Job Search and Migration in a System of Cities*

Benoît Schmutz[†] Modibo Sidibé [‡]

April 8, 2015

Abstract

We build a job search model, where workers engage in both off- and on-the-job search over a set of cities, to quantify the impact of spatial matching frictions and mobility costs on the job search process. Migration decisions, based on a dynamic utility trade-off between locations, can rationalize diverse wage dynamics as part of forward-looking spatial strategies. Our estimation results allow us to characterize each of the largest 200 French cities by a set of city-specific matching parameters and to measure the impact of distance on spatial constraints. We find that after controlling for frictions, mobility cost parameters are significantly lower than previously reported in the literature. Additional results include a robust positive correlation between on-the-job arrival rates and local wage dispersion, which provides new empirical support to the wage-posting framework and suggests an alternative explanation for the city size wage gap.

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

JEL Classification: J2; J3; J6

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

---

*We thank Jim Albrecht, Pat Bayer, Bruno Decreuse, Rafael Dix Carneiro, Florence Goffette-Nagot, John Kennan, Thomas Le Barbanchon, Sébastien Roux, John Rust and Maxime Tô, as well as participants to many seminars and conferences, for their help, comments and discussions. This project was initiated when both authors were junior researchers at CREST and pursued when Schmutz was visiting Georgetown University and Sidibé was a post-doctoral fellow at Duke University. Schmutz also benefited from a Faculty Summer Research Fellowship from Howard University. The generosity of these four institutions is gratefully acknowledged. Data was made available through the Centre d'Accès Sécurisé à Distance.

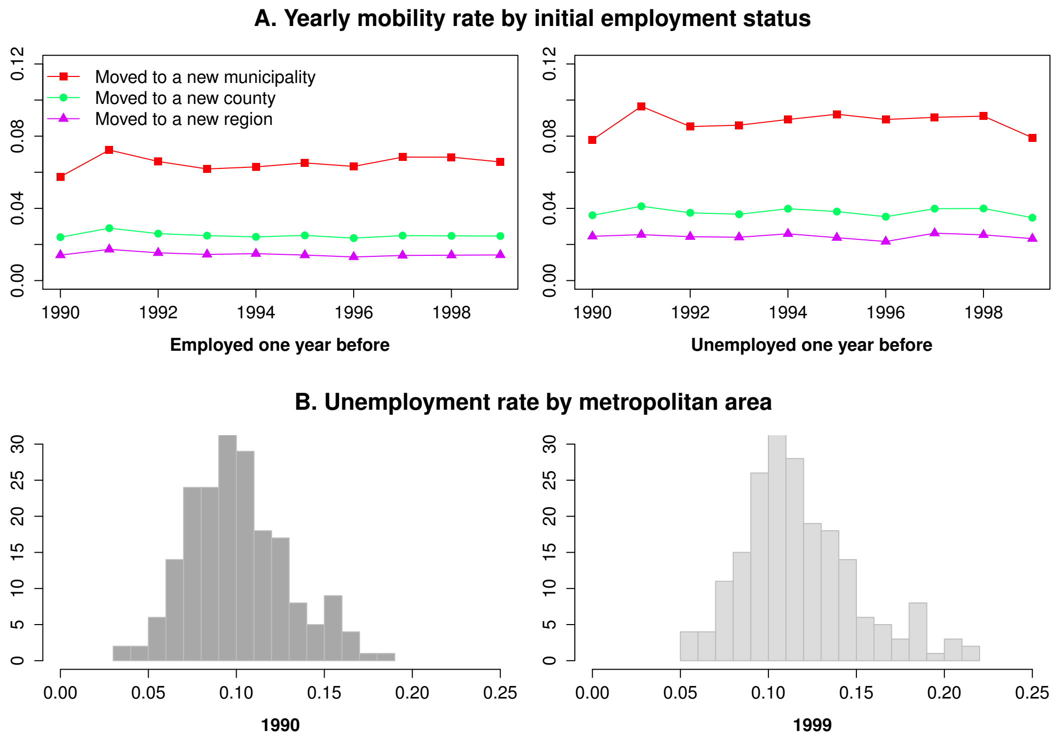
[†]Howard University and CREST; benoit.schmutz@howard.edu

[‡]Duke University and CREST; 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 in the 1990s



Notes: (i) Mobility rates: probability to have changed location in the past year, conditional on previous employment status; (ii) Counties stand for the French "departements"; there are 22 regions, 9 departements and over 36,000 municipalities in France; each of these three levels forms a partition of the French territory, unlike metropolitan areas, which are aggregates of municipalities and may cross county or region boundaries; (iii) Unemployment rates are computed on the 25-54 age bracket and for the 200 largest metropolitan areas in continental France, keeping a constant municipal composition based on the 2010 "Aires Urbaines" definition; (iv) Sources: Labor Force Surveys 1990-1999 (Figure A) and Census 1990 and 1999 (Figure B).

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

[2]See Elhorst (2003) for an overview of the (mostly empirical) literature on the determinants of regional differences in unemployment rates. Following Harris & Todaro (1970), competitive models have first sought to explain rural/urban migration patterns in developing economies (see Lucas (1993) for an overview). While both mean wages and unemployment risk are taken into account in the Harris-Todaro framework, it is assumed that unemployment is confined to one area (cities). 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, even though 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 rates if they were to introduce individual mobility. That is, mobility would lead to a pattern of convergence in unemployment rates as in Phelps' (1969) 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 affluent 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. In particular, it can help us identify and quantify the two kinds of spatial constraints that arise from the physical separation between cities: informational frictions and mobility costs, without resorting to any kind of spatial idiosyncrasy, such as workers being more productive in or more emotionally attached to a location.[3]

We borrow from the wage posting framework of Burdett & Mortensen (1998) to model a job search model that incorporates spatial segmentation between a large number of interconnected local labor markets, or "cities". Matching based on local labor market conditions generates both city-specific job arrival rates and wage offer distributions, and potential local labor shocks are introduced through city-specific layoff rates. 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. A spatial equilibrium is achieved through the mobility of unemployed workers, who generate congestion externalities upon the non-pecuniary component of utility in each location, such that mobility will only be worth its cost in case the worker has been matched with an attractive job offer in another city.

Most previous quantitative studies of migration rest upon a unidimensional conception of spatial constraints based on a black box called "mobility costs", which encompass both impediments to the mobility of workers when it takes place (actual mobility costs) and impediments to the spatial integration of the labor market (workers' ability to learn about remote vacancies). We argue that separating these mechanisms yields new insights, as they do not take place at the same time and they do not affect the same economic outcomes.

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. Sec-

[3] For example, Lkhagvasuren (2012), in a calibrated macroeconomic model, uses stochastic worker-location match productivity, combined with search frictions, to rationalize why workers may not leave high-unemployment areas. In a completely different strand of research, most empirical models of migration posit some degree of idiosyncratic region preference. Even if allowing for location-specific shocks is a convenient tool to replicate observed regularities, this is not entirely satisfactory from a theoretical viewpoint because the source of idiosyncrasy is often left unmodeled.

ond, spatial segmentation introduces city-level heterogeneity in their non-labor market dimension, which will be called a city-specific "amenity"; this amenity will impact agents' willingness to refuse a job somewhere else, even though this would 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 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 their migration decision, conditional on receiving an acceptable offer. As in Schwartz (1973), these costs encompass a fixed cost of losing local ties and connections, and a cost of moving from one place to another, which mostly depends on distance. Since the model is dynamic, the relative position of the city in the distribution of all possible mobility costs, which determines the level of *accessibility* of the city in the system, will also impact whether the offer was deemed acceptable in the first place.

The 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. They define a complex relationship between wages and the model primitives. As a consequence, the model is able to cope with various wage profiles over the life-cycle, including voluntary wage cuts as in Postel-Vinay & Robin (2002). 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 steady state conditions on market size, unemployment level and wage distributions. Our estimation uses the panel version of the French matched employer-employee database *Declaration Annuelles de Donnees 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 only used to identify 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.

Our results consist of a set of vectors of city-specific structural parameters and a set of matrices of parameters measuring spatial constraints between each pair of cities. 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 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. The predictive power of these regressions is higher for the larger cities.

Our estimation of spatial constraints suggests that geographical distance may increase mobility costs by up to 40% and that both geographical distance and sectoral dissimilarity are much stronger deterrents of the efficiency of spatial search for employed jobseekers, than for unemployed jobseekers. 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; 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. Keane & Wolpin (1997) have shown that individuals make sophisticated calculations regarding work-related decisions, both in terms of pure labor market characteristics (industry, occupation, skills requirement) and location characteristics. Authors have proposed structural models to disentangle between the various underlying mechanisms. 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, as migration is an investment, it requires not only a static tradeoff between economic conditions, but also a comparison between expected future economic conditions. This is the argument made by Gallin (2004), who uses a perfect competition model of migration. Despite its interest and obvious links to the present paper, the classic perfect-competition approach cannot fully reconcile the joint existence of 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.[4]

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.[5] 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 appeal-

[4] The original job search literature emerges as an attempt to capture the existence of frictional unemployment. Interestingly, Phelps (1969)'s island parable is, at least metaphorically, related to this paper. The major breakthrough, due to Burdett & Mortensen (1998), allows to generate ex-post wages differential from ex-ante identical workers, and provides an intuitive way to evaluate the individual unemployment probability as well as the wage offer distribution without solving the value functions.

[5] As shown by 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, among others, Blanchard & Katz (1992)). Postel-Vinay & Robin (2002), who restrict their estimation sample to the Paris region, implicitly recognize this problem.

ing 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 difficulties requires additional assumptions. For example, it is assumed that individuals have knowledge over a limited number of local wage distributions, which correspond to where they used to live. In order to learn about another location, workers need to pay a visiting cost. These assumptions may not reflect the recent increase in workers' ability to learn about other locations before a mobility.[6] 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.[7] 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

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

[7] Among many others, three notable papers are: Gould (2007), who studies the determinants of the urban wage premium in the US; Combes, Duranton, Gobillon, Puga & Roux (2012), who show that higher productivity in larger French cities is mostly due to agglomeration economies and technological complementarities between the productivity of firms and that of workers; and De la Roca & Puga (2012), who show that the city size wage premium in Spain does not reflect initial sorting of workers by ability, but is rather the result of a more efficient learning process in larger cities.

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.[8] 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, to study the determinants of these parameters across cities.[9]

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.[10] Our main reference is Meghir, Narita & Robin (2015), 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

[8] 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 of location on job opportunities takes place through commuting costs only.

[9] This ability to estimate our model comes at a cost: we do not explicitly model the housing market, in contrast of Head, Lloyd-Ellis & Sun (2014), 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. 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).

[10] 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. Shephard (2011) distinguishes between part-time and full-time work to assess the impact of UK tax credit reform on individual participation choices.

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.

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. Our theoretical framework shows that this extension is far from conceptually trivial. It also allows us to put the existing frictionless migration literature into perspective.

1.4 Identification

The identification of standard search and matching models is extremely challenging. The inclusion of mobility adds an extra-layer of difficulty. In structural econometrics, Flinn & Heckman (1982) and Magnac & Thesmar (2002) have shown that the theoretical identification of labor market parameters hinges on the transition rates. Our identification relies on this strategy, but in addition, we are able to use data on wages to identify mobility costs. We use transition rates to identify local labor market parameters and spatial frictions parameters. As a consequence, it is only the frequency of transition within and between each pair of cities that identifies the layoff and job arrival rates.

Additional information on wage-accepting patterns between locations conditional on the structural parameters can be used to identify the mobility cost. When a city exhibits high accepted wages regardless of the origin of incoming workers, this information provides identification for a low level of local amenities.[11] Similarly, when workers from the same location accept very heterogeneous wages to go into different destinations, we use this information to identify mobility costs.

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

[11] A downside is that we cannot separately identify local amenities from cost of living. As a consequence, we assume additivity and separability between these elements in the utility of workers.

through workers' mobility; 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 workers' mobility, both on the labor market and across space.

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

The functional definition of a metropolitan area brings together the notions of city and local labor market. A more precise partition of space, for instance based on municipal boundaries, would lead to a confusion between job-related motives for migration and other motives.[12] 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.[13] 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.[14] As shown in Figure 6 in Appendix D, 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.

Population Since we do not model the participation choice of workers, labor force is analogous to population. Data from 1999 and 2006 Census shows that the Paris region accounts for more than 25% of total labor force. As shown in Figure 2, local labor force is Pareto-distributed and absolute variation in local labor force between 1999 and 2006 is negligible.[15]

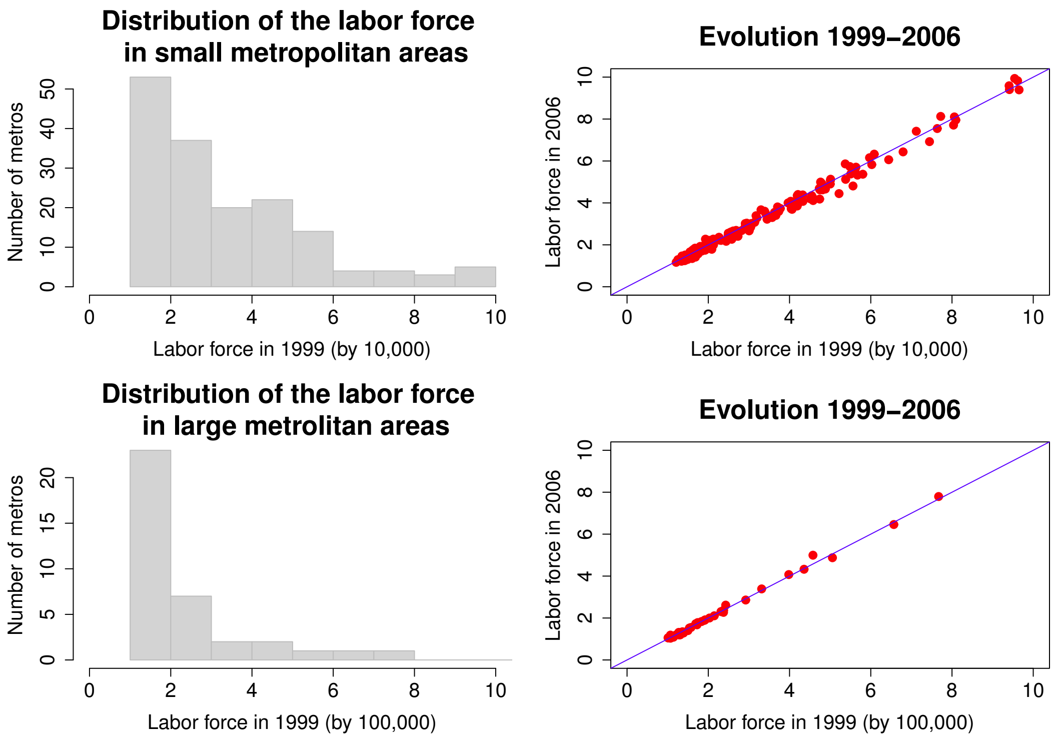
[12] According to the 2006 French Housing Survey, 16% of the households in the labor force who had been mobile in the past four years declared that the main reason for their move was job-related. However, this small proportion hides a large heterogeneity which is correlated with the scale of the migration,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

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

[14] According 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).

[15] This stable distribution of the labor force is at odds with the fact that metropolitan areas face diverse net migration patterns. The explanation lies in the contribution of nonparticipants (retired, young individuals) to the net migration. According to Gobillon & Wolff (2011), 31.5% of French grand-parents aged 68-92 in 1992 declared that they moved out when they retired. Among them, 44.1% moved to another region. Most of these migration decisions are motivated by differences in location-specific amenities or by the desire to live closer to other family members.

Figure 2: Local labor force



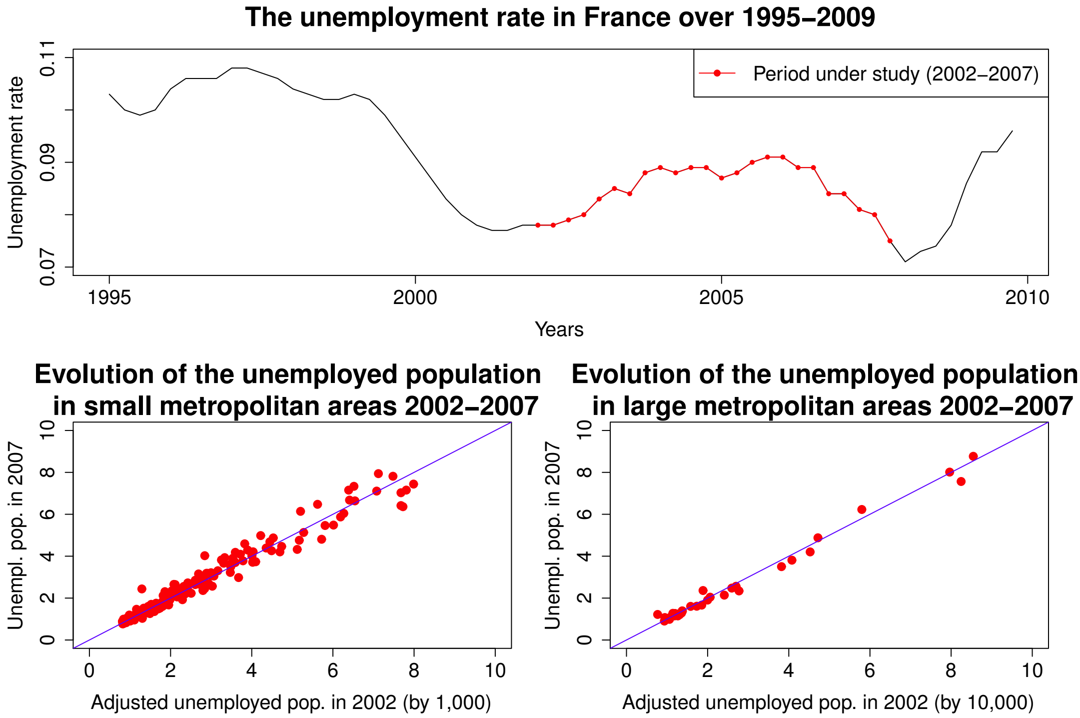
Notes: (i) Labor force is composed of unemployed and employed individuals aged between 15 and 64; the labor force in the 300 largest metropolitan areas in continental France amounts to 19.5 millions in 1999 and to 19.3 millions in 2006; (ii) For the sake of exposition, we do not represent Paris; its 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 3 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. The two bottom graphs show that between 2002 and 2007, city-specific unemployment patterns have remained remarkably stable.[16]

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

[16]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.

Figure 3: Aggregate and local 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 (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).*

DADS are the main source of data used in this paper.[17] 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.[18] Although the wage premium in Paris may be partly offset by the cost of living, there exist persistent wage differentials among cities with comparable size and cost of living. For instance, Oloron is richer than all the other cities of Panel 2, including cities which are far larger.In addition, there is a strong positive correlation of 0.44 between wage dispersion and city size. These trends are supported by the log-difference between the top and bottom decile (or between the 3rd and the 1st

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

[18]Our data selection procedure that excludes part-time workers and civil servants increases the wage gap between Paris and smaller locations. Using all the available payroll data in 2007, the mean wage in Paris is around 22,501€, which is 35% higher than the average wage.

Table 1: Local wage distributions

Panel 1: the nine largest cities

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

Panel 2: nine other cities

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

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

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.

Workers' heterogeneity Apart from the size of the labor force and the unemployment rate, other dimensions, such as the skill and the sectoral composition, are also important drivers of local labor market heterogeneity and dynamics. However, as discussed in Appendix E, we believe that, as a first-order approximation, the assumption of workers' homogeneity is not very costly when focusing on a short time-span, during which the distribution of observable characteristics across cities remains stable.

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 C.1, for more details).

Since the DADS panel is based on firms payroll reports, it does not contain any information on unemployment. However, it reports for each employee the duration of the job, along with the wage. We use this information to construct a potential calendar of unemployment events and, in turn, identify transitions on the labor market.[19] 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.[20] 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

[19] Our algorithm is available at this address.

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

a single draw. Hence, we rule out the possibility of a sequential job search whereby individuals would first change locations before accepting a new job offer. From a theoretical viewpoint, this means that mobility has to be job-related. From a practical viewpoint, in the DADS data, the sequential job search process is observationally equivalent to the joint mobility process.

Labor market transitions Table 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 similar

Table 3: Number and characteristics of transitions

	Characteristics of the spells			
Type of history	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*

to the non-mobility rate of 45% reported by Postel-Vinay & Robin (2002) from 1996 to 1998. Approximately 23% of the sample records at least one job-to-job transition, and the 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 employed 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.[21]

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

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.[22] Table 5

Table 4: Mobility between the largest cities

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

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 originating from 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*

completes this overview by comparing the mobility patterns within the Lyon region (also known as "Rh?ne-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).

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 following a job-to-job transition within the same city.[23] This suggests that mobility costs are high compared to local differences in economic opportunities. Second, if mobility costs are

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.

[22] 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.

[23] It should be noted that this pattern does not preclude the existence of mobility strategy with wage cut. There are numerous cases in the full data where 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.

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

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

Notes: (i) UE stands for transition out of unemployment and EE stands for job-to-job transition; (ii) Reading: 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*

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 offer 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).

3 Job search between many local labor markets: theory

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 steady state. 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.

3.1 Framework

We consider a continuous time model, where infinitely lived, risk neutral agents maximize their expected steady-state discounted future income (at rate r). The economy is organized as a system $\mathscr{J}$ of J interconnected local labor markets, or "cities", where workers both live and work. Whereas the spatial position of each city $j \in \mathscr{J}$ within the system is exogenous, all its other observable characteristics

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

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

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

arise from the job search process: population m_j, unemployment rate u_j/m_j, share of local firms n_j, and distribution of observed earnings $G_j(\cdot)$.

3.1.1 Local search and matching

Workers and firms are both ex ante homogeneous. They form two continua of respective measures M and 1. Workers are fully characterized by their employment status $i = e, u$, their wage level w when employed (they earn uniform benefits b when unemployed) and their location j. Individuals engage in both off-the-job and on-the-job search and a type-i worker faces a job finding rate λ_j^i in city j. Firms create output according to a linear production technology and post a wage w in order to maximize their profit flow $(p_j - w)\ell_j(w)$, where p_j denotes the production of a match, net of city-specific operating cost, and $\ell_j(w)$ is the labor force available to a w-paying firm in city j. In equilibrium, all firms have the same profit rate, regardless of their location and the wage they offer.[24]

The wage posting strategy of homogenous firms generates a city level wage offer distribution $F(\cdot) \equiv \{F_j(\cdot)\}_{j\in\mathscr{J}}$ of support $[\underline{w},\overline{w}]^J \subset (b,\infty)^J$ (with $\overline{F}_j(\cdot) \equiv 1 - F_j(\cdot)$) and the optimal number n_j of firms in city j, which depends on the size of the labor market. City population and unemployment rates m_j and u_j/m_j are determined by workers' work-related and migration-related decisions. Those

[24]The firm side of the model is voluntarily left as simple as possible, in the spirit of van den Berg & Ridder (1998). Its main purpose is to generate endogenous wage offer distributions and matching rates. Allowing for varying production technologies between cities and heterogeneous firms is a promising venue that is beyond the scope of this paper.

three outcomes (m_j, u_j, n_j) combine into a (local) market tightness

$$\theta_j = \frac{n_j}{u_j + \xi_j (m_j - u_j)} \quad (1)$$

where ξ_j measures the city-specific relative search effort of an employed worker compared to an unemployed worker.[25] The likelihood of a match is determined by a meeting function $\boldsymbol{M}(\theta_j) = \Omega\theta_j^{\eta}$. This yields the following expressions for the matching rates: $\lambda_j^u = \theta_j \boldsymbol{M}(\theta_j)$ and $\lambda_j^e = \xi_j \lambda_j^u$.

3.1.2 Spatial segmentation and migration

Cities are heterogeneous, both from in terms of labor market and living conditions. Employed workers in city j face an exogenous, location-specific unemployment risk characterized by the layoff probability δ_j. In addition, all workers in city j face an indirect utility γ_j, which summarizes the difference between amenities and (housing) costs in city j. This value is separable from the level of earnings, such that the instant value of a type-i worker in city j equals $y^i + \gamma_j$, with $y^u = b$ (unemployment benefits) and $y^e = w$.[26]

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

3.1.3 Workers' value functions

Let $(x)^+ \equiv \max\{x, 0\}$. Workers do not bargain over wages. They only decide whether to accept or refuse the job offer which they have received. The respective value functions of unemployed workers living

[25] Alternatively, we could write a country-level market tightness, as in Meghir et al. (2015).

[26] This specification accounts for differences in local costs of living, so that wages can still be expressed in nominal terms.

in city j and of workers employed in city j for a wage w are recursively defined by equations 2 and 3:

$$rV_j^u = b + \gamma_j + \lambda_j^u \int_{\underline{w}}^{\overline{w}} \left(V_j^e(x) - V_j^u\right)^+ dF_j(x) + \sum_{k \in \mathscr{J}_j} s_{jk}^u \lambda_k^u \int_{\underline{w}}^{\overline{w}} \left(V_k^e(x) - c_{jk} - V_j^u\right)^+ dF_k(x) \quad (2)$$

$$\begin{aligned} rV_j^e(w) &= w + \gamma_j + \lambda_j^e \int_{\underline{w}}^{\overline{w}} \left(V_j^e(x) - V_j^e(w)\right)^+ dF_j(x) + \sum_{k \in \mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \left(V_k^e(x) - c_{jk} - V_j^e(w)\right)^+ dF_k(x) \\ &+ \delta_j \left[V_j^u - V_j^e(w)\right] \end{aligned} \quad (3)$$

3.2 Workers' strategies

Accepting a good offer in a city conveys parameters that are city-specific. Jobs are no longer defined by the single attribute of wage, but rather by a non-trivial combination of all the structural parameters of the economy, which determines the offer's option value. 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. This multivariate, and dynamic trade-off allows us to define *spatial strategies*, where workers' decision to accept a job in a given city is not only driven by the offered wage and the primitives of the local labor market, but also by the employment prospects in all the other locations, which depend upon the city's specific position within the system. The sequence of cities where individuals are observed can then be rationalized as part of lifetime mobility-based careers.

3.2.1 Definitions

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.

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 $(j,l) \in \mathscr{J} \times \mathscr{J}_j$. These functions associate the current wage w earned in location j to a wage which would yield the same 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. By definition, static indifference wages have a stationary property, whereby $\chi_{lk}(\chi_{jl}(w)) = \chi_{jk}(w)$. As will be made clear later, the introduction of $\chi_{jl}(w)$ is important to understand the role of mobility costs in the dynamics of the model.

Proposition 1 *OPTIMAL STRATEGIES*

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

$$\phi_j = b + \left(\lambda_j^u - \lambda_j^e\right)\int_{\phi_j}^{\overline{w}} \Xi_j(x)dx + \sum_{k\in\mathscr{J}_j}\left(s_{jk}^u\lambda_k^u - s_{jk}^e\lambda_k^e\right)\left(\int_{q_{jk}(\phi_j)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(\phi_j))c_{jk}\right) \quad (4)$$

$$\begin{aligned} q_{jl}(w) &= \zeta_{jl}w + \left(\zeta_{jl}\gamma_j - \gamma_l\right) + (r+\delta_l)c_{jl} + \left(\zeta_{jl}\delta_j V_j^u - \delta_l V_l^u\right) + \zeta_{jl}\lambda_j^e\int_w^{\overline{w}} \Xi_j(x)dx - \lambda_l^e\int_{q_{jl}(w)}^{\overline{w}} \Xi_l(x)dx \quad (5)\\ &+ \zeta_{jl}\sum_{k\in\mathscr{J}_j} s_{jk}^e\lambda_k^e\left(\int_{q_{jk}(w)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(w))c_{jk}\right) - \sum_{k\in\mathscr{J}_l} s_{lk}^e\lambda_k^e\left(\int_{q_{lk}(q_{jl}((w))}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{lk}(q_{jl}(w)))c_{lk}\right) \end{aligned}$$

with:

$$V_j^u = \frac{1}{r}\left[b + \gamma_j + \lambda_j^u\int_{\phi_j}^{\overline{w}} \Xi_j(x)dx + \sum_{k\in\mathscr{J}_j} s_{jk}^u\lambda_k^u\left(\int_{q_{jk}(\phi_j)}^{\overline{w}} \Xi_k(x)dx - \overline{F}_k(q_{jk}(\phi_j))c_{jk}\right)\right] \quad (6)$$

$$\Xi_j(x) = \frac{\overline{F}_j(x)}{r + \delta_j + \lambda_j^e\overline{F}_j(x) + \sum_{k\in\mathscr{J}_j} s_{jk}^e\lambda_k^e\overline{F}_k(q_{jk}(x))} \quad (7)$$

- *Equations 4 and 5 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 4 and 5 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.*

3.2.2 Interpretation

The interpretation of Equation 4 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.[27]

The interpretation of Equation 5 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 in terms of amenities. 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 risk δ_j or δ_l. The second part is the difference in the expected wage following a local job-to-job transition and the third part is the difference in the expected wages that will be found through mobile job search, net of mobility costs.

This last term introduces the relative centrality and accessibility of city j and city l. Centrality stems from the comparison of the strength of spatial frictions between the two locations j and l and the rest of the world: a worker living in city j who receives an offer from city l must take into account the respective spatial frictions from city j and from city l to any tier location k that she may face in the future, in order to maximize her future job-offer rate. As for accessibility, it stems from the

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

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

3.3 Equilibrium

3.3.1 Spatial equilibrium

Workers in city j are always free to move into city l if they pay the mobility cost c_{jl} and become unemployed. Given the existence of the reservation wage strategy, this type of migration out of the labor market will mostly be an option for unemployed workers. However, the inflow of unemployed workers into an attractive location will generate congestion externalities which will negatively impact local amenities and will also push housing prices upwards. This adjustment mechanism operates through the parameters γ_j and leads to the following proposition:

Proposition 2 *CONGESTION — the vector of city amenities* $\Gamma = \{\gamma_j\}_{j \in \mathcal{J}}$ *satisfies the set of constraints 8:*

$$V_j^u \geq \max_{k \in \mathcal{J}_j} \{V_k^u - c_{jk}\} \tag{8}$$

Because the instant value of unemployment is a linear function of local amenities, the solution is unique to a constant.

3.3.2 Steady state distribution of unemployment rates

As already explained in section 2, a cross-sectional description of the labor market 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. We now describe the theoretical counterparts to these three components.

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

Proposition 3 *STEADY STATE UNEMPLOYMENT — the distribution of unemployment rates is given by* $\mathscr{U} = \left\{\frac{u_j}{m_j}\right\}_{j\in\mathscr{J}}$*, where:*

$$\frac{u_j}{m_j} = \frac{\delta_j}{\delta_j + \lambda_j^u \overline{F}_j(\phi_j) + \sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))} \tag{9}$$

3.3.3 Steady state distribution of city populations

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

$$(m_j - u_j) \sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x) + u_j \sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j)) \equiv \tag{10}$$
$$\lambda_j^e \sum_{k\in\mathscr{J}_j} s_{kj}^e (m_k - u_k) \int_{\underline{w}}^{\overline{w}} \overline{F}_j(q_{kj}(x)) dG_k(x) + \lambda_j^u \sum_{k\in\mathscr{J}_j} s_{kj}^u u_k \overline{F}_j(q_{kj}(\phi_k))$$

Plugging Equation 9 into Equation 10, we recover a closed form solution for the system, written as:

$$\mathscr{A}\mathscr{M} = 0 \tag{11}$$

where $\mathscr{M}$ is the vector of city sizes $\{m_j\}_{j\in\mathscr{J}}$ and $\mathscr{A}$ is the matrix of typical element $\{\mathscr{A}_{jl}\}_{(j,l)\in\mathscr{J}^2}$ defined by:

$$\mathscr{A}_{jj} = \frac{\left[\lambda_j^u \overline{F}_j(\phi_j) + \sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))\right] \times \left[\sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x)) dG_j(x)\right] + \delta_j \left[\sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))\right]}{\delta_j + \lambda_j^u \overline{F}_j(\phi_j) + \sum_{k\in\mathscr{J}_j} s_{jk}^u \lambda_k^u \overline{F}_k(q_{jk}(\phi_j))}$$

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

where off-diagonal elements equal the fraction of the population in the city in column who migrates into the city in row at any point in time, and diagonal elements equal the fraction of the population in the city in question who moves out at any point in time. This yields the following proposition:

Proposition 4 *STEADY STATE POPULATION — The distribution of city sizes is the positive vector* $\mathscr{M} \in \ker \mathscr{A}$ *s.t.* $\sum_{j\in\mathscr{J}} m_j = M$.

Note that Equation 10 defines a relationship between m_j and all the other city sizes in $\mathscr{M}$, whereas it is not the case for u_j, which is determined by a single linear relationship to m_j. The flow of workers into unemployment in city j is only composed of workers previously located in city j, whereas in Equation 10, 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.

3.3.4 Steady state distribution of observed wages

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 need to correspond to a wage higher than $q_{jk}(x)$, where $x < w$ is the wage previously earned in city j. The measure of this flow, which stems from the fact that we consider several separate markets, requires an integration over the distribution of observed wages in city j. Inflows to city j are first composed of previously unemployed workers who find and accept a job in city j with a wage lower than w. These workers may come from city j or from any city $k \neq j$. However, they will only accept such a job if w is higher than their reservation wage ϕ_j or than the mobility-compatible indifference wage of their reservation wage $q_{kj}(\phi_k)$. The second element of inflows is made of workers who were previously employed in any city $k \neq j$ at a wage x lower than the maximum wage such that moving to city j would yield a utility of $V_j^e(w)$ (we denote this wage $q_{jk}^{-1}(w)$) and find a job at a wage between $q_{kj}(x)$ and w.[28]

This is all summarized in Equation 12:

$$(m_j - u_j)\Big[G_j(w)\big(\delta_j + \lambda_j^e \overline{F}_j(w)\big) + \sum_{k\in\mathscr{J}_j} s_{jk}^e \lambda_k^e \int_{\underline{w}}^{w} \overline{F}_k(q_{jk}(x))dG_j(x)\Big] \equiv \tag{12}$$

$$\lambda_j^u\Big[\psi_{jj}(w)u_j\big(F_j(w) - F_j(\phi_j)\big) + \sum_{k\in\mathscr{J}_j} s_{kj}^u \psi_{kj}(w)u_k\big(F_j(w) - F_j(q_{kj}(\phi_k))\big)\Big]$$

$$+\lambda_j^e \sum_{k\in\mathscr{J}_j} s_{kj}^e (m_k - u_k)\int_{\underline{w}}^{q_{kj}^{-1}(w)} [F_j(w) - F_j(q_{kj}(x))]dG_k(x)$$

where $\psi_{kj}(w) = \mathbb{1}_{w>q_{kj}(\phi_k)}$ is a dummy variable indicating whether unemployed jobseekers in city k are willing to accept the job paid at wage w in city j. Similarly, the integral in the last term gives the measure of job offers in city j that are associated with a wage lower than w yet high enough to attract employed workers from any city $k \neq j$ and it is nil if $q_{kj}^{-1}(w) < \underline{w}$. These restrictions mean that very

[28] Because of the existence of mobility costs, $q_{jk}^{-1}(w) \neq q_{kj}(w)$ (see Section A.3 for details).

low values of w will not attract many jobseekers. We can differentiate Equation 12 with respect to w. This yields the following linear system of functional differential equations:

$$f_j(w) = \frac{g_j(w)(m_j - u_j)\left[\delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k\in\mathcal{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(w))\right]}{\lambda_j^u\left(\psi_{jj}(w)u_j + \sum_{k\in\mathcal{J}_j} s_{kj}^u \psi_{kj}(w)u_k\right) + \lambda_j^e\left((m_j - u_j)G_j(w) + \sum_{k\in\mathcal{J}_j} s_{kj}^e (m_k - u_k)G_k(q_{kj}^{-1}(w))\right)} \tag{13}$$

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 11, the uniqueness of the solution is not guaranteed. We defer the question of identification to section 4.2. We can then write the following proposition:

Proposition 5 *STEADY STATE WAGE OFFER DISTRIBUTIONS — The distribution of wage offers by location is solution to the system 13.*

3.3.5 Steady state distribution of the number of firms

In equilibrium, firms are indifferent between all possible wage strategies. In particular, this gives $(p_j - w)\ell_j(w) = (p_j - \underline{w})\ell_j(\underline{w})$.

The measure of jobs paid at wage w in city j can be computed from the worker's side: $(m_j - u_j)dG_j(w)$ and from the firm's side: $n_j \ell_j(w) dF_j(w)$. Since both measures are equal, this yields

$$n_j f_j(w)\ell_j(w) = (m_j - u_j)g_j(w) \tag{14}$$

Using $\int_{\underline{w}}^{\overline{w}} \ell_j(x)dx = 1$, we finally get:

Proposition 6 *STEADY STATE NUMBER OF FIRMS — The spatial distribution of the population of firms across cities is given by the vector $\mathcal{N} = \{n_j\}_{j\in\mathcal{J}}$ as solution to equation 15:*

$$n_j = (m_j - u_j) \times \frac{g_j(w)}{f_j(w)} \times (p_j - w)\ln\left(\frac{p_j - \overline{w}}{p_j - \underline{w}}\right) \tag{15}$$

This equation holds for each level of w: when a firm forgoes additional profit by offering a higher wage, this must be perfectly compensated for by a larger market share.

3.3.6 Summary

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

1. The reservation wage strategy ϕ_j in Equation 4 describes the job acceptation behavior of immobile unemployed workers.

2. The mobility strategy between two locations $q_{jl}(\cdot)$ is defined by the indifference wage described in Equation 5.

3. The set of unemployment rates $\mathscr{U}$ is given by Equation 9.

4. The set of city populations $\mathscr{M}$ is solution to the linear system 11.

5. The set of local amenities Γ satisfies the market clearing constraints described by equation 8.

6. The behaviour of firms, summarized by the set of wage offer distributions $F(\cdot)$ is solution to the system of functional differential equations 13.

7. The set of number of firms in each city $\mathscr{N}$ is given by equation 15.

4 Estimation

The model is estimated by simulated method of moments. The estimator minimizes the distance between a set of empirical moments and their theoretical counterparts, which are constructed by solving the equilibrium of the model. In Appendix B, we present a full set of solutions to solve the indifference wages and the functional equations. We take advantage of the exact structure of the model and use an embedded algorithm that allows us to recover a piecewise approximation of all indifference wages, and wage offer distributions. We detail here our choice of parametrization of spatial constraints and our identification strategy.

4.1 Parametrization

The model is based on a set of parameters $\theta = \{\lambda_j^e, \lambda_j^u, \delta_j, s_{jk}^e, s_{jk}^u, c_{jk}, p_j\}_{(j,k)\in\mathscr{J}\times\mathscr{J}_j}$ such that $|\theta| = 120{,}200$ 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 $\mathscr{J}$. We take an alternative path and we posit and estimate two parsimonious parametric models:

$$s_{jl}^i = \frac{\exp\left(s_{j0}^i + s_{0l}^i + s_1^i d_{jl} + s_2^i d_{jl}^2 + s_3^i h_{jl} + s_4^i h_{jl}^2\right)}{1 + \exp\left(s_{j0}^i + s_{0l}^i + s_1^i d_{jl} + s_2^i d_{jl}^2 + s_3^i h_{jl} + s_4^i h_{jl}^2\right)} \quad (16)$$

$$c_{jl} = c_0 + c_1 d_{jl} + c_2 d_{jl}^2 \quad (17)$$

where s^i_{j0} and s^i_{0l} 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.[29]

The model rests upon the premise that spatial friction parameters take on values in $[0,1]$. Given the lack of existing literature on the explicit structure of spatial frictions, we choose to use a logistic function in Equation 16 because of its analytical properties.[30] 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 16 is akin to a standard gravity equation: the fixed effects measure the relative openness of the local labor markets: either the ability of each city to dispatch its jobseekers to jobs located elsewhere (s_{j0}) or to fill its vacancies with workers coming from other locations (s_{0l}), and the other parameters account for the effect of distance between two locations.[31]

Physical distance is arguably the most important characteristic and both equations 16 and 17 rely on it. 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.[32] Finally, we do not allow for fixed effects in Equation 17. Those fixed effects cannot be identified separately from the local amenity parameters. Also, our estimates of mobility costs will depend on the pair of cities involved, but not on the direction of the move.[33] Under these two specifications, the total number of parameters to be estimated amounts to 2,011.

4.2 Identification

Identification is based on Proposition 7:

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

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

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

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

[33] This symmetry assumption could easily be relaxed, for instance by including an indicator variable on whether the destination city is larger or smaller than the departure city, as in Kennan & Walker (2011). However, as shown by Levy (2010) on US data, this may not be empirically relevant.

Proposition 7 *EXISTENCE AND UNIQUENESS OF WAGE OFFER DISTRIBUTIONS The system of differential equations* $\mathfrak{f}: \mathbb{R}^J \to (0,1)^J$ *has a unique fixed point.*

Proof Existence stems from a direct application of *Schauder fixed-point theorem.* Regarding uniqueness, first note that since each $f_j(\cdot)$ is a probability density function, it is absolutely continuous and its nonparametric kernel estimate is Lipschitz continuous; then, by contradiction, it is easy to show that two candidate solutions $h^0(\cdot)$ and $h^1(\cdot)$ cannot at the same time solve the differential equation, define a contraction, and be Lipschitzian. For more details, see Theorem 2.3 in Hale (1993).

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. 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.

Given the parametrization of s^i_{jl}, the model is over-identified: in particular, the $2J(J-1)$ transition rates at the city-pair level that would be required to identify each parameter s^i_{jl} are no longer needed. In order to identify the fixed-effect components, we use the $2J$ 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 two cities does still requires transition rates at the city-pair level. Given that Equation 16 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.[34]

While spatial friction parameters are identified from transition rates between pairs of cities, mobility costs are identified from wage moments. This strategy is made possible because we do not use

[34] In practice, we use the off-the-job and job-to-job transitions rates from the urban areas ranked fourth to eleventh (Toulouse, Lille, Bordeaux, Nice, Nantes, Strasbourg, Grenoble and Rennes) to the urban areas ranked fifteenth, nineteenth to twenty-second and twenty-fifth (Montpellier, Clermont-Ferrand, Nancy, Orléans, Caen and Dijon). This selection is designed to include locations that are widely scattered across the French territory (see Figure 7 in appendix D for details).

Table 7: Moments and Identification

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

Notes: (i) For details on the construction of the empirical moments, see Appendix C.2; (ii) Alternatively, one could use the transitions into unemployment: $(m_j - u_j)\delta_j$ and the transition out of unemployment within city j: $u_j \lambda_j^u \overline{F}_j(\underline{w})$ to identify δ_j and λ_j^u.

wage data to approximate indifference wages.[35] 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 $w_{e_je_j}^{init}$ denote the average initial wage of agents employed in city j and who will experience a job-to-job transition within city j. Using the fact that $q_{jl}(\cdot)$ is a function, we consider as a theoretical moment, the difference between $q_{jl}(w_{e_je_j}^{init})$ and $w_{e_je_j}^{init}$ (which, by definition, is equal to $q_{jj}(w_{e_je_j}^{init})$). The corresponding empirical moment is the difference between $w_{e_je_l}^{fin}$ (the average wage after a job-to-job transition from city j to city l), and $w_{e_je_j}^{fin}$ (the average wage after a job-to-job transition within city j).[36] 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

[35] This is our main departure from Meghir et al. (2015), who have to use wage data to recover subsequent optimal contracts numerically.

[36] In 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.

with mobility, these differences identify the mobility cost c_{jl}.[37] 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 $|\mathcal{T}_2| = 12$.[38]

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.[39] 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

5.1.1 Presentation

Figure 4 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

[37] If 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.

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

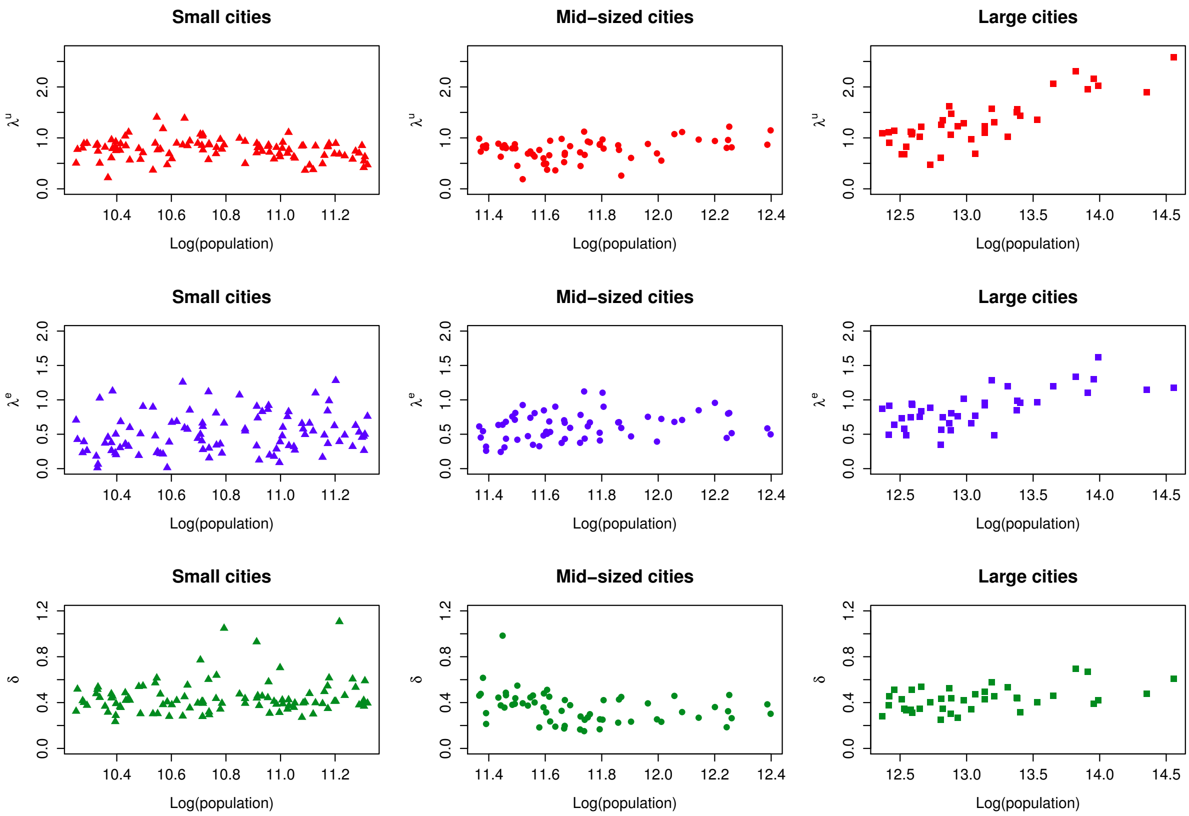
[39] We parametrize b = €6,000 (an approximation of the minimum guaranteed income, which amounts to about half of the minimum wage) and r = 13.4% (the level of inflation between January 2002 and December 2007). The model is optimized using Quasi-Newton algorithm methods. Integrals are evaluated numerically using a Newton sequence on 100 points. 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.

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

Figure 4: The matching parameters along the city size distribution



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

cities. In Paris for example, the job arrival rate of 8.9 implies that offers accrue approximately every 9 months on average.[40] The median value of the job arrival rate, around 0.8, confirms the very low tran-

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

sition 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 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 4 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.

5.1.2 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 the their first two moments.

Table 9: Correlation between the local labor market primitives

Panel 1: All cities

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

Panel 2: Large cities

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

Panel 3: Mid-sized cities

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

Panel 4: Small cities

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

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

The three main findings are the following. First, panel 1 shows that there is a positive correlation between off-the-job and on-the-job arrival rates.[41] 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

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

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 5.3.2 will confirm this hypothesis using a wage decomposition approach.

5.1.3 Observable covariates 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 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 A dataset of city-pair-specific parameters

5.2.1 Presentation

The observation of our results for the spatial friction parameters yields two conclusions. First, relative to the internal job arrival rates λ_j^i, the job arrival rate from other locations $\sum_{k\in\mathscr{J}_j} s_{jk}^i \lambda_k^i$ is rather high, which gives support in favor of our modelling choice to take between-city mobilities into account. For instance, the median value of this rate for unemployed jobseekers is close to half of the median value of the λ_j^u.[42] 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, $\sum_{k\in\mathscr{J}_M} s_{Mk}^u \lambda_k^u = 0.9$, than in Lyon, $\sum_{k\in\mathscr{J}_L} s_{Lk}^u \lambda_k^u = 0.01$. A similar pattern can be observed in the other two groups of cities. In Brive-la-gaillarde, for example, the city with the lowest off-the-job arrival rate (0.189), spatial mobility opportunities accrue at a rate of 1.6, when Cholet, the city just above in the city size distribution (ranks 85 and 84, respectively), faces a rate of 0.69.[43]

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 4.1, these two dimensions are measured in relative terms, by the spatial distance and the sectoral dissimilarity between each pair of cities.[44] 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 16 with respect to d_{jl} and h_{jl}.

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

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

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

5.2.2 The impact of distance: marginal effects

Because of the city fixed effect in spatial frictions, the effect of distance is not uniform. Table 10 reports the distribution of these marginal effects for all city pairs $(j,l) \in \mathscr{J} \times \mathscr{J}_j$ and for the city pairs (j,Paris). All the estimates of $\{s_1^i, s_2^i, s_3^i, s_4^i\}_{i=e,u}$ that enter in the specification of the spatial frictions are significantly different from zero.[45] 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)			
	$\frac{\partial s^u_{jl}}{\partial d_{jl}}$	$\frac{\partial s^e_{jl}}{\partial d_{jl}}$	$\frac{\partial s^u_{jl}}{\partial h_{jl}}$	$\frac{\partial s^e_{jl}}{\partial h_{jl}}$	$\frac{\partial s^u_{jP}}{\partial d_{jP}}$	$\frac{\partial s^e_{jP}}{\partial d_{jP}}$	$\frac{\partial s^u_{jP}}{\partial h_{jP}}$	$\frac{\partial s^e_{jP}}{\partial h_{jP}}$
Min	-0.1475	-2.6985	-0.1824	-2.5095	-0.0791	-1.4863	-0.1201	-1.8084
1st Qu.	-0.0099	-0.1200	-0.0185	-0.1452	-0.0061	-0.1061	-0.0125	-0.1218
Median	-0.0015	-0.0001	-0.0028	-0.0001	-0.0012	-0.0001	-0.0022	-0.0001
Mean	-0.0074	-0.1561	-0.0133	-0.1711	-0.0058	-0.1054	-0.0107	-0.1319
Sd	0.0122	0.3453	0.0207	0.3651	0.0108	0.2357	0.0185	0.2921
3rd	-0.0002	-0.0000	-0.0003	-0.0000	-0.0001	-0.0000	-0.0003	-0.0000
Max	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
N	39601	39601	39601	39601	199	199	199	199

Our estimates for Equation 17 indicate that the mobility cost function is given by $\widehat{c}_{jl} = 9.018 + 27.000 d_{jl} - 2.998 d_{jl}^2$, with distance measured in 10^5 km. This function is positive and increasing for all possible values of d, which means that contrary to Kennan & Walker (2011), we do not find any evidence of negative mobility costs (or relocation subsidies) in the French labor market. Given that log wages are used in the estimation, the monetary equivalent of this cost function amounts to an average value of €9,360, which is approximately equal to the annual minimum wage.[46] As shown in

[45] Standard errors are available upon request.

[46] Although this cost is still high, one must bear in mind that mobility costs also encompass two other features: first, they include relocation costs, and particularly transaction costs on the housing market. Such costs may be high, especially for homeowners. Second, mobility costs, in all generality, must include a measure of psychological costs. Even if those are

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 Decompositions

5.3.1 The mobility rate

[FORTHCOMING]

We can use our parameter estimates for spatial frictions and mobility costs to account for the low mobility puzzle described in Figure 1. According to our model, the migration rate out of city j can be computed as follows:

$$MIG_j = \sum_{k\in\mathcal{J}_j} \left[\left(\frac{u_j}{m_j}\right) s^u_{jk}\lambda^u_k \overline{F}_k(q_{jk}(\phi_j)) + \left(1-\frac{u_j}{m_j}\right) s^e_{jk}\lambda^e_k \int_{\underline{w}}^{\overline{w}} \overline{F}_k(q_{jk}(x))dG_j(x)\right] \tag{18}$$

and the nationwide migration rate is obtained by simple reweighting:

$$MIG = \sum_{j\in\mathcal{J}} \left(\frac{m_j}{M}\right) MIG_j \tag{19}$$

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.

5.3.2 The city size wage premium

We have shown that cities of different sizes exhibit very different features with respect to their internal job matching process, as well as a substantial heterogeneity with respect to their level of openness. More than size, the main drivers of openness are related to the position of each city in the system. We now quantify the respective impact of these two dimensions on the average wage level in each city. In our controlled environment, city-specific wage dispersion can be expressed as a function of the labor market primitives. The expected wage in city j is $\mathbb{E}(w|j) = \int_{\underline{w}}^{\overline{w}} x g_j(x) dx$, where $g_j(\cdot)$ is given by Equation 42:

$$\begin{aligned}\mathbb{E}(w|j) &= \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\daleth_j(x)\lambda_j^u u_j\right)dx}_{\text{Local off-the job search}} + \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\daleth_j(x)\lambda_j^e (m_j - u_j) G_j(x)\right)dx}_{\text{Local on-the-job search}} \\ &+ \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\daleth_j(x)\lambda_j^u \sum_{k\in\mathscr{J}_j} s_{kj}^u \psi_{kj}(x) u_k\right)dx}_{\text{Mobile off-the-job search}} + \underbrace{\int_{\underline{w}}^{\overline{w}} x\left(\daleth_j(x)\lambda_j^e \sum_{k\in\mathscr{J}_j} s_{kj}^e (m_k - u_k) G_k(q_{jk}(x))\right)dx}_{\text{Mobile on-the-job search}}\end{aligned} \tag{20}$$

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

Table 12 reports the decomposition described in Equation 20 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.[47]

5.3.3 Unemployment

XXXXX

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

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

5.4 Experiments

Standard decompositions do not take into account the interdependence between the different parameters of the economy. However we can mitigate this concern, by taking advantage of the dimensionality of our model in order to simulate a matching function.

In the following experiments, distance is fixed: the location of each city remains exogenous.

Other truly exogenous parameter: the matching technology? (η, Ω)

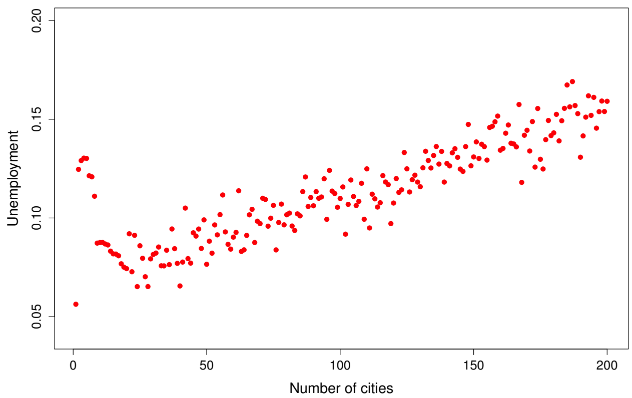
5.4.1 The optimal number of cities

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 express all the city-specific parameters as a polynomial function of the number of firms n_j, population m_j and area a_j:

$$M(n_j, m_j, a_j) \approx (\lambda_j^u, \lambda_j^e, \delta_j, s_{j0}^e, s_{j0}^u, s_{0j}^e, s_{0j}^u, \alpha_j, \beta_j, \gamma_j) \tag{21}$$

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.[48]

Figure 5: Unemployment and the number of cities



Notes: (i) Aggregate unemployment rate as a function of the number of cities, taking city location and relative size fixed; (ii) We use the values of the labor market primitives and fixed effects predicted by the estimates of an OLS regression of Equation 21.

Figure 5 reports the relationship between the number of cities and aggregate unemployment. It is not trivial.[49] 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.[50] As we add more cities, and smaller cities start to fill the vacant space between the large 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 stiffer competition for the most attractive ones.

5.4.2 What is the optimal policy design to change local attractiveness?

Transfers and incentives to leave bad labor markets?

[48] This experiment ignores the externalities (congestion, public good provision) associated with city size. It serves as a baseline to highlight the relationship between spatial segmentation, proxied by the number of cities, and economic performance, proxied by aggregate unemployment.

[49] The 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.

[50] Arguably, this result depends on the shape of the French urban network, where the largest cities are far away from each other. However, this feature is shared by many developed countries.

Conclusion

In this paper, we propose a job search model to study persistent inequalities across local labor markets. Using a panel 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's (2007) 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. We also introduce a new level of complexity in the definition of spatial constraints upon the labor market, by explicitly distinguishing between mobility costs and informational frictions. 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, it cannot be used to analyze the sorting of workers across cities, even though this has been shown to be a major driver of spatial wage differences (Combes, Duranton & Gobillon, 2008). Second, 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, Postel-Vinay & Robin (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. More generally, this paper leaves largely unexplored the firms' side of the dynamic location model. Whereas a mere extension à-la Meghir et al. (2015) would not convey much interest without an explicit theory of location choice, agglomeration economies and wages, we believe such explicit theory to be a promising venue for future research.

References

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

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

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

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

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

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

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

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

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

Eckstein, Z. & van den Berg, G. J. (2007). Empirical labor search: A survey. *Journal of Econometrics*, *136*(2), 531–564.

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.

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

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

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

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., & Sun, H. (2014). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review*, *104*(4), 1172–1210.

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

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.

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

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

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

Lkhagvasuren, D. (2012). Big locational unemployment differences despite high labor mobility. *Journal of Monetary Economics*, *59*(8), 798–814.

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. (2015). Wages and informality in developing countries. *American Economic Review*.

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.

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

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

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

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

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

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.

van den Berg, G. J. & Ridder, G. (1998). An empirical equilibrium search model of the labor market. *Econometrica*, *66*(5), 1183–1222.

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 Theory: proofs and discussions

A.1 Expressions

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

$$V_j^u \equiv V_j^e(\phi_j) \tag{22}$$

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

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

Equations 2 and 3 can be rewritten as:

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

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

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

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

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

where:

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

Finally, using Equations 22 and 24, we find that ϕ_j and $q_{jl}(w)$ are given by Equations 4 and 5.

A.2 Existence and uniqueness

From Equation 5, we derive the following proposition:

Proposition 8 *Let's denote by $\mathscr{W} = [\underline{w}, \overline{w}]$ the support of the wage distribution. $\mathscr{W}$ 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 value w_0. Given that q_{jl} is differentiable, equation 32 can be restated in the differential form as:

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

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

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

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 Discussion: the impact of mobility costs

We discuss here the impact of introducing mobility costs, both from a theoretical viewpoint and in relationship to the frictionless migration literature. We first show that, in order to fully derive a model where workers are only described by their current situation, we need to make an additional assumption regarding the impact of mobility costs on their mobility decisions. Then, we show how we can recover a more classical expression that summarizes the determinants of the migration decision in a frictionless framework.

A.3.1 Past dependence in indifference wages

Mobility costs yield a non-trivial past dependence in the definition of indifference wages: they impact the wage that will be accepted in the new city, which in turn impacts future wage growth prospects in this new city; this difference in terms of option value will have an additional impact on indifference wages, and so forth. As shown in Equation 31, this dynamic feedback effect will mechanically exacerbate the difference between static and dynamic indifference wages:

$$q_{jl}(w) = \chi_{jl}(w) + (r+\delta_l)c_{jl} + \lambda_l^e \int_{\chi_{jl}(w)}^{q_{jl}(w)} \Xi_l(x)dx + \sum_{k\in\mathscr{J}_l} s_{lk}^e \lambda_k^e \int_{q_{lk}(\chi_{jl}(w))}^{q_{lk}(q_{jl}((w))} \left[\Xi_k(x) - f_k(x)c_{lk}\right]dx \tag{31}$$

Because of this feature, indifference wages do not have a tractable closed-form solution, unless workers are characterized by their entire migration history. Our solution is to assume that 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. Under this assumption, dynamic indifference wages can be recovered thanks to the stationary property of static indifference wages. Equation 31 becomes:

$$q_{jl}(w) = \chi_{jl}(w) + (r+\delta_l)c_{jl} \tag{32}$$

with:

$$\begin{aligned}
\chi_{jl}(w) &= \zeta_{jl}w + \zeta_{jl}\gamma_j - \gamma_l + \zeta_{jl}\delta_j V_j^u - \delta_l V_l^u + \zeta_{jl}\lambda_j^e \int_w^{\overline{w}} \Phi_j(x)dx - \lambda_l^e \int_{\chi_{jl}(w)}^{\overline{w}} \Phi_l(x)dx \qquad (33)\\
&+ \zeta_{jl} \sum_{k\in\mathcal{J}_j} s_{jk}^e \lambda_k^e \left(\int_{\chi_{jk}(w)}^{\overline{w}} \Phi_k(x)dx - \overline{F}_k(\chi_{jk}(w))c_{jk} \right) - \sum_{k\in\mathcal{J}_l} s_{lk}^e \lambda_k^e \left(\int_{\chi_{jk}(w)}^{\overline{w}} \Phi_k(x)dx - \overline{F}_k(\chi_{jk}(w))c_{lk} \right)\\
V_j^u &= \frac{1}{r}\left[b + \gamma_j + \lambda_j^u \int_{\phi_j}^{\overline{w}} \Phi_j(x)dx + \sum_{k\in\mathcal{J}_j} s_{jk}^u \lambda_k^u \left(\int_{\chi_{jk}(\phi_j)}^{\overline{w}} \Phi_k(x)dx - \overline{F}_k(\chi_{jk}(\phi_j))c_{jk} \right) \right] \qquad (34)\\
\Phi_j(x) &= \frac{\overline{F}_j(x)}{r + \delta_j + \lambda_j^e \overline{F}_j(x) + \sum_{k\in\mathcal{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(\chi_{jk}(x))} \qquad (35)
\end{aligned}$$

When thinking about subsequent moves from the job in city l that is currently under consideration, a worker in city j takes as a fallback value her initial discounted utility $V_j^e(w)$ and therefore, a wage in city l equal to $\chi_{jl}(w)$.[51] As shown in Equation 33, this assumption preserves the main dynamic effect of mobility costs, based on the relative accessibility of city j and city l and measured by the difference between c_{jk} and c_{lk} for every third city k. A behavioral interpretation 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 $\chi_{jl}(w)$) 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. Note that equation 32 also gives the inverse function $q_{jl}^{-1}(w) = \chi_{lj}\left(w - (r+\delta_l)c_{jl}\right)$.

A.3.2 Mobility costs and frictions: interpretation

The traditional frictionless migration literature appeals to the existence of high mobility costs to account for low migration patterns. Matching frictions alone cannot be a substitute for mobility costs because they cannot reconcile within-city job-finding patterns with between-city mobility rates, even in the presence of differences in local amenities. However, our model makes it possible to assume that the only spatial constraints are mobility costs: in terms of the model, this means that $s_{jk} = 1$. This assumption dramatically affects the computation of transition rates, populations and unemployment rates: under constant matching rates, the predicted mobilities skyrocket, unless indifference wages

[51] See the Pandora stopping problem described in Weitzman (1979) for a similar assumption. This is akin to partial myopia. As discussed in Eckstein & van den Berg (2007), full myopia would be for workers to always consider w as the fallback wage.

become prohibitively high. The equation of indifference wages simplifies to:

$$\begin{aligned} q_{jl}(w) &= \zeta_{jl} w + (r+\delta_l) c_{jl} + \zeta_{jl}\gamma_j - \gamma_l + \zeta_{jl}\delta_j V_j^u - \delta_l V_l^u \\ &+ \sum_{k\in\mathcal{J}} (\zeta_{jl}-1)\lambda_k^e \int_{\chi_{jk}(w)}^{\overline{w}} \Phi_k(x)dx - \sum_{k\in\mathcal{J}} \lambda_k^e \overline{F}_k(\chi_{jk}(w))\left(\zeta_{jl}c_{jk} - c_{lk}\right) \end{aligned} \tag{36}$$

Equation 36 shows that the crucial role of mobility costs in frictionless models is comprised by our model as a special case where unemployment risk can be neglected. Indeed, this is easy to see that:

$$\lim_{(\delta_j,\delta_l)\to(0,0)} q_{jl}(w) = w + rc_{jl} + \gamma_j - \gamma_l - \sum_{k\in\mathcal{J}} \lambda_k^e \overline{F}_k(\chi_{jk}(w))\left(c_{jk} - c_{lk}\right) \tag{37}$$

The first four terms on the right-hand side feature a classical expression where mobility decisions are driven by wage levels, capitalized mobility costs and differences in local amenities. Since differences in local amenities cannot explain the coexistence of low mobility rates both out of and into the same city, the only factor left is mobility costs. Because of on-the-job search, the relative accessibility of cities j and l, which determines the cost of subsequent moves, still comes into play: if city l, in addition to being far from city j, is also not easily accessible to the other cities, this will reduce the migration rate to city l even more.

B Algorithm and numerical solutions

B.1 Algorithm

Let $g(\cdot) \equiv \{g_j(\cdot)\}_{j\in\mathcal{J}}$ 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 5 to recover indifference wages $q(\cdot)$
4. Solve Equation 11 to recover equilibrium population $\mathcal{M}$
5. Solve Equation 13 to update the distribution of job offers $F(\cdot)$
6. Solve Equation 15 to update the distribution of number of firms $\mathcal{N}$
7. Solve Equation 8 to update the distribution of local amenities Γ

8. Update θ using the maximum of $\mathscr{L}(\theta)$.

9. Repeat steps 3 to 8 until convergence.

B.2 Indifference wages

The model raises several numerical challenges, in particular in steps 3 and 5. In step 3, $q(\cdot)$ defines a system of $J^2 - J$ equations, to be solved $\dim(\mathbb{W})$ times. Moreover, since $\Phi_j(\cdot)$ is a function of all $\{\chi_{jk}(\cdot)\}_{k\in\mathscr{J}_j}$, the numerical integration of $\Phi_j(\cdot)$ requires a prior knowledge of the functional form of all $\{\chi_{jk}(\cdot)\}_{k\in\mathscr{J}_j}$. 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.[52] 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.

3.1 Set an initial guess $\chi^C_{jl}(w) = \chi^0_{jl}(w)$ such that all indifference wages on the right-hand side of Equation 33 are equal to the starting wage:

$$\begin{aligned}\chi^0_{jl}(w) &= \zeta_{jl}w + \zeta_{jl}\gamma_j - \gamma_l + \zeta_{jl}\delta_j V^{u0}_j - \delta_l V^{u0}_l + \zeta_{jl}\lambda^e_j \int_w^{\overline{w}} \Phi^0_j(x)dx - \lambda^e_l \int_w^{\overline{w}} \Phi^0_l(x)dx \qquad (38)\\ &+ \zeta_{jl}\sum_{k\in\mathscr{J}_j} s^e_{jk}\lambda^e_k\Big(\int_w^{\overline{w}} \Phi^0_k(x)dx - \overline{F}_k(w)c_{jk}\Big) - \sum_{k\in\mathscr{J}_l} s^e_{lk}\lambda^e_k\Big(\int_w^{\overline{w}} \Phi^0_k(x)dx - \overline{F}_k(w)c_{lk}\Big)\end{aligned}$$

with:

$$rV^{u0}_j = b + \gamma_j + \lambda^u_j \int_{\phi_j}^{\overline{w}} \Phi^0_j(x)dx + \sum_{k\in\mathscr{J}_j} s^u_{jk}\lambda^u_k\Big(\int_{\phi_j}^{\overline{w}} \Phi^0_k(x)dx - \overline{F}_k(\phi_j)c_{jk}\Big) \qquad (39)$$

$$\Phi^0_j(x) = \frac{\overline{F}_j(x)}{r + \delta_j + \lambda^e_j\overline{F}_j(x) + \sum_{k\in\mathscr{J}_j} s^e_{jk}\lambda^e_k\overline{F}_k(x)} \qquad (40)$$

3.2 Use $\chi^C_{jl}(w)$ and interpolation techniques to numerically evaluate the integrals in Equations 33 and 34.

3.3 Update $\chi^C_{jl}(w)$, and repeat step 3.2 until convergence, then recover $q(\cdot)$ through Equation 32.

B.3 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

[52] See section 4.2 for details.

Equation 13. First, for any system of three cities or more, the system can only be solved numerically.[53] 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, and following Meghir et al. (2015), we assume that $F(\cdot)$ follows a parametric distribution:

$$\hat{F}_j(x) = \text{betacdf}\left(\frac{x-b}{\overline{w}-b}, \alpha_j, \beta_j\right) \tag{41}$$

where $\text{betacdf}\left(\cdot, \alpha_j, \beta_j\right)$ is the cdf of a beta distribution with shape parameter α_j and scale parameter β_j. Then, since our empirical counterparts are based on real wages, we treat the empirical cdf $G(\cdot)$ as unknown and we estimate the set of parameters $\alpha \equiv \{\alpha_j\}_{j\in\mathcal{J}}$ and $\beta \equiv \{\beta_j\}_{j\in\mathcal{J}}$ which minimize the distance between the empirical cdf $G(\cdot)$ and its theoretical counterpart. This theoretical counterpart is given as the solution to the following functional equation, derived from Equation 13:

$$g_j(w) = f_j(w) \times \frac{\lambda_j^u\left(\psi_{jj}(w)u_j + \sum_{k\in\mathcal{J}_j} s_{kj}^u \psi_{kj}(w) u_k\right) + \lambda_j^e\left((m_j - u_j)G_j(w) + \sum_{k\in\mathcal{J}_j} s_{kj}^e (m_k - u_k) G_k(q_{kj}^{-1}(w))\right)}{(m_j - u_j)\left(\delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k\in\mathcal{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(q_{jk}(w))\right)} \tag{42}$$

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) such that $\hat{F}_j^0(x) = \hat{G}_j(x)$ the beta approximation of G. At step 5, we need a solution $G(\cdot)$ to Equation 42 in order to update (α, β). We develop a simple iterative process based on Euler's approach. That is,

5.1 At initial iteration, set $q_{jl}(w) = w$. Then, $g_j^0(w)$ becomes:

$$g_j^0(w) = f_j(w) \times \frac{\lambda_j^u\left(u_j + \sum_{k\in\mathcal{J}_j} s_{kj}^u u_k\right) + \lambda_j^e\left((m_j - u_j)G_j^0(w) + \sum_{k\in\mathcal{J}_j} s_{kj}^e (m_k - u_k) G_k^0(w)\right)}{(m_j - u_j)\left(\delta_j + \lambda_j^e \overline{F}_j(w) + \sum_{k\in\mathcal{J}_j} s_{jk}^e \lambda_k^e \overline{F}_k(w)\right)} \tag{43}$$

and equation 13 becomes a standard ODE.

5.2 Set the step size h and use Euler's 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 42.

5.5 Repeat steps 5.3 to 5.4 until convergence.

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

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

$$G_j(w_1) = G_j(w_0) + hg_j(w_0) \tag{44}$$

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

C Data

C.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 900 euros per month (the net minimum wage is around 900 euros): or a monthly wage higher than 8,000 euros: The first case is considered as measurement error; whereas the second case reflects a real situation, it extends the support of wage distributions too dramatically for very few individuals (about 1% of the population).
- 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 Min	Mean	Median	Max	Number of individuals by metro Min	Mean	Median	Max
2002	310,153	332,446	95	1,581	433	84,302	97	1,662	445	90,452
2003	297,697	311,309	99	1,505	406	80,981	101	1,556	412	84,950
2004	308,179	321,557	107	1,558	424	84,104	108	1,607	433	88,027
2005	310,949	325,580	71	1,573	441	84,911	72	1,627	449	89,600
2006	316,613	332,848	105	1,604	436	86,712	106	1,664	449	91,525
2007	313,693	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

C.2 Empirical moments used in the first column in Table 7

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

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

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

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

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

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

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

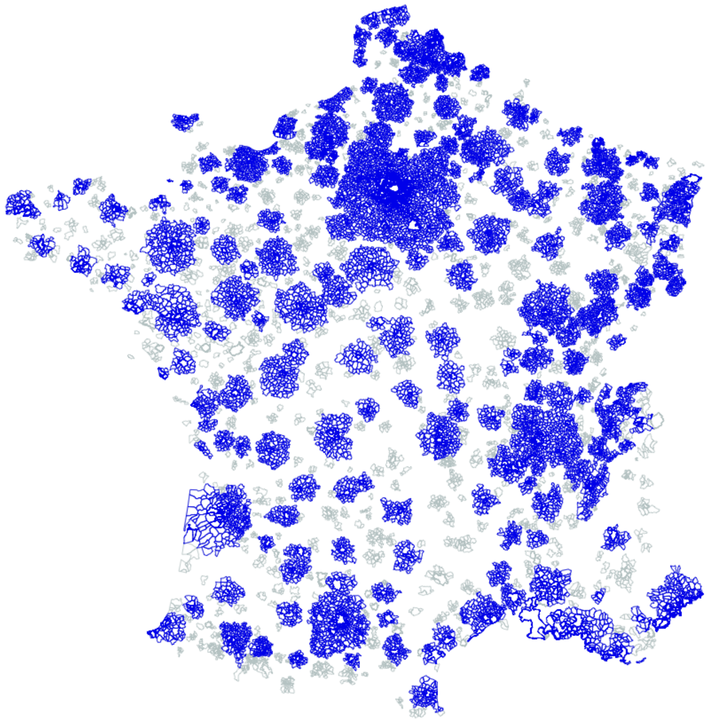
Accepted wages ee into city l: average wage following a job-to-job transition into city l observed over the period; the average is the sum of the accepted wages ee from city j to city l as defined below, weighted by the number of transitions ee from city j to city l

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.

Share of local firms in city j: ratio of the number of firms observed in city j over the period to the total number of firms observed in all cities $l \in \mathscr{J}$ over the period.

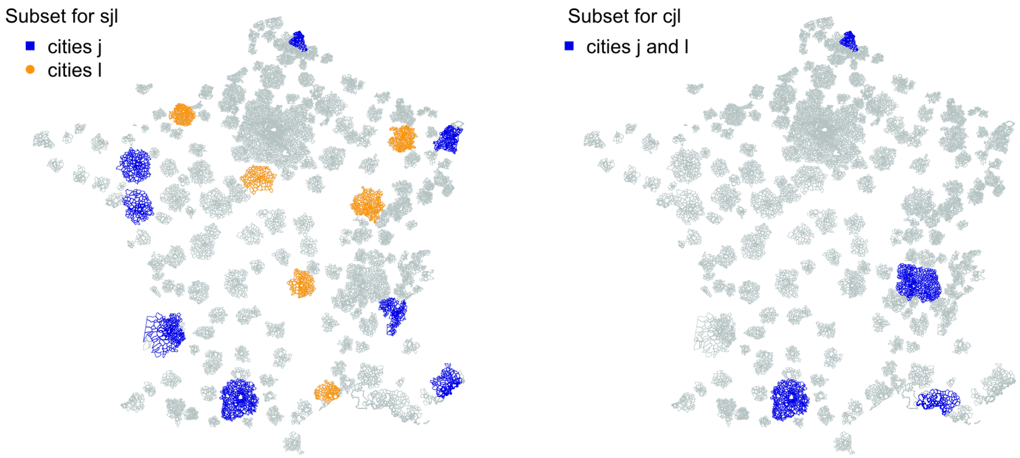
D Figures

Figure 6: 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

Figure 7: The metropolitan areas in subset $\mathcal{T}_1$ (left) and subset $\mathcal{T}_2$ (right)

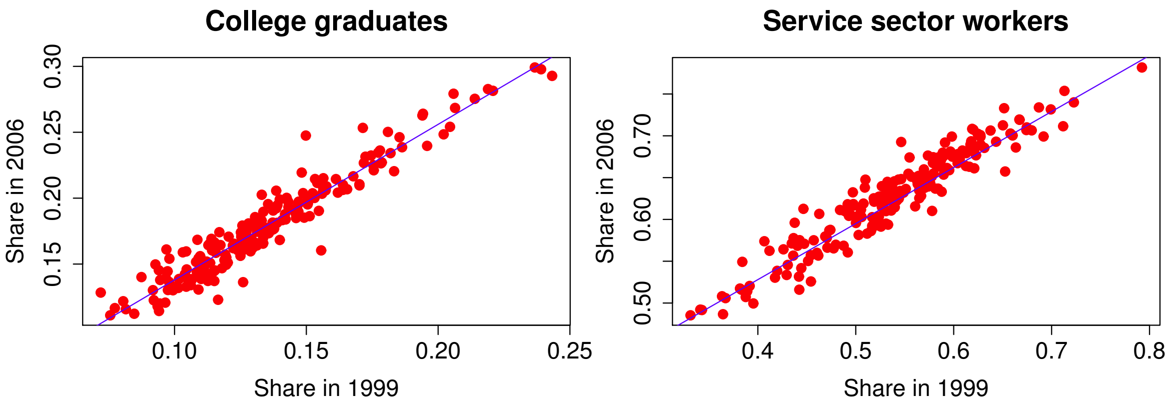


Notes: (i) see Figure 6; (ii) Subset $\mathcal{T}_1$ is used to identify the effect of physical distance and dissimilarity on spatial frictions based on pair-specific out-of-unemployment and job-to-job transition rates; subset $\mathcal{T}_2$ is used to identify the effect of physical distance on moving costs based on pair-specific average accepted wages after a job-to-job transition with mobility.

E Worker heterogeneity

As documented elsewhere, developed countries such as France have witnessed an increase in overall skill level and in the share of the service sector during the past decades. Over a long period, these wide recomposition patterns make it unlikely that an equilibrium model could effectively be used. We do not address this issue in this paper. However, we believe that, as a first-order approximation, the assumption of workers' homogeneity is not very costly when focusing on a short time-span. As shown in Figure 8, these reallocations, roughly described as a linear process, affect all cities in a very similar fashion between 1999 and 2006 and the position of each city in the hierarchy of skill and sectoral composition is very stable across the period.

Figure 8: Heterogeneity and stability in skill and sectoral composition



Notes: (i) Shares are computed on the 25-54 age bracket for the population of men (left) and the population of men workers (right) and for the 200 largest metropolitan areas in continental France, keeping a constant municipal composition based on the 2010 "Aires Urbaines" definition; (ii) The respective equations of the least squares line are $\hat{C}_{06} = 1.18 \times C_{99} + 0.02$ (left) and $\hat{S}_{06} = 0.67 \times S_{99} + 0.26$ (right). *Source: INSEE, Census 1999 and 2006*

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

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