This is not the case of Shephard (2011), who distinguishes between part-time and full-time work to assess the impact of UK tax credit reform on individual participation choices. However, the level of segmentation is limited in this paper, which essentially assumes that the probability of receiving an offer in one sector does not depend on which sector agents are currently employed in. Moreover, the author's "indifference conditions" (the equivalent of our indifference wages) stem from the mere comparison of instantaneous utility flows. One may wonder if the small general equilibrium effect that the author finds do not partly reflect this assumption. Finally, the closest paper to ours is Meghir, Narita & Robin (2012) who study the impact of the existence of an informal sector 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.[9] We offer to generalize some of their insights to any number of markets. In addition, we allow for a more general definition of segmentation, where the option value of unemployment is location-specific and therefore, inherited from past decisions.

The rest of the paper is organized as follows: in the first section, we describe the French labor market as a system of interconnected local urban labor markets; in the second section, we present the model; the third section explains our estimation strategy; estimation results are discussed in the fourth section, whereas a fifth section extends our framework to include moving costs and local amenities.

[9] One noteworthy feature of this paper is that the authors do not need to define indifference conditions between sectors, because they directly focus on labor"contracts", which summarize the entire discounted income flow. Although the optimal contract can be characterized analytically, as we will show, they opt to recover it numerically.

1 Empirical evidence

In this section, we provide descriptive evidence in favor of the modelling 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 mobility of agents; and (iii) stability in key economic variables. We first document the heterogeneity and the stability of the three features which will characterize a local labor market throughout the paper: its population, its unemployment rate and its wage distribution. Then, we describe the mobility patterns of agents on both the labor market and across space.

1.1 France as an equilibrium system of local urban labor markets

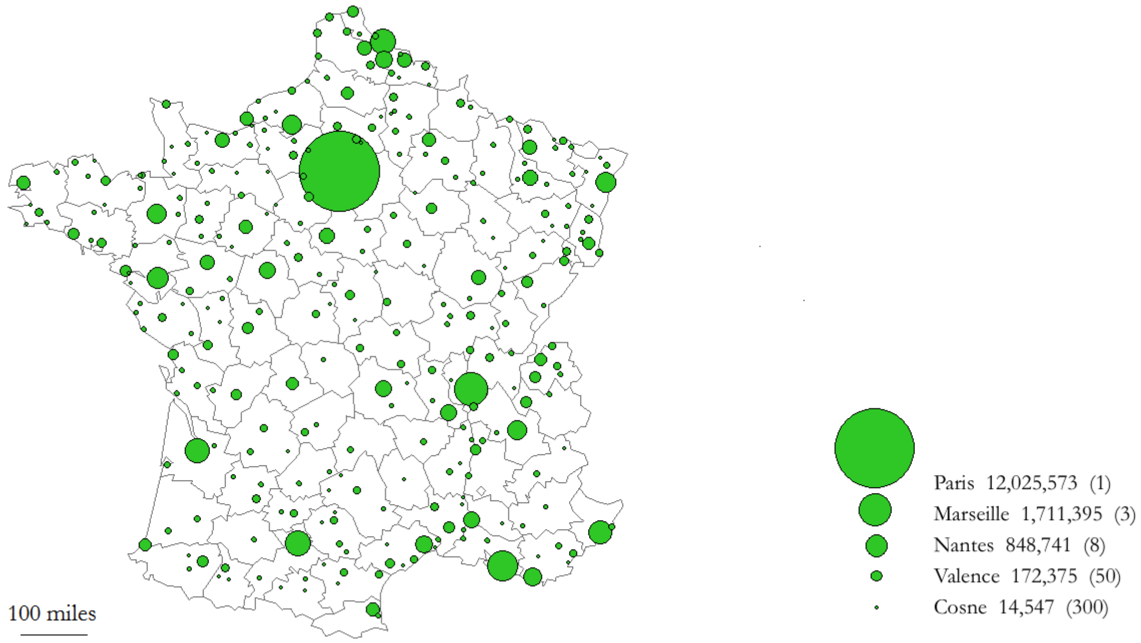
French metropolitan areas (or "aires urbaines") are continuous clusters of municipalities with a main employment center of at least 5,000 jobs and a commuter belt composed of the surrounding municipalities with at least 40% of residents working in the employment center.[10] We consider the 300 largest metropolitan areas in continental France, as defined by the 2010 census. As shown in Figure 2, these metropolitan areas is spread around the entire country and covers a population range from less than fifteen thousand inhabitants to more than twelve millions.

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.[11] French metropolitan areas vary a lot in terms of size, unemployment rates, and wage distributions, yet this heterogeneity is stable over time. We provide here evidence supporting these two claims on the set of the 300 largest metropolitan areas in continental France.

[10] 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. The boundaries of metropolitan areas are subject to dramatic changes because of threshold effects on these statistical criteria. For example, in France, the metropolitan area of Vienne, near Lyon, is a cluster of 13 municipalities in 1999 and becomes a cluster of 40 municipalities after the 2010 revision, based on the 2008 Census; its population jumps accordingly, from 53,787 to 102,474; however, the population of the 13 original municipalities only reaches 55,773 in 2008. In addition, some metropolitan areas disappear into a larger neighbor. This is the case of Saint-Chamond and Elbeuf, respectively 84,925 and 86,162 inhabitants in 1999, which are merged respectively with Saint-Etienne and Rouen. We choose to circumvent this problem by keeping constant the municipal composition of the metropolitan areas.

[11] 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.

Figure 2: The first three-hundred French metropolitan areas



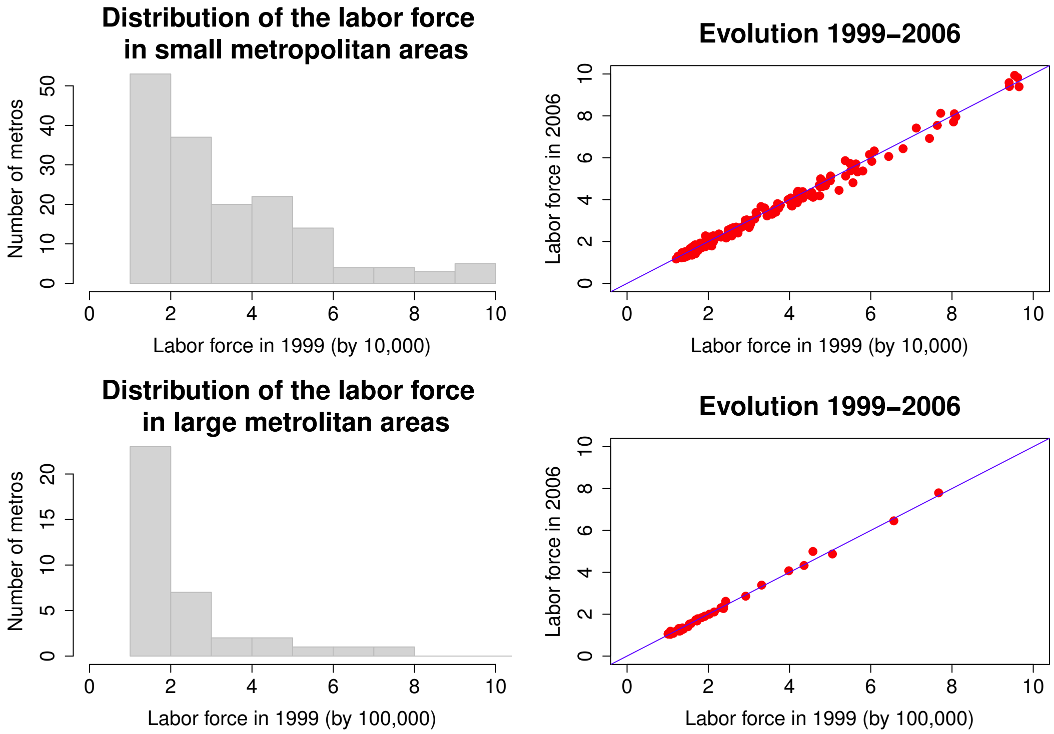
Notes: (i) The circles represent the population in 2007. Legend gives five examples of cities with their name, exact population and rank in the distribution. Population goes from below 15,000 in the six last metropolitan areas to over 12 millions in Paris; the first 7 metropolitan areas have more than one million inhabitants and the first 17, more than 0.5 million; (ii) The lines are the boundaries of the 94 counties ("départements"). *Source: INSEE, Census 2007*

Population Since, we do not model the participation choice of agents, labor force is analogous to total population. To recover precise information about the labor force in each city, we use data from the 1999 and 2006 censuses. Representing approximately 20% of the French population, the Paris region accounts more than 25% of the labor force. As a consequence, the labor force distribution is Pareto-shaped, which is similar to what would have been observed for total population, as shown in the left-hand-side graphs in Figure 3. The right-hand-side graphs in Figure 3 show that absolute variation in the labor force between 1999 and 2006 are negligible.[12]

Unemployment Figure 1 illustrates the dispersion of local unemployment rates in 1999. In addition, Figure 4 establishes that these city-specific unemployment patterns are quite stable over time, especially over a period of stable aggregate unemployment. According to the top graph, stability in aggregate unemployment occurs from 2002 to 2007 both in terms of range and in terms of variation of the annual moving average. For this reason, we will focus on this period throughout the paper. To look at the variation of unemployment at the city level over this period, we use yearly administrative data from the National Unemployment Agency.

[12]This stable distribution of the labor force is at odds with the fact that metropolitan areas face diverse net migration patterns. The explanation lies in the contribution of nonparticipants (retired, young individuals) to the net migration. According to Gobillon & Wolff (2011), 31.5% of French grand-parents 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.

Figure 3: Heterogeneity and stability of the labor force



Notes: (i) Labor force is composed of unemployed and employed individuals aged between 15 and 64; the labor force in the 300 largest metropolitan areas in continental France amounts to 19.5 millions in 1999 and to 19.3 millions in 2006; (ii) For the sake of exposition, we do not represent Paris; its labor force amounts to 5.60 millions in 1999 and 5.55 millions in 2006; moreover, we split the sample according to a 100,000 cut-off: the "small metropolitan areas" are here the metropolitan areas which have a labor force of less than 100,000 people; (iii) The sum of the absolute values of location-by-location changes amounts to 0.57 million,i.e, 3% of total labor force in 1999; (iv) An ordinary-least-squares regression of the 2006 labor force on the 1999 labor force yields a coefficient estimate of 0.99 (t-value of 1318), an estimate of the intercept of 33 (t-value of 0.9) and a R-squared greater than 99.9%. *Source: Census 1999 and 2006.*

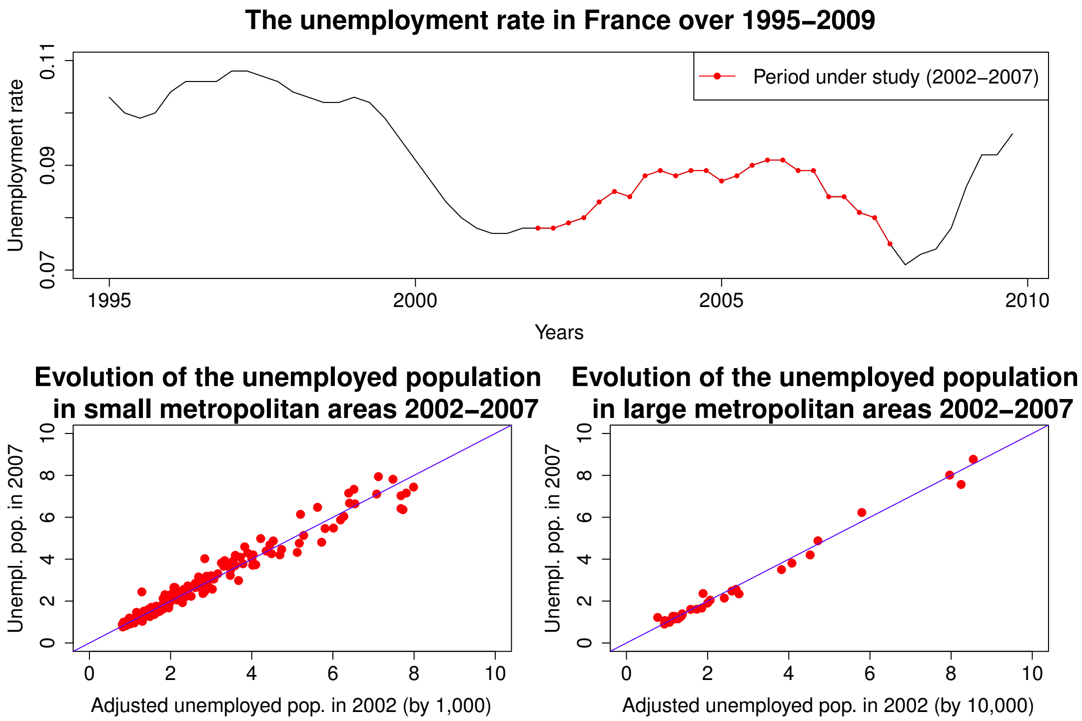
This data cannot be used to compute unemployment rates because it does not provide information on the labor force, but it allows us to look at the absolute changes in the unemployed population. The two bottom graphs show that between 2002 and 2007, city-specific unemployment patterns have remained remarkably stable.

Earning 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.[13] The DADS are the main source of data used in this paper.[14] Table 1 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. Wage distributions in the largest cities stochastically dominate wage distributions in smaller cities. The average wage (33.888 €) in Paris is 51.3% higher than the

[13] See Abowd, Kramarz & Margolis (1999) and Abowd, Kramarz & Roux (2006) for additional information.

[14] We use the longitudinal version of the DADS on a specific subsample of the population (see section 1.2 for details).

Figure 4: Stability of unemployment



Notes: (i) Top graph: quarterly unemployment rate in France; bottom graphs: unemployed population in the 300 largest metropolitan areas in continental France defined in the 2008 Census on December 31, 2001 and 2007; (ii) The unemployed population in 2002 corresponds to the unemployed population of the last day of 2001; it is adjusted to take into account the change in definition that occurred in unemployment statistics in 2005 (see IGF (2007) for more details; conversely, the definition of the unemployment rate used in the top graph is constant); the adjustment factor is set equal to 0.81 to match the ratio of unemployed in France in the last quarter of 2007 as measured by the National Unemployment Agency to its counterpart in the last quarter of 2001; (iii) For the sake of exposition, we do not represent Paris; its adjusted unemployed population in 2002 amounts to 0.390 million and its unemployed population amounts to 0.402 million in 2007; moreover, we split the sample according to a 10,000 cut-off: the "small metropolitan areas" are here the metropolitan areas which have an unemployed population of less than 10,000 people. *Source: Série Longue Trimestrielle INSEE (top) and National Unemployment Agency (bottom).*

city-level average wage. Other large cities have similar wage premia.[15] 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. For instance, Oloron is richer than all the other cities of Panel 2, including cities which are far larger.In addition, there is a strong positive correlation of 0.44 between wage dispersion and city size. These trends are supported by the log-difference between the top and bottom decile (or between the 3rd and the 1st quartiles). They both indicate a higher wage dispersion in Paris, mainly driven by the affluence of high wages.

On a smaller set of moments, Table 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.

[15] 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 22,501€, which is 35% higher than the average wage.

Table 1: Wage distributions

Panel 1: the nine largest cities

Moments	City 1 Paris	City 2 Lyon	City 3 Marseille	City 4 Toulouse	City 5 Lille	City 6 Bordeaux	City 7 Nice	City 8 Nantes	City 9 Strasbourg
P_{10}	14,478	14,264	13,624	13,758	13,407	13,796	13,540	14,166	14,264
Q_1	18,327	17,060	16,141	16,255	15,576	16,117	16,145	16,488	17,074
Q_2	25,815	21,774	20,854	21,093	19,701	20,236	20,768	20,396	21,640
$\mathbb{E}$	33,888	27,221	25,486	25,971	24,751	24,628	26,296	25,143	25,565
Q_3	39,351	30,686	29,223	29,868	27,711	27,820	30,244	27,807	28,792
P_{90}	59,755	45,591	41,450	43,150	40,890	39,759	45,592	40,871	40,113
$\sqrt{\mathbb{V}}$	29,589	18,410	17,370	17,361	17,290	16,403	17,233	17,227	15,948
Q_3/Q_1	2.14	1.80	1.81	1.83	1.78	1.72	1.87	1.68	1.68
P_{90}/P_{10}	4.12	3.19	3.04	3.13	3.04	2.88	3.36	2.88	2.81

Panel 2: nine other cities

Moments	City 20 Nancy	City 30 Brest	City 40 Nimes	City 120 Marsan	City 130 Saintes	City 140 Rochefort	City 220 Luneville	City 230 Oloron	City 240 Lourdes
P_{10}	13,826	13,490	12,823	13,350	13,287	13,192	13,663	13,760	13,154
Q_1	16,311	15,628	14,615	14,599	15,130	14,857	15,547	17,200	15,321
Q_2	20,329	19,123	17,817	16,946	18,125	17,635	18,249	22,763	17,648
$\mathbb{E}$	24,554	23,240	21,207	21,442	21,467	20,451	22,224	24,630	19,704
Q_3	27,384	25,584	23,332	22,564	23,568	23,385	22,408	29,868	21,175
P_{90}	39,290	38,530	32,973	32,151	32,379	31,982	31,360	33,586	30,842
$\sqrt{\mathbb{V}}$	15,419	13,112	12,082	14,434	10,750	8,558	19,302	9,953	7,373
Q_3/Q_1	1.68	1.63	1.59	1.54	1.55	1.57	1.44	1.73	1.38
P_{90}/P_{10}	2.84	2.85	2.57	2.41	2.43	2.42	2.29	2.44	2.28

Notes: (i) Wages are in 2002 Euros and wage distributions are evaluated over the six-year span 2002-2007 *Source: Panel DADS 2002-2007*

Table 2: Stability of the wage distributions

Moments	P_{10}	P_{20}	P_{30}	P_{40}	P_{50}	P_{60}	P_{70}	P_{80}	P_{90}
Q_1^{2007}/Q_1^{2002}	1.005	1.006	1.007	1.008	1.008	1.009	1.009	1.010	1.011
Q_2^{2007}/Q_2^{2002}	1.005	1.006	1.007	1.007	1.008	1.009	1.009	1.010	1.011
$\mathbb{E}^{2007}/\mathbb{E}^{2002}$	1.005	1.007	1.007	1.007	1.008	1.009	1.009	1.010	1.011
Q_3^{2007}/Q_3^{2002}	1.003	1.006	1.006	1.007	1.008	1.009	1.009	1.011	1.013

Notes: (i) Deciles of the distributions of ratios of the moments of the city-specific log-wage distributions in 2007 and in 2002 *Source: Panel DADS 2002-2007*; for details on the sample, see Table 1.

1.2 Labor and geographical mobility

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, which corresponds to the second half of the Chirac presidency. The French economy is in an intermediate state, between a short boom in the last years of the twentieth century, which witnessed a sharp decrease of unemployment and the 2008 financial crisis. The main restrictions over our 2002-2007 sample are the following: first, to mitigate the risk of confusion between non-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 375,000 individuals and 1.5 millions observations (see appendix B.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, to identify transitions on the labor market.[16] 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.[17] 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 where individuals first change locations before accepting a new job offer. From a practical viewpoint, in the DADS data, the sequential job search process is observationally equivalent to the joint mobility process.

Table 3 describes the 719,601 transitions of the 375,276 individuals in our sample. Over our period of study, a third of the sample has recorded no mobility. This figure is comparable with the non-mobility rate of 45% reported by Postel-Vinay & Robin (2002) from 1996 to 1998. Approximately 23% of the sample records

[16] Our algorithm is available at this address

[17] For a recent example of a similar assumption, see Bagger, Fontaine, Postel-Vinay & Robin (2013).

at least one job-to-job transition, and the number of transitions into unemployment and the number of transitions out of unemployment are almost identical. Average wages are almost constant over time, as

Table 3: Number and characteristics of transitions

Type of history	Number of events	Spell Duration	Initial Wage	Final Wage
	Characteristics of Spells			
No transitions while employed	126,157	68.297	25919.515	
Out of unemployment	273,599	13.055		23829.320
with mobility	51,595	13.123		23756.225
without mobility	222,003	12.761		24143.834
Job to job mobility	102,421	22.349	29657.728	31768.862
with mobility	20,970	22.849	29555.303	31669.750
without mobility	81,450	20.404	30055.559	32153.825
Into unemployment	282,177	18.034	24164.645	
Full sample	364,462	24.945	26800.610	27634.238

Notes: (i) Wages are in 2002 Euros and spell durations in months; (ii) Time begins on January 1st 2002. *Source: Panel DADS 2002-2007*

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, 7% lower than the wage of employed agents who do not make any transition, 25% lower than the final wage of agents who have experienced a job-to-job transition, and roughly equal to the initial wage of individuals who will fall into unemployment. For this latter group, note that their initial employment spell is notably shorter than for the rest of the population, which suggests more instability.[18]

Finally, we consider geographical mobility, which accounts for 19.8% of transitions out of unemployment and 22.9% of job-to-job transitions. As shown by Table 4, Paris is both the most prominent destination and the city with the highest rate of transition (90.4%) with no associated mobility.[19] Table 5 completes this overview by comparing the mobility patterns within the Lyon region (also known as "Rhône-Alpes") and between the Lyon region and Paris. Although Paris is the destination of a sizable share of mobile agents, 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. As a consequence, we will incorporate distance between locations as a determinant of spatial frictions (see section 3 for details).

[18] 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 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.

[19] 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 interregional mobility.

Table 4: Mobility between the largest cities

Origin		Destination					
		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	-

Notes: (i) UE stands for transition out of unemployment and EE stands for job-to-job transition; (ii) Reading: among the transitions out of unemployment that started in the city of Lyon, 81.1% led to a job in Lyon, 4.8% led to a job in Paris and 0.7% led to a job in Marseille. *Source: Panel DADS 2002-2007*

Table 5: Distance vs size: mobility within the Lyon region and between the Lyon region and Paris

Origin		Destination					
		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: 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*

We have shown that the French labor market can be considered as a system of interconnected local labor markets, each of which being close to a situation of equilibrium. Three main questions arise: what are the structural determinants of the heterogeneity between these local labor markets? Why is there apparently so little convergence between them? And are spatial frictions the main determinants of jobseekers' geographical mobiliy? In the next sections, we draw upon these various observations and questions to construct and estimate an equilibrium model of job search in a system of cities.

2 A model of job search in a system of cities

2.1 Framework

We consider a system $\mathscr{J}$ of J interconnected local labor markets, or "cities", where agents both live and work. Agents are ex ante identical. They are fully characterized by their employment status $i = e, u$ and their location $j \in \mathscr{J}$. The measures of total population, unemployed population and employed population in each city j are respectively noted m_j, u_j and e_j, with $m_j \equiv e_j + u_j$. The total number of agents in the economy is exogenous and equal to M.

Agents engage in both off-the-job and on-the-job search. Their probability to receive a new job offer depends on their current employment status, their location, as well as on the location associated with the job offer itself. We introduce inverse spatial friction parameters $s^i_{jl} \in (0,1)$ such that a state-i jobseeker living in location j will receive job offers from location j at rate λ^i_j and from location $l \neq j$ at rate $s^i_{jl}\lambda^i_l$ to capture the fact that search process might be less efficient between cities than within cities. In addition, employed agents in city j face a location-specific unemployment risk characterized by the layoff probability δ_j. When they become unemployed, agents receive uniform unemployment benefits b.

Jobseekers sample wages x from a wage offer distribution $F(\cdot)$. The resulting distribution of earnings, or accepted wages, is $G(\cdot)$. Both $F(\cdot)$ and $G(\cdot)$ are location-specific and therefore denoted by $F_j(\cdot)$ and $G_j(\cdot)$. The survival function associated with wage offers in city j is denoted $\overline{F}_j(\cdot) \equiv 1 - F_j(\cdot)$. These distributions are assumed to share a common support $[\underline{w}, \overline{w}]$ over all locations, with $\underline{w} \geq b$ and $\overline{w} < \infty$. Agents do not bargain over wages. They maximize expected lifetime income, discounted at constant, exogenous rate r, by deciding whether to accept or refuse the job offer which they have received. The respective value functions of unemployed agents living in city j and agents living in city j and employed at wage w are recursively defined by equations 1 and 2:

$$rV^u_j = b + \lambda^u_j \int_{\underline{w}}^{\overline{w}} \max\left\{V^e_j(x) - V^u_j, 0\right\} dF_j(x) + \sum_{k \in \mathscr{J}_j} s^u_{jk}\lambda^u_k \int_{\underline{w}}^{\overline{w}} \max\left\{V^e_k(x) - V^u_j, 0\right\} dF_k(x) \quad (1)$$

$$\begin{aligned} rV^e_j(w) &= w + \lambda^e_j \int_{\underline{w}}^{\overline{w}} \max\left\{V^e_j(x) - V^e_j(w), 0\right\} dF_j(x) + \sum_{k \in \mathscr{J}_j} s^e_{jk}\lambda^e_k \int_{\underline{w}}^{\overline{w}} \max\left\{V^e_k(x) - V^e_j(w), 0\right\} dF_k(x) \\ &+ \delta_j \left[V^u_j - V^e_j(w)\right] \end{aligned} \quad (2)$$

where $\mathscr{J}_j \equiv \mathscr{J} - \{j\}$ denotes the set of all cities $k \neq j$.

Mechanisms In the simplest possible job-search model, unemployed jobseekers already face an optimal stopping problem between current and prospective offers. The same mechanism remains valid when one allows for on-the-job search, provided that the probability to receive a job offer while employed is sufficiently lower than its counterpart for unemployed jobseekers. In a model of job search between cities, additional mechanisms come into play, since accepting a good offer in a city conveys parameters that are city-specific. The existence of multiple markets increases the likelihood of a strategic unemployment since an agent may be better off when unemployed in a promising location, than in employment in a depressed location. Such a mechanism occurs when the wage premium associated to a job offer does not compensate for the increase in unemployment risk or the decrease in the expected future wage offers. By refusing an offer, agents would, in a sense, bet on their current unemployment against their future unemployment probability. The same kind of reasoning applies to job-to-job transitions. If agents are willing to accept a wage cut in another location, this decision is somewhat analogous to the purchase of an unemployment insurance contract.

In our system of cities, jobs are no longer defined by the single attribute of wage, but rather by a non-trivial combination of wage, matching parameters and the distribution of wage offers, which determines the offer's option value. This multivariate, and dynamic trade-off allows us to define *spatial strategies*, where agents' 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 of cities. In this case, a job offer is also defined by the city-specific spatial friction parameters and the sequence of cities where individuals are observed can then be rationalized as part of lifetime mobility-based careers.

2.2 Optimal strategies

In order to formalize the previous statements, we now describe the agents' optimal strategies. These strategies are determined by agent'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 agents are ex-ante identical. The optimal strategy consists 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 agent will be willing to accept in her location. Reservation wages, which are therefore location-specific, are denoted ϕ_j and verify $V_j^u \equiv V_j^e(\phi_j)$. Mobility-compatible indifference wages are functions of wage which are specific to any ordered pair of locations $(j,l) \in J \times J$. These functions associate the current wage w earned in location j to a wage which would yield the same dynamic utility in location l. They are denoted $q_{jl}(\cdot)$ and verify $V_j^e(w) \equiv V_l^e(q_{jl}(w))$. By definition, $q_{kl}[q_{jk}(\cdot)] \equiv q_{jl}(\cdot)$. A corollary is that $q_{lj}[q_{jl}(\cdot)] \equiv q_{jj}(\cdot) \equiv \text{id}$. The definition of $q_{jl}(\cdot)$ extends to unemployed agents in city j who receive a job offer in city l: we have $V_j^u \equiv V_l^e(q_{jl}(\phi_j)$.

Proposition 1 *OPTIMAL STRATEGIES*

- *The reservation wage for unemployed workers in city j is defined by the following equation:*

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

which admits a closed-form expression:

$$\phi_j = b + \left(\lambda_j^u - \lambda_j^e\right) \int_{\phi_j}^{\overline{w}} \Phi_j(x) dx + \sum_{k \in \mathscr{J}_j} \left(s_{jk}^u \lambda_k^u - s_{jk}^e \lambda_k^e\right) \int_{q_{jk}(\phi_j)}^{\overline{w}} \Phi_k(x) dx \tag{4}$$

with $\Phi_j(x) \equiv \overline{F}_j dV_j^e(x) = \dfrac{\overline{F}_j(x)}{r + \delta_j + \lambda_j^e \overline{F}_j(x) + \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(x))}$.

- *The mobility-compatible wage in city l for an agent employed in city j at wage w is defined by the following equation:*

$$\begin{aligned} q_{jl}(w) \equiv\ & w + \delta_j V_j^u - \delta_l V_l^u + (1 - s_{lj}^e)\lambda_j^e \int_w^{\overline{w}} \overline{F}_j(x) dV_j^e(x) - (1 - s_{jl}^e)\lambda_l^e \int_{q_{jl}(w)}^{\overline{w}} \overline{F}_l(x) dV_l^e(x) \\ & + \sum_{k \in \mathscr{J}_{jl}} (s_{jk}^e - s_{lk}^e)\lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} \overline{F}_k(x) dV_k^e(x) \end{aligned} \tag{5}$$

where $\mathscr{J}_{jl} \equiv \mathscr{J} - \{j,l\}$, which admits a closed-form expression:

$$\begin{aligned} q_{jl}(w) =\ & w + \frac{\delta_j}{r}\left(b + \lambda_j^u \int_{\phi_j}^{\overline{w}} \Phi_j(x) dx + \sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \int_{q_{jk}(\phi_j)}^{\overline{w}} \Phi_k(x) dx\right) \\ & - \frac{\delta_l}{r}\left(b + \lambda_l^u \int_{\phi_l}^{\overline{w}} \Phi_l(x) dx + \sum_{k \in \mathscr{J}_l} s_{lk}^u \lambda_k^u \int_{q_{lk}(\phi_l)}^{\overline{w}} \Phi_k(x) dx\right) \\ & + (1 - s_{lj}^e)\lambda_j^e \int_w^{\overline{w}} \Phi_j(x) dx - (1 - s_{jl}^e)\lambda_l^e \int_{q_{jl}(w)}^{\overline{w}} \Phi_l(x) dx + \sum_{k \in \mathscr{J}_{jl}} (s_{jk}^e - s_{lk}^e)\lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} \Phi_k(x) dx \end{aligned} \tag{6}$$

- *Equations 4 and 6 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 to 6 using the definitions of ϕ_j and $q_{jl}(\cdot)$, a closed-form expression for $dV_j^e(\cdot)$ and integration by parts; in appendix A.2 we demonstrate the existence and uniqueness of the solution through an application of the *Banach fixed-point theorem*.

The interpretation of equation 3 is straightforward: the difference in the instantaneous values of unemployment and employment $(\phi_j - b)$ can be understood as a difference in opportunity cost, which must be perfectly compensated for by the difference in the option values of unemployment and employment, ie the wages associated with the jobs that would be found, either while unemployed, or through on-the-job search. The interpretation of equation 5 is similar to the previous one. Here, the difference in the instant values of employed agents in location l and location j is $(q_{jl}(w) - w)$. As for the option values, they are a bit more complicated. The first part is given by the difference in the value of unemployment, weighted by unemployment risks δ_j and δ_l. The second part is given by the difference in the value of on-the-job search without mobility. Finally, the third part is given by the difference in the value of on-the-job search with mobility. It stems from the comparison of the strength of spatial frictions between the two locations j and l and the rest of the world, which captures the relative *centrality* of cities j and l, for a given wage level.

From equation 6, we get the following expression for the derivative of the mobility-compatible indifference wage:

$$dq_{jl}(w) = \frac{1 - (1 - s_{lj}^e)\lambda_j^e \Phi_j(w) - \sum_{k \in \mathcal{J}_{jl}} (s_{jk}^e - s_{lk}^e)\lambda_k^e \Phi_k(q_{jk}(w)) dq_{jk}(w)}{1 - (1 - s_{jl}^e)\lambda_l^e \Phi_l(q_{jl}(w))} \tag{7}$$

Given that $dV_j^e(\cdot)$ is always strictly positive for any city $j \in \mathcal{J}$, $dq_{jl}(\cdot)$ is always strictly positive for any pair of cities $(j, l) \in \mathcal{J} \times \mathcal{J}_j$.This feature ensures the uniqueness of the solution to the system formed by equation 7 for all pairs of locations (see appendix A.2 for details). Note, however, that the sign of $q_{jl}(w) - w$ remains ambiguous: it depends on the pair of cities (j, l) and the wage level w. More precisely, one can see that $q_{jl}(w) \geq w \Leftrightarrow q_{lj}(w) \leq w$. The model predicts that half of the potential between-cities job-to-job transitions allow for wage cuts.[20]

[20] This symmetry assumption can be relaxed if we take mobility costs into account. See Section 5 for details.

Discussion: minimum wage and voluntary unemployment In principle, spatial segmentation would allow for an unusual individual behavior: a geographical mobility with strategic transition into unemployment. Whereas the reservation wage strategy ensures that voluntary unemployment is never an optimal strategy when there only is one labor market, this is not a sufficient condition in our model. In the (λ^u, λ^e)-plane, it is possible to find a pair of cities (j, l) such that $\phi_j < q_{lj}(\phi_l)$. In this case, workers employed in city j at a wage $w \in [\phi_j, q_{lj}(\phi_l)]$ should accept any job offer coming from city l lower than ϕ_l, provided it is greater than $q_{jl}(w)$. However, once settled in city l, these newcomers would be worse-off than their unemployed neighbors and would take advantage from quitting their job. Although this strategy is plausible between very unequal markets with low mobility, it represents an unsustainable off-equilibrium path: in the long-run, the inflow of workers willing to accept low wages would drive down the reservation wage. Therefore, we rule it out by assuming that $\max_{j \in \mathscr{J}}\{\phi_j\} \leq \underline{w}$. Such assumption is more plausible for low-skilled workers than for the entire population. However, empirically, it has been shown to be almost costless in the French context, characterized by a high minimum wage (see online appendix in Cahuc et al. (2006)).

2.3 Steady-State Equilibrium

As already explained in section 1, a cross-section description of the economy as a system of cities is fully characterized by a set of city-specific populations, unemployment rates and earning distributions. If all these multi-dimensional outcome variables are constant, the economy can be said to have reached a steady-state equilibrium.

Unemployment rate At each point in time, the number of unemployed agents in a city j is constant. A measure $u_j \lambda_j^u \overline{F}_j(\phi_j)$ of agents 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 agents who were previously employed in city j but have just lost their job. This equilibrium condition leads to the following proposition:

Proposition 2 *EQUILIBRIUM UNEMPLOYMENT*

In equilibrium, the unemployment rate in each location j is given by:

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

Population 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 agents in city j who find and accept another job in any city $k \neq j$; conversely, inflows are composed by employed and unemployed agents 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:

$$
\begin{aligned}
(m_j - u_j) \sum_{k \in \mathscr{J}_j} s^e_{jk} \lambda^e_k \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x) + u_j \sum_{k \in \mathscr{J}_j} s^u_{jk} \lambda^u_k \overline{F}_k(q_{jk}(\phi_j) \equiv \\
\lambda^e_j \sum_{k \in \mathscr{J}_j} s^e_{kj} (m_k - u_k) \int_{\underline{w}}^{\overline{w}} \overline{F}_j(q_{kj}(x) dG_k(x) + \lambda^u_j \sum_{k \in \mathscr{J}_j} s^u_{kj} u_k \overline{F}_j(q_{kj}(\phi_k))
\end{aligned}
\tag{9}
$$

After plugging equation 8 into equation 9, one can recover a closed form solution for the system, which can be written as:

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

where $\mathscr{M}$ is the vector of $\{m_j\}_{j \in \mathscr{J}}$ and A is the matrix of typical element (A_{jl}) defined by equations 11 and 12:

$$A_{jj} = \frac{\left(\lambda^u_j \overline{F}_j(\phi_j) + \sum_{k \in \mathscr{J}_j} s^u_{jk} \lambda^u_k \overline{F}_k(q_{jk}(\phi_j))\right) \cdot \left(\sum_{k \in \mathscr{J}_j} s^e_{jk} \lambda^e_k \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x)\right) + \delta_j (\sum_{k \in \mathscr{J}_j} s^u_{jk} \lambda^u_k \overline{F}_k(q_{jk}(\phi_j)))}{\delta_j + \lambda^u_j \overline{F}_j(\phi_j) + \sum_{k \in \mathscr{J}_j} s^u_{jk} \lambda^u_k \overline{F}_k(q_{jk}(\phi_j))} \tag{11}$$

$$A_{jl} = -\frac{\left(\lambda^u_j \overline{F}_j(\phi_j) + \sum_{k \in \mathscr{J}_j} s^u_{jk} \lambda^u_k \overline{F}_k(q_{jk}(\phi_j))\right) \cdot \left(s^e_{lj} \lambda^e_j \int_{\underline{w}}^{\overline{w}} \overline{F}_j(q_{lj}(x)) dG_l(x)\right) + \delta_l s^u_{lj} \left(\lambda^u_j \overline{F}_j(q_{lj}(\phi_l))\right)}{\delta_j + \lambda^u_j \overline{F}_j(\phi_j) + \sum_{k \in \mathscr{J}_j} s^u_{jk} \lambda^u_k \overline{F}_k(q_{jk}(\phi_j))} \qquad \text{if } j \neq l \tag{12}$$

Rounding issues aside, we can therefore write the following proposition:

Proposition 3 *EQUILIBRIUM POPULATION*

In equilibrium, the distribution of city sizes is the positive vector $\mathscr{M} \in \ker A$ *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. This difference stems from the assumption whereby agents do not change locations without finding a job. For this reason, the flow of agents into unemployment in city j is only composed of agents previously located in city j, whereas in equation 9, the population in city j is also determined by the flow of agents who come from everywhere else and have found a job in city j.

Earning distributions Finally, the distribution of observed wages is considered. 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 agents 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 have 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 agents who find and accept a job in city j with a wage lower than w. These agents 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 agents who were previously employed in any city $k \neq j$ at a wage x lower than the mobility-compatible indifference wage of w (that is, $q_{jk}(w)$) and find a job at a wage between w and $q_{kj}(w)$. As before, the measure of this flow requires an integration over the distributions of observed wages in any city $k \neq j$.

This is all summarized in equation 13:

$$
\begin{aligned}
(m_j - u_j)\Big[G_j(w)\big(\delta_j + \lambda_j^e \overline{F}_j(w)\big) + \sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{w} \overline{F}_k(q_{jk}(x))dG_j(x)\Big] \equiv & \quad (13)\\
\lambda_j^u\Big[\psi_{jj}(w)u_j\big(F_j(w) - F_j(\phi_j)\big) + \sum_{k\in\mathscr{J}_j} s_{kj}^u \psi_{kj}(w)u_k\big(F_j(w) - F_j(q_{kj}(\phi_k))\big)\Big] & \\
+\lambda_j^e \sum_{k\in\mathscr{J}_j} s_{kj}^e (m_k - u_k)\int_{\underline{w}}^{q_{jk}(w)} [F_j(w) - F_j(q_{kj}(x))]dG_k(x) &
\end{aligned}
$$

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 agents from any city $k \neq j$ and it is nil if $q_{jk}(w) < \underline{w}$. These restrictions mean that very low values of w will not attract many jobseekers. We can differentiate equation 13 with respect to w. This yields the following linear system of functional differential equations:

$$
f_j(w) = \frac{g_j(w)(m_j - u_j)\Big[\delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(w))\Big]}{\lambda_j^u\Big(\psi_{jj}(w)u_j + \sum_{k\in\mathscr{J}_j} s_{kj}^u \psi_{kj}(w)u_k\Big) + \lambda_j^e\Big((m_j - u_j)G_j(w) + \sum_{k\in\mathscr{J}_j} s_{kj}^e (m_k - u_k)G_k(q_{jk}(w))\Big)} \quad (14)
$$

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. Unlike the system 10, the uniqueness of the solution is not guaranteed. We defer the question of identification to section 3.4. We can then write the following proposition:

Proposition 4 *EQUILIBRIUM WAGE OFFER DISTRIBUTIONS*

In equilibrium, the distribution of wage offers by location is solution to the system 14.

Note that Equation 14 also defines a system of differential equations for the earning distribution:

$$g_j(w) = \frac{f_j(w)\left[\lambda_j^u\left(\psi_{jj}(w)u_j + \sum_{k\in\mathscr{J}_j} s_{kj}^u \psi_{kj}(w)u_k\right) + \lambda_j^e\left((m_j-u_j)G_j(w) + \sum_{k\in\mathscr{J}_j} s_{kj}^e(m_k-u_k)G_k(q_{jk}(w))\right)\right]}{(m_j-u_j)\left[\delta_j + \lambda_j^e\overline{F}_j(w) + \sum_{k\in\mathscr{J}_j} s_{jk}^e\lambda_k^e\overline{F}_k(q_{jk}(w))\right]} \tag{15}$$

Definition of the steady-state equilibrium *A steady-state equilibrium* for this economy is a household value function, $V_j^u(\cdot)$ and $V_j^e(\cdot)$; structural parameters $(b, r, \{s_{jk}^u, s_{jk}^e, \delta_j, \lambda_j^e, \lambda_j^u\}_{(j,k)\in\mathscr{J}\times\mathscr{J}_j})$ and wage offer distributions $F(\cdot) \equiv \{F_j(\cdot)\}_{j\in\mathscr{J}}$, such that:

1. The reservation wage strategy in equation 3 describes the optimal job acceptation behavior of immobile unemployed agents.
2. The optimal mobility strategy between two locations is defined by the indifference condition described in equation 5.
3. The optimal set of unemployment rates is given by equation 8.
4. The optimal set of city sizes is solution to the linear system 10.
5. The optimal wage posting behaviour of firms, summarized by the set of wage offer distributions $F(\cdot)$ is given as the solution of the system of functional differential equations 14.

3 Estimation

3.1 Simulated Method of moments

The model is estimated by simulated method of moments (hereafter, SMM).[21] We restrict the number of cities to 200. 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.[22]

The SMM estimator minimizes the distance between a set of empirical moments and their theoretical counterparts. Let θ denote the set of parameters to be estimated, $m(\theta)$ and $\hat{m}$ the theoretical and empirical

[21] See Gourieroux, Monfort & Renault (1993) for details.

[22] As a consequence, the smallest metropolitan area which is isolated in our analysis is Redon, with 28,706 inhabitants in 2009. A population threshold of 28,000 remains very low.

moments. The SMM consists in finding the parameters $\hat{\theta}$ that minimize the criterion function $\mathscr{L}(\theta)$ given by:

$$\mathscr{L}(\theta) = -\frac{1}{2}\big(\widehat{m} - m(\theta)\big)^T \widehat{W}^{-1}\big(\widehat{m} - m(\theta)\big) \tag{16}$$

where $\widehat{W}$ is the diagonal of the estimated covariance matrix of $\widehat{m}$.

3.2 Algorithm and numerical solutions

For notational convenience, let $G(\cdot) \equiv \{G_j(\cdot)\}_{j\in\mathscr{J}}$, $g(\cdot) \equiv \{g_j(\cdot)\}_{j\in\mathscr{J}}$ and $q(\cdot) \equiv \{q_{jl}(\cdot)\}_{(j,l)\in\mathscr{J}\times\mathscr{J}_j}$. 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 5 to recover indifference wages $q(\cdot)$
4. Solve equation 10 to recover equilibrium population $\mathscr{M}$
5. Solve equation 14 to update the distribution of job offers $F(\cdot)$
6. Update θ using the maximum of $\mathscr{L}(\theta, t, w)$.
7. Repeat steps 3 to 6 until convergence.

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 $\dim(w)$ times. Moreover, since $\Phi_j(\cdot)$ is a function of all $\{q_{jk}(\cdot)\}_{k\in\mathscr{J}_j}$, the numerical integration of $\Phi_j(\cdot)$ requires a prior knowledge of the functional form of all $\{q_{jk}(\cdot)\}_{k\in\mathscr{J}_j}$. We circumvent this problem by using an embedded algorithm that allows us to recover a piecewise approximation of all indifference wages. Note that for any $w^* > w$ and a relatively small h, Newton's formula yields:

$$q_{jl}(w^*) = q_{jl}(w) + h d q_{jl}(w) \tag{17}$$

We further assume that the derivatives of the indifference wages are independent, such that equation 7 becomes:

$$d\tilde{q}_{jl}(w) = \frac{1 - (1 - s^e_{lj})\lambda^e_j \Phi_j(w) - \sum_{k\in\mathscr{J}_{jl}} (s^e_{jk} - s^e_{lk})\lambda^e_k \Phi_k(q_{jk}(w))}{1 - (1 - s^e_{jl})\lambda^e_l \Phi_l(q_{jl}(w))} \tag{18}$$

This assumption is not necessary, since we could update the value of all $dq_{jk}(w)$ at each step; however, this would dramatically increase the computational burden, for no real gains, given the smoothness of these functions. The indifference wage can be solved using the following sequential process:

3.1 Declare an initial guess for $q(\underline{w})$

3.2 Use the values of $q(\underline{w})$ to iteratively recover $q(w_1)$, then $q(w_2)$, up to $q(\overline{w})$.

3.3 Use the sequence $q(w_1),...,q(\overline{w})$ to approximate the function $q(\cdot)$.

3.4 Use the function $q(\cdot)$ to update the value of $q(\underline{w})$

3.5 Repeat steps 3.2 to 3.4 until convergence.

In step 3.1, $q(\underline{w})$ is determined by setting the value of each integral in equation 6 to 1:

$$q_{jl}(w) = w + \frac{\delta_j}{r}\left(b + \lambda_j^u + \sum_{k\in\mathcal{J}_j} s_{jk}^u \lambda_k^u\right) - \frac{\delta_l}{r}\left(b + \lambda_l^u + \sum_{k\in\mathcal{J}_l} s_{lk}^u \lambda_k^u\right) + (1 - s_{lj}^e)\lambda_j^e - (1 - s_{jl}^e)\lambda_l^e + \sum_{k\in\mathcal{J}_{jl}} (s_{jk}^e - s_{lk}^e)\lambda_k^e \tag{19}$$

Once the indifference wages are recovered, we can turn to the evalution of the wage distributions (step 5 in the general algorithm). There are two difficulties when solving for the system defined by equation 15. The first one stems from the dimensionality of the problem: for any $J \geq 3$, the system can only be solved numerically.[23] Second, the system is composed of functional equations, which standard differential solvers are not designed to handle. For simplicity, we first assume that $F(\cdot)$ follows a parametric distribution:

$$\hat{F}_j(x) = \text{gammacdf}\left(x, \alpha_j, \beta_j\right) \tag{20}$$

where $\text{gammacdf}\left(\cdot, \alpha_j, \beta_j\right)$ is the CDF of a gamma distribution with shape α_j and scale β_j parameters. Let $\alpha \equiv \{\alpha_j\}_{j\in\mathcal{J}}$ and $\beta \equiv \{\beta_j\}_{j\in\mathcal{J}}$. Our empirical counterparts are based on real wages, so we treat the empirical cdf G as unknown and the set of parameters α and β are estimated such that they minimize the distance between the empirical cdf G and its empirical counterpart as given by equation 15. 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 15 in order to update (α, β). We develop a simple iterative process based on Euler approach. That is,

5.1 At initial iteration, set $q_{jl}(w) = w$, such that equation 14 becomes a standard ODE.

[23] Two-sector models, such as the one presented in Meghir et al. (2012), yield systems of two ordinary differential equations. These systems can be rewritten in a way such that they still admit a closed-form solution.

5.2 Set the step size $h = 0.1$, and use Euler method to approximate the sequence of $G_j(\cdot)$.

5.3 Derive estimate for $G_l(q_{jl}(w))$ for all $j \in \mathscr{J}$.

5.4 Use estimates of $G_l(q_{jl}(w))$ to solve the functional differential equation 15.

5.5 Repeat steps 5.3 to 5.4 until convergence.

In practice, for an initial value $w_0 = \underline{w} - \epsilon$, we set $G_j(w_0) = 0$ for all $j \in \mathscr{J}$.[24] Hence, for any $w_1 = w_0 + h$, we can write:

$$G_j(w_1) = G_j(w_0) + hg_j(w_0) \tag{21}$$

and we iterate until reaching the maximum wage, $\overline{w}$. Once a solution for $G_j(\cdot)$ is recovered, we update α and β by minimizing the distance between $G(\cdot)$ and $\tilde{G}(\cdot)$ over the space of gamma distributions.

3.3 Parametrization

The model is based on a set of parameters $\theta = \left\{\lambda_j^e, \lambda_j^u, \delta_j, s_{jk}^e, s_{jk}^u, \alpha_j, \beta_j\right\}_{(j,k)\in\mathscr{J}\times\mathscr{J}_j}$ such that $|\theta| = 81{,}000$ with $J = 200$. However, in practice, estimating a separate parameter s_{jl} for each pair of cities would be too computationally demanding and would require to drastically restrict $\mathscr{J}$. We take an alternative path and we posit and estimate a parsimonious parametric model which is akin to a constrained gravity equation:[25]

$$s_{jl}^i = \frac{\exp\left(s_{0l}^i + s_1^i d_{jl} + s_2^i d_{jl}^2 + s_3^i h_{jl} + s_4^i h_{jl}^2\right)}{1 + \exp\left(s_{0l}^i + s_1^i d_{jl} + s_2^i d_{jl}^2 + s_3^i h_{jl} + s_4^i h_{jl}^2\right)} \tag{22}$$

where s_{0l} is a city fixed effect, d_{jl} is the measure of physical distance between city j and city l and h_{jl} is a dissimilarity index based on the sectoral composition of the workforce between 35 sectors.[26] The fixed effects measure the relative openness of the local labor markets and more precisely, the ability of each city to fill its vacancies with workers coming from other locations. The other parameters account for the effect of distance between two locations. First, physical distance is arguably the most important characteristic.[27] Second, sectoral dissimilarity gives a proxy for potential coordination frictions between the two locations. The sectoral dissimilarity is particularly important to rationalize job-to-job mobility rates between highly

[24] See Judd (1998) for details.

[25] See Head & Mayer (2013) for the current state of the art about gravity equations.

[26] We use the traditional Duncan index: if v is a categorical variable defined by categories k in proportion v_j^k in city j and in proportion v_l^k in city l, $h_{jl} = 0.5\sum_k |v_j^k - v_l^k|$. In order to construct this variable, we use the 2007 SIRENE database. SIRENE is a firm level census.

[27] As shown, among others, by Combes & Lafourcade (2005) on trade between French cities, physical distance is a very good proxy for transport cost in a cross-section analysis.

specialized cities (for example, biotechnologies in Lyon and Strasbourg). We allow for variability in the returns to these two measures of distance by considering a second-order polynomial. Given the lack of existing literature on the explicit structure of spatial frictions, we choose to use a logistic function for s_{jl} because of its analytical properties.[28] Under this specification, the total number of parameters to be estimated amounts to 1,408.

3.4 Identification

Our main identification challenge comes from the wage offer distribution, which is defined by a functional differential equation. Assuming that a reasonable initial condition exists, we provide the following proposition:

Proposition 5 *EXISTENCE AND UNIQUENESS OF WAGE OFFER DISTRIBUTIONS The system of differential equations* $\mathfrak{f} : \mathbb{R}^J \to (0,1)^J$ *has a unique fixed point.*

Proof 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. For more details, see Theorem 2.3 in Hale (1993).

As shown by Flinn & Heckman (1982) and Magnac & Thesmar (2002), structural parameters are identified from transition rates. The within and between transitions from unemployment to employment identify λ^u and s^u. The same reasoning applies to the on-the-job search rates λ^e and s^e. Finally, job destruction rates δ are identified from transitions into unemployment. However, instead of using the raw transitions between employment and unemployment, we choose to identify λ^u and δ using the city-specific populations and unemployment rates. Since these two distributions are the most relevant dimensions of our model, we want to make sure that our estimation reproduces them as accurately as possible.

Given the parametrization 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 total transitions probabilities into a given city. The identification of the parameters relating to the distance and the dissimilarity between two cities requires transition

[28] See Zenou (2009a) for a theoretical approach in terms of endogenous search intensity.

rates at the city-pair level. We select a subset of city pairs $\mathscr{T} \subset \mathscr{J}^2$, with $|\mathscr{T}| = 30$, which we use in the estimation. In practice, we use the off-the-job and job-to-job transitions rates from the urban areas ranked fifth to tenth to the urban areas ranked twentieth to twenty-fifth (see Figure 8 in appendix C for details).[29] Our identification strategy is summarized in Table 6.

Table 6: Moments and Identification

Empirical moments	Theoretical moments	Identifying Parameters
Unemployment rate in city j	u_j/m_j	δ_j, λ_j^u
Labor force in city j	m_j	δ_j, λ_j^u
Earning distribution in city j	$G_j(\varpi)$	α_j, β_j
Transition rate EE within city j	$\lambda_j^e \int_{\underline{w}}^{\overline{w}} \overline{F}_j(x) dG_j(x)$	λ_j^e
Transition rate UE into city l (total)	$\lambda_l^u \sum_{k\in\mathscr{J}_l} s_{kl}^u \overline{F}_l(q_{kl}(\underline{w}))$	s_{0l}^u
Transition rate EE into city l (total)	$\lambda_l^e \sum_{k\in\mathscr{J}_l} s_{kl}^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{kl}(x)) dG_k(x)$	s_{0l}^e
Transition rate UE between city j and city l (subset)	$\left\{s_{jl}^u \lambda_l^u \int_{\underline{w}}^{\overline{w}} \overline{F}_l(q_{jl}(x)) dG_j(x)\right\}_{(j,l)\in\mathscr{T}}$	s_1^u, s_2^u, s_3^u, s_4^u
Transition rate EE between city j and city l (subset)	$\left\{s_{jl}^e \lambda_l^e \int_{\underline{w}}^{\overline{w}} \overline{F}_l(q_{jl}(x)) dG_j(x)\right\}_{(j,l)\in\mathscr{T}}$	s_1^e, s_2^e, s_3^e, s_4^e

Notes: ϖ is a grid of 17 levels of wages.

4 Results

In this section, we first present our structural estimation results and in particular, the distribution of the matching parameters, the impact of distance on spatial frictions and the impact of on-the-job search and of city openness on the wage level in each city. We then propose a linear model to understand the determinants of the local matching parameters. Finally, we estimate a parsimonious matching function that only depends on city's area, population and number of firms and use it to study the impact of the number of cities on aggregate unemployment.

4.1 A dataset of city-specific matching parameters

Figure 5 describes the matching parameters $(\lambda^u, \lambda^e, \delta)$ that characterize the 200 largest French cities. For clarity in exposition, we split these cities into three size groups: 40 large cities (from Paris to Valence), 60 mid-sized cities (from Saint-Brieuc to Sète) and 100 smaller cities (from Thonon-Les-Bains to Redon). As will be shown in several instances, these three groups of cities do not follow the same logics, which makes

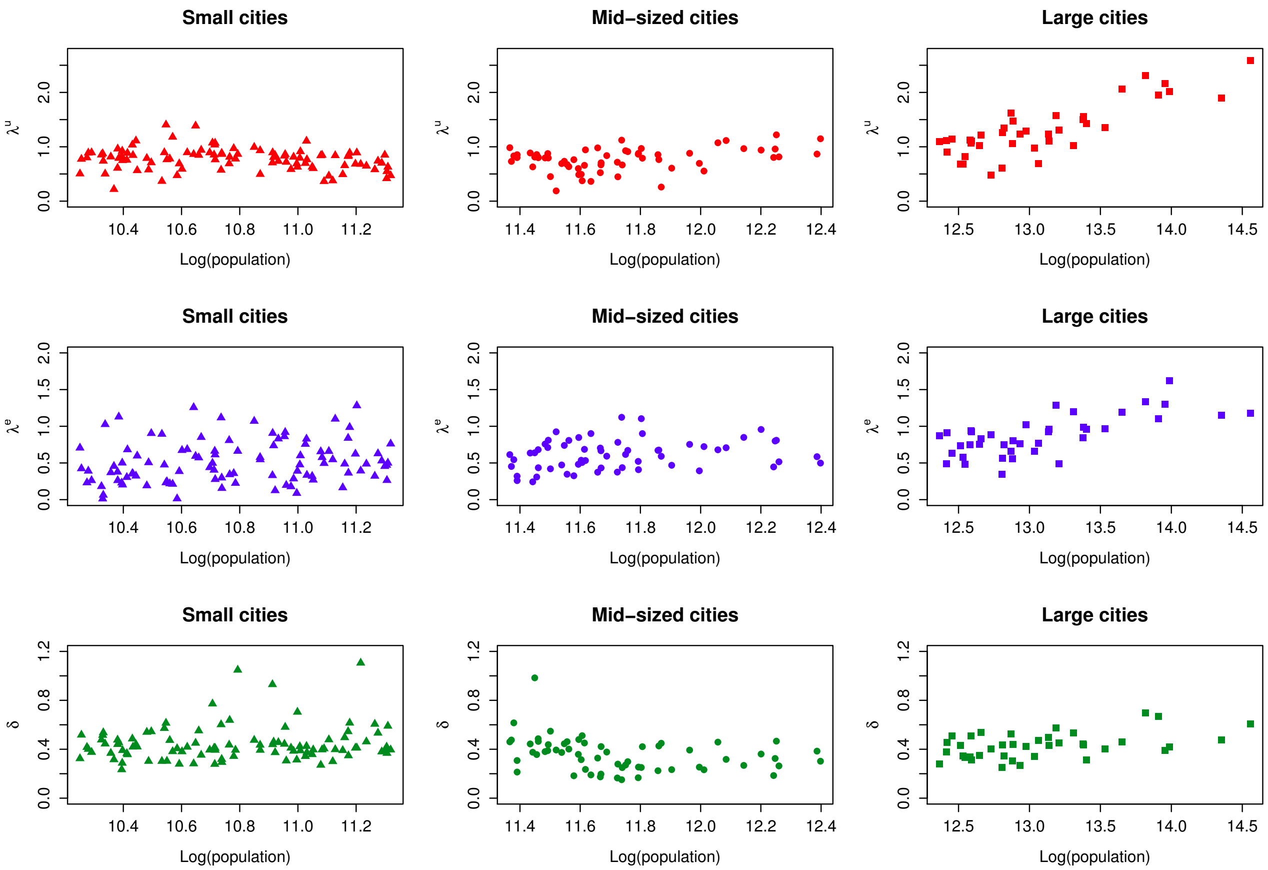
[29]This choice, somewhat random, is meant to include locations that are widely scattered across the French territory. Datasets and program files are available upon request and the reader is invited to check if this selection affects the results.

their separate study interesting. However, in contrast to Baum-Snow & Pavan (2012), we allow for heterogeneity in the structural parameters within each subset of cities. Table 7 complements this presentation by providing the summary statistics of these matching parameters and the between-city job arrival rate for the unemployed.

Table 7: The matching parameters: summary statistics

	λ_j^u	λ_j^e	δ_j	$\sum_{k\in\mathcal{J}_j} s_{jk}^u \lambda_k^u$
Minimum	0.189	0.013	0.151	0.000
1st Quartile	0.688	0.391	0.326	0.017
Mean	0.918	0.618	0.413	0.343
Sd	0.671	0.302	0.141	0.032
Median	0.833	0.593	0.401	0.381
3rd Quartile	0.972	0.809	0.462	0.624
Maximum	8.931	1.660	1.106	1.945

Figure 5: The matching parameters along the city size distribution



Notes: (i) Estimated values of the structural parameters $(\lambda^u, \lambda^e, \delta)$ for the 40 largest cities, the 100 smallest cities and the 60 cities in between; (ii) For the sake of exposition, we do not represent Paris; its log(population) amounts to 16.30, $\lambda^u_{Paris} = 8.93$, $\lambda^e_{Paris} = 1.66$ and $\delta_{Paris} = 0.15$; (iii) The city 200 is made of all remaining metropolitan areas. It is included in the estimation but its parameters are not meaningful and therefore are not represented here.

The estimated values of λ^u, which range from 0.2 to 8.9, show substantial heterogeneity across cities. In Paris for example, the job arrival rate of 8.9 implies that offers accrue approximately every 9 months on average.[30] The median value of the job arrival rate, around 0.8, confirms the very low transition rate of the French economy as documented by Jolivet et al. (2006). There is also considerable heterogeneity in both voluntary and involuntary job separation rates. As shown in Table 7, on-the-job search is a crucial component of the French labor market. Even though it is often very low, this feature is critical in many local labor markets (in 46 cities, λ^e is even higher than λ^u), such that the unweighted average of λ^e is no less than two thirds of its counterpart for λ^u. Figure 5 shows that λ^e is strongly correlated with city size. Seemingly, the job destruction rate δ is not. We come back to this issue in more details when trying to infer on the determinants of our structural parameters.

We now turn to the correlation between our parameters and between our parameters and the local wage distributions. In Table 8, these distributions are summarized by the their first two moments. The three main findings are the following. First, panel 1 shows that there is a positive correlation between off-the-job and on-the-job arrival rates.[31] This may be interpreted as indirect evidence that the labor market is not segmented between "insiders" and "outsiders". However, as shown in panels 2 to 4, this correlation is driven by the largest cities. Second, the strong correlation between on-the-job search rate and wage dispersion at the city level provides a direct test of the wage posting theory, as outlined by Burdett & Mortensen (1998). However, this correlation is also mostly driven the group of large cities. Finally, the negative correlation between the average wage and the job separation rate suggests that agents do accept lower wages when they face a higher unemployment risk. This piece of evidence provides another assessment of the wage posting theory. Interestingly, this correlation is driven by the groups of mid-sized and small cities. These three observations, which are only valid, either for large cities, or for small and mid-sized cities, suggest the existence of very different wage dynamics according to city size. Section 4.3 will confirm this hypothesis using a wage decomposition approach.

Finally, we adopt a least squares approach to study the determinants of the structural matching parameters which characterize each local labor market. Unlike previous studies, the large number of parameter estimates allows us to draw this kind of inference, both for the total population, and for the three groups of cities taken separately. We model the job arrival rates and the job separation rate as linear functions of the number of firms in the city, the population density, the share of the population below thirty years old, the share of the population without qualifications, the share of males, the share of blue-collar jobs and the share

[30] Recall that these parameters define a matching process that takes place on a six-year span, between 2002 and 2007.

[31] Note that this correlation may partly reflect a co-dependence to a third variable, such as city size.

Table 8: Correlation between the local labor market primitives

Panel 1: All cities

	λ^u	λ^e	δ	$\overline{w}$
λ^e	0.44***			
δ	-0.03	0.04		
$\overline{w}$	0.46***	0.29***	-0.23**	
σ_w	0.42***	0.42***	-0.02	0.73***

Panel 2: Large cities

	λ^u	λ^e	δ	$\overline{w}$
λ^e	0.61***			
δ	-0.18	0.25		
$\overline{w}$	0.74***	0.64***	-0.14	
σ_w	0.76***	0.78***	0.06	0.84***

Panel 3: Mid-sized cities

	λ^u	λ^e	δ	$\overline{w}$
λ^e	0.09			
δ	-0.06	0.02		
$\overline{w}$	0.14	0.07	-0.31*	
σ_w	0.00	0.20	-0.06	0.55***

Panel 4: Small cities

	λ^u	λ^e	δ	$\overline{w}$
λ^e	0.01			
δ	0.09	0.04		
$\overline{w}$	0.09	-0.10	-0.21*	
σ_w	-0.11	0.12	0.00	0.68***

Notes: (i) Significance: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$; (ii) $\overline{w}$ is the average wage in each city and σ_w is the standard deviation of wages in each city: (iii) The parameters for the city 200 are not included in panels 1 and 4

of jobs in the manufacturing sector. Results are presented in Table 9. This parsimonious linear specification explains 90% of the variation in off-the-job arrival rates, ie three times more than for on-the-job arrival rates and five times more than for job separation rates. This very high explanatory power is driven by the subset of large cities. More generally, whereas the R-squared of the regressions for both λ^u and λ^e decreases with city size, this is not the case for δ.

Several coefficients are significant with the expected signs. The number of firms, which proxies the supply side of the matching function, is positively correlated with job arrival rates, and negatively correlated with separation rates, at least in larger cities. The share of young people has a similar effect. Less educated and less dense cities witness less on-the-job search, whereas these two characteristics do not affect the job finding rate for the unemployed. To summarize, while large and mid-sized cities share roughly the same patterns in the determination of job search parameters, small cities have a distinctive mechanism. In particular, blue-collar and manufacturing small cities are characterized by lower separation rates. The positive interplay between population density and separation rates in the group of small cities may be explained by the lack of heterogeneity in this subset of cities.

Table 9: Explaining the primitives of local labor markets

	Panel 1: All cities			Panel 2: Large cities			Panel 3: Mid-sized cities			Panel 4: Small cities		
	λ^u	λ^e	δ	λ^u	λ^e	δ	λ^u	λ^e	δ	λ^u	λ^e	δ
(Intercept)	−0.237	2.894	1.588	−2.600	−0.058	4.623	1.044	4.146	1.364	−1.354	1.522	−0.547
	(1.427)	(1.617)	(0.837)	(5.683)	(4.390)	(2.324)	(2.785)	(2.758)	(1.701)	(1.641)	(2.381)	(0.976)
Number of firms	0.837***	0.101***	−0.024	0.803***	0.059*	−0.027*	2.194**	0.752	−0.442	−1.379	3.222	1.495*
	(0.024)	(0.027)	(0.014)	(0.030)	(0.023)	(0.012)	(0.711)	(0.704)	(0.435)	(1.155)	(1.675)	(0.686)
Density	0.041	0.229	0.125	0.323	0.281	0.030	−0.072	0.227	0.116	0.172	0.061	0.659***
	(0.112)	(0.127)	(0.066)	(0.213)	(0.165)	(0.087)	(0.186)	(0.184)	(0.114)	(0.225)	(0.326)	(0.134)
Young	2.892***	3.512***	−1.054*	5.979**	−0.704	−0.696	1.661	1.286	−2.068	−2.805	3.416	−2.826*
	(0.852)	(0.966)	(0.500)	(2.036)	(1.573)	(0.833)	(1.925)	(1.906)	(1.176)	(1.812)	(2.628)	(1.077)
Males	0.782	−4.129	−1.542	5.280	7.313	−7.312	−3.807	−7.933	0.159	4.510	−2.889	2.279
	(3.159)	(3.578)	(1.852)	(12.414)	(9.589)	(5.076)	(6.274)	(6.213)	(3.833)	(3.542)	(5.137)	(2.105)
Drop out	0.567	−0.675	0.046	−0.127	−0.424	−0.569	1.112	−0.707	0.190	0.152	0.504	0.128
	(0.443)	(0.502)	(0.260)	(1.360)	(1.050)	(0.556)	(0.827)	(0.819)	(0.505)	(0.606)	(0.878)	(0.360)
Blue collar	−0.723	−2.504**	−0.568	−0.379	−7.912***	−0.815	1.729	0.438	−2.388*	1.060	−1.950	0.970
	(0.762)	(0.863)	(0.447)	(2.530)	(1.954)	(1.034)	(1.586)	(1.571)	(0.969)	(1.076)	(1.561)	(0.640)
Manufacturing	0.444	0.507	−0.654*	−1.958	−0.052	−0.535	−0.036	0.300	0.187	0.496	0.260	−1.199***
	(0.439)	(0.497)	(0.257)	(1.684)	(1.301)	(0.689)	(0.759)	(0.752)	(0.464)	(0.529)	(0.768)	(0.315)
R^2	0.896	0.340	0.182	0.970	0.664	0.329	0.190	0.124	0.284	0.113	0.078	0.385
Num. obs.	199	199	199	40	40	40	60	60	60	99	99	99

Notes: (i) Ordinary-least-square regressions of the structural parameters. The dependent variable is the estimated parameter; (ii) Standard errors in Parentheses; Significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, · $p < 0.1$; (ii) Young, Males, Drop outs, Blue Collar and Manufacturing are shares reported to total population or total number of jobs; Young refers to people under 30; Manufacturing refers to all manufacturing jobs; (iii) The parameters for the city 200 are not included in panels 1 and 4

4.2 The impact of distance on spatial frictions

The observation of our results for the spatial friction parameters yields two conclusions. First, relative to the internal job arrival rates λ_j^i, the job arrival rate from other locations $\sum_{k\in\mathscr{J}_j} s_{jk}^i\lambda_k^i$ is rather high, which gives support in favor of our modelling choice to take between-city mobilities into account. For instance, the median value of this rate for unemployed jobseekers is close to half of the median value of the λ_j^u.[32] Second, there is substantial heterogeneity within the three previous groups of cities regarding their level of connection to the other cities in the system. This heterogeneity cannot be captured in a three-type model à-la Baum-Snow & Pavan (2012). For instance, the second and third cities, Lyon (L) and Marseille (M) differ substantially. The internal job prospects for unemployed agents is substantially higher in Lyon($\lambda_L^u = 2.6$ and $\lambda_M^u = 1.9$). On the other hand, the external job prospects are much higher in Marseille, $\sum_{k\in\mathscr{J}_M} s_{Mk}^u\lambda_k^u = 0.9$, than in Lyon, $\sum_{k\in\mathscr{J}_L} s_{Lk}^u\lambda_k^u = 0.01$. A similar pattern can be observed in the other two groups of cities. In Brive-la-gaillarde, for example, the city with the lowest off-the-job arrival rate (0.189), spatial mobility opportunities accrue at a rate of 1.6, when Cholet, the city just above in the city size distribution (ranks 85 and 84, respectively), faces a rate of 0.69.[33]

This within-group heterogeneity is due to the impact of the other characteristics of the cities, besides size, and in particular their location and their level of specialization. As explained in Section 3.3, these two dimensions are measured in relative terms, by the spatial distance and the sectoral dissimilarity between each pair of cities.[34] Given our specification, we can recover the estimated impact of on the level of spatial frictions between each pair of city. It is given by the first-order conditions on equation 22 with respect to d_{jl} and h_{jl}. Because of the city fixed effect in spatial frictions, the effect of distance is not uniform. Table 10 reports the distribution of these marginal effects for all city pairs $(j,l)\in\mathscr{J}\times\mathscr{J}_j$ and for the city pairs (j,Paris). All the estimates of $\{s_1^i, s_2^i, s_3^i, s_4^i\}_{i=e,u}$ that enter in the specification of the spatial frictions are significantly different from zero.[35] Both physical distance and sectoral dissimilarity increase the level of spatial frictions as expected. Moreover, the effect is much stronger for employed agents. This is easy to understand for sectoral dissimilarity, since a large share of job-to-job transitions take place within the same sector. The differential impact of distance is a little less straightforward. It may be due to the fact that unemployed

[32] Note that the weighted average of this rate predicts an annual between-city mobility rate of 0.15 for the unemployed population, which closely matches the between-municipality rate measured in the Labor Force Surveys (see Figure 1)

[33] In light of these observations, one may wonder if there is a substitution effect between local and outside offers. A correlation test suggests that such a substitution effect does holds for the small and mid-sized cities.

[34] The correlation between these two measures is positive and significant at the 5% confidence level. Only 5% of the pairs of cities are in the first quartile of spatial distance and the last quartile of sectoral dissimilarity, and 5% are in the reverse situation. One notable feature is that a stronger sectoral similarity between the largest cities often partially compensates for the distance between them. For instance, the distance between Nice and Nantes (respectively, the seventh and the eighth city) is at the 95th percentile of the distance matrix and their level of sectoral dissimilarity corresponds to the first percentile of the dissimilarity matrix.

[35] Standard errors are available upon request.

jobseekers are more often linked with more formal matchmakers, such as unemployment agencies, which may have information regarding employment opportunities all over the country, while employed jobseekers have to rely more on unofficial networks which are more sensitive to distance.

Table 10: The effect of distance of spatial frictions

	Panel 1: All city pairs				Panel 2: City pairs to Paris (P)			
	$\frac{\partial s^u_{jl}}{\partial d_{jl}}$	$\frac{\partial s^e_{jl}}{\partial d_{jl}}$	$\frac{\partial s^u_{jl}}{\partial h_{jl}}$	$\frac{\partial s^e_{jl}}{\partial h_{jl}}$	$\frac{\partial s^u_{jP}}{\partial d_{jP}}$	$\frac{\partial s^e_{jP}}{\partial d_{jP}}$	$\frac{\partial s^u_{jP}}{\partial h_{jP}}$	$\frac{\partial s^e_{jP}}{\partial h_{jP}}$
Min	-0.1475	-2.6985	-0.1824	-2.5095	-0.0791	-1.4863	-0.1201	-1.8084
1st Qu.	-0.0099	-0.1200	-0.0185	-0.1452	-0.0061	-0.1061	-0.0125	-0.1218
Median	-0.0015	-0.0001	-0.0028	-0.0001	-0.0012	-0.0001	-0.0022	-0.0001
Mean	-0.0074	-0.1561	-0.0133	-0.1711	-0.0058	-0.1054	-0.0107	-0.1319
Sd	0.0122	0.3453	0.0207	0.3651	0.0108	0.2357	0.0185	0.2921
3rd	-0.0002	-0.0000	-0.0003	-0.0000	-0.0001	-0.0000	-0.0003	-0.0000
Max	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
N	39601	39601	39601	39601	199	199	199	199

4.3 Decomposition: the city size wage premium

Section 4.1 has shown that cities of different size exhibit very different features with respect to their internal job matching process. Section 4.2 has documented a substantial heterogeneity with respect to the level of openness of each city. This heterogeneity is somewhat transversal to size groups and pertains to other dimensions, such as the level of centrality of each city. We now quantify the respective impact of these two dimensions on the average wage level in each city. In our controlled environment, city-specific wage dispersion can be expressed as a function of the labor market primitives. The expected wage in city j is $\mathbb{E}(w|j) = \int_{\underline{w}}^{\overline{w}} x g_j(x) dx$, where $g_j(\cdot)$ is given by equation 15:

$$\begin{aligned}
\mathbb{E}w|j &= \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\urcorner_j(x)\left[\lambda^u_j u_j\right]\right)dx}_{\text{Local off-the job search}} + \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\urcorner_j(x)\left[\lambda^e_j (m_j-u_j)G_j(x)\right]\right)dx}_{\text{Local on-the-job search}} \\
&+ \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\urcorner_j(x)\left[\lambda^u_j \sum_{k\in\mathscr{J}_j} s^u_{kj}\psi_{kj}(x)u_k\right]\right)dx}_{\text{Off-the-job search between cities}} + \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\urcorner_j(x)\left[\lambda^e_j \sum_{k\in\mathscr{J}_j} s^e_{kj}(m_k-u_k)G_k(q_{jk}(x))\right]\right)dx}_{\text{On-the-job search between cities}}
\end{aligned} \tag{23}$$

where $\urcorner_j(w) = \dfrac{f_j(w)}{(m_j-u_j)\left[\delta_j + \lambda^e_j\overline{F}_j(w) + \sum_{k\in\mathscr{J}_j} s^e_{jk}\lambda^e_k\overline{F}_k(q_{jk}(w))\right]}$.

Table 11 reports the decomposition described in equation 23 for a subset of large, mid-sized and small cities. Whereas most of the local wage level can be imputed to local on-the-job search in the larger cities, local exit from unemployment is the main driver in the smaller cities. This finding confirms the hypothesis that stemmed from the observation of the correlations displayed in Table 8. Differences are extreme between Paris, where 99% of the local wage level is explained by on-the-job search, and Tulle or Dinard, where 99% of the local wage level is driven by off-the-job search. In addition, there is a large within-group variability in the role of mobility: from 16% in Marseille or Thann to about 0% in Nice or Tournon. Note that in a few cities, mostly in the mid-sized group, wages are largely determined by mobile jobseekers, to the extent of 43% in Bourg-en-Bresse or even the extreme 95% in Tarbes.[36]

Table 11: City-level wage decomposition

Panel 1: largest cities

	City 1 Paris	City 2 Lyon	City 3 Marseille	City 4 Toulouse	City 5 Lille	City 6 Bordeaux	City 7 Nice
Local UtoJ	0.52	5.82	5.10	4.31	5.64	17.40	17.54
Local JtoJ	95.50	93.64	78.61	82.36	80.08	82.02	82.00
Mobile UtoJ	0.49	0.02	1.19	0.96	1.32	0.06	0.06
Mobile JtoJ	3.49	0.51	15.09	12.36	12.95	0.52	0.40

Panel 2: mid-sized cities

	City 71 Bourg-en-Bresse	City 72 Tarbes	City 73 Belfort	City 74 St-Quentin	City 75 La Roche	City 76 Vienne	City 77 Évreux
Local UtoJ	37.34	4.13	29.78	63.99	53.23	53.70	50.68
Local JtoJ	19.35	1.16	69.82	35.99	40.47	23.45	21.40
Mobile UtoJ	27.89	67.66	0.40	0.01	3.99	11.48	14.79
Mobile JtoJ	15.42	27.05	0.00	0.00	2.31	11.37	13.12

Panel 3: smallest cities

	City 191 Tulle	City 192 Thann	City 193 Dinard	City 194 Tournon	City 195 Sable	City 196 Pontarlier	City 197 St-Gaudens
Local UtoJ	98.64	60.42	99.73	96.71	87.49	89.49	89.84
Local JtoJ	1.04	23.99	0.21	3.22	4.51	5.99	8.79
Mobile UtoJ	0.32	9.27	0.06	0.06	7.40	4.18	1.23
Mobile JtoJ	0.01	6.32	0.00	0.00	0.60	0.35	0.14

Notes: (i) Local UtoJ is local off-the-job search, Local JtoJ is local on-the-job search, Mobile UtoJ is between-cities off-the-job search and Mobile JtoJ is between-cities on-the-job search

4.4 An urban engineering experiment

In this section, we investigate the optimal design of a country for minimizing unemployment. While section 4.1 has shown that observed characteristics only explain a small fraction of some of the labor market primitives, especially the job separation rates, Table 15 in Appendix D demonstrates that a simple func-

[36] This last example refers to an isolated city in the Pyrénées mountains. In Gap, a similar city located in the Alps, mobility accounts for 83%. These pieces of evidence suggest a "mountain effect".

tion of the log number of firms, log population and log area of the location can explain most of the variation in the off-the-job arrival rates ($R^2 = 0.68$), on-the-job arrival rates ($R^2 = 0.87$) and the separation rates ($R^2 = 0.90$). It can also explain a sizeable part of the location fixed effects s^i_{0l}. Although this specification does not have any clear economic interpretation, it can be used to simulate a counterfactual experiment that does not rely on an independence assumption of the parameters.[37]

Using the parameters of this regression, we analyze whether there is an optimal country structure, keeping cities' location and relative size fixed.[38] Figure 6 reports the relationship between the number of cities and aggregate unemployment. It is not trivial. The smallest unemployment rate, 5.6%, is obtained with a single city established in the current geographical setting of Paris. Starting to increment the number of cities, one early nightmare scenario consists of a three-city country, with Paris, Lyon and Marseille. Under this scenario, the effect of physical distance impacts both spatial frictions and unemployment. As we add more cities, and smaller cities start to fill the vacant space between the largest cities, the unemployment rate decreases, reaching 6.5% with 28 cities. After this threshold, the relation between the number of cities and unemployment is unambiguous. That is, as the number of cities increases, local labor markets with low job arrival rates emerge. In addition, the spatial frictions are strengthened by the increasing share of unattractive locations, and the stiff competition for the most attractive ones.

5 Extension: local amenities and moving costs

So far we have focused on the equilibrium properties of a job search model featuring spatial frictions between heterogeneous local labor markets. In order to model a more realistic mobility process within a system of cities, one has to incorporate two additional dimensions: the non-labor value of each city measured by city-specific amenities, and the spatial dimension of mobility. Local amenities and moving costs are key in any attempt to model jointly of the housing market and the labor market in this multi-dimensional framework.

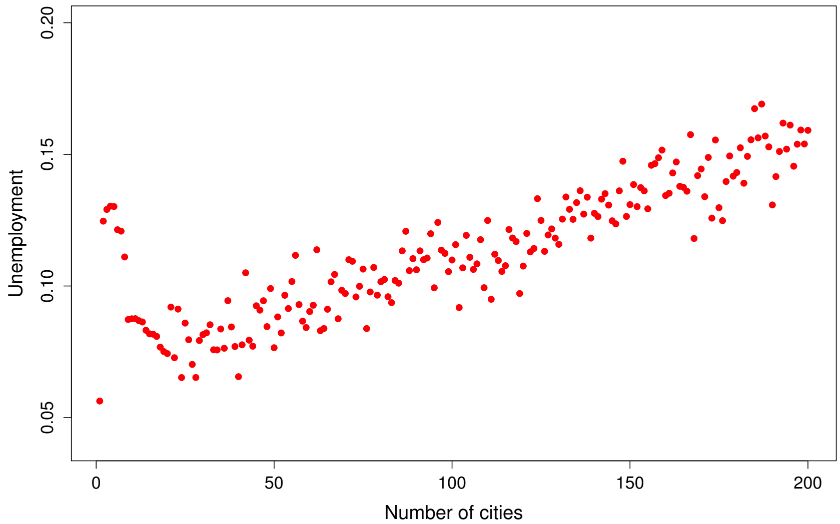
The identification of both local amenities and mobility cost in the utility of agents is highly controversial in the literature. In the location choice literature, due to the scarcity of identification sources, the mobility cost parameter is calibrated to 3-10% of the housing price.[39] Regarding local amenities, recent hedonic

[37] In this regression, we need to include the parameters of the residual city 200.

[38] The experiment is the following: consider two cities 1 and 2, defined by $f(1)$ and $f(2)$ firms, $p(1)$ and $p(2)$ inhabitants and of respective areas $a(1)$ and $a(2)$. Assume city 2 is the smaller, such that its rank in the city distribution is $rk_2 > rk_1$. Any conterfactual system of at least rk_2 cities is made of two fictional cities I and II, which take the location of cities 1 and 2 and which respect the following constraints: $f(I)/f(II) = f(1)/f(2)$, $pop(I)/pop(II) = pop(1)/pop(2)$, $a(I) = a(1)$ and $a(II) = a(2)$. The level of $(f(I), f(II), pop(I), pop(II))$ is determined by the number of cities considered in the experiment.

[39] See for example, Bayer, McMillan, Murphy & Timmins (2011).

Figure 6: Unemployment vs Number of cities



Notes: (i) Aggregate unemployment rate as a function of the number of cities, taking city location and relative size fixed; (ii) We use the values of the labor market primitives and fixed effects predicted by the estimation results provided in Table 15.

literature, based on very detailed datasets and spatial discontinuities has made enormous progress towards estimates of the willingness to pay for location characteristics (Bayer, Ferreira & McMillan, 2007). In our setup, once controlled for the location specific effect as measured by the bundle $(\lambda^u, \lambda^e, \delta, F)$, the mobility cost will be captured by the variation in the wages accepted by individuals after a mobility between each given pair of cities. Similarly, given a set of labor market primitives, the differential between local amenities in two locations will be captured by the variation in the average wage accepted by agents coming from elsewhere.

5.1 Empirical puzzles

The baseline model generates a complete ordering of the cities. This ordering is determined by the primitives of each local labor market and by each city's relative *centrality*. If an agent in location j is willing to accept a wage cut to move to location l, then an agent in location l will only accept a wage increase to move to location j. However, as shown in Table 12, this symmetry assumption does not hold. In almost all job-to-job transitions between and within the five largest French metropolitan areas, the average post-transition wage is higher after a transition with mobility than after a transition without mobility.[40] This feature suggests the existence of mobility costs.

Moreover, a close analysis of Table 12, in light of the results presented in section 4, suggests a wage

[40] It should be noted that this pattern does not preclude the existence of mobility strategy with wage cut. There are numerous cases in the full data where agents do accept lower wages in between-cities on-the-job search than in within-cities on-the-job search. Between-cities on-the-job search with wage cut strategy involves mainly young agents.

Table 12: Identification of the mobility cost: average wages following a job-to-job transition

	Paris	Lyon	Marseille	Toulouse	Lille
Paris	38,576	47,824	43,800	36,470	36,406
	(36,739)	(351)	(221)	(166)	(164)
Lyon	44,602	30,234	33,358	31,607	45,155
	(338)	(3,945)	(56)	(17)	(24)
Marseille	45,981	37,778	27,983	43,255	46,776
	(217)	(62)	(2,162)	(15)	(12)
Toulouse	36,926	37,196	34,720	28,454	41,579
	(147)	(19)	(18)	(2,386)	(4)
Lille	41,139	42,253	40,504	29,346	27,753
	(177)	(23)	(8)	(12)	(2,466)

Notes: (i) Average final wage after a job-to-job transition, by city of origin (in line) and city of destination (in column); (ii) In parentheses: the number of observations. *Source: Panel DADS 2002-2007*

setting mechanism that cannot be fully rationalized by the existence of mobility costs. For example, even if Paris offers many more opportunities than Toulouse, agents who are leaving Lille require a higher wage in Paris (average earnings of €41,139) than in Toulouse (average earnings of €29,346). Since Paris is about four times closer to Lille than Toulouse, the addition of mobility costs alone cannot cope with this simple observation, unless we allow for heterogeneous local amenities.

5.2 Theory

Cost of living differs substantially across cities and the correlation between housing prices and the size of cities contributes to the city size premium (Combes, Duranton & Gobillon, 2012). Although our model features homogenous agents, a system of heterogeneous cities will deliver a partition with positive population, as long as prices compensate for local amenities, as in Epple & Sieg (1999). In order to keep the framework tractable, we now assume that there exists an indirect utility γ_j that summarizes the difference between local amenities and housing prices in city j, in the spirit of Bayer et al. (2007). This value is separable from the level of earnings, such that the instant value of an unemployed (resp., employed at wage w) agent living in city j is now equal to $b+\gamma_j$ (resp, $w+\gamma_j$). As shown in equation 26, the introduction of local amenities adds a direct term $\gamma_j - \gamma_l$ into the mobility-compatible indifference wage that captures the relative *attractivity* of city j compared to city l: job offers coming from a less endowed city can only attract jobseekers if they are high enough to overcome the difference in local amenities. Conversely, the introduction of local amenities may shrink the acceptable wages of the less attractive locations, because of lower housing prices.

This feature adds another degree of segmentation to our framework, where not only the option value of unemployment is location-specific, but the instant value of unemployment is as well.

Allowing for moving costs adds a substantial layer of complexity to the problem. Let c_{jl} denote the lump-sum mobility cost associated with a transition from city j to city l. The value functions of the problem now let agents who have received a job offer x in city $k \neq j$ maximize between $V_k^e(x) - c_{jk}$ and their current value in city j. The problem of an agent is summarized by:

$$rV_j^u = b + \gamma_j + \lambda_j^u \int_{\underline{w}}^{\overline{w}} \max\left\{V_j^e(x) - V_j^u, 0\right\} dF_j(x) + \sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \int_{\underline{w}}^{\overline{w}} \max\left\{V_k^e(x) - c_{jk} - V_j^u, 0\right\} dF_k(x) \tag{24}$$

$$\begin{aligned} rV_j^e(w) &= w + \gamma_j + \lambda_j^e \int_{\underline{w}}^{\overline{w}} \max\left\{V_j^e(x) - V_j^e(w), 0\right\} dF_j(x) + \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \max\left\{V_k^e(x) - c_{jk} - V_j^e(w), 0\right\} dF_k(x) \\ &+ \delta_j [V_j^u - V_j^e(w)] \end{aligned} \tag{25}$$

The main differences are captured in the new expressions for the mobility-compatible indifference wage $q_{jl}^{ac}(w)$ (see Appendix A.3 for details):

$$\begin{aligned} q_{jl}^{ac}(w) &= w + \gamma_j - \gamma_l + c_{jl} + [\delta_j V_j^u - \delta_l V_l^u] + (1 - s_{lj}^e)\lambda_j^e \int_w^{\overline{w}} \Xi_j(x) dx - (1 - s_{jl}^e)\lambda_l^e \int_{q_{jl}^{ac}(w)}^{\overline{w}} \Xi_l(x) dx \\ &+ \sum_{k \in \mathscr{J}_{jl}} (s_{jk}^e - s_{lk}^e)\lambda_k^e \int_{q_{jk}^{ac}(w)}^{\overline{w}} \Xi_k(x) dx - \sum_{k \in \mathscr{J}_{jl}} \lambda_k^e \Big(s_{jk}^e \overline{F}_k(q_{jk}^{ac}(w)) c_{jk} - s_{lk}^e \overline{F}_k(q_{lk}^{ac}(w)) c_{lk} \Big) \end{aligned} \tag{26}$$

where $\Xi_j(w) = \dfrac{\overline{F}_j(w)\Big(1 + \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e c_{jk} f_k(q_{jk}^{ac}(w)) dq_{jk}^{ac}(w)\Big)}{r + \delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}^{ac}(w))}$.

The introduction of mobility costs adds a direct term c_{jl} into the mobility-compatible indifference wage, which simply states that job offers can only attract jobseekers from elsewhere if they are high enough to overcome the mobility costs. However, there also is a dynamic effect, shown in the final term of equation **??**. This last term captures the relative *accessibility* of locations j and l. That is, 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 city l to any tier location k that she may face in the future. If, for instance, the mobility cost from city j to a city k, weighted by the probability of realization of an attractive offer, is higher than its counterpart from city l, the mobility cost component will reduce the agent's indifference wage $q_{jl}^{ac}(w)$. Note that, contrary to the direct effect, the sign of this dynamic effect may revert along the wage distribution. Finally, the introduction of mobility costs adds another effect through the value of on-the-job search, with $\Phi(\cdot)$ becoming $\Xi(\cdot)$. This effect is a direct consequence of the absence of ordering between cities. The curvature of the value of working in each location j no longer only features the weighted probability of drawing a higher wage, but also the product between this probability and the variation in the potential cost generated by between-cities on-the-job search.

The difference between the models without and with local amenities and mobility costs is summarized by the empirical content of the indifference wages q_{jl} and q_{jl}^{ac}, for the instant and the option values. Whereas the difference in instant values is straightforward, the difference in option values is slightly more convoluted. In the baseline model, the term $\sum_{k\in\mathscr{J}_j}(s_{jk}^e - s_{lk}^e)\lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} \Phi_k(x)dx$ captures the fact that individuals seek to maximize the value of the next offer. The additional term due to mobility costs, $\sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e [\overline{F}_k(q_{jk}^{ac}(w))c_{jk} - \overline{F}_k(q_{lk}^{ac}(w))c_{lk}]$, states that individuals also seek to minimize the cost of their next move. It is not clear whether such a generality is required to capture the mobility behaviour of agents. Our empirical analysis will provide insights into the relative importance of each of these two components.

5.3 Estimation

The estimation of local amenities parameters does not raise additional difficulties. In contrast, the introduction of mobility costs deepens the relationship between the indifference wage and its derivatives. Each derivative is now given by:

$$dq_{jl}^{ac}(w) = \frac{1-(1-s_{lj}^e)\lambda_j^e \Xi_j(w) - \sum_{k\in\mathscr{J}_{jl}} \lambda_k^e \Big[(s_{jk}^e - s_{lk}^e)\Xi_k(q_{jk}^{ac}(w))dq_{jk}^{ac}(w) + \big(s_{jk}^e c_{jk} dq_{jk}^{ac}(w) f_k(q_{jk}^{ac}(w)) - s_{lk}^e c_{lk} dq_{lk}^{ac}(w) f_k(q_{lk}^{ac}(w))\big)\Big]}{1-(1-s_{jl}^e)\lambda_l^e \Xi_l(q_{jl}^{ac}(w))} \tag{27}$$

In order to proceed with the same algorithm as before, one only needs to make a similar (maybe stronger) assumption regarding the independence of the derivates of the indifference wages (see Appendix A.3 for details).

While spatial friction parameters are identified from transition rates between pairs of cities, mobility costs and local amenities parameters are identified from the average wages accepted by incoming agents. γ_l is identified by the average wage accepted by agents who have experienced a job to job mobility with mobility that led them into city l. As for c_{jl}, it is identified by the average wage accepted by agents who have experienced a job to job mobility from city j to city l. To be more specific, let $w_{e_je_j}^{init}$ denote the average initial wage of agents employed in city j and who will experience a job-to-job transition within city j. Using the fact that $q_{jl}(\cdot)$ is a function, we consider as a theoretical moment, the difference between $q_{jl}(w_{e_je_j}^{init})$ and $w_{e_je_j}^{init}$ (which is equal to $q_{jj}(w_{e_je_j}^{init})$). Under the assumption that, conditional on the local parameter values, jobseekers are as likely to draw a wage above their indifference wage, no matter the type of job-to-job transition, this difference identifies the mobility cost c_{jl}. The corresponding empirical moment is the difference between $w_{e_je_l}^{fin}$ (the average wage after a job-to-job transition from city j to city l), and $w_{e_je_j}^{fin}$ (the average wage after a job-to-job transition within city j).[41]

We parametrize the mobility cost as a simple quadratic function of physical distance:

$$c_{jl} = c_0 + c_1 d_{jl} + c_2 d_{jl}^2 \tag{28}$$

Unlike with spatial frictions parameters, we do not impose any functional form to the value of moving c_{jl}, allowing possible negative mobility cost. One possible interpretation of negative moving costs would be relocation subsidies (see Kennan & Walker (2011) for an example of the coexistence of both highly positive and highly negative moving costs). Finally, note that we assume that moving costs do not vary with labor market status. For the identification of mobility cost, we select a subset of city pairs $\mathscr{W} \subset \mathscr{J} \times \mathscr{J}_j$. We use all the average accepted wages between Marseille, Nice and Lille, so that $|\mathscr{W}| = 6$.[42]

5.4 Results

Our estimates for equation 28 indicate that the mobility cost function is given by $\widehat{c}_{jl} = 9.018 + 27.000 d_{jl} - 2.998 d_{jl}^2$, with distance measured in 10^5 km. This function is positive and increasing for all possible values of d, which means that contrary to Kennan & Walker (2011), we do not find any evidence of negative mobility costs (or relocation subsidies) in the French labor market. Given that log wages are used in the estimation, the monetary equivalent of this cost function amounts to an average value of €9,360, which is approximately equal to the annual minimum wage.[43] As shown in Table 13, the physical distance between locations generates a lot of variability in this cost, up to a 40% distance penalty for the most peripheral locations.

These estimates are substantially lower than the mobility cost found by Kennan & Walker (2011), who estimate a cost of $312,000 for the average mover. We believe that the introduction of spatial frictions allows us to obtain such a result. That is, the low mobility rate is not rationalized by extremely high mobility costs but rather, by the existence of spatial frictions. As a consequence, our identification of mobility costs, which relies on the spatial variation in accepted wages, is less affected by other imperfections.

[41] In Table 12, this corresponds to the difference between the value in the off-diagonal cases and the value in the diagonal case on the same line.

[42] See Figure 8 in appendix C for details. Six moments seem enough to identify (c_0, c_1, c_2). Our choice was driven by the necessity to include both a pair of cities that were physically close to each other and a third one which is more remote. In addition, the number of pair-specific job-to-job transitions needed to be high enough for the average accepted wage to be meaningful. In effect, there are 12 transitions between Marseille and Lille, 8 between Lille and Marseille, 6 between Lille and Nice, 5 between Nice and Lille, 173 between Nice and Marseille and 12 between Marseille and Nice.

[43] Although, this cost is still high, one must bear in mind that mobility costs also encompass two other features: first, they include relocation costs, and particularly transaction costs on the housing market. Such costs may be high, especially for homeowners. Second, mobility costs, in all generality, must include a measure of psychological costs. Even if those are difficult to quantify, they are likely to be substantial. These two features explain why the fixed component accounts for a sizable share of the mobility cost.

Table 13: Distribution of the mobility costs involving all cities or one of the eight first cities

	All	Paris	Lyon	Marseille	Toulouse	Lille	Bordeaux	Nice	Nantes
Min.	8250	8349	8310	8344	8351	8307	8346	8335	8381
1st Qu.	8920	8791	8812	9037	9044	9016	9043	9170	8990
Median	9330	9177	9150	9556	9465	9409	9448	9700	9503
Mean	9363	9139	9191	9532	9409	9379	9446	9715	9485
3rd Qu.	9746	9425	9524	9979	9804	9765	9851	10210	9935
Max.	11650	10000	10740	11140	10320	10380	10610	11530	10700
N	39601	199	199	199	199	199	199	199	199

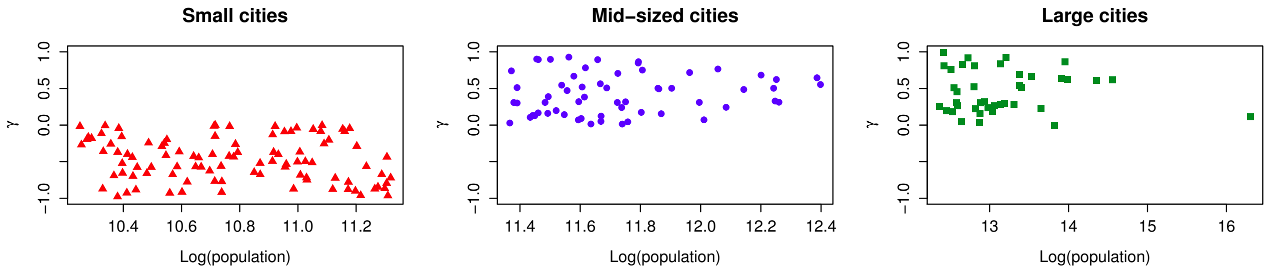
The estimated values of the (net-of-cost) amenities parameters γ display several interesting patterns. They are almost uniformly distributed on the $[-1,1]$ segment. As shown in Figure 7, large and mid-sized cities benefit from positive amenities, despite their higher housing prices, whereas the opposite is true for small cities. At the median value for the group of large and mid-sized cities, local amenities increase the instant valuation of a worker employed at the minimum wage by one half. In small cities, they reduce the instant valuation of the same worker by one third. The partition between these two groups of 100 cities is almost perfect, since only one small city (Saumur) has a positive (yet almost zero) parameter, and no mid-sized or large city has a negative parameter. In addition to the results presented in Section 4, we see this partition as another striking piece of evidence in favor of the separate study of our three groups of cities. However, there is also a lot of heterogeneity within these groups. For example, 11 mid-sized or large cities and 16 small cities are characterized by very low values of $|\gamma|$ (<0.1), whereas, on the other hand, 5 mid-sized cities and 7 small cities are characterized by extreme values, below -0.9 or above 0.9. Seven out of the ten largest cities have a parameter above the 80% percentile. The exceptions are Nantes, Nice and Paris. The low value for Paris (0.112), which probably reflects high housing prices, may provide an explanation for the second empirical puzzle described in section 5.1.[44]

Conclusion

In this paper, we propose a job search model to study persistent inequalities across local labor markets. Using data from a French matched employer-employee dataset, we recover the local determinants of job creation and job destruction that rationalize both unemployment and wage differentials across cities. In contrast to the former literature, we can generate much more wage dispersion using a search framework. A

[44] One might be tempted to interpret the extremely low value for Nice ($< 10^{-3}$) as a "Riviera effect": Nice is the largest city of the French Sunbelt, where housing properties are driven upwards by the continuous inflow of retirees.

Figure 7: The amenities parameters along the city size distribution



Notes: (i) Estimated values of the amenity parameters γ for the 40 largest cities, the 100 smallest cities and the 60 cities in between; (ii) The city 200 is made of all remaining metropolitan areas. It is included in the estimation but its parameters are not meaningful and therefore are not represented here.

decomposition of wages show that the differential in on-the-job search rate can explain most of the city size wage premium, without resorting to a differential in the return to skills. We overcome the main limitations of Kennan & Walker (2011): estimating a model at city level with spatial frictions. The random search technology avoids the state space complications, and significantly reduces the computational burden. Counterfactual simulations using the matching function show a non-monotonic relationship between aggregate unemployment and the number of cities.

Despite its many achievements, our model has several limitations. First, a more precise decomposition of the city size wage premium is called for, that would incorporate the possibility of on-the-job wage bargaining à-la Cahuc et al. (2006) as a third dimension, in addition to city-specific on-the-job search and openness. Second, this paper leaves unexplored the firms' side of the dynamic location model. A mere extension à-la Meghir et al. (2012) would be trivial and we believe that the interest of a spatial steady state distribution of firms is limited without an explicit theory of location choice, agglomeration economies and wages. We leave these extensions for future work. Another interesting venue for research would allow unemployed workers to move across locations. Relaxing the immobility assumption for unemployed workers is not easily compatible with the equilibrium assumptions of the model. One should think of a price mechanism that would counterbalance the labor market mechanisms in order to maintain an empirically reasonable scattering of the unemployed population. Another way of dealing with this issue could also be to consider that individuals have heterogeneous preferences over local amenities, which would generate sorting.

References

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–334.

Abowd, J. M., Kramarz, F., & Roux, S. (2006). Wages, mobility and firm performance: Advantages and insights from using matched worker-firm data. *Economic Journal*, *116*(512), F245–F285.

Bagger, J., Fontaine, F., Postel-Vinay, F., & Robin, J.-M. (2013). Tenure, experience, human capital and wages: A tractable equilibrium search model of wage dynamics. Sciences po publications, Sciences Po.

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

Bayer, P., Ferreira, F., & McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, *115*(4), 588–638.

Bayer, P., McMillan, R., Murphy, A., & Timmins, C. (2011). A dynamic model of demand for houses and neighborhoods. NBER Working Papers 17250, National Bureau of Economic Research, Inc.

Blanchard, O. J. & Katz, L. F. (1992). Regional evolutions. *Brookings Papers on Economic Activity*, *23*(1), 1–76.

Bontemps, C., Robin, J.-M., & Van den Berg, G. J. (1999). An empirical equilibrium job search model with search on the job and heterogeneous workers and firms. *International Economic Review*, *40*(4), 1039–74.

Brueckner, J. K. & Zenou, Y. (2003). Space and unemployment: The labor-market effects of spatial mismatch. *Journal of Labor Economics*, *21*(1), 242–241.

Burda, M. C. & Profit, S. (1996). Matching across space: evidence on mobility in the czech republic. *Labour Economics*, *3*(3), 255–278.

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.

Combes, P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, *80*(6), 2543–2594.

Combes, P.-P., Duranton, G., & Gobillon, L. (2012). The costs of agglomeration: Land prices in french cities. IZA Discussion Papers 7027, Institute for the Study of Labor (IZA).

Combes, P.-P. & Lafourcade, M. (2005). Transport costs: measures, determinants, and regional policy implications for france. *Journal of Economic Geography*, *5*(3), 319–349.

Coulson, N. E., Laing, D., & Wang, P. (2001). Spatial mismatch in search equilibrium. *Journal of Labor Economics*, *19*(4), 949–72.

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

De la Roca, J. & Puga, D. (2012). Learning by working in big cities. CEPR Discussion Papers 9243, C.E.P.R. Discussion Papers.

Dey, M. S. & Flinn, C. J. (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica*, *73*(2), 571–627.

Elhorst, J. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of Economic Surveys*, *17*(5), 709–748.

Epple, D. & Sieg, H. (1999). Estimating equilibrium models of local jurisdictions. *The Journal of Political Economy*, *107*(4), 645–681.

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. (2006). Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates. *Econometrica*, *74*(4), 1013–1062.

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.

Gould, E. D. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *Review of Economic Studies*, *74*(2), 477–506.

Gourieroux, C., Monfort, A., & Renault, E. (1993). Indirect inference. *Journal of Applied Econometrics*, *8*(S), S85–118.

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

Harris, J. R. & Todaro, M. P. (1970). Migration, unemployment & development: A two-sector analysis. *American Economic Review*, *60*(1), 126–42.

Head, A. & Lloyd-Ellis, H. (2012). Housing liquidity, mobility, and the labour market. *Review of Economic Studies*, *79*(4), 1559–1589.

Head, K. & Mayer, T. (2013). Gravity equations: Workhorse, toolkit, and cookbook. CEPR Discussion Papers 9322, C.E.P.R. Discussion Papers.

IGF (2007). Rapport sur les méthodes statistiques d'évaluation du chômage. La Documentation Française 2007-M-066-01, Inspection Générale des Finances.

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.

Judd, K. (1998). *Numerical Methods in Economics*. MIT Press.

Kaplan, G. & Schulhofer-Wohl, S. (2012). Understanding the long-run decline in interstate migration. Working Paper 18507, National Bureau of Economic Research.

Karahan, F. & Rhee, S. (2013). Geographical reallocation and unemployment during the great recession: the role of the housing bust. Staff Reports 605, Federal Reserve Bank of New York.

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.

Lentz, R. (2010). Sorting by search intensity. *Journal of Economic Theory*, *145*(4), 1436–1452.

Lucas, R. E. (1993). Internal migration in developing countries. In M. R. Rosenzweig & O. Stark (Eds.), *Handbook of Population and Family Economics*, volume 1 of *Handbook of Population and Family Economics* chapter 13, (pp. 721–798). Elsevier.

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

Manning, A. & Petrongolo, B. (2011). How local are labor markets? evidence from a spatial job search model. IZA Discussion Papers 6178, Institute for the Study of Labor (IZA).

Meghir, C., Narita, R., & Robin, J.-M. (2012). Wages and informality in developing countries.

Menzio, G. & Shi, S. (2010). Block recursive equilibria for stochastic models of search on the job. *Journal of Economic Theory*, *145*(4), 1453–1494.

Moretti, E. (2011). Chapter 14 - local labor markets. volume 4, Part B of *Handbook of Labor Economics* (pp. 1237 – 1313). Elsevier.

Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

Mortensen, D. T. & Pissarides, C. A. (1999). New developments in models of search in the labor market. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics* chapter 39, (pp. 2567–2627). Elsevier.

Moscarini, G. & Postel-Vinay, F. (2013). Stochastic search equilibrium. *Review of Economic Studies*.

Neal, D. (1999). The complexity of job mobility among young men. *Journal of Labor Economics*, *17*(2), pp. 237–261.

Patacchini, E. & Zenou, Y. (2007). Spatial dependence in local unemployment rates. *Journal of Economic Geography*, *7*(2), 169–191.

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

Pissarides, C. A. & Wadsworth, J. (1989). Unemployment and the inter-regional mobility of labour. *Economic Journal*, *99*(397), 739–55.

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.

Robin, J. (2011). On the dynamics of unemployment and wage distributions. *Econometrica*, *79*(5), 1327–1355.

Rupert, P. & Wasmer, E. (2012). Housing and the labor market: Time to move and aggregate unemployment. *Journal of Monetary Economics*, *59*(1), 24–36.

Shephard, A. (2011). Equilibrium search and tax credit reform. Working Papers 1336, Princeton University, Department of Economics, Center for Economic Policy Studies.

Topel, R. H. (1986). Local labor markets. *Journal of Political Economy*, *94*(3), S111–43.

Wasmer, E. & Zenou, Y. (2006). Equilibrium search unemployment with explicit spatial frictions. *Labour Economics*, *13*(2), 143–165.

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

Zenou, Y. (2009b). *Urban Labor Economics*. Number 9780521875387 in Cambridge Books. Cambridge University Press.

A Additional expressions and proofs

A.1 Reservation wages and mobility-compatible indifference wages

In this appendix, we compute the expressions for the reservation wages ϕ_j and the mobility-compatible indifference wages $q_{jl}(w)$. Equations 2 and 1 can be rewritten as:

$$r V_j^u = b + \lambda_j^u \int_{\phi_j}^{\overline{w}} [V_j^e(x) - V_j^u] dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \int_{q_{jk}(\phi_j)}^{\overline{w}} [V_k^e(x) - V_j^u] dF_k(x) \tag{29}$$

$$\begin{aligned} r V_j^e(w) &= w + \lambda_j^e \int_w^{\overline{w}} [V_j^e(x) - V_j^e(w)] dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} [V_k^e(x) - V_j^e(w)] dF_k(x) \\ &+ \delta_j [V_j^u - V_j^e(w)] \end{aligned} \tag{30}$$

Rearranging 30 we get:

$$V_j^e(w) = \frac{w + \delta_j V_j^u + \lambda_j^e \int_w^{\overline{w}} V_j^e(x) dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} V_k^e(x) dF_k(x)}{r + \delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(w))} \tag{31}$$

Note from equation 31 that:

$$dV_j^e(w) = \frac{1}{r + \delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(w))} \tag{32}$$

Integrating by part, equations 29 and 30, we obtain:

$$r V_j^u = b + \lambda_j^u \int_{\phi_j}^{\overline{w}} \overline{F}_j(x) dV_j^e(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \int_{q_{jk}(\phi_j)}^{\overline{w}} \overline{F}_k(x) dV_k^e(x) \tag{33}$$

$$r V_j^e(w) = w + \lambda_j^e \int_w^{\overline{w}} \overline{F}_j(x) dV_j^e(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} \overline{F}_k(x) dV_k^e(x) + \delta_j [V_j^u - V_j^e(w)] \tag{34}$$

Plugging 32 into 33, we obtain the following closed form for V_j^u:

$$r V_j^u = b + \lambda_j^u \int_{\phi_j}^{\overline{w}} \Phi_j(x) dx + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \int_{q_{jk}(\phi_j)}^{\overline{w}} \Phi_k(x) dx \tag{35}$$

with $\Phi_j(x) = \frac{\overline{F}_j(x)}{\left(r + \delta_j + \sum_{k \in \mathcal{J}} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(x))\right)}$.

Plugging 32 into 31, we obtain the following expression for $V_j^e(w)$:

$$r V_j^e(w) = w + \delta_j [V_j^u - V_j^e(w)] + \lambda_j^e \int_w^{\overline{w}} \Phi_j(x) dx + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{q_{jk}(w)}^{\overline{w}} \Phi_k(x) dx \tag{36}$$

Using equations 35 and 36 with $w = \phi_j$, we obtain the closed form for the reservation wage ϕ_j given in equation 4 and the closed form for the mobility-compatible indifference wage $q_{jl}(w)$ given in equation 6.

A.2 Existence and uniqueness of the optimal strategies

In this section, we prove that the optimal steady-state strategies described in section 2.2 exist and are unique. We focus on the mobility-compatible indifference wages $q_{jl}(\cdot)$. The existence and uniqueness of reservation wages ϕ_j can be proven along the same lines.

Equation 5 defines a system of $J^2 - J$ equations with $J^2 - J$ unknowns. The existence and uniqueness is established more easily using the differential form defined by equation 7. With an initial condition on a wage level w_0, the differential equation can be written:

$$q_{jl}(w) = q_{jl}(w_0) + \int_{\underline{w}}^{w} h_{jl}(x, q_j(x))dx \tag{37}$$

where $q_j(x) \equiv \left\{q_{jk}(x)\right\}_{k\in\mathscr{J}_j}$.

Starting from initial value $\underline{w}$, we can use Picard's iterative process $\left(q_{jl}^{(1)},...q_{jl}^{(k)}\right)$ to show that :

$$q_{jl}^{(m)}(w) = K^{(m)}(w_0)(w) \tag{38}$$

with

$$K(q_{jl})(w) = q_{jl}(w_0) + \int_{\underline{w}}^{w} h(x, q_j(x))dx \tag{39}$$

Note that $h_{jl}(x, q_j(x)) = dq_{jl}(x, q_j(x))$. As already noted, each function $dq_{jl}(\cdot)$ is strictly positive; moreover, given that the structural matching parameters $\{\lambda_j^u, \lambda_j^e, \delta_j\}_{j\in\mathscr{J}}$ are all positive and $r > 0$, it 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 equation 37 has a unique solution.

A.3 Mobility costs and Local amenities: expressions

The problem of an agent is summarized by:

$$rV_j^u = b + \gamma_j + \lambda_j^u \int_{\underline{w}}^{\overline{w}} \max\left\{V_j^e(x) - V_j^u, 0\right\} dF_j(x) + \sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \int_{\underline{w}}^{\overline{w}} \max\left\{V_k^e(x) - c_{jk} - V_j^u, 0\right\} dF_k(x) \tag{40}$$

$$\begin{aligned} rV_j^e(w) &= w + \gamma_j + \lambda_j^e \int_{\underline{w}}^{\overline{w}} \max\left\{V_j^e(x) - V_j^e(w), 0\right\} dF_j(x) + \sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \max\left\{V_k^e(x) - c_{jk} - V_j^e(w), 0\right\} dF_k(x) \\ &+ \delta_j\left[V_j^u - V_j^e(w)\right] \end{aligned} \tag{41}$$

This yields:

$$
\begin{aligned}
rV_j^u &= b+\gamma_j+\lambda_j^u\int_{\phi_j}^{\overline{w}}\left(V_j^e(x)-V_j^u\right)dF_j(x)+\sum_{k\in\mathcal{J}_j}s_{jk}^u\lambda_k^u\int_{q_{jk}(\phi_j)}^{\overline{w}}\left(V_k^e(x)-c_{jk}-V_j^u\right)dF_k(x) &(42)\\
rV_j^e(w) &= w+\gamma_j+\lambda_j^e\int_{w}^{\overline{w}}\left(V_j^e(x)-V_j^e(w)\right)dF_j(x)+\sum_{k\in\mathcal{J}_j}s_{jk}^e\lambda_k^e\int_{q_{jk}(w)}^{\overline{w}}\left(V_k^e(x)-c_{jk}-V_j^e(w)\right)dF_k(x)\\
&+ \delta_j\left[V_j^u-V_j^e(w)\right] &(43)
\end{aligned}
$$

Rearranging equation 43, we obtain:

$$
\begin{aligned}
\left[r+\delta_j+\lambda_j^e\overline{F}_j(w)+\sum_{k\in\mathcal{J}_j}s_{jk}^e\overline{F}_k(q_{jk}(w))\right]V_j^e(w) &= w+\gamma_j-\sum_{k\in\mathcal{J}_j}s_{jk}^e\overline{F}_k(q_{jk}(w))c_{jk}+\lambda_j^e\int_w^{\overline{w}}V_j^e(x)dF_j(x)\\
&+\sum_{k\in\mathcal{J}_j}s_{jk}^e\lambda_k^e\int_{q_{jk}(w)}^{\overline{w}}V_k^e(x)dF_k(x)+\delta_jV_j^u &(44)
\end{aligned}
$$

After derivation:

$$
dV_j^e(w)=\frac{1+\sum_{k\in\mathcal{J}_j}s_{jk}^e\lambda_k^e c_{jk}f_k(q_{jk}(w))dq_{jk}(w)}{r+\delta_j+\lambda_j^e\overline{F}_j(w)+\sum_{k\in\mathcal{J}_j}s_{jk}^e\overline{F}_k(q_{jk}(w))} \tag{45}
$$

Integrating by part, equations 42 and 43, we obtain:

$$
\begin{aligned}
rV_j^u &= b+\gamma_j-\sum_{k\in\mathcal{J}_j}s_{jk}^u\lambda_k^u\overline{F}_k(q_{jk}(\phi_j))c_{jk}+\lambda_j^u\int_{\phi_j}^{\overline{w}}\Xi_j(x)dx+\sum_{k\in\mathcal{J}_j}s_{jk}^u\lambda_k^u\int_{q_{jk}(\phi_j)}^{\overline{w}}\Xi_k(x)dx\\
rV_j^e(w) &= w+\gamma_j-\sum_{k\in\mathcal{J}_j}s_{jk}^e\lambda_k^e\overline{F}_k(q_{jk}(w))c_{jk}+\delta_j[V_j^u-V_j^e(w)]+\lambda_j^e\int_w^{\overline{w}}\Xi_j(x)dx+\sum_{k\in\mathcal{J}_j}s_{jk}^e\lambda_k^e\int_{q_{jk}(w)}^{\overline{w}}\Xi_k(x)dx
\end{aligned}
$$

where $\Xi_j(w)=\frac{\overline{F}_j(w)\left(1+\sum_{k\in\mathcal{J}_j}s_{jk}^e\lambda_k^e c_{jk}f_k(q_{jk}(w))dq_{jk}(w)\right)}{r+\delta_j+\lambda_j^e\overline{F}_j(w)+\sum_{k\in\mathcal{J}_j}s_{jk}^e\lambda_k^e\overline{F}_k(q_{jk}(w))}$

The reservation wage compatible with mobility cost, is given by:

$$
\begin{aligned}
\phi_j &= b+\left(\lambda_j^u-\lambda_j^e\right)\int_{\phi_j}^{\overline{w}}\Xi_j(x)dx-\sum_{k\in\mathcal{J}_j}\left(s_{jk}^u\lambda_k^u-s_{jk}^e\lambda_k^e\right)\overline{F}_k(q_{jk}(\phi_j))c_{jk} &(46)\\
&+\sum_{k\in\mathcal{J}_j}\left(s_{jk}^u\lambda_k^u-s_{jk}^e\lambda_k^e\right)\int_{q_{jk}(\phi_j)}^{\overline{w}}\Xi_k(x)dx
\end{aligned}
$$

Similarly, the mobility-cost-compatible indifference wage solves $V_j^e(w)=V_l^e(q_{jl}(w))$ and is given by:

$$
\begin{aligned}
q_{jl}(w) &\equiv w+\gamma_j-\gamma_l-\sum_{k\in\mathcal{J}_j}\lambda_k^e\left(s_{jk}^e\overline{F}_k(q_{jk}(w))c_{jk}-s_{lk}^e\overline{F}_k(q_{lk}(w))c_{lk}\right)+(1-s_{lj}^e)\lambda_j^e\int_w^{\overline{w}}\Xi_j(x)dx &(47)\\
&- (1-s_{jl}^e)\lambda_l^e\int_{q_{jl}(w)}^{\overline{w}}\Xi_l(x)dx+\delta_jV_j^u-\delta_lV_l^u+\sum_{k\in\mathcal{J}_j}(s_{jk}^e-s_{lk}^e)\lambda_k^e\int_{q_{jk}(w)}^{\overline{w}}\Xi_k(x)dx
\end{aligned}
$$

For the estimation, we consider the following approximation of $dq^c_{jl}(w)$:

$$d\tilde{q}^c_{jl}(w) = \frac{1-(1-s^e_{lj})\lambda^e_j\tilde{\Xi}_j(w) - \sum_{k\in\mathscr{J}_{jl}}\lambda^e_k\Big[(s^e_{jk}-s^e_{lk})\tilde{\Xi}_k(q^c_{jk}(w)) + \big(s^e_{jk}c_{jk}f_k(q^c_{jk}(w)) - s^e_{lk}c_{lk}f_k(q^c_{lk}(w))\big)\Big]}{1-(1-s^e_{jl})\lambda^e_l\tilde{\Xi}_l(q^c_{jl}(w))}$$

with: $\tilde{\Xi}_j(w) = \dfrac{\overline{F}_j(w)\Big(1+\sum_{k\in\mathscr{J}_j}s^e_{jk}\lambda^e_k c_{jk}f_k(q^c_{jk}(w))\Big)}{r+\delta_j+\lambda^e_j\overline{F}_j(w)+\sum_{k\in\mathscr{J}_j}s^e_{jk}\lambda^e_k\overline{F}_k(q^c_{jk}(w))}$.

B Dataset, additional descriptive statistics and lesser estimation results

B.1 Data selection and structure of the dataset

The initial sample is composed of 43,010,827 observations over the period 1976-2008. Our sample selection is as follows:

- We restrict the sample to observations recorded between 2002 to 2007, related to the main job of individuals in urban continental France
- We dispose of female workers as well as individuals who at some point were older than 58 years, and younger than 15 years.
- We drop individuals who at some point were working: in the public sector , as apprentice, as home workers, and part time workers.
- We drop individuals who at some point had a reported wage that is inferior to the 850 euros per month (the net minimum wage is around 900 euros): or a monthly wage higher than 65,000 euros: The first case is considered as measurement error; the second case may reflect a real situation but it extends the support of wage distributions too dramatically for very few individuals.
- Finally, for computational reasons, we get rid of individuals observed only once and we do not use 40% of the sample in Paris.

Finally, we end up with the following dataset:

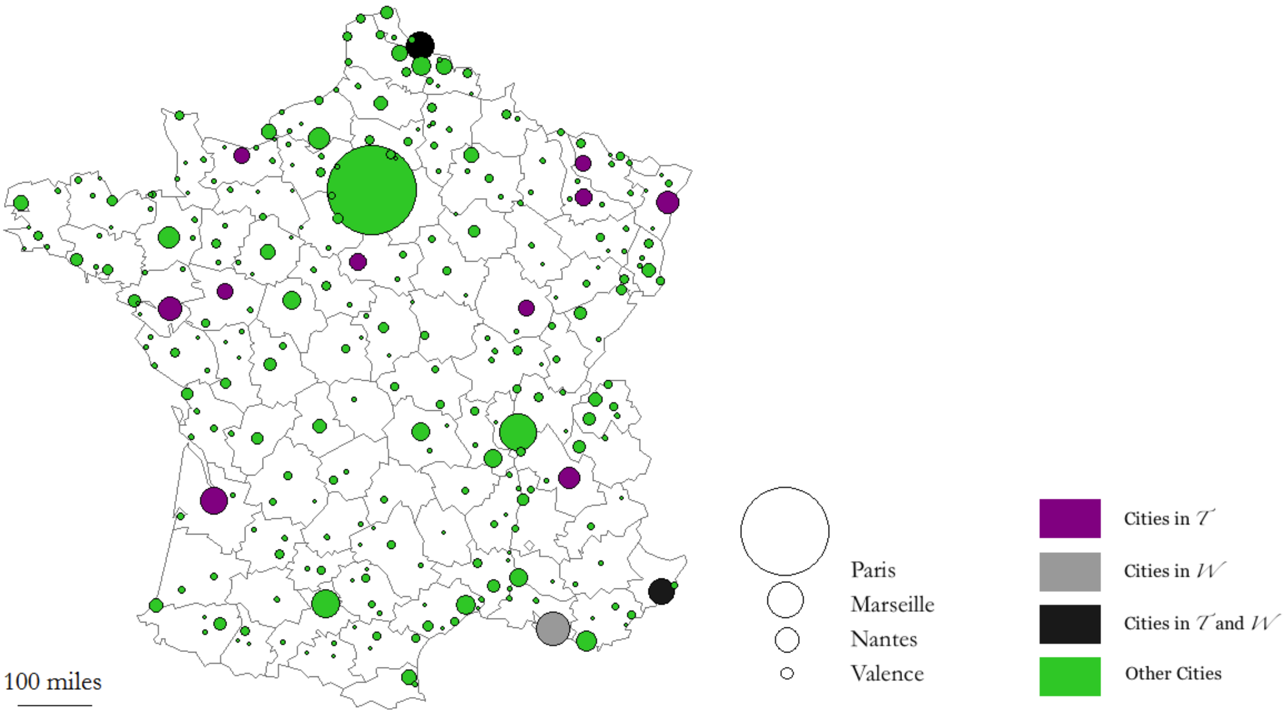
Table 14: Structure of the dataset

Year	Number of Individuals	Number of Observations	Number of obs. by metro Min	Mean	Median	Max	Number of individuals by metro Min	Mean	Median	Max
2002	310153	332,446	95	1,581	433	84,302	97	1,662	445	90,452
2003	297697	311,309	99	1,505	406	80,981	101	1,556	412	84,950
2004	308179	321,557	107	1,558	424	84,104	108	1,607	433	88,027
2005	310949	325,580	71	1,573	441	84,911	72	1,627	449	89,600
2006	316613	332,848	105	1,604	436	86,712	106	1,664	449	91,525
2007	313693	335,460	111	1,597	432	86,029	112	1,677	455	92,169
Total	477,068	2,548,719	260	8,467	1,877	650,010	65	1,917	454	135,460

Notes: (i) Metros are here the clusters of municipalities forming the 300 largest metropolitan areas in 2010; (ii) Source: Panel DADS 2002-2007

C Figures

Figure 8: The 12 metropolitan areas in subset $\mathscr{T}$ and the 3 metropolitan areas in subset $\mathscr{W}$



Notes: (i) see Figure 2; (ii) Subset $\mathscr{T}$ is used to identify the effect of physical distance and dissimilarity on spatial frictions based on pair-specific UE and EE transition rates; subset $\mathscr{W}$ is used to identify the effect of physical distance on moving costs based on pair-specific average accepted wages after a EE transition with mobility.

D Estimation of the matching function

Table 15: The matching function

	λ^u	λ^e	δ	s^u	s^e
log number of firms	1.060^{***}	0.314^{***}	-0.018	13.519^{**}	2.397
	(0.141)	(0.045)	(0.026)	(5.050)	(5.634)
log area	0.215^{*}	-0.039	-0.066^{***}	4.772	0.860
	(0.096)	(0.031)	(0.018)	(3.448)	(3.846)
log pop	-0.679^{***}	-0.115^{***}	0.083^{***}	-12.127^{***}	-4.035
	(0.087)	(0.028)	(0.016)	(3.123)	(3.484)
R^2	0.680	0.873	0.895	0.250	0.386
Num. obs.	200	200	200	200	200

Notes: (i) Significance: $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$; (ii) Contrary to Table 9, the parameters for the city 200 are included