

Job Search and Migration in a System of Cities^{*}

Benoît Schmutz[†]

Modibo Sidibé[‡]

April 1, 2014

Abstract

We build an equilibrium job search model, where workers engage in both off-the-job and on-the-job search over a set of cities, to quantify the impact of matching frictions, spatial frictions and mobility costs on the job search process. The mobility problem of workers introduces the concept of “mobility-compatible indifference wages”, based on a dynamic utility trade-off between locations, which can rationalize very diverse wage dynamics as part of forward-looking spatial strategies. The model is estimated 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 and amenity parameters and to quantify the level of spatial constraints upon each possible move. Our findings include: (i) 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 and an alternative explanation for the city size wage gap; (ii) an inference on the determinants of the matching parameters which shows that most of the variance in job arrival rates for the unemployed can be explained by a few basic city characteristics.

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, Sébastien Roux and John Rust, as well as participants to seminars and conferences in Duke University, Sciences-Po, GATE-Lyon St-Etienne, Paris School of Economics, 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.

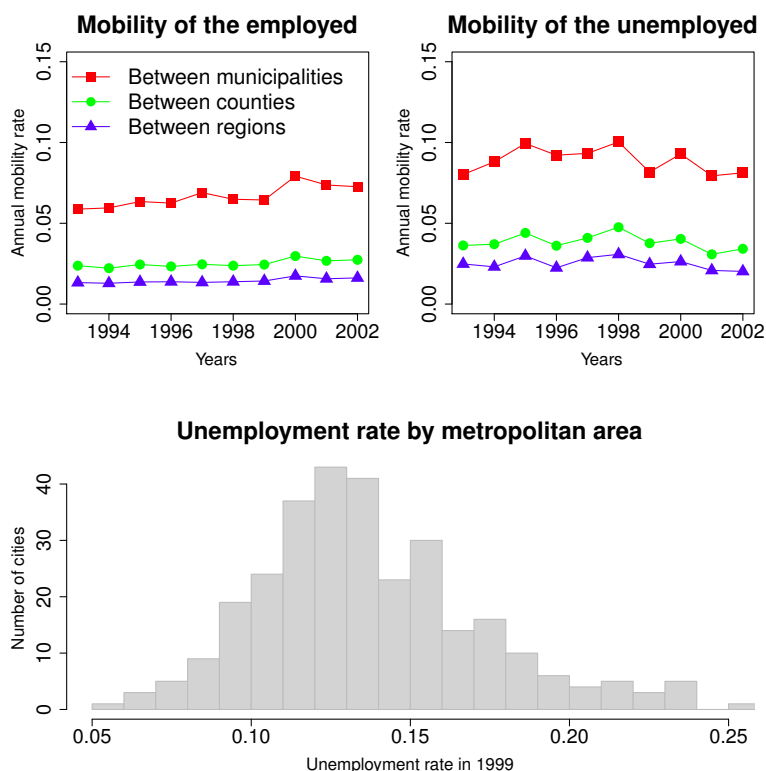
[†]Howard University; benoit.schmutz@howard.edu

[‡]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.

¹See [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).

²See [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.

This paper aims to provide a micro-foundation to the coexistence of these aggregate features and understand why despite similar cultural and labor market institutions, and the rapid progress of transportation and communication technologies, individuals do not take advantage of the opportunity to move into more prosperous cities. We argue that search theory, with the usual forward-looking, risk-neutral workers, can rationalize this low mobility rate when we account for the spatial structure of a country. We extend the wage posting framework of [Burdett & Mortensen \(1998\)](#) to incorporate spatial segmentation between a large number of interconnected local labor markets, or “cities”. Each city is characterized by specific job arrival rate, layoff probability and a wage offer distribution. We consider the optimal strategy of ex-ante identical workers, who engage in both off-the-job and on-the-job search, both within and between cities, and who may only migrate after receiving a job offer from another city. A dynamic search and matching model of migration allows us to define spatial constraints more precisely than most previous quantitative studies of migration. Those studies rest upon a unidimensional conception of spatial constraints based on a black box called “mobility costs”, which encompass both impediments to the mobility of workers when it takes place (actual mobility costs) and impediments to the spatial integration of the labor market (workers’ ability to learn about remote vacancies and their willingness to fill those vacancies). We will show that a lot is to be gained by clearly distinguishing between these different features, because they do not take place at the same time and they do not have the same consequences on workers’ mobility decisions.

Spatial constraints are threefold. First and foremost, they may disconnect workers from the labor market. Job search between cities is subject to an information loss that will lower the probability that a jobseeker will hear about a vacancy posted in another city, that is, the efficiency of the job search process between locations. This dimension determines the *centrality* of each city in the system. Second, spatial segmentation allows for the possibility that cities will be heterogeneous in their non-labor market dimension, which will be called a city-specific “amenity”; this amenity will impact agents’ willingness to refuse (accept) a job somewhere else, even though this would (not) be a sound decision from a pure labor-market standpoint. The ranking of each city according to this dimension contributes to the *attractiveness* of each city in the system, in addition to the local labor market conditions the city has to offer. Finally, workers do face classical mobility costs, which are a lump sum that they will need to pay to be able to move and that will ultimately determine what their final decision will be, conditional on receiving an acceptable offer. Since the model is dynamic, the relative position of the city in the distribution of all possible mobility costs, which determines the level of *accessibility* of the city in the system, will also impact whether the offer was deemed acceptable in the first place.

The key innovation of our model is the definition of “mobility-compatible indifference wages”, based on a dynamic utility trade-off between locations. These functions of wage, which are specific to each pair of cities, are defined by the worker’s indifference condition between her current state (a given wage in a given city) and a potential offer in a different city. The indifference wages define a complex relationship between wages and the model primitives and can be decomposed between: the current wage, the difference in the value of local on-the-job search and on-the-job search with mobility, the difference in the average cost of the next potential move, and three constant terms: the mobility cost, the difference in local amenities, and the difference in the value of unemployment, weighted by local unemployment risk. As a consequence, the model is able to cope with various wage profiles over the life-cycle, including voluntary wage cuts as in [Postel-Vinay & Robin \(2002\)](#). 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 non-homogenous functional differential equations.

The model is solved using equilibrium conditions on market size, unemployment level and wage distributions. It is estimated on the panel version of the French matched employer-employee database *Déclaration Annuelles de Données Sociales* (DADS) from 2002 to 2007, with local labor markets defined at the metropolitan area level. The identification strategy is a major novelty of the paper. In contrast to most of the literature, the identification of local labor market parameters and spatial friction parameters is based on the frequency of labor and geographical mobility. Data on wages are used only to identify local amenity parameters and mobility costs. Therefore, we can fully disentangle between the impact of mobility costs and the impact of spatial frictions on the mobility rate. The other breakthrough is computational. The model is based on a partition between submarkets which can, in theory, be made as detailed as possible: we address the challenges raised by the high dimensionality and we allow the final level of precision to only depend on the research question. In our case, we consider that cities make up for plausible intermediaries between the micro level of the individual jobseekers and the macro level of the nationwide labor market. Yet, the model is fractal and may apply to the analysis of spatial segmentation at the neighborhood level within a single metropolitan labor market, or even to international migration. It is also transferable to occupational mismatch at little cost.

Our results characterize each local labor market primitives by a set of three matching parameters (lay-off rates, job arrival rates for the unemployed, job arrival rates for the employed) and one local amenity parameter, as well as three matrices of parameters measuring spatial constraints between each pair of cities (spatial search efficiency for unemployed and employed jobseekers and mobility costs). We use the dataset

of matching 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 30, share of males, share of people without qualification, share of blue-collar jobs and share of manufacturing jobs) accounts for 90% of the variation in the job arrival rate for unemployed workers, against 31% for on-the-job arrival rates and 18% for job separation rates.

Our estimation of spatial constraints suggests that geographical distance may increase mobility costs by up to 40% and that both geographical and sectoral distances are much stronger deterrents of the efficiency of spatial search for employed jobseekers, than for unemployed jobseekers. In addition, we show that local amenities are higher in larger cities, even net of the higher local costs. 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. Finally, using a matching function, we run a counterfactual experiment 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.

The rest of the paper is organized as follows. In the first section, we provide an overview of the related literature and detail our contributions; in a second section, we describe the French labor market as a system of interconnected local urban labor markets; the third section is the presentation of the model; the fourth section explains our estimation strategy and the results are discussed in a fifth section.

1 Contributions to the literature

In this section, we discuss the various contributions of our paper. First, we review the literature on the determinants of migration. Second, we examine the literature on the city size premium. Third, we pursue with the applied-theoretical question of the interactions between competing submarkets. Finally, we conclude with the econometric issue of the identification strategy used in the estimation of migration models.

1.1 Migration

The career choice of workers has long been investigated by economists. Works by [Keane & Wolpin \(1997\)](#) and [Neal \(1999\)](#) have shown that individuals make sophisticated calculations regarding work-related de-

cisions, both in terms of pure labor market characteristics (industry, occupation, skills requirement) and location choice. To understand the underlying mechanism, [Dahl \(2002\)](#) proposes a model of mobility and earnings over the US states and shows that higher educated individuals self-select into states with higher returns to education; however, 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. However, his model does not describe the individual mobility problem. Finally, despite its interest and obvious links to the present paper, the classic perfect-competition approach cannot fully reconcile the low mobility rate with local labor market differences.

In this paper, we argue that friction-based search and matching models can tackle this puzzle. 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 local labor markets requires to handle multiple high-dimensional objects such as wage distributions. 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 model which includes several appealing features 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.

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 US states and provide many interesting insights with respect to the mobility decision of workers, including mobility costs. However, the computational difficul-

³The 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 workers, 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](#)).

⁴As 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.

ties requires additional assumptions. For example, it is assumed that individuals have knowledge over a limited number of wage distributions, which correspond to locations they lived in. In order to learn about another location, workers need to pay a visiting cost. These assumptions may not reflect the recent increase in workers' ability to learn about other locations before a mobility ([Kaplan & Schulhofer-Wohl, 2012](#)).⁵ 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, which are better proxied by metropolitan areas ([Moretti, 2011](#)). In this paper, we try to overcome these shortcomings by considering a search and matching model with mobility cost and a more detailed partition of space at the metropolitan level.

1.2 Economic geography and urban labor markets

Our estimation results shed new light on the determinants of the city size wage premium. The frictionless economic geography literature has focused on the determinants of the wage growth across cities or the differences between cities and rural areas ([Gould, 2007](#)). [Combes, Duranton, Gobillon, Puga & Roux \(2012\)](#), on French data, show that higher productivity in larger cities is mostly due to agglomeration economies and technological complementarities between the productivity of firms and that of workers. Using Spanish data, [De la Roca & Puga \(2012\)](#) show that the city size wage premium does not reflect initial sorting of workers by ability, but is rather the result of a more efficient learning process in larger cities. Although individual wages are disconnected from productivity in our setup, the existence of search frictions allows us to reproduce both the upwards shift and the greater variability of the earning distributions, without resorting, neither to human capital accumulation, nor to production externalities. The optimal strategy of a worker 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. These simple Markovian dynamics between 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.⁶ [Rupert & Wasmer \(2012\)](#) have recently incorporated endogenous mobility decisions into a job search model with an explicit housing market. However, the location of an agent does not affect its job-finding rate. Therefore, the impact

⁵This experience good perspective is however more justified in their model, which allows for heterogeneous match productivity between the worker and the location.

⁶See, for example, [Coulson, Laing & Wang \(2001\)](#), [Brueckner & Zenou \(2003\)](#) or [Wasmer & Zenou \(2006\)](#).

of location on job opportunities takes place through commuting costs only. 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 as a consequence study the determinants of these parameters across cities.

This ability to estimate our model comes at a cost, we do not explicitly model the housing market, in contrast of [Head & Lloyd-Ellis \(2012\)](#), who 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 still be very challenging.⁷ However, the fact remains that mobility costs and local amenities are difficult to interpret in the absence of a separate housing market. Modeling the housing market would constitute an interesting extension, but it would require to restrict the number of markets and to merge the French matched employer-employee dataset with other data sources (like census).

1.3 Job search and frictions between competing submarkets

There is a notable effort in the recent empirical job search literature to look at search patterns in competing submarkets. A few papers seek to provide new dynamic micro-foundations to the old concept of dualism in the labor market. The underlying idea is that jobs are not only defined by wages, but also by a set of benefits that are only available within some submarkets. This creates potential tradeoffs between a more regulated sector, which offers more employment protection (in terms of unemployment risk and insurance) and a less regulated sector, which allows for more flexibility and possibly better wage paths. 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

⁷Similarly, [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, they consider a binary partition of space in their calibration where cities are classified according to the rate of decline in house prices.

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 workers 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 effects 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. 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. 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, whereas papers on dualism assume that the job finding rate for the unemployed is not impacted by the sector where workers were working before losing their job. From the separate estimation of their model on two local labor markets (Sao Paulo, the economic center of the country, and the poorer and more marginal Salvador), [Meghir et al. \(2012\)](#) show that high search frictions explain why informality persists in equilibrium. They also hint that marginal increases in regulation would benefit workers in dynamic Sao Paulo, whereas they would (slightly) decrease workers' welfare in Salvador. This focus on two cities, driven by data limitation, is a little frustrating, since one would like to be able to draw inference on this kind of important relationships using more than two points. Despite some simplifications, we offer a generalization of their framework to a large number of markets.

We provide another generalization of the previous models by explicitly allowing for mobility costs, which are analogous to what has been known as switching costs in the dynamic discrete choice literature. To the best of our knowledge, this is the first paper to consider the problem of deterministic, move-specific switching costs within a dynamic search and matching model.⁸ Whereas the modelling of sectoral mobility may do without switching costs, this is clearly not the case of large-scale geographical mobility. In addition, if one seeks to assess the relevance of the high mobility cost estimates of [Kennan & Walker \(2011\)](#), one cannot

⁸In a continuous environment, one of the very few examples of switching costs is [Lamadon, Lise, Meghir & Robin \(2013\)](#), who allow for stochastic training costs for unemployed workers. In a discrete environment, [Baum-Snow & Pavan \(2012\)](#) allow for a deterministic component in the mobility cost but this component does not depend on the type of city of origin and of destination.

solely replace mobility costs by spatial frictions: both need to be allowed for. Our theoretical framework shows that this extension is far from conceptually trivial and requires to make a (possibly strong) behavioral assumption regarding the impact of mobility costs on mobility decisions.

1.4 Identification

It is well recognized that the trade off between local amenities and living cost is the most important component of static location choice, while mobility cost is the main hurdle to the adjustment of residential location over the life-cycle. Despite this consensus, there are very few empirical estimations of the mobility cost and local amenities. Regarding local amenities, recent hedonic literature, based on detailed datasets and spatial discontinuities has made enormous progress towards estimates of the willingness to pay for location characteristics (Bayer, Ferreira & McMillan, 2007). 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.⁹ In competitive structural models of the labor market, estimation of mobility cost based on transition rates yield extremely high values (Lee & Wolpin, 2006; Kennan & Walker, 2011).

In structural econometrics, it is well-known that the theoretical identification of parameters hinges on the transition rates. Our identification relies on this strategy, but in addition, we are able to use wage data to identify local amenities and mobility. We use transition rates to identify local labor market parameters and spatial frictions parameters. As a consequence, it is only the frequency of transition from one state to another one that identifies the structural parameters. However, we have additional information in the dataset that consists of wage-accepting patterns between locations conditional on the structural parameters. When a city exhibits high (resp., low) accepted wages regardless of the origin of incoming workers, this information provides identification for a low (resp., high) level of local amenities. A downside is that we cannot separately identify local amenities from cost of living. As a consequence, we assume additivity and separability between between cost of living and local amenities in the utility of workers. Similarly, when workers from the same location accept very heterogeneous wages to go into different destinations, we use this information to identify mobility costs.

⁹See for example, Bayer, Fang & McMillan (2011).

2 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 workers; 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 workers on both the labor market and across space.

2.1 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 200 largest metropolitan areas in continental France, as defined by the 2010 census. Below a certain population threshold, the assumption that each of these metropolitan areas is an accurate proxy of a local labor market becomes difficult to support. As a consequence, the smallest metropolitan area which is isolated in our analysis is Redon, with 28,706 inhabitants in 2009. Such population threshold remains very low.¹¹ As shown in Figure 2, these metropolitan areas cover a very large fraction of the country. Paris and its 12 millions inhabitants stand out, before six other millionaire cities and eleven other metropolitan areas with more than 0.5 million inhabitants.

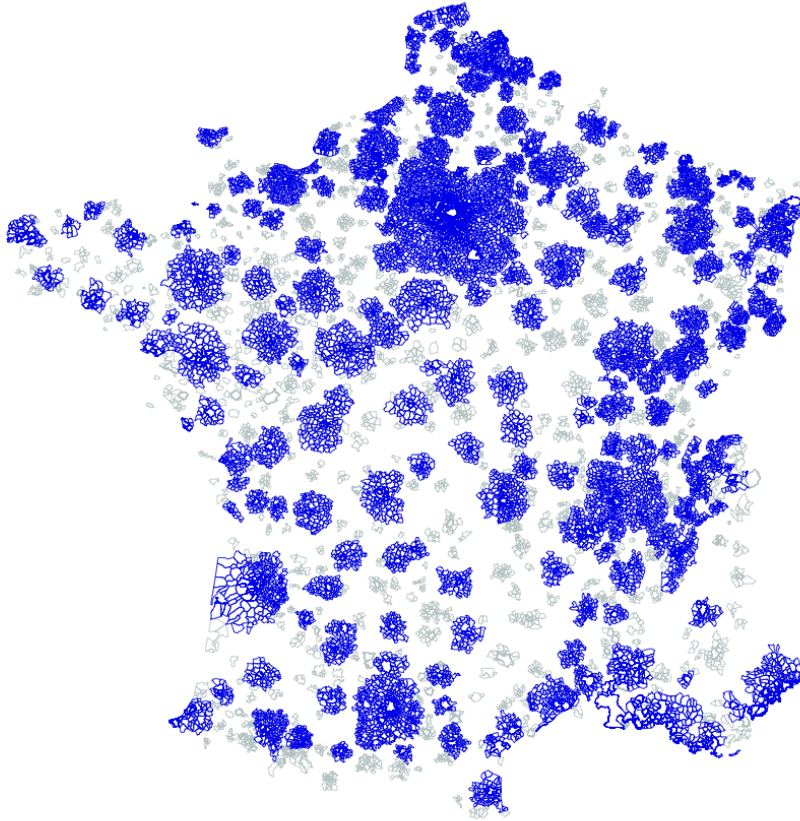
The functional definition of a metropolitan area brings together the notions of city and local labor market. A more precise partition of space, for instance based on municipal boundaries, would lead to a confusion between job-related motives for migration and other motives.¹² French metropolitan areas vary a

¹⁰US 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.

¹¹According to the 2010 US census, matching this level of precision on the US would require to distinguish between more than 800 cities (either metropolitan, or micropolitan statistical areas).

¹²According 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%

Figure 2: The French urban archipelago



Notes: the spatial unit is the municipality. There are more than 700 metropolitan areas according to the 2010 definition. In dark, the border of the municipalities that constitute the largest 200 metropolitan areas. In light, the border of all the other municipalities within a metropolitan area. *Source: INSEE, Census 2007*

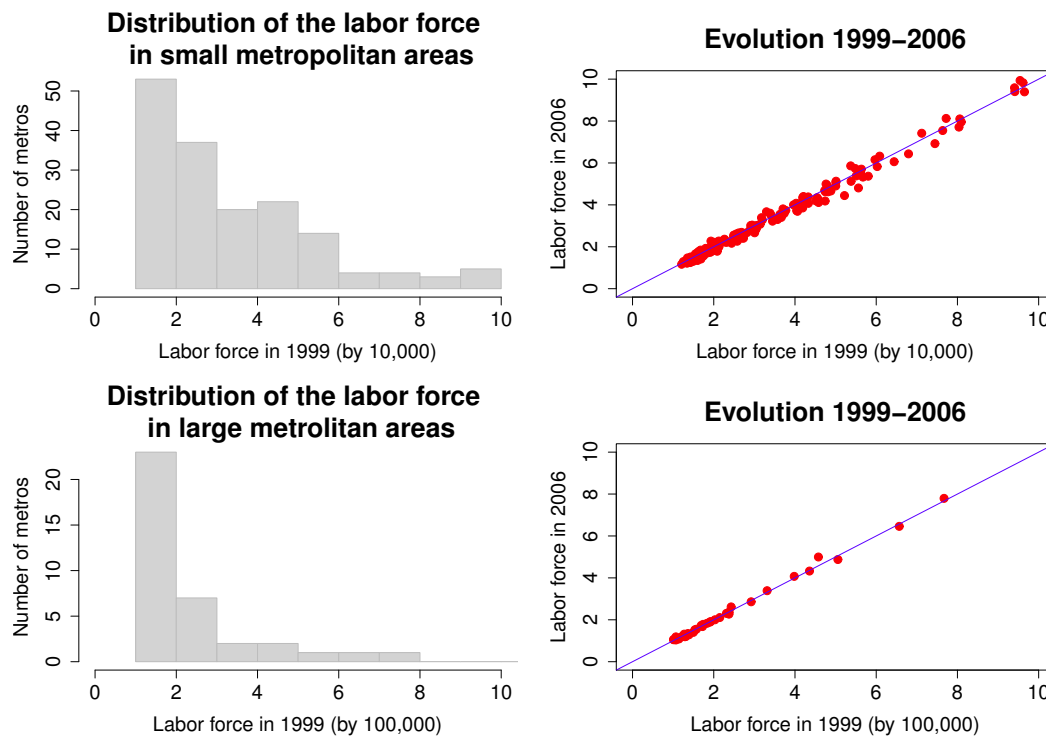
lot in terms of size, unemployment rates, and wage distributions, yet this heterogeneity is stable over time. We provide here evidence supporting these two claims on the set of the 300 largest metropolitan areas in continental France.

Population Since, we do not model the participation choice of workers, 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.¹³

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.

¹³This 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

Figure 3: Heterogeneity and stability of the labor force

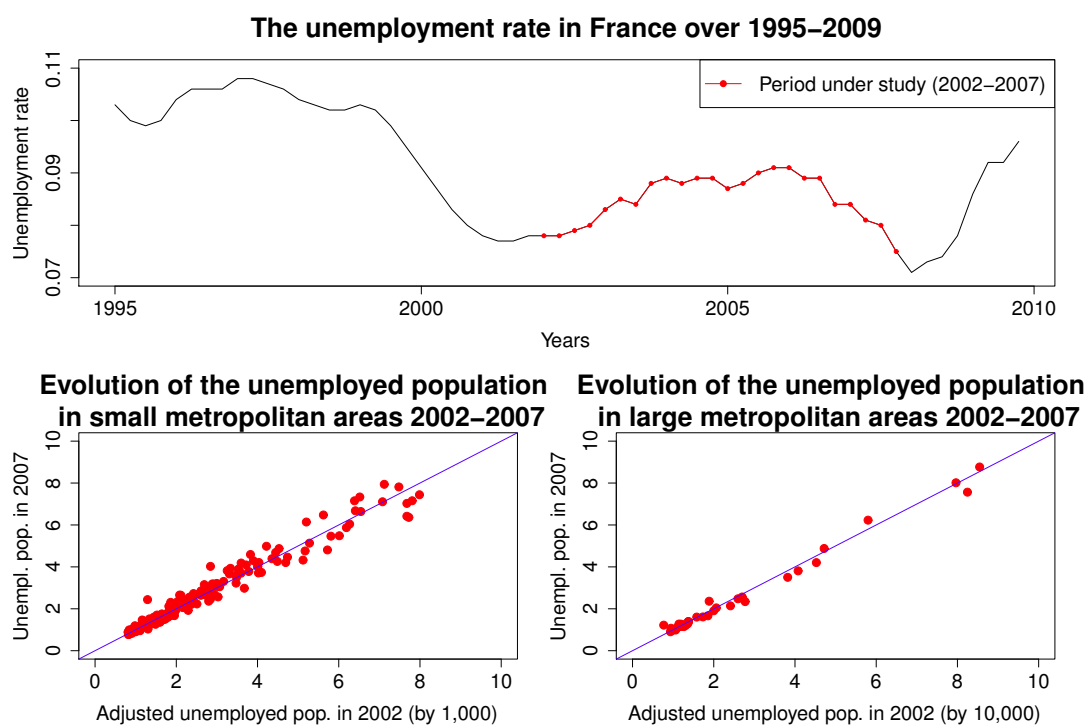


Notes: (i) Labor force is composed of unemployed and employed individuals aged between 15 and 64; the labor force in the 300 largest metropolitan areas in continental France amounts to 19.5 millions in 1999 and to 19.3 millions in 2006; (ii) For the sake of exposition, we do not represent Paris; its 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.*

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. 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.

to [Gobillon, Magnac & Selod \(2011\)](#), 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 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).*

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.¹⁴ The DADS are the main source of data used in this paper.¹⁵ 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 city-level average wage. Other large cities have similar wage premia.¹⁶ Although the wage premium in Paris may be partly offset by the cost of living, there exist persistent wage differentials among cities with compa-

¹⁴See Abowd, Kramarz & Margolis (1999) and Abowd, Kramarz & Roux (2006) for additional information.

¹⁵We use the longitudinal version of the DADS on a specific subsample of the population (see section 2.2 for details).

¹⁶Our 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{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{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*

able 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.

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.

2.2 Labor and geographical mobility

Data We now turn to the mobility patterns of jobseekers across France. To make a precise assessment regarding geographical transitions between each pair of cities, we use a specific subsample of the DADS data. Since 1976, a yearly longitudinal version of the DADS has been following all employed individuals born in October of even-numbered years. Since 2002, the panel includes all individuals born in October. Due to the methodological change introduced in 2002, and amid concerns about the stability of the business cycle, we focus on a six-year span between 2002 and 2007, 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.¹⁷ 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.¹⁸ Finally, we need to make an important assumption regarding the geographical transitions of unemployed individuals: we attribute all the duration of unemployment to the initial location, assuming therefore that any transition from unemployment to employment with migration is a single draw. Hence, we rule out the possibility of a sequential job search whereby individuals would first change locations before accepting a new job offer.¹⁹ From a theoretical viewpoint, this means that mobility has to be job-related. From a practical viewpoint, in the DADS data, the sequential job search process is observationally equivalent to the joint mobility process.

¹⁷Our algorithm is available at [this address](#).

¹⁸For a recent example of a similar assumption, see Bagger, Fontaine, Postel-Vinay & Robin (2013).

¹⁹A very similar assumption is found in Baum-Snow & Pavan (2012).

Labor market transitions 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

Table 3: Number and characteristics of transitions

Type of history	Characteristics of the spells			
	Number of events	Share	Initial Wage	Final Wage
No transitions while employed	126,227		26,088	-
Out of unemployment	302,024		-	24,303
with mobility	59,605	19.8	-	24,793
without mobility	242,418	80.2	-	24,182
Job to job mobility	114,659		30,814	32,936
with mobility	26,199	22.9	30,464	32,343
without mobility	88,459	87.1	30,914	33,111
Into unemployment	302,918		24,555	-
Full sample	719,601		27,956	28,255
Individuals	375,276			

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

the non-mobility rate of 45% reported by [Postel-Vinay & Robin \(2002\)](#) from 1996 to 1998. Approximately 23% of the sample records at least one job-to-job transition, and the number of transitions into unemployment and the number of transitions out of unemployment are almost identical. Average wages are almost constant over time, as shown in the last line of the table. Job-to-job transitions are accompanied by a substantial wage increase (around 7%). Transitions out of unemployment lead to a wage that is 7% lower than the wage of employed workers who do not make any transition, 25% lower than the final wage of workers 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.²⁰

Geographical transitions Geographical mobility 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.²¹ Table 5 completes this

²⁰In 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.

²¹[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. We conclude that this low rate deserves an explanation.

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 France	UE	0.041	0.012	0.005	0.007	0.008	-
	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*

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 workers, geographical proximity can overcome this attractiveness, as shown for the cities of Grenoble, Saint-Etienne and Bourg-en-Bresse that are located less than 60 miles away from Lyon. As a consequence, we will incorporate distance between locations as a determinant of spatial frictions (see section 4 for details).

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*

Wage dynamics within and between cities As shown in Table 6, wage dynamics following a job-to-job transition are characterized by two noteworthy features. First, they are not symmetrical: average wages following a job-to-job transition with mobility into a given city are almost always higher than average wages

Table 6: Average wages following a job-to-job transition

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

following a job-to-job transition within the same city.²² This suggests that mobility costs are high compared to local differences in economic opportunities. Second, if mobility costs are mostly determined by the physical distance between two locations, wage dynamics cannot be fully rationalized by them. For example, as will be shown in section 5, Paris does offers many more opportunities than Toulouse, yet workers 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 (or, equivalently, local costs of living).

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 workers' geographical mobility? 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.

²²It 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 workers 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 workers.

3 A spatial model of job search in a system of cities

3.1 Framework

We consider a system \mathcal{J} of J interconnected local labor markets, or “cities”, where workers both live and work. Workers are ex ante identical. They are fully characterized by their employment status $i = e, u$ and their location $j \in \mathcal{J}$. 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 workers in the economy is exogenous and equal to M . Workers 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 reverse spatial friction parameters $s_{jl}^i \in (0, 1)$ such that a state- i worker living in location j will receive job offers from location j at rate λ_j^i and from location $l \in \mathcal{J}_j \equiv \mathcal{J} - \{j\}$ at rate $s_{jl}^i \lambda_l^i$. These spatial friction parameters aim to capture the fact that search might be less efficient between cities than within cities, most likely because of information loss. If they accept a job offer in another city, workers have to pay a lump-sum mobility cost. Let c_{jl} denote the mobility cost associated with a transition from city j to city l .

Employed workers in city j face a location-specific unemployment risk characterized by the layoff probability δ_j . When they become unemployed, workers receive uniform unemployment benefits b . We 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) worker living in city j is equal to $b + \gamma_j$ (resp., $w + \gamma_j$). 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 $\bar{F}_j(\cdot) \equiv 1 - F_j(\cdot)$. These distributions are assumed to share a common support $[\underline{w}, \bar{w}] \subset [b, \infty)$ over all locations. Workers 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 workers living in city j and workers living in city j and

employed at wage w are recursively defined by equations 1 and 2:

$$rV_j^u = b + \gamma_j + \lambda_j^u \int_{\underline{w}}^{\bar{w}} \max\{V_j^e(x) - V_j^u, 0\} dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \int_{\underline{w}}^{\bar{w}} \max\{V_k^e(x) - c_{jk} - V_j^u, 0\} dF_k(x) \quad (1)$$

$$\begin{aligned} rV_j^e(w) &= w + \gamma_j + \lambda_j^e \int_{\underline{w}}^{\bar{w}} \max\{V_j^e(x) - V_j^e(w), 0\} dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\bar{w}} \max\{V_k^e(x) - c_{jk} - V_j^e(w), 0\} dF_k(x) \\ &+ \delta_j [V_j^u - V_j^e(w)] \end{aligned} \quad (2)$$

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 a worker 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, workers 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 workers are willing to accept a wage cut in another location, this decision is somewhat analogous to the purchase of an unemployment insurance contract.

In a 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 workers' decision to accept a job in a given city is not only driven by the offered wage and the primitives of the local labor market, but also by the employment prospects in all the other locations, which depend upon the city's specific position within the system 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.

3.2 Optimal strategies

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

Definition A reservation wage corresponds to the lowest wage an unemployed worker 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 $(j, l) \in \mathcal{J} \times \mathcal{J}_j$. These functions associate the current wage w earned in location j to a wage which would yield the same dynamic utility in location l , once the mobility cost c_{jl} taken into account. They are denoted $q_{jl}(\cdot)$ and verify $V_j^e(w) \equiv V_l^e(q_{jl}(w)) - c_{jl}$. The definition of $q_{jl}(\cdot)$ extends to unemployed workers in city j who receive a job offer in city l : we have $V_j^u \equiv V_l^e(q_{jl}(\phi_j) - c_{jl})$. Finally, let $\chi_{jl}(w)$ denote another indifference wage, verifying $V_j^e(w) \equiv V_l^e(\chi_{jl}(w))$. This indifference wage equalizes the utility levels between two individuals located in cities j and l . We shall therefore refer to it as the “static” indifference wage, unlike the “dynamic” indifference wage $q_{jl}(w)$, which equalizes the utility level between one worker located in city j and the same worker after a move into city l . The introduction of $\chi_{jl}(w)$ is important to understand the role of mobility costs in the dynamics of the model.

Proposition 1 OPTIMAL STRATEGIES

- The reservation wage for unemployed workers in city j and the dynamic mobility-compatible wage in city l for a worker employed in city j at wage w are defined by the following equations:

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

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

where:

$$\zeta_{jl} = \frac{r + \delta_l}{r + \delta_j}$$

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

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

• Equations 3 and 4 define a system of J^2 contractions and admit a unique fixed point.

• The optimal strategy when unemployed in city j is:

1. accept any offer φ in city j strictly greater than the reservation wage ϕ_j
2. accept any offer φ in city $l \neq j$ strictly greater than $q_{jl}(\phi_j)$.

The optimal strategy when employed in city j at wage w is:

1. accept any offer φ in city j strictly greater than the present wage w
2. accept any offer φ in city $l \neq j$ strictly greater than $q_{jl}(w)$.

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

Interpretation The interpretation of Equation 3 is straightforward: the difference in the instantaneous values of unemployment and employment ($\phi_j - b$) 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. Those are made of two elements: the expected wages that will be found through local job search and the expected wages that will be found through mobile job search, net of mobility costs.

The interpretation of Equation 4 is similar. Here, the difference in the instant values of employed workers in location l and location j is $[q_{jl}(w) + \gamma_l] - \zeta_{jl}[w + \gamma_j]$. The term $[\zeta_{jl}\gamma_j - \gamma_l]$ is a measure of the relative attractiveness of city j and city l . The third term states that job offers can only attract jobseekers from elsewhere if they are high enough to overcome the mobility costs: this is the direct effect of mobility costs. As for the difference in the option values of employment in city j and employment in city l , it is threefold. The first part is independent of the wage level and given by the difference in the value of unemployment, weighted by unemployment risks δ_j and δ_l . The second part is the difference in the expected wage following a local job-to-job transition and the third part is the difference in the expected wages that will be found through mobile job search, net of mobility costs.

This last term introduces the relative centrality and the relative accessibility of city j and city l . Centrality stems from the comparison of the strength of spatial frictions between the two locations j and l and the rest of the world: a worker living in city j who receives an offer from city l must take into account the respective

spatial frictions from city j and from city l to any tier location k that she may face in the future, in order to maximize her future job-offer rate. As for accessibility, it stems from the difference in the expected costs associated with mobile on-the-job search from city j and from city l : an individual living in city j who receives an offer from city l must take into account the respective mobility cost from city j and from city l to any tier location k that she may face in the future, in order to minimize the cost associated with the next move. Note that both the relative centrality and the relative accessibility measures depend on the current wage level w : cities may be more or less central and accessible depending on where workers stand in the earning distribution.

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 $(j, l) \in \mathcal{J} \times \mathcal{J}_j$. 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 . Despite the existence of mobility costs, some between-cities job-to-job transitions may allow for wage cuts.

Mobility costs and past dependence Switching costs seriously complicate the resolution of dynamic discrete choice models.²³ Despite the Poisson technology, the same is true here, because the introduction of mobility costs yields a non-trivial past dependence in the definition of indifference wages. To see why, one can compare the expressions for static and dynamic mobility costs. The difference between $q_{jl}(w)$ and $\chi_{jl}(w)$ includes a dynamic component, which can be seen in Equation 5:

$$q_{jl}(w) = \chi_{jl}(w) + (r + \delta_l)c_{jl} + \lambda_l^e \int_{\chi_{jl}(w)}^{q_{jl}(w)} \Xi_l(x) dx + \sum_{k \in \mathcal{J}_l} s_{lk}^e \lambda_k^e \int_{q_{lk}(\chi_{jl}(w))}^{q_{lk}(q_{jl}(w))} (\Xi_k(x) - f_k(x)c_{lk}) dx \quad (5)$$

Positive mobility costs will deter future wage growth prospects by imposing a higher wage level to start with. Pushing the argument to the extreme, this means that a very mobile jobseeker will eventually become unable to find a job. This also means that the mobility decision of a worker who has had one back-and-forth mobility experience, even for an infinitesimal length of time, will differ from the mobility decision of a worker without this experience. We believe that such extreme patterns should not drive the mobility decision. In addition, they prevent from finding a tractable closed-form expression for $q_{jl}(w)$.

To overcome these issues, we assume that jobseekers facing on-the-job-search prospects between cities do not take into account the wage supplement associated with past mobility costs.²⁴ Under this assump-

²³As shown by Banks & Sundaram (1994), a Gittins index solution of the multiarmed bandit problem cannot be applied when switching between two arms is costly, even when the cost is uniform across all possible moves.

²⁴The mobility cost c_{jl} is to be considered as an advance which will be paid back by the worker at rate $(r + \delta_l)$ once settled in the new city l . However, if the same worker moves again to a city m , she will stop paying for the first move and will only start reimbursing the second move. A worker paid at wage w in city j is characterized by a spatial wage path made of the set of static indifference wages starting from city j and the hypothetical moves considered by the worker do not alter the structure of this wage

tion, Equation 5 becomes:

$$q_{jl}(w) = \chi_{jl}(w) + (r + \delta_l)c_{jl} \quad (6)$$

with:

$$\begin{aligned} \chi_{jl}(w) &= \zeta_{jl}w + [\zeta_{jl}\gamma_j - \gamma_l] + [\zeta_{jl}\delta_j V_j^u - \delta_l V_l^u] + [\zeta_{jl}\lambda_j^e \int_{\bar{w}}^{\bar{w}} \Phi_j(x) dx - \lambda_l^e \int_{\chi_{jl}(w)}^{\bar{w}} \Phi_l(x) dx] \\ &\quad + \zeta_{jl} \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \left(\int_{\chi_{jk}(w)}^{\bar{w}} \Phi_k(x) dx - \bar{F}_k(\chi_{jk}(w))c_{jk} \right) - \sum_{k \in \mathcal{J}_l} s_{lk}^e \lambda_k^e \left(\int_{\chi_{lk}(w)}^{\bar{w}} \Phi_k(x) dx - \bar{F}_k(\chi_{lk}(w))c_{lk} \right) \\ V_j^u &= \frac{1}{r} \left[b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\bar{w}} \Phi_j(x) dx + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \left(\int_{\chi_{jk}(\phi_j)}^{\bar{w}} \Phi_k(x) dx - \bar{F}_k(\chi_{jk}(\phi_j))c_{jk} \right) \right] \\ \Phi_j(x) &= \frac{\bar{F}_j(x)}{r + \delta_j + \lambda_j^e \bar{F}_j(w) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \bar{F}_k(\chi_{jk}(w))} \end{aligned}$$

The tractability of these modified indifference wages stems from the stationarity property of static indifference wages, whereby $\chi_{lk}(\chi_{jl}(w)) = \chi_{jk}(w)$ (for complete derivation, see Appendix A.3).

Strategic unemployment and the minimum wage Another issue with the definition of optimal strategies arises at the lower bound of the wage distribution. In principle, spatial segmentation would allow for an unusual individual behavior: a geographic 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 $\max_{j \in \mathcal{J}} \{\phi_j\} \leq \underline{w}$. 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.²⁵

path. This assumption reminds of the Pandora stopping problem described in Weitzman (1979): a worker who has moved from city j to city l but is now considering the next move into city m , takes into account as a fallback value her initial discounted utility $V_j^e(w)$, i.e, a starting wage equal to $\chi_{jl}(w)$ in city l . This does not mean that mobility costs do not impact each static indifference wage separately: for $\chi_{jl}(w)$, the relative accessibility of city j and city l still matters, through the comparison between future mobility costs c_{jk} and c_{lk} . A realistic behavioral interpretation of this assumption is that workers paid w in city j may be able to gather information about the prospects of their counterparts in another city l (other workers paid at $\chi_{jl}(w)$) but they cannot gather information about workers just like them who would have experienced the exact mobility from a wage w in city j to a wage $q_{jl}(w)$ in city l .

²⁵See online appendix in Cahuc et al. (2006).

3.3 Steady-State Equilibrium

As already explained in section 2, 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. We now describe the theoretical counterparts to these three equilibrium components.

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

Proposition 2 *EQUILIBRIUM UNEMPLOYMENT* In equilibrium, the unemployment rate in each location j is given by:

$$\frac{u_j}{m_j} = \frac{\delta_j}{\delta_j + \lambda_j^u \bar{F}_j(\phi_j) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\phi_j))} \quad (7)$$

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 workers in city j who find and accept another job in any city $k \neq j$; conversely, inflows are composed by employed and unemployed workers in any city $k \neq j$ who find and accept a job in city j . The equality between population inflow and outflow defines the following equation:

$$(m_j - u_j) \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\bar{w}} \bar{F}_k(q_{jk}(x)) dG_j(x) + u_j \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\phi_j)) \equiv \lambda_j^e \sum_{k \in \mathcal{J}_j} s_{kj}^e (m_k - u_k) \int_{\underline{w}}^{\bar{w}} \bar{F}_j(q_{kj}(x)) dG_k(x) + \lambda_j^u \sum_{k \in \mathcal{J}_j} s_{kj}^u u_k \bar{F}_j(q_{kj}(\phi_k)) \quad (8)$$

After plugging Equation 7 into Equation 8, one can recover a closed form solution for the system, which can be written as:

$$\mathcal{A} \mathbf{m} = 0 \quad (9)$$

where \mathbf{m} is the vector of $\{m_j\}_{j \in \mathcal{J}}$ and \mathcal{A} is the matrix of typical element (\mathcal{A}_{jl}) defined by:

$$\mathcal{A}_{jl} = \begin{cases} \frac{\left(\lambda_j^u \bar{F}_j(\phi_j) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\phi_j)) \right) \times \left(\sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\bar{w}} \bar{F}_k(q_{jk}(x)) dG_j(x) \right) + \delta_j \left(\sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\phi_j)) \right)}{\delta_j + \lambda_j^u \bar{F}_j(\phi_j) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\phi_j))} & \text{if } j = l \\ \frac{\left(\lambda_j^u \bar{F}_j(\phi_j) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\phi_j)) \right) \times \left(s_{lj}^e \lambda_j^e \int_{\underline{w}}^{\bar{w}} \bar{F}_j(q_{lj}(x)) dG_l(x) \right) + \delta_l s_{lj}^u \left(\lambda_j^u \bar{F}_j(q_{lj}(\phi_l)) \right)}{\delta_j + \lambda_j^u \bar{F}_j(\phi_j) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\phi_j))} & \text{if } j \neq l \end{cases}$$

This yields the following proposition:

Proposition 3 EQUILIBRIUM POPULATION

In equilibrium, the distribution of city sizes is the positive vector $\mathbf{m} \in \ker \mathcal{A}$ s.t. $\sum_{j \in \mathcal{J}} m_j = M$.

Note that Equation 8 defines a relationship between m_j and all the other city sizes in \mathbf{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 workers do not change locations without finding a job. For this reason, the flow of workers into unemployment in city j is only composed of workers previously located in city j , whereas in Equation 8, the population in city j is also determined by the flow of workers who come from everywhere else and have found a job in city j .

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 workers who found a better match. If it is located in city j , such match will correspond to a wage higher than w . However, if it is located in any city $k \neq j$, this match will only 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 workers who find and accept a job in city j with a wage lower than w . These workers may come from city j or from any city $k \neq j$. However, they will only accept such a job if w is higher than their reservation wage ϕ_j or than the mobility-compatible indifference wage of their reservation wage $q_{kj}(\phi_k)$. The second element of inflows is made of workers who were previously employed in any city $k \neq j$ at a wage x lower than the 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 10:

$$\begin{aligned}
(m_j - u_j) \left[G_j(w) (\delta_j + \lambda_j^e \bar{F}_j(w)) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^w \bar{F}_k(q_{jk}(x)) dG_j(x) \right] \equiv & \quad (10) \\
\lambda_j^u \left[\psi_{jj}(w) u_j (F_j(w) - F_j(\phi_j)) + \sum_{k \in \mathcal{J}_j} s_{kj}^u \psi_{kj}(w) u_k (F_j(w) - F_j(q_{kj}(\phi_k))) \right] \\
+ \lambda_j^e \sum_{k \in \mathcal{J}_j} s_{kj}^e (m_k - u_k) \int_{\underline{w}}^{q_{jk}(w)} [F_j(w) - F_j(q_{jk}(x))] dG_k(x)
\end{aligned}$$

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 workers 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 10 with respect to w . This yields the following linear system of functional differential equations:

$$f_j(w) = \frac{g_j(w) (m_j - u_j) \left[\delta_j + \lambda_j^e \bar{F}_j(w) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \bar{F}_k(q_{jk}(w)) \right]}{\lambda_j^u \left(\psi_{jj}(w) u_j + \sum_{k \in \mathcal{J}_j} s_{kj}^u \psi_{kj}(w) u_k \right) + \lambda_j^e \left((m_j - u_j) G_j(w) + \sum_{k \in \mathcal{J}_j} s_{kj}^e (m_k - u_k) G_k(q_{jk}(w)) \right)} \quad (11)$$

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 9, the uniqueness of the solution is not guaranteed. We defer the question of identification to section 4.4. 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 11.

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

1. The reservation wage strategy in Equation 3 describes the optimal job acceptance behavior of immobile unemployed workers.
2. The optimal mobility strategy between two locations is defined by the indifference wage described in Equation 6.
3. The optimal set of unemployment rates is given by Equation 7.

4. The optimal set of city sizes is solution to the linear system 9.
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 11.

4 Estimation

4.1 Simulated Method of moments

The model is estimated by simulated method of moments (hereafter, SMM).²⁶ 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 moments. The SMM consists in finding the parameters $\hat{\theta}$ that minimize the criterion function $\mathcal{L}(\theta)$ given by:

$$\mathcal{L}(\theta) = -\frac{1}{2}(\hat{m} - m(\theta))^T \widehat{W}^{-1}(\hat{m} - m(\theta)) \quad (12)$$

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

4.2 Algorithm and numerical solutions

For notational convenience, let $G(\cdot) \equiv \{G_j(\cdot)\}_{j \in \mathcal{J}}$, $g(\cdot) \equiv \{g_j(\cdot)\}_{j \in \mathcal{J}}$ and $q(\cdot) \equiv \{q_{jl}(\cdot)\}_{(j,l) \in \mathcal{J} \times \mathcal{J}_j}$. 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 4 to recover indifference wages $q(\cdot)$
4. Solve Equation 9 to recover equilibrium population \mathcal{M}
5. Solve Equation 11 to update the distribution of job offers $F(\cdot)$
6. Update θ using the maximum of $\mathcal{L}(\theta)$.
7. Repeat steps 3 to 6 until convergence.

²⁶See [Gourieroux, Monfort & Renault \(1993\)](#) for details.

Indifference wages The model raises several numerical challenges, in particular in steps 3 and 5. In step 3, $q(\cdot)$ defines a system of $J^2 - J$ equations, to be solved $\dim(w)$ times. Moreover, since $\Phi_j(\cdot)$ is a function of all $\{\chi_{jk}(\cdot)\}_{k \in \mathcal{J}_j}$, the numerical integration of $\Phi_j(\cdot)$ requires a prior knowledge of the functional form of all $\{\chi_{jk}(\cdot)\}_{k \in \mathcal{J}_j}$. A potential solution to this problem would be to parametrize $q(\cdot)$ as a polynomial function of wages and structural parameters. However, this would obliterate any prospect to identify separately mobility costs, amenities and labor market matching parameters.²⁷ Instead, we take advantage of the exact structure of the model and we use 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:

$$\chi_{jl}(w^*) = \chi_{jl}(w) + h d\chi_{jl}(w) \quad (13)$$

Indifference wages can then be recovered using the following sequential process, on a grid of wages w_i :

- 3.1 Declare an initial guess for $\chi(w)$
- 3.2 Use the values of $\chi(w)$ to iteratively recover $\chi(w_1)$, then $\chi(w_2)$, up to $\chi(\bar{w})$.
- 3.3 Use the sequence $\chi(w_1), \dots, \chi(\bar{w})$ to approximate the function $\chi(\cdot)$.
- 3.4 Use the function $\chi(\cdot)$ to update the value of $\chi(w)$
- 3.5 Repeat steps 3.2 to 3.4 until convergence.

In step 3.1, $\chi(w)$ is determined by setting the value of each integral in Equation 6 to the discounted average wage $\{\frac{\bar{w}_j}{r}\}_{j \in \mathcal{J}}$ observed in each city:

$$\begin{aligned} \chi_{jl}^0(w) &= \zeta_{jl} w + [\zeta_{jl} \gamma_j - \gamma_l] + \frac{\zeta_{jl} \delta_j}{r} \left[b + \gamma_j + \lambda_j^u \frac{\bar{w}_j}{r} + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \left(\frac{\bar{w}_k}{r} - c_{jk} \right) \right] \\ &- \frac{\delta_l}{r} \left[b + \gamma_l + \lambda_l^u \frac{\bar{w}_l}{r} + \sum_{k \in \mathcal{J}_l} s_{lk}^u \lambda_k^u \left(\frac{\bar{w}_k}{r} - c_{lk} \right) \right] + \zeta_{jl} \left[\lambda_j^e \frac{\bar{w}_j}{r} + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \left(\frac{\bar{w}_k}{r} - c_{jk} \right) \right] - \left[\lambda_l^e \frac{\bar{w}_l}{r} + \sum_{k \in \mathcal{J}_l} s_{lk}^e \lambda_k^e \left(\frac{\bar{w}_k}{r} - c_{lk} \right) \right] \end{aligned} \quad (14)$$

In step 3.2, use Equation 13 to move up the wage ladder. For each value w of the grid, and for each city j , $d\chi_{jl}(w)$ are solutions to the linear system 15, given by:

$$\mathcal{B}^j(w) d\chi_j(w) = \mathcal{C}^j(w) \quad (15)$$

where $d\chi_j(w)$ is the $(J-1)$ -vector of $\{d\chi_{jl}(w)\}_{l \in \mathcal{J}_j}$, $\mathcal{B}^j(w)$ is the $(J-1, J-1)$ -matrix of typical element

²⁷See section 4.4 for details.

$\{\mathcal{B}_{lk}^j(w)\}_{(l,k) \in \mathcal{J}_j^2}$ and $\mathcal{C}^j(\mathbf{w})$ is the $(J-1)$ -vector of typical element $\{\mathcal{C}_l^j(w)\}_{l \in \mathcal{J}_j}$ defined by:

$$\begin{aligned} \mathcal{B}_{lk}^j(w) &= \begin{cases} 1 - \lambda_l^e \left[\Phi_l(\chi_{jl}(w)) - \zeta_{jl} s_{jl}^e \left[\Phi_l(\chi_{jl}(w)) - f_l(\chi_{jl}(w)) c_{jl} \right] \right] & \text{if } k = l \\ \lambda_k^e \left[\zeta_{jl} s_{jk}^e \left[\Phi_k(\chi_{jk}(w)) - f_k(\chi_{jk}(w)) c_{jk} \right] - s_{lk}^e \left[\Phi_k(\chi_{jk}(w)) - f_k(\chi_{jk}(w)) c_{lk} \right] \right] & \text{if } k \neq l \end{cases} \\ \mathcal{C}_l^j(w) &= \zeta_{jl} - \lambda_j^e \left[\zeta_{jl} \Phi_j(w) - s_{lj}^e \left[\Phi_j(w) - f_j(w) c_{lj} \right] \right] \end{aligned}$$

Steps 3.3 and 3.4 use standard linear interpolation and extrapolation techniques. Finally, $q_{jl}(\cdot)$ is recovered through Equation 6.

Functional equations Once the indifference wages are recovered, we can turn to the evaluation of the wage distributions (step 5 in the general algorithm). There are two difficulties when solving for the system defined by Equation 11. First, for any $J \geq 3$, the system can only be solved numerically.²⁸ Second, the system is composed of functional equations, which standard differential solvers are not designed to handle. Our solution is twofold. First, in order to reduce the computational burden and ensure the smoothness of the density functions, we assume that $F(\cdot)$ follows a parametric distribution:

$$\hat{F}_j(x) = \text{gammacdf}(x, \alpha_j, \beta_j) \quad (16)$$

where $\text{gammacdf}(\cdot, \alpha_j, \beta_j)$ is the CDF of a gamma distribution with shape parameter α_j and scale parameter β_j .²⁹ Then, since our empirical counterparts are based on real wages, we treat the empirical cdf G as unknown and we estimate the set of parameters $\alpha \equiv \{\alpha_j\}_{j \in \mathcal{J}}$ and $\beta \equiv \{\beta_j\}_{j \in \mathcal{J}}$ which minimize the distance between the empirical cdf G and its theoretical counterpart. This theoretical counterpart is given as the solution to the following functional equation, derived from Equation 11:

$$g_j(w) = \frac{f_j(w) \left[\lambda_j^u \left(\psi_{jj}(w) u_j + \sum_{k \in \mathcal{J}_j} s_{kj}^u \psi_{kj}(w) u_k \right) + \lambda_j^e \left((m_j - u_j) G_j(w) + \sum_{k \in \mathcal{J}_j} s_{kj}^e (m_k - u_k) G_k(q_{jk}(w)) \right) \right]}{(m_j - u_j) \left[\delta_j + \lambda_j^e \bar{F}_j(w) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \bar{F}_k(q_{jk}(w)) \right]} \quad (17)$$

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) which corresponds to \tilde{G} , the gamma approximation of G . At step 5, we need a solution $G(\cdot)$ to Equation 17 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 11 becomes a standard ODE.

²⁸Two-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.

²⁹Meghir et al. (2012) use beta distributions.

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 $j \in \mathcal{J}$.

5.4 Use estimates of $G_l(q_{jl}(w))$ to solve the functional differential equation 17.

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 $j \in \mathcal{J}$.³⁰ Hence, for any $w_1 = w_0 + h$, we can write:

$$G_j(w_1) = G_j(w_0) + hg_j(w_0) \quad (18)$$

and we iterate until reaching the maximum wage, \bar{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.

4.3 Parametrization

The model is based on a set of parameters $\theta = \{\lambda_j^e, \lambda_j^u, \delta_j, s_{jk}^e, s_{jk}^u, c_{jk}, \gamma_j, \alpha_j, \beta_j\}_{(j,k) \in \mathcal{J} \times \mathcal{J}}$ such that $|\theta| = 120,600$ with $J = 200$. In practice, estimating parameters s_{jl}^i and c_{jl} for each pair of cities would be too computationally demanding and would require to drastically restrict \mathcal{J} . We take an alternative path and we posit and estimate two parsimonious parametric models:

$$s_{jl}^i = \frac{\exp(s_{j0}^i + s_{0l}^i + s_1^i d_{jl} + s_2^i d_{jl}^2 + s_3^i h_{jl} + s_4^i h_{jl}^2)}{1 + \exp(s_{j0}^i + s_{0l}^i + s_1^i d_{jl} + s_2^i d_{jl}^2 + s_3^i h_{jl} + s_4^i h_{jl}^2)} \quad (19)$$

$$c_{jl} = c_0 + c_1 d_{jl} + c_2 d_{jl}^2 \quad (20)$$

where s_{j0}^i and s_{0l}^i are city-position (either on the sending or the receiving end of the job offer) fixed effects, 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.³¹

The model rests upon the premise that spatial friction parameters take on values between 0 and 1. Given the lack of existing literature on the explicit structure of spatial frictions, we choose to use a logistic function in Equation 19 because of its analytical properties.³² On the contrary, we do not constraint the range of possible values taken by mobility costs. One plausible interpretation of negative mobility costs would be relocation subsidies. Equation 19 is akin to a standard gravity equation: the fixed effects measure the

³⁰See Judd (1998) for details.

³¹We use the traditional Duncan index: if v is a categorical variable defined by categories k in proportions $v_j(k)$ and $v_l(k)$ in cities j and l , $h_{jl} = \sum_k |v_j(k) - v_l(k)|$. In order to construct this variable, we use the 2007 version of a firm-level census called SIRENE.

³²See Zenou (2009a) for a theoretical approach in terms of endogenous search intensity.

relative openness of the local labor markets: either the ability of each city to dispatch its jobseekers to jobs located elsewhere (s_{j0}) or to fill its vacancies with workers coming from other locations (s_{0j}), and the other parameters account for the effect of distance between two locations.³³

Physical distance is arguably the most important characteristic and both equations 19 and 20 rely on physical distance.³⁴ In addition, we allow spatial frictions to be also impacted by another measure of distance: sectoral dissimilarity, which proxies potential coordination frictions between the two locations. This feature is particularly important to rationalize job-to-job mobility rates between highly specialized cities (for example, biotechnologies in Lyon and Strasbourg). We let returns to these two measures of distance vary by considering a second-order polynomial. Note that, in order to ensure continuity at the reservation wage, we assume that moving costs do not vary with labor market status, unlike spatial frictions.³⁵ Finally, we do not allow for fixed effects in Equation 20: as will be made clear in Section 4.4, those fixed effects cannot be identified separately from the local amenity parameters. For this reason, our estimates of mobility costs will depend on the pair of cities involved, but not on the direction of the move. Under these two specifications, the total number of parameters to be estimated amounts to 2,011.

4.4 Identification

Identification is based on Proposition 5 and Table 7.

Proposition 5 *EXISTENCE AND UNIQUENESS OF WAGE OFFER DISTRIBUTIONS* *The system of differential equations $f: \mathbb{R}^J \rightarrow (0, 1)^J$ 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 \(1993\)](#).

Table 7 describes the empirical and theoretical moments used in the estimation. In the third column, identifying parameters must be understood as the main parameters involved in the comparison of the two moments, even though all parameters are, obviously, related to each other, in particular through the indifference wages.

³³See [Head & Mayer \(2013\)](#) for the current state of the art about gravity equations.

³⁴As shown, among others, by [Combes & Lafourcade \(2005\)](#) on trade between French cities, physical distance is a very good proxy

Table 7: Moments and Identification

Empirical moments	Theoretical moments	Identifying Parameters
Unemployment rate in city $j \in \mathcal{J}$	u_j/m_j	δ_j, λ_j^u
Labor force in city $j \in \mathcal{J}$	m_j	δ_j, λ_j^u
Transition rate ee within city $j \in \mathcal{J}$	$\lambda_j^e \int_{\underline{w}}^{\bar{w}} \bar{F}_j(x) dG_j(x)$	λ_j^e
Earning distribution in city $j \in \mathcal{J}$	G_j	α_j, β_j
Transition rate ue out of city $j \in \mathcal{J}$	$\sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(\underline{w}))$	s_{j0}^u
Transition rate ue into city $l \in \mathcal{J}$	$\lambda_l^u \sum_{k \in \mathcal{J}_l} s_{kl}^u \bar{F}_l(q_{kl}(\underline{w}))$	s_{0l}^u
Transition rate ee out of city $j \in \mathcal{J}$	$\sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\bar{w}} \bar{F}_k(q_{jk}(x)) dG_j(x)$	s_{j0}^e
Transition rate ee into city $l \in \mathcal{J}$	$\lambda_l^e \sum_{k \in \mathcal{J}_l} s_{kl}^e \int_{\underline{w}}^{\bar{w}} \bar{F}_k(q_{kl}(x)) dG_k(x)$	s_{0l}^e
Transition rate ue from city j to city l , $(j, l) \in \mathcal{T}_1$	$s_{jl}^u \lambda_l^u \bar{F}_l(q_{jl}(\underline{w}))$	$s_1^u, s_2^u, s_3^u, s_4^u$
Transition rate ee from city j to city l , $(j, l) \in \mathcal{T}_1$	$s_{jl}^e \lambda_l^e \int_{\underline{w}}^{\bar{w}} \bar{F}_l(q_{jl}(x)) dG_j(x)$	$s_1^e, s_2^e, s_3^e, s_4^e$
Accepted wage ee into city $l \in \mathcal{J}$	$\frac{\sum_{k \in \mathcal{J}_l} s_{kl}^e \int_{\underline{w}}^{\bar{w}} \bar{F}_k(q_{kl}(x)) dG_k(x) q_{kl}(w_{e_k}^{init})}{\sum_{k \in \mathcal{J}_l} s_{kl}^e \int_{\underline{w}}^{\bar{w}} \bar{F}_k(q_{kl}(x)) dG_k(x)}$	γ_l
Accepted wage ee between city j and city l , $(j, l) \in \mathcal{T}_2$	$q_{jl}(w_{e_j}^{init})$	c_0, c_1, c_2

Notes: For details on the construction of the empirical moments, see Appendix B.2.

As shown by [Flinn & Heckman \(1982\)](#) and [Magnac & Thesmar \(2002\)](#), structural parameters are identified from transition rates. Transitions out of 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. Moreover, given the structure of the DADS data, we think that the measurement of transitions out of and into unemployment may sometimes lack precision.

Given the parametrization of s_{jl}^i , 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_{jl}^i are no longer needed. In order to identify the fixed-effect components, we use the $2J$ total transitions rates into and out of any given city. On the other hand, the identification of the parameters related to the distance and the dissimilarity between

for transport cost in a cross-section analysis.

³⁵This assumption may not be fully innocuous if unemployed jobseekers have access to some specific segments of the housing market, such as public housing.

two cities does still requires transition rates at the city-pair level. Given that Equation 19 only specifies four parameters for each labor market status, we drastically restrict the set of city pairs, down to a subset $\mathcal{T}_1 \subset \mathcal{J} \times \mathcal{J}_j$, with $|\mathcal{T}_1| = 48$, which we use in the estimation.³⁶

While spatial friction parameters are identified from transition rates between pairs of cities, mobility costs and local amenities parameters are identified from wage moments. This strategy is made possible because we do not use wage data to approximate indifference wages.³⁷ The amenity parameter γ_l is identified by the average wage accepted by agents who have experienced a job to job transition with migration into city l . The mechanism is simple: if this average accepted wage into city l is higher (resp. lower) than predicted by the labor market parameters and the level of centrality of city l , this means that city l is less (resp. more) attractive, on average, than the other cities in the system. The identification of c_{jl} goes along the same lines: if the average wage accepted in city l by jobseekers initially located in city j differs from what is predicted by the labor market parameters, the level of centrality and the level of attractiveness of city j and city l , this difference will be attributed to the specific distance between the two cities.

To be more specific, let $w_{e_j e_j}^{init}$ 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 $q_{jl}(w_{e_j e_j}^{init})$ and $w_{e_j e_j}^{init}$ (which, by definition, is equal to $q_{jj}(w_{e_j e_j}^{init})$). The corresponding empirical moment is the difference between $w_{e_j e_l}^{fin}$ (the average wage after a job-to-job transition from city j to city l), and $w_{e_j e_j}^{fin}$ (the average wage after a job-to-job transition within city j).³⁸ Under the assumption that, conditional on the local parameter values, jobseekers are as likely to draw a wage above their indifference wage when they do a job-to-job transition without mobility and when they do a job-to-job transition with mobility, these differences identify the mobility cost c_{jl} .³⁹ Given there only are three parameters to estimate, we select a subset of city pairs $\mathcal{T}_2 \subset \mathcal{J} \times \mathcal{J}_j$ such that $|\mathcal{T}_2| = 12$.⁴⁰ Finally, the identification of γ_l is based on the average of these previous differences in accepted wages, for all job-to-job transitions into city l .

³⁶In practice, we use the off-the-job and job-to-job transitions rates from the urban areas ranked fourth to eleventh (Toulouse, Lille, Bordeaux, Nice, Nantes, Strasbourg, Grenoble and Rennes) to the urban areas ranked fifteenth, nineteenth to twenty-second and twenty-fifth (Montpellier, Clermont-Ferrand, Nancy, Orléans, Caen and Dijon). This selection is designed to include locations that are widely scattered across the French territory (see Figure 8 in appendix C for details).

³⁷This is our main departure from Meghir et al. (2012), who have to use wage data to recover subsequent optimal contracts numerically.

³⁸In Table 6, this corresponds to the difference between the value in the off-diagonal cases and the value in the diagonal case on the same line.

³⁹If we did not use this differential approach, we would have to use the minimum observed values of accepted wages, which is not as well-behaved and would allow for more sampling error.

⁴⁰In practice, we use the average accepted wages following a job-to-job transition between the cities ranked second to fifth (Lyon, Marseille, Toulouse and Lille). This subset has to be more restrictive than \mathcal{T}_1 because, while very low transitions rates convey reliable information, they do not allow to compute accurate measures of average accepted wages. Note that for homogeneity concerns, we do not include Paris, because its size is too large compared to the other cities.

4.5 Optimization and fit

We parametrize $b = \text{€}500$ and $r = 13.4\%$.⁴¹ An iteration takes 18 seconds. The model is optimized using a mix of derivative-free and Quasi Newton algorithm (BFGS). The optimum is reached at 3.98. Standard deviations are obtained using a Laplace-Based MCMC starting from the optimum. The model allows to reproduce many features of the data, particularly the distribution of unemployment rates and the transition rates.

5 Results

In this section, we first present our structural estimation results and in particular, the distribution of the city specific parameters and the impact of distance on spatial frictions and mobility costs. We run inference on the determinants of the local matching parameters. We provide a decomposition of city-specific average wages between the impact of on-the-job search and the impact of openness and a decomposition of the aggregate mobility rate between spatial frictions, mobility costs and local amenities. Finally, we present an experiment on the optimal number of cities to minimize aggregate unemployment.

5.1 A dataset of city-specific 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 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 8 complements this presentation by providing the summary statistics of these matching parameters and the between-city job arrival rates.

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.⁴² 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 8, on-the-job search is a crucial

⁴¹The minimum guaranteed income (Revenu Minimum d'Insertion) was about €420 for a single individual during the period under study. Similarly, inflation amounted to 13.4% between January 2002 and December 2007.

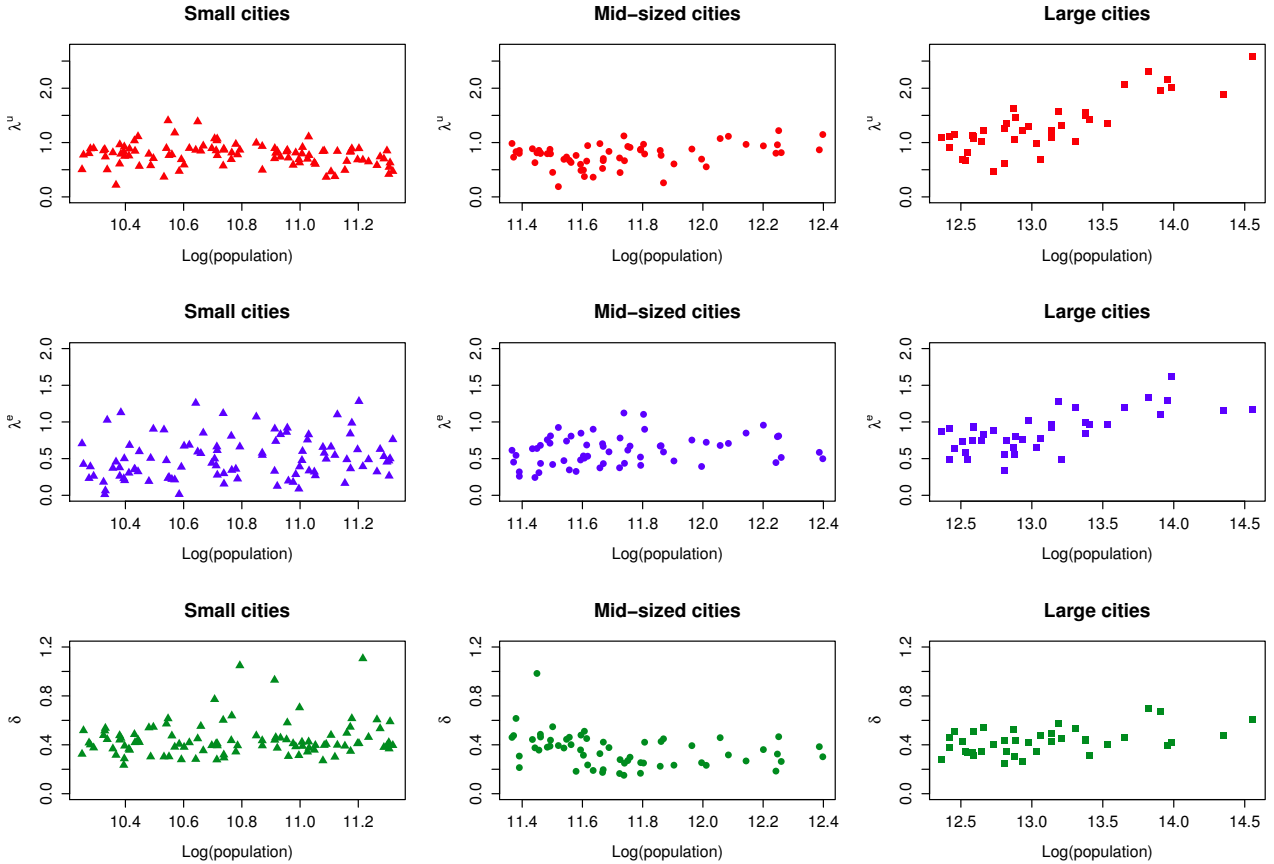
⁴²Recall that these parameters define a matching process that takes place on a six-year span, between 2002 and 2007.

Table 8: The matching parameters: summary statistics

	λ_j^u	λ_j^e	δ_j	$u_j e_l$	$e_j e_l$
Minimum	0.189	0.013	0.000	0.000	0.000
1st Quartile	0.688	0.391	0.001	0.017	0.000
Mean	0.918	0.618	0.413	0.104	0.104
Sd	0.671	0.302	0.141	0.151	0.164
Median	0.833	0.593	0.401	0.02	0.001
3rd Quartile	0.972	0.809	0.462	0.169	0.174
Maximum	8.931	1.660	1.106	0.869	0.643

Notes: (i) $u_j e_l = \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \bar{F}_k(q_{jk}(w))$ is the transition rate out of unemployment with geographical mobility; (ii) $e_j e_l = \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_w^{\bar{w}} \bar{F}_k(q_{jk}(x)) dG_j(x)$ is the job-to-job transition rate with geographical mobility.

Figure 5: The matching parameters along the city size distribution



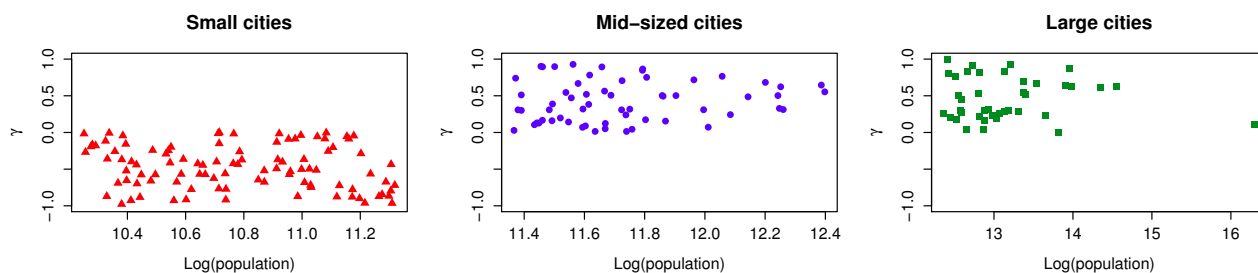
Notes: (i) Estimated values of the structural parameters ($\lambda^u, \lambda^e, \delta$) for the 40 largest cities, the 100 smallest cities and the 60 cities in between; (ii) For the sake of exposition, we do not represent Paris; its $\log(\text{population})$ amounts to 16.30, $\lambda_{Paris}^u = 8.93$, $\lambda_{Paris}^e = 1.66$ 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.

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.

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 6, 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 5, 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 the end of Section 2.2.⁴³

Figure 6: 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.

The structure of a local labor market We now turn to the correlation between our parameters and between our parameters and the local wage distributions. In Table 9, these distributions are summarized by their first two moments. The three main findings are the following. First, panel 1 shows that there is a

⁴³One 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.

positive correlation between off-the-job and on-the-job arrival rates.⁴⁴ 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 workers 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 5.3 will confirm this hypothesis using a wage decomposition approach.

Table 9: Correlation between the local labor market primitives

Panel 1: All cities					Panel 2: Large cities				
	λ^u	λ^e	δ	\bar{w}		λ^u	λ^e	δ	\bar{w}
λ^e	0.44***				λ^e	0.61***			
δ	-0.03	0.04			δ	-0.18	0.25		
\bar{w}	0.46***	0.29***	-0.23**		\bar{w}	0.74***	0.64***	-0.14	
σ_w	0.42***	0.42***	-0.02	0.73***	σ_w	0.76***	0.78***	0.06	0.84***

Panel 3: Mid-sized cities					Panel 4: Small cities				
	λ^u	λ^e	δ	\bar{w}		λ^u	λ^e	δ	\bar{w}
λ^e	0.09				λ^e	0.01			
δ	-0.06	0.02			δ	0.09	0.04		
\bar{w}	0.14	0.07	-0.31*		\bar{w}	0.09	-0.10	-0.21*	
σ_w	0.00	0.20	-0.06	0.55***	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) \bar{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

The observable determinants of local labor market primitives In order to complete the previous observations, 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

⁴⁴Note that this correlation may partly reflect a co-dependence to a third variable, such as city size.

share of the population without qualifications, the share of males, the share of blue-collar jobs and the share of jobs in the manufacturing sector. Results are presented in Table 14. 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.

5.2 The impact of distance on spatial constraints

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 $\sum_{k \in \mathcal{J}_j} s_{jk}^i \lambda_k^i$ 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 .⁴⁵ 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 workers 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, $\sum_{k \in \mathcal{J}_M} s_{Mk}^u \lambda_k^u = 0.9$, than in Lyon, $\sum_{k \in \mathcal{J}_L} s_{Lk}^u \lambda_k^u = 0.01$. 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

⁴⁵Note 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)

and 84, respectively), faces a rate of 0.69.⁴⁶

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 4.3, these two dimensions are measured in relative terms, by the spatial distance and the sectoral dissimilarity between each pair of cities.⁴⁷ 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 19 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 $(j, l) \in \mathcal{J} \times \mathcal{J}_j$ 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.⁴⁸ Both physical distance and sectoral dissimilarity increase the level of spatial frictions as expected. Moreover, the effect is much stronger for employed workers. 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 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)			
	$\frac{\partial s_{jl}^u}{\partial d_{jl}}$	$\frac{\partial s_{jl}^e}{\partial d_{jl}}$	$\frac{\partial s_{jl}^u}{\partial h_{jl}}$	$\frac{\partial s_{jl}^e}{\partial h_{jl}}$	$\frac{\partial s_{jP}^u}{\partial d_{jP}}$	$\frac{\partial s_{jP}^e}{\partial d_{jP}}$	$\frac{\partial s_{jP}^u}{\partial h_{jP}}$	$\frac{\partial s_{jP}^e}{\partial h_{jP}}$
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

Our estimates for Equation 20 indicate that the mobility cost function is given by $\hat{c}_{jl} = 9.018 + 27.000d_{jl} -$

⁴⁶In 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.

⁴⁷The 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.

⁴⁸Standard errors are available upon request.

$2.998d_{j,l}^2$, 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.⁴⁹ As shown in Table 11, 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.

Table 11: 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

5.3 On-the-job search, openness and the city size wage premium

Section 5.1 has shown that cities of different size exhibit very different features with respect to their internal job matching process. Section 5.2 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

⁴⁹Although, 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.

$\mathbb{E}(w|j) = \int_{\underline{w}}^{\bar{w}} x g_j(x) dx$, where $g_j(\cdot)$ is given by Equation 17:

$$\begin{aligned} \mathbb{E}w|j &= \underbrace{\int_{\underline{w}}^{\bar{w}} x \left(\tau_j(x) \lambda_j^u u_j \right) dx}_{\text{Local off-the-job search}} + \underbrace{\int_{\underline{w}}^{\bar{w}} x \left(\tau_j(x) \lambda_j^e (m_j - u_j) G_j(x) \right) dx}_{\text{Local on-the-job search}} \\ &+ \underbrace{\int_{\underline{w}}^{\bar{w}} x \left(\tau_j(x) \lambda_j^u \sum_{k \in \mathcal{J}_j} s_{kj}^u \psi_{kj}(x) u_k \right) dx}_{\text{Mobile off-the-job search}} + \underbrace{\int_{\underline{w}}^{\bar{w}} x \left(\tau_j(x) \lambda_j^e \sum_{k \in \mathcal{J}_j} s_{kj}^e (m_k - u_k) G_k(q_{jk}(x)) \right) dx}_{\text{Mobile on-the-job search}} \end{aligned} \quad (21)$$

$$\text{where } \tau_j(w) = \frac{f_j(w)}{(m_j - u_j) \left[\delta_j + \lambda_j^e \bar{F}_j(w) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \bar{F}_k(q_{jk}(w)) \right]}.$$

Table 12 reports the decomposition described in Equation 21 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 9 and it is in sharp contrast with Baum-Snow & Pavan (2012), who find that search frictions do not really matter for generating city size wage premia in the US. 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.⁵⁰

5.4 Mobility and unemployment

(forthcoming) We evaluate the respective contributions of spatial frictions and mobility costs to the mobility rate and the unemployment rate.

On the optimal number and size of cities: an urban engineering experiment The previous decompositions do not take into account the interdependence between the different parameters of the economy. We propose a way of overcoming this issue, by taking advantage of the dimensionality of our model. We simulate the general equilibrium consequences of a policy whereby population (and firms) in the smaller cities would be optimally reshuffled into the larger ones in order to minimize aggregate unemployment. We ex-

⁵⁰This 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”.

Table 12: 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 off-the job search	0.52	5.82	5.10	4.31	5.64	17.40	17.54
Local on-the job search	95.50	93.64	78.61	82.36	80.08	82.02	82.00
Mobile off-the job search	0.49	0.02	1.19	0.96	1.32	0.06	0.06
Mobile on-the job search	3.49	0.51	15.09	12.36	12.95	0.52	0.40

Panel 2: mid-sized cities							
	City 71 Bourg	City 72 Tarbes	City 73 Belfort	City 74 St-Quentin	City 75 La Roche	City 76 Vienne	City 77 Évreux
Local off-the job search	37.34	4.13	29.78	63.99	53.23	53.70	50.68
Local on-the job search	19.35	1.16	69.82	35.99	40.47	23.45	21.40
Mobile off-the job search	27.89	67.66	0.40	0.01	3.99	11.48	14.79
Mobile on-the job search	15.42	27.05	0.00	0.00	2.31	11.37	13.12

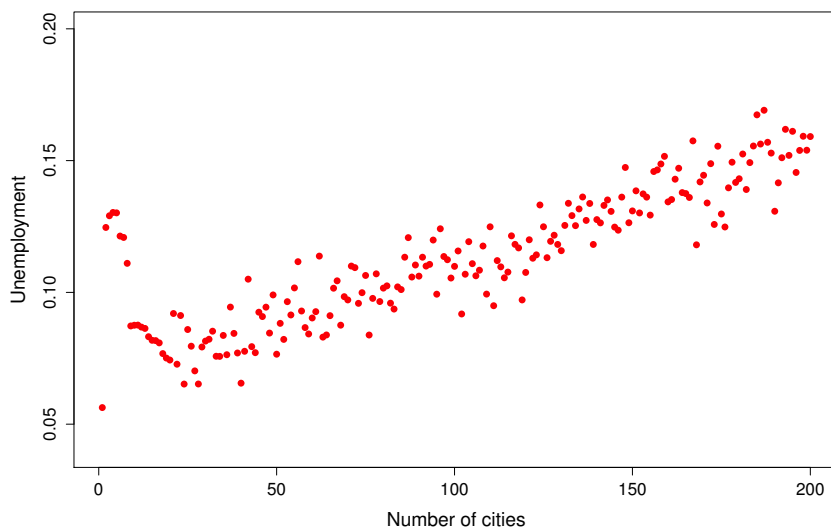
Panel 3: smallest cities							
	City 191 Tulle	City 192 Thann	City 193 Dinard	City 194 Tournon	City 195 Sable	City 196 Pontarlier	City 197 St-Gaudens
Local off-the job search	98.64	60.42	99.73	96.71	87.49	89.49	89.84
Local on-the job search	1.04	23.99	0.21	3.22	4.51	5.99	8.79
Mobile off-the job search	0.32	9.27	0.06	0.06	7.40	4.18	1.23
Mobile on-the job search	0.01	6.32	0.00	0.00	0.60	0.35	0.14

press all the city-specific parameters as a function of the log number of firms, log population and log area of the location. 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. Using the parameters of this regression, we analyze whether there is an optimal country structure, keeping cities' location and relative size fixed.

Figure 7 reports the relationship between the number of cities and aggregate unemployment. It is not trivial.⁵¹ 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.

⁵¹The smallest unemployment rate, 5.6%, is obtained with a single city established in the current geographical setting of Paris. However, this situation is not very meaningful because it relies on the lack of spatial constraints.

Figure 7: 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 vaLocal off-the-job searches of the labor market primitives and fixed effects predicted by the estimation results provided in Table ??.

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. From a theoretical standpoint, in contrast to [Shimer \(2007\)](#)'s mismatch theory, whereby migration decisions are driven by the irrational belief that local economic downturns will eventually reverse, we show that forward-looking profit-maximizers may remain stuck in inauspicious locations. From an empirical standpoint, in contrast to the frictionless economic geography literature, we show that a mere differential in on-the-job search rates can explain most of the city size wage premium, without resorting to a differential in the return to skills. Finally, from a computational standpoint, in contrast to the reference work by [Kennan & Walker \(2011\)](#), we show that the random search technology makes it possible to consider the full state space of a discrete choice model at the city level.

Notwithstanding, our model has several important 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. Cities vary in the number and size of firms and therefore, in the possibilities of wage bargaining they offer. In order to truly understand the contribution of location in the variation of lifetime inequalities, this third dimension cannot be overlooked. However, this extension is far from trivial. In particular, while a single

bargaining parameter may not be very realistic, allowing for spatial heterogeneity in workers' bargaining power would seriously complicate the resolution of the bargaining game. Second, this paper leaves unexplored the firms' side of the dynamic location model. Whereas a mere extension à-la [Meghir et al. \(2012\)](#) would not convey much interest without an explicit theory of location choice, agglomeration economies and wages, we believe such explicit theory to be a promising venue for future research. Finally, another interesting extension would be to allow unemployed workers to move across locations. However, relaxing the immobility assumption for unemployed workers is not easily compatible with the equilibrium assumptions of the model. The introduction of an endogenous housing market might be able to counterbalance the labor market mechanisms, and therefore, to maintain an empirically reasonable scattering of the unemployed population. However, the data requirements of such an augmented model would be very high. Alternatively, a model allowing for heterogeneous preferences over local amenities would be able to generate sorting across locations.

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.
- Banks, J. S. & Sundaram, R. K. (1994). Switching costs and the gittins index. *Econometrica*, 62(3), 687–94.
- Baum-Snow, N. & Pavan, R. (2012). Understanding the city size wage gap. *Review of Economic Studies*, 79(1), 88–127.
- Bayer, P., Fang, H., & McMillan, R. (2011). Separate When Equal? Racial Inequality and Residential Segregation. Working Papers 11-09, Duke University, Department of Economics.
- 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.
- 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. & 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.
- 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., Magnac, T., & Selod, H. (2011). The effect of location on finding a job in the paris region. *Journal of Applied Econometrics*, 26(7), 1079–1112.
- 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.
- Lamadon, T., Lise, J., Meghir, C., & Robin, J.-M. (2013). Matching, sorting and wages. Mimeo.
- Lee, D. & Wolpin, K. I. (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica*, 74(1), 1–46.
- 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.
- Shimer, R. (2007). Mismatch. *American Economic Review*, 97(4), 1074–1101.
- 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.
- Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica*, 47(3), 641–54.
- 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 Reservation wages and mobility-compatible indifference wages

A.1 Expressions

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

$$V_j^u \equiv V_j^e(\phi_j) \quad (22)$$

$$V_j^e(w) \equiv V_l^e(\chi_{jl}(w)) \quad (23)$$

$$V_j^e(w) \equiv V_l^e(q_{jl}(w)) - c_{jl} \quad (24)$$

Equations 1 and 2 can be rewritten as:

$$rV_j^u = b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\bar{w}} (V_j^e(x) - V_j^u) dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \int_{q_{jk}(\phi_j)}^{\bar{w}} (V_k^e(x) - c_{jk} - V_j^u) dF_k(x) \quad (25)$$

$$\begin{aligned} rV_j^e(w) &= w + \gamma_j + \lambda_j^e \int_w^{\bar{w}} (V_j^e(x) - V_j^e(w)) dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{q_{jk}(w)}^{\bar{w}} (V_k^e(x) - c_{jk} - V_j^e(w)) dF_k(x) \\ &+ \delta_j [V_j^u - V_j^e(w)] \end{aligned} \quad (26)$$

After integration by parts of equations 25 and 26, we get:

$$V_j^u = \frac{1}{r} \left[b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\bar{w}} \Xi_j(x) dx + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \left(\int_{q_{jk}(\phi_j)}^{\bar{w}} \Xi_k(x) dx - \bar{F}_k(q_{jk}(\phi_j)) c_{jk} \right) \right] \quad (27)$$

$$V_j^e(w) = \frac{1}{r + \delta_j} \left[w + \gamma_j + \delta_j V_j^u + \lambda_j^e \int_w^{\bar{w}} \Xi_j(x) dx + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \left(\int_{q_{jk}(w)}^{\bar{w}} \Xi_k(x) dx - \bar{F}_k(q_{jk}(w)) c_{jk} \right) \right] \quad (28)$$

where:

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

Finally, using Equations 22 and 24, we find that ϕ_j and $q_{jl}(w)$ are given by Equations 3 and 4.

A.2 Existence and uniqueness

From Equation 4, we derive the following proposition:

Proposition 6 *Let's denote by $\mathcal{W} = [w, \bar{w}]$ the support of the wage distribution. \mathcal{W} is a closed subset of a Banach space. The set of functions $q_{jl}(\cdot)$ defines a contraction. In addition, they have a unique fixed point.*

Proof Consider a grid with minimal vaLocal off-the job search w_0 . Given that q_{jl} is differentiable, equation 6 can be restated in the differential form as:

$$q_{jl}(w) = q_{jl}(w_0) + \int_{w_0}^w h_{jl}(x, q_j(x)) dx, \quad (29)$$

where $q_j(x) \equiv \{q_{jk}(x)\}_{k \in \mathcal{J}_j}$. Starting from initial value local off-the-job search \underline{w} , we can use Picard's iterative process $(q_{jl}^{(1)}, \dots, q_{jl}^{(k)})$ to show that :

$$q_{jl}^{(m)}(w) = K^{(m)}(w_0)(w) \quad (30)$$

with $K(q_{jl})(w) = q_{jl}(w_0) + \int_{\underline{w}}^w h(x, q_j(x)) dx$ and $h_{jl}(x, q_j(x)) = dq_{jl}(x, q_j(x))$. Since $dV_j^e(\cdot) > 0$ and $dV_l^e(\cdot) > 0$, we have $dq_{jl}(\cdot) > 0$; moreover, given that all the structural matching parameters (s^i, λ^i, δ) are positive and the interest rate r is strictly positive, $dq_{jl}(\cdot)$ can be bounded. Therefore, it is easy to see that $dh_{jl}(\cdot) = d^2q_{jl}(\cdot)$ is also bounded. As a consequence, $dq_{jl}(x, q_j(x))$ is *Lipschitz continuous*. The *Banach fixed-point theorem* states that equation 29 has a unique solution.

A.3 Approximate solution for dealing with past dependence

Assumption 1 *MYOPIC MOBILITY*—*Jobseekers facing a mobility decision evaluate the on-the-job-search prospects in the future location without taking into account the wage supplement associated with past mobility costs.*

Given assumption 1, the recursive problem of agents (equations 25 and 26) can be restated as:

$$rV_j^u = b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\bar{w}} (V_j^e(x) - V_j^u) dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \int_{\chi_{jk}(\phi_j)}^{\bar{w}} (V_k^e(x) - c_{jk} - V_j^u) dF_k(x) \quad (31)$$

$$\begin{aligned} rV_j^e(w) &= w + \gamma_j + \lambda_j^e \int_w^{\bar{w}} (V_j^e(x) - V_j^e(w)) dF_j(x) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \int_{\chi_{jk}(w)}^{\bar{w}} (V_k^e(x) - c_{jk} - V_j^e(w)) dF_k(x) \\ &+ \delta_j [V_j^u - V_j^e(w)] \end{aligned} \quad (32)$$

Integrating by part Equations 31 and 32, we obtain:

$$V_j^u = \frac{1}{r} \left[b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\bar{w}} \Phi_j(x) dx + \sum_{k \in \mathcal{J}_j} s_{jk}^u \lambda_k^u \left(\int_{\chi_{jk}(\phi_j)}^{\bar{w}} \Phi_k(x) dx - \bar{F}_k(\chi_{jk}(\phi_j)) c_{jk} \right) \right] \quad (33)$$

$$V_j^e(w) = \frac{1}{r + \delta_j} \left[w + \gamma_j + \delta_j V_j^u + \lambda_j^e \int_w^{\bar{w}} \Phi_j(x) dx + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \left(\int_{\chi_{jk}(w)}^{\bar{w}} \Phi_k(x) dx - \bar{F}_k(\chi_{jk}(w)) c_{jk} \right) \right] \quad (34)$$

where:

$$\Phi_j(x) \equiv \bar{F}_j(x) dV_j^e(x) = \frac{\bar{F}_j(x)}{r + \delta_j + \lambda_j^e \bar{F}_j(w) + \sum_{k \in \mathcal{J}_j} s_{jk}^e \lambda_k^e \bar{F}_k(\chi_{jk}(w))}$$

Using Equations 23 and 24, we find $\chi_{jl}(w)$ and $q_{jl}(w)$, as given in Equation 6.

B Data

B.1 Data selection

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

Finally, we end up with the dataset described in Table 13.

Table 13: Structure of the dataset

Year	Number of Individuals	Number of Observations	Number of obs. by metro				Number of individuals by metro			
			Min	Mean	Median	Max	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 199+1 metropolitan areas in 2010; (ii) Source: Panel DADS 2002-2007

B.2 Empirical moments

The first column in Table 7 is constructed as follows:

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

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

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

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

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

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

Transition rate ee out of city j : ratio of the number of job-to-job transitions out of city j observed over the period, to the potentially-employed population in city j

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

Transition rate ue from city j to city l : ratio of the number of transitions out of unemployment from city j to city l observed over the period, to the potentially-unemployed population in city j

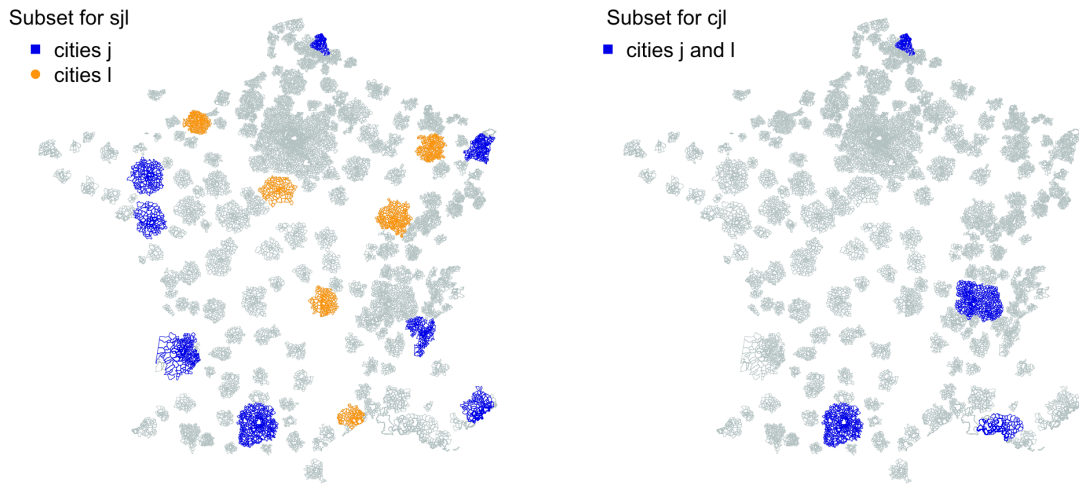
Transition rate ee from city j to city l : ratio of the number of job-to-job transitions from city j to city l observed over the period, to the potentially-employed population in city j

Accepted wages ee into city l : average wage following a job-to-job transition into city l observed over the period; the average is the sum of the accepted wages ee from city j to city l as defined below, weighted by the number of transitions ee from city j to city l

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

C Figures

Figure 8: The metropolitan areas in subset \mathcal{T}_1 (left) and subset \mathcal{T}_2 (right)



Notes: (i) see Figure 2; (ii) Subset \mathcal{T}_1 is used to identify the effect of physical distance and dissimilarity on spatial frictions based on pair-specific out-of-unemployment and job-to-job transition rates; subset \mathcal{T}_2 is used to identify the effect of physical distance on moving costs based on pair-specific average accepted wages after a job-to-job transition with mobility.

Table 14: 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 (1.427)	2.894 (1.617)	1.588 (0.837)	-2.600 (5.683)	-0.058 (4.390)	4.623 (2.324)	1.044 (2.785)	4.146 (2.758)	1.364 (1.701)	-1.354 (1.641)	1.522 (2.381)	-0.547 (0.976)
Number of firms	0.837*** (0.024)	0.101*** (0.027)	-0.024 (0.014)	0.803*** (0.030)	0.059* (0.023)	-0.027* (0.012)	2.194** (0.711)	0.752 (0.704)	-0.442 (0.435)	-1.379 (1.155)	3.222 (1.675)	1.495* (0.686)
Density	0.041 (0.112)	0.229 (0.127)	0.125 (0.066)	0.323 (0.213)	0.281 (0.165)	0.030 (0.087)	-0.072 (0.186)	0.227 (0.184)	0.116 (0.114)	0.172 (0.225)	0.061 (0.326)	0.659*** (0.134)
Young	2.892*** (0.852)	3.512*** (0.966)	-1.054* (0.500)	5.979** (2.036)	-0.704 (1.573)	-0.696 (0.833)	1.661 (1.925)	1.286 (1.906)	-2.068 (1.176)	-2.805 (1.812)	3.416 (2.628)	-2.826* (1.077)
Males	0.782 (3.159)	-4.129 (3.578)	-1.542 (1.852)	5.280 (12.414)	7.313 (9.589)	-7.312 (5.076)	-3.807 (6.274)	-7.933 (6.213)	0.159 (3.833)	4.510 (3.542)	-2.889 (5.137)	2.279 (2.105)
Drop out	0.567 (0.443)	-0.675 (0.502)	0.046 (0.260)	-0.127 (1.360)	-0.424 (1.050)	-0.569 (0.556)	1.112 (0.827)	-0.707 (0.819)	0.190 (0.505)	0.152 (0.606)	0.504 (0.878)	0.128 (0.360)
Blue collar	-0.723 (0.762)	-2.504** (0.863)	-0.568 (0.447)	-0.379 (2.530)	-7.912*** (1.954)	-0.815 (1.034)	1.729 (1.586)	0.438 (1.571)	-2.388* (0.969)	1.060 (1.076)	-1.950 (1.561)	0.970 (0.640)
Manufacturing	0.444 (0.439)	0.507 (0.497)	-0.654* (0.257)	-1.958 (1.684)	-0.052 (1.301)	-0.535 (0.689)	-0.036 (0.759)	0.300 (0.752)	0.187 (0.464)	0.496 (0.529)	0.260 (0.768)	-1.199*** (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: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$; (iii) 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