The effect of location on finding a job in the Paris region*

L. Gobillon[†], T. Magnac[‡], H. Selod[§]

March 5, 2009

*We are grateful to a co-editor and two anonymous referees for their comments and we thank as well the participants at seminars at INRA, CREST, University of Tokyo, Louvain-la-Neuve and the North American Meetings of the Regional Science Association International 2006. We thank the French National Employment Agency (ANPE) and the Ministry of Transport's Regional Directorate (DREIF) for providing us with the data. All remaining errors are ours. The findings, interpretations, and conclusions expressed in this paper are ours and do not represent the view of our employers, including the World Bank, its executive directors, or the countries they represent.

[†]Institut National d'Etudes Démographiques (INED), PSE-INRA and CREST. Address: INED, 133 boulevard Davout, 75980 Paris Cedex 20, France. E-mail: laurent.gobillon@ined.fr.

[‡] Toulouse School of Economics (GREMAQ & IDEI) Address: Manufacture des Tabacs, 21 allée de Brienne, 31000 Toulouse, France. E-mail: magnac@cict.fr.

[§]The World Bank and CREST. Address: 1818 H Street NW, Washington DC 20433, United States of America: hselod@worldbank.org.

Abstract

An important but empirically debated issue in spatial economics is whether spatial differences in unemployment reflect residential sorting on invividividual characteristics or a true effect of location. We investigate this issue in the 1,300 French muncipalities that constitute the Paris region and across which there is overwhelming evidence of spatial disparities in unemployment durations. We resort to a methodology that enables us to disentangle individual and unspecified local effects. In order to control for individual determinants, we estimate a proportional hazard model stratified by municipality using an exhaustive dataset of all unemployment spells starting in the first semester of 1996. This model allows us to recover a survival function for each municipality that is purged of individual observed heterogeneity. We show that only around 30% of the disparities in the observed determinants of the survival rates relate to individual variables. Nearly 70% of the remaining disparities are captured by local indicators which we show to be mainly correlated with local measures of residential segregation. We are also able to show that local and individual characteristics reinforce one another in their contribution to spatial disparities in unemployment duration.

Keywords: Unemployment, Duration models, Economic geography, Urban economics.

JEL Codes: C41, J64, R23

1 Introduction

The determinants of urban unemployment have raised the interest of economists for decades. In the US, two major trends of literature have tried to explain how location could have an adverse impact on employment, involving a variety of mechanisms. The first set of works is the so-called *spatial mismatch* literature which investigates how the physical disconnection from jobs can exacerbate unemployment among low-skilled minority workers (see Ihlanfeldt and Sjoquist, 1998, for an empirical survey, and Gobillon, Selod and Zenou, 2007, for a theoretical one). The second set of works investigates the impact of *residential segregation* on the poor labor-market outcomes of ghetto residents (see e.g. Wilson, 1996, Cutler and Glaeser, 1997). In both literatures, papers usually resort to cross-section methods and try to explain individual unemployment probabilities or local unemployment rates (see e.g. Ihlanfeldt, 1993, Conley and Topa, 2002, Weinberg, 2000 and 2004).

In this paper, we focus on the local determinants of unemployment *duration*. Only a few papers, mainly on the US, have studied unemployment dynamics at the individual level with a spatial perspective (Holzer, Ihlanfeldt and Sjoquist, 1994, Rogers, 1997, Dawkins, Shen and Sanchez, 2005, Johnson, 2006). In these works, authors usually investigate the impact of local indicators proxying for spatial mismatch or residential segregation in an unemployment duration model. They typically estimate a proportional hazard model with a single baseline hazard common to all locations, a set of individual variables and local indicators. We adopt a much broader approach that consists in estimating a baseline hazard function for each location while controlling for individual characteristics in a proportional hazard model. This key methodological innovation, known as the Stratified Partial Likelihood Estimator (SPLE), was first proposed by Ridder and Tunali (1999) in another context. We apply it in this paper to a large administrative dataset containing records of unemployment spells and adapt it to include some new econometric features.

Compared to the previous literature, the advantages of our empirical strategy are threefold. First, we do not need to choose a specific function for the local hazard functions. We can thus measure the overall effects of location without only focusing on a few arbitrarily selected mechanisms, proxied by criticizable local indicators. Second, we allow the effect of location to vary depending on the time spent unemployed. We can thus assess the effect of location on the short run (say, after 6 months) and on the long run (say, after two years). Third, the model is sufficiently versatile to allow us to further restrict local hazard rates while controling for the generality of these restrictions.

The estimation procedure has three steps. The first step consists in estimating a proportional hazard model with an unspecified municipality-specific hazard baseline hazard. In the second step of the estimation procedure, we impose that the municipality effects are multiplicative in the hazard rate. This multiplicative component model summarizes the local effects through a single indicator. In the third stage, we assess how this local indicator may capture local determinants reflecting the different mechanisms put forward by the literature. This is done by regressing municipality effects on these variables and computing their explanatory power. We do not interpret this last stage as a causal regression because omitted variable or reverse causality concerns cannot be dismissed. Our procedure however ensures that the previous stages are immune to these endogeneity issues so that we can separate the robust estimation of local effects from less robust results obtained in the third stage.

Yet, we cannot easily deal with individual unobserved heterogeneity in our proportional hazard specification. This is in line with Baker and Melino (2000)'s finding that identification of both flexible hazard rates and the distribution of unobserved heterogeneity is fragile in empirical studies. Moreover, individual unobserved heterogeneity like, for instance, omitted variables related to preferences, is partially captured since we model the hazard rate function at the level of the municipality. Furthermore, we apply goodness of fit test procedures under the form of a Kolmogorov statistic as developed in Andrews (1997) and show that the model fits the data very well at the level of each municipality.

Our approach requires a very large dataset comprising enough unemployment spells in each location. We use a unique exhaustive administrative dataset available for the Paris region from which we extract unemployment spells that started in the first semester of 1996. Unemployment spells can end in three different ways: finding a job, dropping out of the labor force, and right-censorship (including exits for unknown reasons). We model the first two exits in an independent competing risk framework.

Our main empirical results are as follows. We find that controling for individual characteristics explains around 30% of the spatial disparities across municipalities in unemployment durations until finding a job. Among the main individual determinants of unemployment duration as education, it should be stressed that some nationalities, Africans in particular, experience significantly much larger durations until finding a job. Furthermore, the association between average individual characteristics at the municipality level and baseline hazards is positive. Presumably because of sorting effects, individual and local effects reinforce each other in their contribution to spatial disparities of unemployment duration. In other words, durations are not only larger because of an adverse individual characteristic but also because the average of this adverse characteristic is larger at the local level. Finally, nearly 70% of the remaining local disparities are captured by local indicators, mainly segregation indices.

The rest of the paper is structured as follows. In section 2, we provide a short survey of the litterature on how segregation and bad physical accessibility to jobs can increase unemployment duration. Section 3 presents the data and a selection of descriptive statistics to measure spatial disparities. Section 4 details the SPLE method. Section 5 discusses the results. Finally, section 6 concludes.

2 Why should location influence unemployment duration?

The duration of unemployment depends on many factors. To discuss this issue in an orderly manner, it is useful to adopt a job-search perspective considering that exit from unemployment can occur at the end of a three-stage process. In the first stage, workers must wait some time before coming into *contact* with a job opportunity. In the second stage, an *offer* from an employer may materialize. Finally, workers may *accept or reject* the offer depending on whether the offered wage is greater or smaller than their reservation wage. With this framework in mind, job seekers who, on average, wait long before experiencing contacts with employers and who have few chances to transform their contacts into offers and matches should experience long unemployment spells. For instance, educated workers could be advantaged in the first stage if they are more efficient in obtaining information about jobs and in contacting firms, or if labor demand is biased in their favor. They may also have an advantage in the second stage if they write better application letters and resumes, and fare better during interviews. However, educated workers may be more likely to reject an offer when they face or anticipate many well-paid outside offers. Other individual and family characteristics such as gender, race/ethnicity, age, experience, marital status or the number and age of children and dependants should also be expected to affect unemployment duration through one or several stages of the job-acquisition process.

This section describes how *location*, i.e. the *disconnection from job opportunities* (in cases where job opportunities are unevenly distributed within a metropolitan area) and/or *residential segregation* (in terms of education, race/ethnicity/nationality or employment status), can also influence the duration of unemployment. We decompose the effects on each stage of the job-acquisition process.

Disconnection from job opportunities may directly affect the time spent searching for a job in the first stage of the process. Indeed, job-seekers residing in areas with few local job vacancies or in areas located far away from employment centers are exposed only to a small pool of vacancies. Residing in loose local labor markets, they should spend more time searching before getting into contact with a potential employer. Of course, job-seekers also have the possibility to search for jobs in other areas. But having to search away from one's area of residence penalizes job seekers. At least three reasons may come into view. Firstly, because of informational frictions, *job-seekers may not search efficiently far away from their residences.* For instance, workers residing far away from job opportunities may not hear about job offers when firms resort to recruiting methods that favor the local labor force (i.e. by posting 'wanted' signs in retail shops, or by choosing not to publicize job offers beyond a certain distance). Alternatively, job-seekers may obtain only partial information on the location of distant jobs or may have only a vague idea about the types of jobs offered in parts of the metropolitan area they are not familiar with. They may end up searching in the wrong places (Ihlanfeldt, 1997, Stoll and Raphael, 2000). Secondly, because search is costly, *workers may restrict their search horizon at the vicinity of their neighborhood.* They may search less often in order to reduce the number of job-search trips or may not search at all for jobs located in distant places. In this context, access to public transport or car ownership can reduce job-search costs and expand the job-search horizon (Stoll, 1999). Thirdly, the individual search effort may depend on the local cost of living so that *workers residing in areas disconnected from job opportunities may not search intensively.* It has been argued that workers residing in such areas usually incur low housing costs and thus may feel relatively little less pressure to actively search for a job in order to pay their rent (Smith and Zenou, 2003, Pattachini and Zenou, 2006).

Disconnection from job opportunities may also reduce the frequency of job proposals in the second stage. *Employers may then be reluctant to propose jobs to distant workers* because commuting long distances would make these workers less productive (they would show up late or be tired due to excessive commuting, see Zenou, 2002).

Distance to job opportunities may also reduce the probability of a job acceptance in the third stage. Indeed, *workers may reject a job offer that would involve commutes that are too long* if commuting to that job would be too costly in view of the proposed wage (Zax and Kain, 1996). In other words, distance is likely to make the offered wage net of commuting costs drop below a worker's given reservation wage.

The effect of *residential segregation* on the first stage of the job-acquisition process is also likely to be harmful to the extent that job contacts often occur through friends and relatives (Mortensen and Vishwanath, 1994). Because *social networks are at least partly localized*, when the unemployment rate is high in a given area, workers are less likely to know employed neighbors that can let them know about existing vacancies (Calvó and Jackson, 2004, Selod and Zenou, 2006).

Residential segregation is also likely to reduce the probability for a worker residing in a segregated area to receive a job offer. This is because employers may discriminate against residentially segregated workers, a practice known as *redlining* (see Wilson, 1996, for stories of firms not hiring workers located in 'bad' neighborhoods). For employers, the motivation can hinge upon the stigma or prejudice associated with the residential location of candidates (*sheer discrimination*), or because they consider that, on average, workers from stigmatized areas have bad work habits or are more likely to be criminal (*statistical discrimination*). In industries and jobs in which workers are in contact with customers, employers may discriminate against residentially-segregated workers in order to satisfy the perceived prejudices of their clients, a practice known as *customer discrimination* (Holzer and Ihlanfeldt, 1998). In France, the issue of *redlining* is increasingly being put forward in the public debate to account for the unemployment of the young adults that reside in distressed areas. To our knowledge, however, the issue has not yet been studied empirically.

All these economic mechanisms suggest that the rate at which workers leave unemployment, and thus the duration of unemployment, depends on both individual characteristics and local features. In the present paper, we propose a methodology to disentangle individual and unspecified local effects. We explore the nature of local effects by regressing them on indices of segregation and distance to job opportunities. We assess the overall impact of these indices on finding a job, but we do not try to identify through which specific mechanisms they percolate.

3 Description of the Data

3.1 The area of study

The paper focuses on the Ile-de-France region (the Paris region hereafter), an administrative unit of 10.9 million inhabitants distributed over 1,280 municipalities centered around the city of Paris and the 20 administrative subdistricts of Paris (which will be treated as municipalities in the analysis). These 1,300 spatial units may have very different population sizes which range from 225,000 in the most populous Parisian subdistrict to small villages located some 80 km away from the center of Paris. They correspond more or less to the Paris Metropolitan Area as can be seen from Graph 1 which represents the population density in each municipality.

[*Insert Graph* 1]

Graph 2 provides evidence that the studied area exhibits large spatial disparities in the local unemployment rates across municipalities. In particular, the unemployment rates in municipalities located to the North-East of Paris are more than four times higher than in most municipalities located to the West.

[*Insert Graph* 2]

3.2 The ANPE historical file

We use the historical file of job applicants to the National Agency for Employment (*Agence Nationale pour l'Emploi* or *ANPE* hereafter) for the Paris region to study spatial disparities in unemployment durations. This database provides a quasi-exhaustive list of unemployment spells in the region as it has been estimated that 90% of job seekers in France are indeed registered with the ANPE (Chardon and Goux, 2003). The reason is that registering with the ANPE is a prerequisite for unemployment workers to be able to claim their unemployment benefits. A significant share of those not eligible for unemployment benefits—as for instance first time job seekers—also register with the ANPE to assist them in their job search.

The ANPE is organized in hundreds of local agencies and unemployed workers register in the agency closest to their residence.[1] The exhaustive dataset that we have for the Paris region contains information on the exact date of an application (the very day), the unemployment duration (in days), and the reason for which the application came to an end. Along with the municipality where the individual lives and registers, a set of socio-economic characteristics were reported upon registration with the employment agency: age, gender, nationality, diploma, marital status, number of children and disabilities. To build our working sample, we select individuals who applied to the employment agency between January 1 and June 30, 1996 and who lived in the Paris region at that time. As we have information on unemployment spells until 2003, starting as early as 1996 enables us to follow unemployed workers over a long period of time and to minimize the number of incomplete spells due to the end of the observation period (which only concerns 0.6% of the exits in our sample. After deleting the very few observations for which socio-economic characteristics are missing, we end up with $430,695$ observations on individual unemployment durations. More details on the construction and the contents of the dataset are given in Appendix A.

We group the reasons given for the termination of the application with the agency into three types: (1) finding a job, (2) exiting to non-employment, and (3) right censoring, which groups together unknown destinations and incomplete spells.[2] In the following, we assume that right-censoring is independent of the durations until exiting to a job or non-employment and that these two exits are independent, conditional on all observed characteristics including the municipality of residence. A large proportion of exits are right-censored (55.3%), of which 29.5% correspond to an

[1] Except in very specific occupations (artists, ...).

[2] An exit to non-employment covers the following situations: a training period, an illness, a pregnancy, a job accident (as some unemployed workers can in fact work for a very small number of hours), an exemption from the rule imposing to actively search for a job, retirement, or military service. Unknown destinations can result from mobility between four subregions (see text below), an absence at a control, an expulsion for some misbehavior, an absence after a notification, a training or job refusal, a fake statement, the lack of a positive action to search for a job, and other unspecified cases.

absence at a control.[3] The remaining unemployment spells mostly end up with a job (28%) even if exit to non-employment is far from negligible (16.7%). The average unemployment duration for individuals finding a job is 269 days whereas it stands at a higher level of 368 days for individuals who exit to non-employment. The higher unemployment duration for exits to non-employment could possibly reflect the discouragement of workers that could not find a job after a long time.

A crucial issue in this data arises because residential mobility across municipalities could blur our estimation of spatial disparities. In particular, because mobility can change the way unemployment spells are recorded, it could give rise to (*i*) right and left censoring, to (*ii*) measurement errors of local effects, and (*iii*) to departures from the independence assumption because of right censoring. To see how these issues may emerge and why we believe, however, that they may be relatively minor, consider the following lines of reasoning:

First, the French Employment Agency is organized geographically in large subregions called ASSEDIC and when residential mobility brings about a change in ASSEDIC, the unemployment spell is right-censored and a (mistakenly fresh) spell is started in the new ASSEDIC. The unemployment spell is thus cut into two halves, the first spell being right censored, and the second spell being left-censored. Fortunately, there are only 4 ASSEDIC located in the Paris region (West, East, South-East and Paris) and in our date only 4.83% of unemployed workers change ASSEDIC as stated in the reasons for exits. Moreover, double counting is also mitigated by the fact that we consider a flow-sampling window of only six months starting at the beginning of 1996.

Second, even if mobility takes place within the same ASSEDIC, it might bring a change in municipality and local agency. Although the spell is registered as uninterrupted, the stated place of residence may either correspond to that recorded at the origin or at the destination agency.[4] Yet, measurement errors of local effects, if anything, would likely attenuate our measures of spatial disparities.

[3]There is no evidence that these absences would mainly concern unemployed workers that neglected to report they found a job. Indeed, a 2005 follow-up survey on a small random sample of unemployed workers having left the ANPE showed that only approximately half of absentees at controls did find a job, which is not in contradiction with the assumption of independence between right censoring and finding a job.

[4]The administrative treatment by ANPE of spells which ended in a local agency different from the one where it started is very obscure. The two spells are registered and one of them is deleted apparently without following any precise written rule.

Finally, in many aspects, the mobility decision could be partly independent of exiting to a job or to non-employment. For instance, the prospect of saving money by living with one's parents might be an exogenous factor conditional on the individual and local effects. On the other hand, mobility is clearly endogenous if a job seeker moves with the prospect of waiting until finding a job near the new location. Given that mobility is low in France with respect to the US (Baccaini, Courgeau and Desplanques, 1993), we believe that these issues have second order effects compared to the main substantive effects of infrastructure, redlining or discrimination to which we now turn.

3.3 Spatial disparities

The Paris region exhibits stark socio-economic disparities which can broadly be depicted as follows. In the North-East, the population is usually little educated, poor, and composed of blue collar workers. Recent migrant minorities are over-represented. In the West, the population is very educated, rich, and comprises mostly white collars. Minorities of recent immigration waves are under-represented.

To further characterize disparities across municipalities and differences in municipality environments, we compute municipality-specific segregation and job-accessibility variables using several sources.

3.3.1 Census measures of segregation and job accessibility

Segregation is accounted for by the municipality proportion of education and nationality groups computed from the 1999 Population Census. Job accessibility is measured by the job density around each municipality. More precisely, for each municipality we are able to identify all the other municipalities than can be reached within 45 minutes for a given transport mode (private vehicles or public transport). The 45-minute cut-off has been chosen just above the average commuting time of 34 minutes in the Paris region (DREIF-INSEE, 1997). This defines a group of accessible municipalities for which we can calculate the overall job density (the ratio of the number of jobs located in the area to the number of occupied and unoccupied workers residing in the same area).[5] Data on the location of jobs and workers are from the 1999 census. Travel times between

[5] For a discussion of alternative indicators see Gobillon and Selod (2007).

municipalities are estimated at morning peak hours by the French Department of Transportation for 2000 using a transport survey on the Paris region (*Enquête Générale des Transports*).

We compute indices of spatial disparities on these local segregation and job-accessibility variables. The indices we compute are the inter-decile ratio, the inter-decile range, the Gini index and the coefficient of variation, and results are reported in Table 1.[6] We find that spatial disparities across muncipalities are very pronounced for the percentage of African nationalities as the inter-decile ratio is over 9 for the percentages of citizens from North Africa and sub-Saharan Africa. This means that the frequency of African citizens is 9 times larger in municipalities at the 9th decile than at the first. Spatial disparities are also large for segregation in terms of education and stands near 4 for the percentage of individuals with a university degree and around 2.5 for the percentage of individuals with a technical degree. Measures of job accessibility also exhibit significant spatial disparities. The inter-decile ratio for job densities by public transport is 3.

[*Insert Table* 1]

3.3.2 Spatial Disparities for the Unemployed

Spatial disparities in the characteristics of unemployed workers also exhibit a similar pattern over the Paris region. Table 2 reports similar indices of spatial disparities across municipalities for several variables of the ANPE historical file. We measure the spatial disparities in the occurrence of exit types, the unemployment duration conditionally on the type of exit, and the individual variables that we use in our empirical analysis below. As with the census data, the indices we compute are the inter-decile ratio, the inter-decile range, the Gini index and the coefficient of variation.

We first comment the spatial disparities in the proportions of individuals who respectively experience an exit to a job, an exit to non-employment, and right-censoring. For simplicity, we restrict our comments to the inter-decile ratio but other indicators give qualitatively similar results. The inter-decile ratio is fairly large for the probability that unemployment finishes with an exit to

[6] To compute the spatial inter-decile index of a variable, we construct the empirical distribution function of the local average of the variable. Observations are weighted by the population in each municipality. We smooth the empirical distribution by a Gaussian kernel with a Silverman's rule of thumb bandwidth and deciles are retrieved using a very fine grid ($1,000,000$ points).

a job as it reaches 1.73. This means that, if we order municipalities with respect to the proportion of unemployment spells ending with an exit to a job, an unemployment spell has 73 percent more chances to end with an exit to a job in the municipality at the ninth decile than in the municipality at the first decile. The inter-decile ratio is smaller for the probability of an exit to non-employment (1.37) and for right-censoring (1.32). These variations in spatial disparities across exit types calls for a careful conditioning on local effects.

If we now look at unemployment durations conditionally on the type of exit, the inter-decile ratio for unemployment spells ending with an exit to a job reaches 1.37. This means that an unemployment spell ending with an exit to a job lasts 37 percent longer in the municipality at the ninth decile than in the municipality at the first decile. For unemployment spells ending with an exit to non-employment, the inter-decile ratio is even greater and stands at 1.43.

[*Insert Table* 2]

As the above data is right-censored because of exits to other states, these statistics are difficult to interpret. This is why we also assess disparities between municipalities with the help of duration models. For each type of exit and for each municipality, we compute the Kaplan-Meier estimator of the survival function (which takes into account right-censorship). Disparities by exit type can then be assessed by comparing the survival function across municipalities for any chosen duration. As survival functions are well estimated only when the number of unemployed workers is large enough, we restrict our attention to municipalities with a population greater than 5,000 inhabitants in 1999. Graph 3 represents the probability of finding a job before 24 months for each municipality of the Paris region. Disparities are large: the probability of finding a job before 24 months is below 40% in many municipalities of the North-East, whereas it is above 55% in many municipalities of the West. Graph 4 represents the probability of exiting to non-employment before 24 months for each municipality. Contrary to the graph for exit to a job, no specific pattern emerges. This contrast suggests that while job search outcomes strongly depend on location, this is less the case for labor-market participation decisions.

[*Insert Graph* 3 *and* 4]

There are also noticeable spatial disparities in some of the socio-demographic characteristics of unemployed workers. Whereas the spatial disparities in age, sex or marital status are small (see

Table 2), there are much larger disparities for some categories of nationality, education, and family size, as well as for disability. The inter-decile ratio for instance is greater than 5 for the proportion of Africans. In other words, the proportion of Africans among unemployed workers in municipalities at the ninth decile is 5 times greater than in municipalities at the first decile. The inter-decile ratio is above 5 for unemployed workers having three children or more, around 4 for unemployed workers with no diploma, and above 2.5 for disabled unemployed workers.

In conclusion, spatial disparities of individual characteristics are large. Results in Tables 1 and 2 are hard to compare, data sources being so different. Likely origins of differences between them could be reporting errors and composition effects although we did not investigate the point further. It very much depends on the spatial disparities in entries into unemployment which we are not in a position to assess using the data that are available to us.

4 The econometric model

We analyze the empirical associations between unemployment durations and the local context (segregation and job accessibility) using a three-stage procedure. First, we specify a proportional hazard model (PH model hereafter) with individual covariates and a municipality-specific baseline hazard. Parameters related to individual variables are estimated using the stratified partial likelihood estimator (SPLE hereafter) as proposed by Ridder and Tunali (1999). Municipality-specific integrated baseline hazards are then recovered using the Breslow estimator. Second, municipal baseline hazards are restricted to be a multiplicative function of an aggregate baseline hazard function and of municipality effects which are both estimated using the first-stage outputs. A third and final descriptive stage consists in regressing the municipality effects on local indicators of segregation (municipality composition) and job accessibility.

Our approach can be justified as follows. The first two stages allow to estimate municipality effects. For computational reasons, this would be unfeasible by maximum likelihood estimation in one stage only since the number of municipalities $(1,300)$ is too large. The final stage consists in regressing those municipality effects on aggregate variables. It enables us to analyse the correlation between spatial effects and segregation or job accessibility indices, although we do not claim to estimate causal effects in the last stage. Our procedure in three steps guards us against specification errors at each stage. First stage estimates are robust to misspecification of the multiplicative model and of the descriptive regression of municipality effects. Second stage estimates are robust to errors at the descriptive regression stage. Furthermore, all steps contribute to the empirical analysis of spatial disparities in unemployment durations.

4.1 Model Specification

Consider an individual i who enters unemployment (i.e. who enters the ANPE file). His unemployment spell lasts until he finds a job (exit labeled e) or drops out of the labor force (exit labeled ne). The unemployment spell is right-censored if the individual disappears from the records during the observation period or has not experienced an exit before the last day of observation in the panel. A latent duration T_k is associated with each exit $k \in \{e, ne\}$. For an individual i, we denote $\lambda_k\left(\bullet\,|X_i, j(i)\right)$ the conditional hazard rate for exit k where X_i is a set of individual explanatory variables – that are fixed over time in our application – and $j\left(i\right)$, where $j(i) \in \{1, ..., J\}$, is the municipality where the individual is located. We adopt the proportional hazard assumption separating the effect of individual characteristics and the effect of local clusters by writing:

$$\lambda_k\left(t\,|X_i, j(i)\right) = \theta_k^{j(i)}\left(t\right) \exp\left(X_i \beta_k\right) \text{ for } k \in \{e, ne\}, \tag{1}$$

where $\theta_k^j\left(t\right)$ is the baseline hazard rate function for municipality j and exit k. Observe that the effect of local variables is not of the proportional hazard type at this stage since the municipality-specific baseline hazard rate is fully flexible. Additionally, the two latent durations and right-censorship are assumed to be independent so that our framework is an independent competing risk model where observations are clustered.

Observe also that the above specification features come at the expense of overlooking unobserved individual heterogeneity whose presence can bias the estimation of the hazard rates and parameters. Latent durations associated with different types of exit might also be dependent if the effect of individual unobserved heterogeneity influencing the different types of exit are correlated. Lancaster (1990) proposes to introduce individual unobserved heterogeneity in a partial likelihood model by modeling it as a gamma distribution and to estimate parameters using an EM algorithm. Yet, the procedure is burdensome and unfeasible in samples where the number of observations is as large as in ours. An alternative way to proceed would be to difference out individual unobserved terms using multiple spells. In theory, this could be done by redefining clusters as couples (municipality, individual) but the number of applicants appearing twice or more is very small (about 8%) and the issue about biases caused by residential mobility could then become serious. Also note that Baker and Melino (2000) argue that it is difficult to empirically identify the unobserved heterogeneity distribution and flexible hazard rates, and that in the current application, the hazards are fully flexible at the municipality levels. For all these reasons, we decided not to incorporate individual unobserved heterogeneity in our econometric specification (1). We nevertheless discuss below the consequences of its presence on our empirical results. Specifically, we will pay attention to the effects of the sorting of individuals across municipalities according to their unobserved characteristics.

4.2 Stratified Partial Likelihood Estimation (SPLE)

Our estimation follows Ridder and Tunali (1999). Start with the estimation of the effects of individual explanatory variables using the SPLE. Denote $\Omega^j(t)$ the set of individuals at risk of exiting unemployment in municipality j at time t. The probability of individual i experiencing a type-k exit at time t conditionally on someone in the same municipality experiencing a type-k exit is:[7]

$$P_i(t,k) = \frac{\exp(X_i\beta_k)}{\sum_{n\in\Omega^{j(i)}(t)} \exp(X_n\beta_k)} \tag{2}$$

Observe that conditioning on the municipality population at risk (instead of the whole population at risk) makes all municipality-specific baseline hazards cancel out so that we do not need to specify its functional form. The stratified partial likelihood function (calculated on all unemployed workers who experience an exit to a job or to non-employment) is:

$$L = \prod_i P_i(t_i,k_i) = \prod_k L_k(\beta_k) \tag{3}$$

[7] This formula is exact only when time is continuous. In our data where time is expressed in days, several individuals may exit the same day and it is impossible to order them depending on their time of exit. Nevertheless, following Breslow (1974), we consider (2) as an approximation of the conditional probability of exit. In practice, when an individual exits a given day, the risk set includes all the other individuals who exit the same day.

where t_i is the time of exit of individual i, k_i is the type of exit of individual i, and $L_k(\beta_k) = \prod_{i|k_i=k} P_i(t_i, k_i)$ is constructed from all unemployment spells that end with a type-k exit. $L_k(\beta_k)$ is the partial likelihood obtained in the hypothetical context where there is only one possible exit k and where unemployment spells are censored if they end up with the other exit. Notice that each set of parameters β_k can be separately estimated by maximizing the corresponding term L_k under the independent competing risks assumption. Denote $\widehat{\beta}_k$ the estimator.

We can now turn to the estimation of the municipality baseline hazard function. For exit k, the Breslow estimator of the integrated baseline hazard of municipality j, $\Theta_k^j(t)$, is defined as:

$$\widehat{\Theta}_k^j(t) = \int_0^t \frac{I(C^j(s) > 0)}{\sum_{i \in \Omega^j(s)} \exp(X_i \widehat{\beta}_k)} dN_k^j(s), \tag{4}$$

where $I(\bullet)$ is the indicator function, $C^j(s) = card\ \Omega^j(s)$, and $dN_k^j(s)$ is a dummy that equals one if someone in municipality j experiences a type k-exit in an arbitrarily short period of time before date s (and zero otherwise). For each t, the variance of $\widehat{\Theta}_k^j(t)$ can be recovered from Ridder and Tunali's formulas.[8]

4.3 Estimation of Spatial Effects

In the second stage, for each type of exit, we estimate municipality effects that summarize the municipality-specific baseline hazard rates by a single quantity. This is a restriction of the general model. Since the estimation procedure can be applied to each type of exit separately, we restrict our attention to a given exit k and drop subscript k for readability. It should be kept in mind that all parameters analyzed below are exit-specific.

We assume that the municipality-specific baseline hazard rates take a multiplicative form:

$$\theta^j(t) = \alpha^j \theta(t) \tag{5}$$

where α^j is a municipality fixed effect and $\theta(t)$ is a general baseline hazard function. Here, we depart from Ridder and Tunali who adopt an additive form. Indeed, we find it more natural to use

[8] This can be done using their equations $A25$, $A27$ and $A29$ and setting $K = 1$, $t_0 = 0$ and $t_1 = T$ in their equation (22).

a multiplicative specification since, when combining (1) and (5), we obtain a proportional hazard model where spatial effects enter multiplicatively. This is presumably quite restrictive though worth investigating.

Instead of directly implementing the functional estimation of (5), we divide the period $[0,\infty)$ into M intervals whose lower (resp. upper) bound is t_{m-1} (resp. t_m), for $m = 1, ..., M$ (where $t_0 = 0$ and $t_M = \infty$). If we denote $\theta_m = \int_{t_{m-1}}^{t_m} \theta(s)\, ds$ the increment of the integrated aggregate baseline hazard over the interval m, the increment of the integrated baseline hazard rate over a time interval m in a municipality j is given by

$$y_m^j = \Theta^j(t_m) - \Theta^j(t_{m-1}) = \alpha^j \theta_m.$$

An estimate of the average hazard rate y_m^j can be constructed from equation (4) and is:[9]

$$\hat{y}_m^j = \widehat{\Theta}^j(t_m) - \widehat{\Theta}^j(t_{m-1}).$$

Using equation (5), we can now set up the estimated model as a minimum distance problem (or *asymptotic least squares*, see Gouriéroux *et al.*, 1985) by writing that:

$$\ln\left(\hat{y}_m^j\right) = \ln\left(\alpha^j\right) + \ln\left(\theta_m\right) + \varepsilon_m^j \tag{6}$$

where $\varepsilon_m^j = \ln(\hat{y}_m^j) - \ln(y_m^j)$ is the residual due to the sampling variability of estimated hazard rates (see Appendix B.2.2 for the computation of the covariance matrix).

There are two statistical issues of importance. First, note that (6) is ill-defined when $\widehat{y}_m^j$ takes the value zero. This happens when there is no exit of type k in municipality j in the time interval $[t_{m-1}, t_m]$. Corresponding observations are ignored in the estimation. It is a small sample issue that can be safely ignored if the number of observations is large as in most municipalities. In practice, there is a trade-off between small sample biases and precision when choosing the intervals. Trading off optimally bias and precision by constructing optimal data-driven intervals is out of the scope of this paper.

Second, equation (6) is a two-component model that can be estimated using weighted least squares where the weights are given by the square root of the inverse of the covariance matrix of residuals ε_m^j. However, this minimum distance estimator is known to perform badly in small samples as

[9] Dealing with the last interval is specific and detailed in the Appendix B.2.1.

shown by Altonji and Segal (1996).[10] We chose to use a slight modification of their equally weighted estimator which is simpler and better behaved. We simply weigh the estimation by the number of unemployed workers at risk at the beginning of the intervals in the municipalities (see Appendix B.2.3). Indeed, the average hazard rate computed for any given time interval (the dependent variable in (6)) is usually computed with more accuracy when the number of unemployed workers at risk is large. We also used other weighting schemes that yielded minor changes to the results.

The final descriptive stage consists in regressing municipality effects on aggregate explanatory variables at the municipality level. We specify:

$$\ln\left(\alpha^j\right) = Z^j\gamma + \eta^j \tag{7}$$

where Z^j are municipality variables and η^j are random terms. As municipality fixed effects are estimated in the previous stages, their exact value is not observed. Introducing these estimators in equation (7), we obtain:

$$\widehat{\ln\left(\alpha^j\right)} = Z^j\gamma + \eta^j + \xi^j \tag{8}$$

where $\xi^j = \widehat{\ln\left(\alpha^j\right)} - \ln\left(\alpha^j\right)$ is a sampling error. Equation (8) is estimated using weighted least squares where the weight is the initial number of unemployed workers in the municipality (see Appendix B.3). This weighting has two justifications. As above, the sampling error decreases with the number of unemployed workers. Second, weighting by the number of unemployed workers can be justified if we assume that municipalities can be decomposed into smaller areas of fixed population in which exit from unemployment is subject to an idiosyncratic shock with variance σ^2 (but affected in the same way by municipality variables). In this context, the aggregate random term η^j at the municipality level in equation (7) is an average of the smaller areas' idiosyncratic shocks. We thus assume that the terms η^j have a variance of the form $\sigma^2/C^j\left(0\right)$ where $C^j\left(0\right)$ is the initial number of unemployed workers in municipality j.

5 Results

We now comment the results of the empirical analysis whose estimation stages were described in the previous section. We first examine the estimated coefficients of the individual explanatory

[10] Correcting small sampling biases by bootstrap or jackknife does not perform better (Horowitz, 1998).

variables obtained using the stratified partial likelihood estimator (stage 1). We then describe the spatial disparities in municipality survival functions obtained from the model. We finally turn to the results concerning the municipality fixed effects (stage 2) and regress them on local variables measuring residential segregation and job accessibility (stage 3).

5.1 Individual Determinants of Unemployment Durations

Table 3 reports the coefficients estimated using SPLE for each type of exit (job and non-employment). Remember that the effects of individual variables should be interpreted as affecting multiplicatively the hazard rates (through the term $\exp(X_i\beta)$ in (1)).

[*Insert Table* 3]

Results are as expected although the magnitude of the effects of some variables is surprisingly large. First, for both exits, younger people have shorter unemployment spells. Although negative and significant, the effect of age is marginally decreasing (in absolute value) as evidenced by the square term. Note that it is never positive in any reasonable age range. Second, women exit significantly more slowly to a job than men (around -18%) while their exit rate to non-employment is much larger (around $+35\%$). Similarly, having children (whatever their number) decreases the exit rate to a job and increases the exit rate to non-employment. Being in a couple significantly increases exit rates both to a job and to non-employment.

The strongest effects are for nationality. Africans and other non-European citizens have an exit rate to a job that is between 45% and 66% lower than the French. In contrast, the effect of nationality variables on the hazard rate to non-employment is significant only for North Africans and the magnitude of the coefficient is much lower than for exit to work.

Education variables also have a strong effect. Overall, education affects the exit rate to a job more than the exit rate to non-employment. For instance, compared to a university degree, a basic degree lowers the exit rate to a job by 59% while it decreases the exit rate to non-employment by “only” 42%. The shadow wage (i.e. the opportunity cost of time in non participation) is less affected by education than market wages.

We perform two specification checks. First, it is interesting to compare our results with the results of the estimation of a standard Cox model where the baseline hazard function is restricted to be constant across municipalities. We use the same individual variables X_i as covariates to which we add the third-stage municipality variables such as segregation indices and job accessibility measures Z_i. Under these assumptions, our three-stage procedure collapses into one stage only and results are not robust to misspecification in the second and third stage of our procedure. The estimated coefficients of individual variables (and their standard errors) are very close to those obtained using our estimation method (SPLE) in stage 1.[11] The estimated coefficients of geographic variables are also quite close to what we will obtain in stage 3 (see below). Nevertheless, the standard errors of the estimated coefficients of aggregate variables obtained with a standard Cox model are at least one-third smaller than those obtained in our third-stage estimation. There are two explanations for this difference. First, the standard Cox model does not account for aggregate unobserved effects whereas our estimation method does. It is widely known that this can lead to very biased standard errors (see Moulton, 1990). Second, there can be an efficiency loss when using our three stage approach since we are not estimating all equations at the same time.

As a second specification check, we compare the Kaplan Meyer estimates of the survival function until an exit to a job with the estimates derived in our model for each municipality.[12] We use a Kolmogorov statistic as developed in Andrews (1997). The difficulty in the procedure – as detailed in Appendix B.1 – is the estimation of the variance of the test statistic evaluated by bootstrap. The results show that our model provides a very good fit at the municipality level. In Graph 5 we report the empirical frequency across municipalities of the p-values associated with the null hypothesis of a good fit. The empirical frequency of municipalities in which we reject a good fit is indeed lower than the level of the test. It shows that we are able to describe unemployment durations at the municipality level in a very satisfactory way. This justifies our specification and, in particular, our choice of not modeling unobserved heterogeneity.

[*Insert Graph* 5]

[11] Complete results are available upon request.

[12] We are thankful to a referee for this suggestion.

5.2 Describing spatial disparities in unemployment duration

We now assess the magnitude of spatial disparities in unemployment duration until finding a job or leaving to non-employment once the effect of individual variables has been controlled for. This is done by looking at the disparities between the municipality survival functions at 24 months. These functions are computed from the Breslow estimator where all individual variables are centered, substracting the corresponding mean in the whole region. Thus, these functions can be interpreted as the municipality survival functions of an "average" unemployed worker.

5.2.1 Comparing the Explanatory Power of Individual and Spatial Effects

We also want to assess the relative importance of individual characteristics and that of spatial effects in explaining the spatial disparities of unemployment durations. To do that, we resort to two complementary approaches.

First, a direct approach is to compare indices of spatial disparities obtained from the Kaplan-Meier estimators and from our model. While Kaplan-Meier estimators represent the raw data and do not control for observed individual determinants of durations, the survival functions obtained from the model (as computed from the integrated hazard functions in equation (4)) do control for individual determinants. In Table 4, we report various disparity indices (inter-decile range and ratio, Gini index and coefficient of variation) of the survival functions after 6 and 24 months both for Kaplan-Meier and for the model. For exits to a job, we find that individual variables explain only around 24% of spatial disparities at 6 months and around 15% at 24 months. To see how this is calculated, consider for instance that the inter-decile ratio at 24 months is 1.503 for the Kaplan-Meier estimate, and 1.426 for the survival function from the model. A coefficient of determination could thus be defined as $(.503 - .426)/.503$, which is equal to 15.3%. This shows that even after controling for the characteristics of local unemployed workers, spatial disparities in finding a job remain large. This is a common theme in the literature (see Maurin, 2004).

$[Insert\ Table\ 4]$

Note that the comparisons, which rely on the usual estimators of the survival functions, are only heuristic. Indeed, they are not based on an analytical relationship between the Kaplan-Meier estimators and the effect of individual variables and the municipality survival functions of the model.

Our second approach to estimate the relative importance of the contributions of individual and local characteristics to spatial disparities in unemployment durations does not make use of the Kaplan-Meier estimator and has firmer analytical grounds. It is a variance analysis of the average integrated hazard at the municipality level. To see why such an analysis is feasible, write the log-integrated hazard of a given individual i as the sum of the effect of individual observed characteristics $X_i\beta$ and the logarithm of the municipality integrated hazard $\Theta^{j(i)}(t)$ and a random term ζ_i:

$$\ln \Lambda\left(t \,|X_i, j\,(i)\right) = X_i\beta + \ln \Theta^{j(i)}(t) + \zeta_i. \tag{9}$$

Here, ζ_i is the intrinsic randomness in durations and $\exp(\zeta_i)$ is exponential(1) distributed (Lancaster, 1990). When we take the expectation of this equation in the population of unemployed workers in a given municipality j, we obtain the following decomposition of the log-integrated hazard of that municipality:

$$E(\ln \Lambda\left(t \,|X_i, j\,\right) \mid i \in \Omega^j(0)) = X^j\beta + \ln \Theta^j(t) + c_0 \tag{10}$$

where $\Omega^j(0)$ is the population of unemployed workers in the municipality, where X^j is the expectation of individual characteristics in that population and where the constant c_0 is the expectation of ζ_i.[13] In practice, the two right-hand side terms can be consistently estimated via the first stage estimations and their sum yields an estimate of the left-hand side term up to the constant term, c_0.

In Table 5, we report the results of a variance analysis across municipalities using equation (10) for durations of various lengths: short (6 months), intermediate (12 months) and long (24 months). Averages of individual observable characteristics explain around 30% of spatial effects in job exits. This confers to individual variables slightly more explanatory power than what we obtained in Table 4, although it remains quite low. For instance at 12 months, the spatial variance of the log-integrated hazard is equal to .0508 while the spatial variance of the average log-integrated hazard in the municipality (LHS of (10)) is .0755. A pseudo-coefficient of determination is thus

[13] As $\exp(\zeta_i)$ is exponential(1) distributed, $c_0 = 1$.

$(.0755 - .0508)/.0755$, which is equal to 33%.

$[Insert\ Table\ 5]$

5.2.2 Spatial Sorting and Spatial Effects

To understand what the remaining spatial disparities capture, it is useful to rewrite equation (10) by decomposing the log municipality integrated hazard function into linear components at time t:

$$\ln \Theta^j (t) = \ln \Theta (t) + Z^j \gamma(t) + \eta^j (t) + E(\varepsilon_i \mid j) \tag{11}$$

where $\Theta (t)$ is the integrated baseline hazard, Z^j are observed municipality characteristics whose coefficients vary with time and $\eta^j (t)$ are unobserved municipality effects due to geographic features such as infrastructures. The last term $E(\varepsilon_i \mid j)$ stands for spatial sorting of individuals into different municipalities. Indeed, the expectation of individual unobserved heterogeneity in the ability of exiting unemployment varies across municipalities because individuals might sort themselves into municipalities, partly because of omitted factors included in the unobserved heterogeneity term.[14] After controlling for individual observed characteristics, (11) shows that the remaining spatial disparities can be due not only to local characteristics (observed or not) but also to variations in the local average of individual unobserved characteristics. The lack of identification of these different effects is one form of the so-called reflection problem of Manski (1993).

5.2.3 The correlation between spatial and individual effects

Returning to the analysis of equation (10), it is also meaningful to calculate the correlations between the municipality composition effects ($X^j \beta$) and the logarithm of the municipality integrated hazard ($\ln \Theta^j (t)$) at 6, 12 and 24 months. This correlation can be interpreted in three ways. It can reflect some sorting on observable municipality effects ($Z^j \gamma(t)$), some sorting on unobservable municipality characteristics ($\eta^j (t)$), or a correlation between the local average of individual observed variables and the local average of individual unobserved variables ($E(\varepsilon_i \mid j)$).

Correlations are shown in Table 5. For exits to jobs, the correlation between the municipality composition effects and the municipality integrated hazard is high (for instance .49 at 12 months),

[14] The relationship between ε_i and the already defined ζ_i is $\zeta_i = \varepsilon_i - E(\varepsilon_i \mid j)$.

whereas, for exits to non-employment, it is very small ($-.05$ at 12 months). In order to assess the robustness of these findings, Graph 6 plots, for exits to jobs, the locally aggregated predictor $X^j\beta$ as a function of the logarithm of the municipality integrated hazard $\ln \Theta^j$ at 24 months for municipalities with more than 10,000 inhabitants. The positive association between these variables appears from these plots. It means that individual and local variables reinforce each other when affecting unemployment exits.

$[Insert\ Graph\ 6]$

Using equation (11), this correlation can be interpreted in three ways due to sorting. First, unemployed workers who are less likely to find a job because of their observable characteristics could sort in municipalities with bad observable neighborhood attributes (for instance where there are many foreigners, as these neighborhoods could be redlined by xenophobic employers). Second, they could also sort themselves in municipalities with bad unobservable attributes (for instance in municipalities which have a bad reputation among employers for some unobserved reason). Third, municipality aggregate of observed and unobserved individual characterics could be positively correlated (for instance workers with no diploma may be less efficient in job-search). In our opinion, one of the main result of this paper is thus to show that disparities in individual characteristics are reinforced by disparities in local characteristics due to residential sorting.

Finally, we investigate whether places that enhance job finding do slow down the exit to non-employment. To do that, we compute the correlations between the municipality integrated hazards for finding a job and for exit to non-employment at 6, 12 and 24 months (weighing by the number of unemployed workers at risk). We find that for short and medium horizons (6 and 12 months), there is little correlation between the two types of local effects (resp. $-.028$ and $.033$). However, in the long run (24 months), the correlation is positive and stands at .176. In municipalities where job exits are more likely to occur, exits to non-employment are also more likely to take place at least in the long run. This result can be understood by comparing reservation wages, shadow wages and job offers. In this framework, unemployed workers exit to non employment when their reservation wage falls below their shadow wage. Our result suggests that this is more likely to happen in the long run in municipalities where unemployed workers are more likely to exit to a job. This could happen if municipalities where residents can easily find a job are also those where holding a job is more likely. Spouses could thus more likely become non participant because their opportunity cost of time increases with their spouse's income.

5.3 Municipality Effects and Spatial Characteristics

5.3.1 Multiplicative Component Model

As explained above, we further restrict the hazard function and consider a multiplicative municipality hazard specified as the product of a municipality effect and an aggregate baseline hazard (see Equation (5)). To implement this approach, we divide the time line into $M = 9$ intervals, with the first eight intervals lasting 90 days and the remaining one lasting the rest of the period. To assess whether the multiplicative specification is too restrictive, we compare the value of disparity indices obtained with the unspecified municipality hazard with those obtained with the multiplicative hazard (see Table 4 and Table 5). We find that the multiplicative hazard reproduces well spatial disparities for finding a job although it performs poorly for exit to non-employment at 6 months (though not for the Gini). This good fit justifies the use of municipality effects as an adequate summary to study the determinants of spatial disparities for finding a job.

In line with the theories presented in Section 2, we investigate how municipality fixed effects can be explained by segregation and job accessibility. Segregation is measured here by the composition of the municipality population by education and by nationality. Job accessibility is measured by local job density (as defined in Section 3.2).

5.3.2 Partial Correlations with Spatial Caracteristics

Table 6 reports various regressions of municipality effects on those spatial characteristics. We computed a pseudo-R^2 to assess the explanatory power of the model taking into account the sampling error (see Appendix B.3). When using only segregation indices as explanatory variables (column 1), we are able to explain 72.4% of the variance of municipality fixed effects. Job accessibility indices (column 2) have a much lower explanatory power since the pseudo R^2 is only 25.9%. This suggests that spatial disparities in finding a job are more strongly associated with differences in the local level of segregation than with variations in job accessibility. When using both segregation and job accessibility indices (column 3), the pseudo R^2 reaches 73.0%.

We now comment on the coefficient of the latter regression (column 3). Large municipality effects in finding a job are associated with a large proportion of unskilled workers and of non-French citizens (especially citizens from sub-Saharan Africa). This is consistent with the existence of redlining (according to nationality and skill) as well as with a social network effect. Municipality effects in finding a job are also correlated with local job accessibility, especially by private transport, but the coefficients of both private and public job accessibility measures are negative in contradiction with the spatial mismatch theory.

[*Insert Table* 6]

Of course, there are some other interpretations of the results which are based on possible omitted local variables, reverse causality or sorting on individual unobservables. There can be omitted local variables correlated with segregation or job-accessibility measures. Our surprising result for job accessibility could be explained if the job density indices captured the low quality and high congestion of transports for instance.

Reverse causality can occur if local unemployment acts as an attraction or a repulsion force on population and jobs. This could affect the job accessibility measure and the segregation indices (provided that the population categories are differentially attracted or repulsed). To take the example of segregation for instance, French people may flee municipalities where unemployment exits to a job are more difficult. This would increase the local proportion of foreigners, especially Africans and could explain the negati coefficient of the municipality proportion of Africans on finding a job.

Municipality explanatory variables can capture the local average of individual unobserved variables if there is a correlation between $Z^j\gamma$ and the omitted term $E(\varepsilon_i \mid j)$ as defined in equation (11) above. This is the case for instance when individuals with a given unobserved attribute (such as motivation to search for a job) choose their location depending on observable municipality variables (attractive residential neighborhoods where jobs are not easily accessible). This may explain the negative effect of the job accessibility index by private transport.

6 Conclusion

In this paper, we study the spatial disparities in exits from unemployment across municipalities in the Paris region. We use a unique and exhaustive administrative dataset which contains all registered unemployment spells over the 1996-2003 period. This dataset contains some individual characteristics of unemployed workers as well as their residential location. It is merged with spatial indices of segregation and job accessibility computed from the census and a transport survey.

Our methodology is based on the estimation of independent competing risk duration models with two exits (finding a job and dropping out of the labor force). We constructed measures of raw spatial disparities across municipalities from the local survival functions after 24 months. We find that there are very large disparities. The local composition of workers' characteristics can explain around 30% of the disparities in finding a job. Our local indices (especially residential segregation measures) capture nearly 70% of the remaining differences. Furthermore, we showed that disparities in individual characteristics are reinforced by disparities in local characteristics due to residential sorting. The latter finding lends credit to the idea that spatial factors exacerbate non spatial factors in the determination of unemployment, or that the most fragile unemployed workers tend to cumulate local and invididual disadvantages.

Our work nevertheless considered broad local effects related to segregation and job accessibility without trying to investigate and disentangle the specific mechanisms at work, which could be the basis for furture research. Another extension of this work could be to compute municipality survival functions by nationality group or class of diploma. This would enable us to assess the extent to which the effect of local factors may differ for these groups. It would also be interesting to study spatial disparities at a much finer scale were the data available. Indeed, our accessibility measures are only at the municipality level whereas accessibility can differ even between two small neighborhoods (e.g. when they are separated by a railroad). Working at a finer geographic scale may also allow for an investigation of other important issues such as the role of spatialized social networks, which are likely to occur within a limited geographic area (see Bayer, Ross and Topa, 2008; Gobillon and Selod, 2007).

A Data Appendix

Over the 1993-2003 period, the panel contains 10,290,225 unemployment spells. We selected unemployment spells beginning between January 1 and June 30, 1996 which form a subsample of 451,191 unemployment spells. By keeping observations corresponding to unemployed workers between 16 and 54 years old only, we ended up with a dataset comprising 433,802 unemployment spells. After deleting observations with missing values and coding problems, the final sample is composed of 430,695 observations. Descriptive statistics about variables in our final dataset are given in Table A1.

B Computational details

B.1 First-stage estimation

We want to test for each municipality that the empirical survival function as estimated by Kaplan Meier estimation is equal to the survival function predicted by the model.

B.1.1 Construction of the Kolmogorov test statistic

Let $S(t,k)$ be the survival function of the data for exit k and let $S(t,k \mid x,\delta)$ be the conditional survival function of the model. Here, δ denotes parameters β_k and the municipality baseline hazard. The null hypothesis writes:

$$H_0 : S(t,k) = \int S(t,k \mid x,\delta) dF(x)$$

where $dF(x)$ is the probability measure of covariates. This is an adaptation of Andrews (1997) with some differences since the latter paper considers null hypotheses of the form:

$$H_0 : S(t,k\,|x)F(x) = S(t,k\,|x,\delta)F(x) \text{ a.s. } F(x).$$

where $S(t,k\,|x)$ is the conditional survival function of the data. The adaptation of the proofs of Andrews (1997) are out of the scope of this paper.

Our sample consists in individuals, $i = 1,...,N$, of characteristics X_i for whom we observe unemploment duration t_i and type of exit. We restrict our attention to exits to job and we drop index k from the survival functions. Computing the test statistic for exits to non-employment follows the same principles.

Let $\hat{H}_N^j(t)$ be the Kaplan Meier estimator of the survival function of duration t in municipality j. Let $\hat{S}_N^j(t;\delta)$ be the survival function until an exit to a job in municipality j, as predicted by the model i.e.

$$\hat{S}_N^j(t) = \frac{1}{N_j} \sum_{i=1, i \in j}^{N} S_i(t \mid X_i, \hat{\delta}),$$

where N_j denotes the number of individuals in municipality j, and where $\hat{\delta}$ denotes the SPLE estimator and the functional Breslow estimator of the baseline hazard rate. The conditional Kolmogorov statistic for municipality j is:

$$CK_N^j = \sqrt{N_j} \max_{i \in j} \left| \hat{H}_N^j(t_i) - \hat{S}_N^j(t_i) \right|$$

In practice, we trim out durations in the last percentile when computing our test statistic. This is because the survival functions are not estimated with accuracy at that percentile and the test statistic takes artificially large values at finite distance.

Alternatively, we could also consider another statistic which is the mean square difference of these two survival functions, in analogy with a Cramer-von Mises statistic:

$$QM_N^j = \frac{1}{\sqrt{N_j}} \sum_{i \in j} \left(\hat{H}_N^j(t_i) - \hat{S}_N^j(t_i) \right)^2$$

where N_j is the number of unemployed workers in municipality j.

For these two test statistics, we need to compute the distribution under the null hypothesis. We proceed by bootstrap as proposed by Andrews (1997).

B.1.2 Computation of the distribution of the test statistic.

We now explain how to compute the distribution of the test statistic for a given municipality. We drop index j for simplicity. For an individual i having a censored unemployment spell, the duration before censorship, $t_{ic} = t_i$, is an exogenous characteristic of the individual. If the unemployment spell is not censored, censorship is not relevant and its duration is not taken into account. We denote $\widetilde{X}_i$ the exogenous information i.e. $\widetilde{X}_i = (X_i, t_{ic})$ for censored individuals and $\widetilde{X}_i = (X_i, \cdot)$ for uncensored ones.

The asymptotic distribution of Andrews' test statistic is computed by (semi)-parametric bootstrap for B replications, $b = 1, ..., B$. For each replication, we generate a duration for each individual using the proportional hazard model. The data are generated conditionally on the exogenous characteristics $\widetilde{X}_i$, the estimated parameters for exit to job $\widehat{\beta}_e$ and the Breslow estimator. The procedure to simulate durations is the following.

For an individual i, the integrated hazard for exit to job at duration t writes:

$$\Lambda (t \,|X_i) = \Lambda (t) \exp (X_i \beta_e)$$

where $\Lambda (t)$ is the integrated baseline hazard. As the survival function is $S (t \,|X_i) = \exp [-\Lambda (t \,|X_i)]$, we have:

$$\Lambda (t) = \frac{-1}{\exp (X_i \beta_e)} \ln S (t \,|X_i)$$

To obtain simulated durations, we draw the value of the survival function in a uniform distribution $[0, 1]$, $\widetilde{S}_i^b$, and replace unknown parameters by their estimates:

$$\widehat{\Lambda}_i^b = \frac{-1}{\exp \left(X_i \widehat{\beta}_e \right)} \ln \widetilde{S}_i^b$$

There remains to invert function $\widehat{\Lambda}$ at point $\widehat{\Lambda}_i^b$ to recover duration $\widetilde{t}_i^b$ for individual i. In practice, $\widehat{\Lambda}_i^b$ increases piecewise and there is a duration at which the function $\widehat{\Lambda}$ makes a jump from $\underline{\Lambda}_i^b$ to $\overline{\Lambda}_i^b$ such that $\underline{\Lambda}_i^b \leqslant \widehat{\Lambda}_i^b < \overline{\Lambda}_i^b$. We define $\widetilde{t}_i^b$ as the duration at which function $\widehat{\Lambda}$ makes this jump.

A practical issue is that $\widehat{\Lambda}$ cannot be computed above the upper bound $\widehat{\Lambda} (t^{\max})$, where $t_j^{\max}$ is the largest duration in the sample. If a simulated value is such that $\widehat{\Lambda}_i^b > \widehat{\Lambda} (t^{\max})$, we set duration to $\widetilde{t}_i^b = t^{\max}$. This small sample bias should disappear asymptotically when the number of individuals in each municipality tends to infinity.

For individuals who were not censored, the generated duration is $t_i^b = \widetilde{t}_i^b$. For individuals who were right-censored, the generated duration is $t_i^b = \min \left(\widetilde{t}_i^b, t_{ic} \right)$. The generated exit is finding a job if $\widetilde{t}_i^b < t_{ic}$, and it is censorship if $\widetilde{t}_i^b > t_{ic}$. We then construct the Kaplan-Meier's estimator denoted $\hat{H}_N^{j,b} (t)$. For individual i, we compute the value of the Kaplan-Meier's estimator of his municipality at the generated duration as: $\hat{H}_i^b = \hat{H}_N^{j,b} \left(t_i^b \right)$. In the same way, we use the Breslow's estimator of any municipality j to construct a survival function which is denoted $\hat{S}_N^{j,b} (t)$. For individual i, we compute the value of the survival function of his municipality at the generated duration as: $\hat{S}_i^b = \hat{S}_N^{j,b} \left(t_i^b \right)$.

The test statistic of municipality j computed for the b^{th} replication can be written as:

$$CK_N^{j,b} = \sqrt{N_j} \sup_{i \in j} \left| \hat{H}_i^b - \hat{S}_i^b \right|$$

As previously, we trim out durations in the last percentile when computing our test statistic. B bootstrap samples of size N_j are simulated. The p-value of the test statistic is defined as:

$$p_{jN}^B = \frac{1}{B} \sum_{b=1}^{B} \mathbf{1}\{CK_N^j > CK_N^{j,b}\}.$$

B.2 Second-stage estimation

B.2.1 Finite-sample issues

We first explain how we take into account finite sample issues when establishing equation (6). For that purpose, we refined appropriatly the quantities involved in (6). We divide the period into M intervals $[t_{m-1}, t_m]$, $m = 1, ..., M$. We denote $\theta_m = \frac{1}{t_m - t_{m-1}} \int_{t_{m-1}}^{t_m} \theta(s)\, ds$ the average baseline hazard over the interval m and $d_m^j = \int_{t_{m-1}}^{t_m} I\left(C^j(s) > 0\right) ds$ the length of time within interval m when some individuals in municipality j are at risk. In particular, $d_m^j < t_m - t_{m-1}$ in the last time interval in which there are some unemployed workers at risk in municipality j. The average hazard rate over a time interval m in municipality j where some people are at risk ($d_m^j > 0$) is given by $y_m^j = \frac{1}{d_m^j}\left[\Theta^j(t_m) - \Theta^j(t_{m-1})\right]$. An estimator of this average hazard rate can be constructed from equation (4) and writes: $\hat{y}_m^j = \frac{1}{d_m^j}[\widehat{\Theta}^j(t_m) - \widehat{\Theta}^j(t_{m-1})]$. We can then re-establish formula (6) where the quantities have been redefined.

B.2.2 Covariance matrix of the sampling errors

We now give the formulas to compute the covariance matrix of $(\varepsilon_m^j)_{j,m}$,which are the sampling errors in equation (6), using Ridder and Tunali's appendix (RT hereafter). We first introduce the following notations that will be used below:

$$\begin{aligned} S_j^0(\beta, s) &= \sum_{i \in \Omega^j(s)} \exp(X_i \beta) \\ S_j^1(\beta, s) &= \sum_{i \in \Omega^j(s)} X_i \exp(X_i \beta) \end{aligned}$$

where $\Omega^j(s)$ is the set of unemployed workers still at risk in municipality j at time s. Note that whereas $S_j^0(\beta, s)$ is a 1×1 matrix, $S_j^1(\beta, s)$ is a $1 \times K$ matrix, where K is the number of explanatory variables in the first stage. We also denote $C^j(s) = card\ \Omega^j(s)$ the number of unemployed workers still at risk in municipality j at time s. According to RT (A28), we have:

$$\exp \varepsilon_m^j = \eta_m^j + \frac{1}{\sqrt{N}} c'_{jm} \xi \tag{12}$$

where $N = \sum_j C^j(0)$ is the number of unemployed workers in the Paris region and:

$$\eta_m^j = \tfrac{1}{d_m^j} \int_{t_{m-1}}^{t_m} I\left(C^j(s) > 0\right) \left[\tfrac{1}{S_j^0(\beta,s)} dN^j(s) - \theta^j(s)\, ds\right] \quad \text{(RT A22)}$$

$$c_{jm} = -\tfrac{1}{d_m^j} \int_{t_{m-1}}^{t_m} I\left(C^j(s) > 0\right) \tfrac{S_j^1(\beta^*,s)}{\left[S_j^0(\beta^*,s)\right]^2} dN^j(s) \quad \text{(RT A27)}$$

$$\xi = \sqrt{N}\left(\widehat{\beta} - \beta\right)$$

where β^* is a value between β and $\widehat{\beta}$ (coming from a Taylor expansion not detailed here), $dN^j(s)$ is a dummy that equals one if someone in municipality j experiences an exit in an arbitrarily short period of time before date s (and zero otherwise), and $d_m^j = \int_{t_{m-1}}^{t_m} I\left(C^j(s) > 0\right) ds$. Here, ξ is uncorrelated with η_m^j. From equation (12), it is possible to get:

$$V\left(\exp \varepsilon_m^j\right) = V\left(\eta_m^j\right) + c'_{jm} V c_{jm} \quad \text{(RT A29)}$$

$$cov\left(\exp \varepsilon_m^j, \exp \varepsilon_n^k\right) = c'_{jm} V c_{kn} \text{ for } j \neq k \text{ or } m \neq n \quad \text{(RT A30)}$$

where $V = V\left(\widehat{\beta}\right)$. These covariance-matrix terms of $\left(\exp \varepsilon_m^j\right)_{j,m}$ can be estimated computing estimators of all terms on the right-hand sides. An estimator of V is obtained from the Fisher information matrix of SPLE. In practice, there is no need to have the theoretical formula to get this estimator as it is directly recovered from the estimation software. Some estimators of $V\left(\eta_m^j\right)$ and c_{jm} are:

$$\widehat{V}\left(\eta_m^j\right) = \tfrac{1}{\left(d_m^j\right)^2} \int_{t_{m-1}}^{t_m} I\left(C^j(s) > 0\right) \tfrac{1}{\left[S_j^0\left(\widehat{\beta},s\right)\right]^2} dN^j(s) \quad \text{(from RT A25)}$$

$$\widehat{c}_{jm} = -\tfrac{1}{d_m^j} \int_{t_{m-1}}^{t_m} I\left(C^j(s) > 0\right) \tfrac{S_j^1\left(\widehat{\beta},s\right)}{\left[S_j^0\left(\widehat{\beta},s\right)\right]^2} dN^j(s) \quad \text{(from RT A27)}$$

These estimators have to be programmed to be computed. From the covariance matrix of $(\exp \varepsilon_m^j)_{j,m}$, we get the covariance matrix of $(\varepsilon_m^j)_{j,m}$ using the delta method.

B.2.3 Estimation

Formulas

We first give some notations we use in this section. Denote J the number of municipalities and M the number of time intervals. For any $JM \times 1$ matrix X, X_j refers to the $M \times 1$ matrix defined by $X_{(j-1)M+[1:M],1}$. For any given $JM \times JM$ matrix X, $X_{j,k}$ refers to the $M \times M$ submatrix defined by $X_{(j-1)M+[1:M],(k-1)M+[1:M]}$ where $[1:M]$ is the vector of integers from 1 to M.
The equation to estimate is (6) where we fix $\theta_1 = 1$ to secure identification. We stack the observations of (6) and obtain:

$$Y = A\alpha + G\theta + \varepsilon \tag{13}$$

where A is a $JM \times J$ matrix such that $A_{(j-1)M+m,k} = 1$ if $j = k$ and $A_{(j-1)M+m,k} = 0$ otherwise, G is a $JM \times (M-1)$ matrix such that $G_{(j-1)M+m,l} = 1$ if $m = l$ and $A_{(j-1)M+m,l} = 0$ otherwise, $Y = \left(\ln y_1^1, ..., \ln y_M^J\right)'$ and $\varepsilon = \left(\varepsilon_1^1, ..., \varepsilon_M^J\right)'$ are some $JM \times 1$ vectors, $\alpha = (\ln \alpha_1, ..., \ln \alpha_J)'$ is a $J \times 1$ vector and $\theta = (\ln \theta_2, ..., \ln \theta_M)'$ is a $(M-1) \times 1$ vector.

Denote $\Delta = diag\,(N_{11}, ..., N_{JM})$ the $JM \times JM$ diagonal matrix where N_{jm} is the number of unemployed workers in municipality j still at risk at the beginning of interval m. After weighting equation (13) with $\Delta^{1/2}$, it becomes:

$$\Delta^{1/2} Y = \Delta^{1/2} A\alpha + \Delta^{1/2} G\theta + \Delta^{1/2}\varepsilon$$

Denote W the projector in the dimension orthogonal to $\Delta^{1/2}A$. Using the first stage of Frisch-Waugh theorem, we obtain the WLS estimator of θ:

$$\begin{aligned} \widehat{\theta} &= (G'\Delta^{1/2}W\Delta^{1/2}G)^{-1}G'\Delta^{1/2}W\Delta^{1/2}Y \\ &= \theta + (G'\Delta^{1/2}W\Delta^{1/2}G)^{-1}G'\Delta^{1/2}W\Delta^{1/2}\varepsilon \end{aligned} \tag{14}$$

The second stage of the Frisch-Waugh theorem gives the WLS estimator of α:

$$\begin{aligned}
\widehat{\alpha} &= (A'\Delta A)^{-1}A'\Delta\left[Y - G\widehat{\theta}\right] \\
&= (A'\Delta A)^{-1}A'\Delta\left[Y - G\theta - G\left(G'\Delta^{1/2}W\Delta^{1/2}G\right)^{-1}G'\Delta^{1/2}W\Delta^{1/2}\varepsilon\right] \\
&= \alpha + (A'\Delta A)^{-1}A'\Delta\left[\varepsilon - G\left(G'\Delta^{1/2}W\Delta^{1/2}G\right)^{-1}G'\Delta^{1/2}W\Delta^{1/2}\varepsilon\right]
\end{aligned}$$

Denote $\Gamma = A'\Delta A$, $\Phi = G'\Delta^{1/2}W\Delta^{1/2}G$ and $\Psi = G'\Delta^{1/2}W\Delta^{1/2}V\Delta^{1/2}W\Delta^{1/2}G$, where $V = V(\varepsilon)$. We have:

$$V\left(\widehat{\theta}\right) = \Phi^{-1}\Psi\Phi^{-1} \tag{15}$$

Also, we get:

$$\begin{aligned}
V(\widehat{\alpha}) &= \Gamma^{-1}A'\Delta V\Delta A\Gamma^{-1} \\
&\quad + \Gamma^{-1}A'\Delta GV\left(\widehat{\theta}\right)G'\Delta A\Gamma^{-1} \\
&\quad - \Gamma^{-1}A'\Delta\left(V\Delta^{1/2}W\Delta^{1/2}G\Phi^{-1}G' + G\Phi^{-1}G'\Delta^{1/2}W\Delta^{1/2}V\right)\Delta A\Gamma^{-1}
\end{aligned} \tag{16}$$

Computation

We have:

$$\begin{aligned}
\Phi &= \sum_{j=1}^{J}\left(W\Delta^{1/2}G\right)_j'\left(W\Delta^{1/2}G\right)_j = \sum_{j=1}^{J}\overline{G}_j'\Delta_{j,j}\overline{G}_j \\
\Psi &= \sum_{j,k=1}^{J}\left(W\Delta^{1/2}G\right)_j'\Delta_{j,j}^{1/2}V_{j,k}\Delta_{k,k}^{1/2}\left(W\Delta G\right)_k = \sum_{j,k=1}^{J}\overline{G}_j'\Delta_{j,j}V_{j,k}\Delta_{k,k}\overline{G}_k \\
\Gamma &= diag\left[tr\left(\Delta_{1,1}\right), ..., tr\left(\Delta_{J,J}\right)\right]
\end{aligned}$$

where for any given variable Z_j of dimension $M \times 1$, $\overline{Z}_j$ is its counterpart centered with its weighted average: $\overline{Z}_j = Z_j - \frac{1}{tr(\Delta_{j,j})}tr\left(\Delta_{j,j}Z_j\right)$.

We also have:

$$\begin{aligned}
\left(A'\Delta V\Delta A\right)_{j,k} &= N_j'V_{j,k}N_k \\
\left(A'\Delta GV\left(\widehat{\theta}\right)G'\Delta A\right)_{j,k} &= N_{j-}'V\left(\widehat{\theta}\right)N_{k-}
\end{aligned}$$

where $N_j = (N_{j,1}, ..., N_{j,M})'$ and $N_{j-} = (N_{j,2}, ..., N_{j,M})'$.

Moreover, $V\Delta^{1/2}W\Delta^{1/2} = \overline{V}\Delta$ where $\overline{V}$ is defined such that any of its given submatrix $\overline{V}_{j,k}$ writes: $\overline{V}_{j,k} = V_{j,k} - \frac{1}{tr(\Delta_{j,j})}\left(1_M \otimes N_j'\right)V_{j,k}$ with $\otimes$ the Kronecker product and 1_M a $M \times 1$ matrix filled with the value 1. Hence, we have:

$$\left(A'\Delta V\Delta^{1/2}W\Delta^{1/2}G\Phi^{-1}G'\Delta A\right)_{j,k} = N_j'\left(\overline{V}'\Delta G\Phi^{-1}G'\right)_{j,k}N_k$$

Moreover, $G\Phi^{-1}G' = J.J_J \otimes \begin{pmatrix} 0 & 0 \\ 0 & \Phi^{-1} \end{pmatrix}$ with J_J the $J \times J$ matrix filled with the value $1/J$.

Hence, $\left(\overline{V}\Delta G\Phi^{-1}G'\right)_{j,k} = \sum\limits_l \left(\overline{V}\Delta\right)_{j,l}\begin{pmatrix} 0 & 0 \\ 0 & \Phi^{-1} \end{pmatrix} = \left(\sum\limits_l \overline{V}_{j,l}\Delta_{l,l}\right)\begin{pmatrix} 0 & 0 \\ 0 & \Phi^{-1} \end{pmatrix}$.

B.3 Third-stage estimation

The third-stage equation to estimate is given by (8). When we stack the observations, we obtain:

$$\widehat{\alpha} = Z\gamma + \eta + \xi \tag{17}$$

where $\widehat{\alpha} = \left(\widehat{\ln\alpha_1}, ..., \widehat{\ln\alpha_J}\right)'$, $\eta = (\eta_1, ..., \eta_J)'$ and $\xi = (\xi_1, ..., \xi_J)'$ are some $J \times 1$ vectors, and $Z = (Z_1', ..., Z_J')'$ is a $J \times K$ matrix. We suppose that $\left(\eta_j\right)_{1,...,J}$ have a covariance matrix $\upsilon^2 Q^{-1}$ where $Q = diag(N_{11}, ..., N_{J1})$. Equation (17) is estimated with weighted least squares where the weights are the square-roots of the numbers of unemployed workers at the initial date ($Q^{1/2}$). The estimated coefficients write:

$$\widehat{\gamma} = \left(Z'QZ\right)^{-1}Z'Q\widehat{\alpha}$$

and their covariance matrix is:

$$\begin{aligned} V\left(\widehat{\gamma}\right) &= \left(Z'QZ\right)^{-1}Z'Q\left[V\left(\xi\right) + \upsilon^2 Q^{-1}\right]QZ\left(Z'QZ\right)^{-1} \\ &= \left(Z'QZ\right)^{-1}Z'QV\left(\xi\right)QZ\left(Z'QZ\right)^{-1} + \upsilon^2\left(Z'QZ\right)^{-1} \end{aligned}$$

It is possible to construct a consistent estimator of υ^2 using the residuals $\widehat{\eta+\xi} = Q^{1/2}\widehat{\alpha} - Q^{1/2}Z\widehat{\gamma}$. This estimator is found from the following calculation sequence:

$$\widehat{\eta+\xi}'\widehat{\eta+\xi} = (\eta+\xi)'\left[I - Q^{1/2}Z\left(Z'QZ\right)^{-1}Z'\right]'Q\left[I - Z\left(Z'QZ\right)^{-1}Z'Q^{1/2}\right](\eta+\xi)$$

where we made the approximation (for N large enough) that:

$$\widehat{\eta+\xi}'\widehat{\eta+\xi} \approx (\eta+\xi)' Q (\eta+\xi)$$

We thus have:

$$E\left[\widehat{\eta+\xi}'\widehat{\eta+\xi}\right] \approx \upsilon^2 J + tr\left[QV(\xi)\right]$$

when $V(\xi)$ has been computed from the first-stage estimation. An estimator of υ^2 can then be defined as:

$$\widehat{\upsilon}^2 = \left[\widehat{\eta+\xi}'\widehat{\eta+\xi} - tr\left[QV(\xi)\right]\right]/J$$

We introduce an error rate coming from sampling error as:

$$err = \frac{tr\left[QV(\xi)\right]}{\widehat{\eta+\xi}'\widehat{\eta+\xi}}$$

We also construct a pseudo-R^2 defined as:

$$R_p^2 = \frac{V_Q^e(Z\widehat{\gamma})}{V_Q^e(Z\widehat{\gamma}) + \widehat{\upsilon}^2 J}$$

where $V_Q^e(\cdot) = \left(Z\widehat{\gamma} - \overline{Z}\widehat{\gamma}\right)' Q \left(Z\widehat{\gamma} - \overline{Z}\widehat{\gamma}\right)/tr(Q)$ is the empirical variance obtained when weighting observations with weights Q (where $\overline{Z} = tr(QZ)/trQ$). Note that when there is no sampling error, this pseudo-R^2 is equal to the usual R^2.

References

[1] Altonji J.G. and L.M. Segal (1996), "Small Sample Bias in GMM estimation of Covariance Structures", Journal of Business & Economic Statistics, 14(3), 353-65.

[2] Andrews D. (1997), "A Conditional Kolmogorov test", Econometrica, 65(5), 1097-1128

[3] Baccaini B., Courgeau D. and G. Desplanques (1993), "Les migrations internes en France de 1982 à 1990. Comparaison avec les périodes antérieures", 48(6), 1771-89.

[4] Baker M. and A. Melino (2000), "Duration Dependence and Non-Parametric Heterogeneity: A Monte-Carlo Study", Journal of Econometrics, 96, 357-93.

[5] Bayer P., Ross S. and G. Topa (2008), "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes", Journal of Political Economy, 116(6), 1150-96.

[6] Breslow N.E. (1974), "Covariance Analysis of Censored Survival Data", Biometrics, 30, pp. 89-99.

[7] Calvo-Armengol A. and M. Jackson (2004), "Social networks in determining unemployment: Patterns, dynamics and inequality", American Economic Review, 94, 191-206.

[8] Chardon O. and D. Goux (2003), "The new european definition of BIT unemployment" (in French), Economie et Statistiques, 362, 67-83.

[9] Conley T. and G. Topa (2002), "Socio-economic distance and spatial patterns in unemployment", Journal of Applied Econometrics, 17(4), 303-327.

[10] Cutler D. and E. Glaeser (1997), "Are Ghettos Good or Bad?", Quarterly Journal of Economics, 112, 827-872.

[11] Dawkins C., Shen Q. and T. Sanchez (2005), "Race, Space and Unemployment Duration", Journal of Urban Economics, 58, 91-113.

[12] DREIF (Direction Régionale de l'Equipement Ile-de-France) / INSEE Ile-de-France (1997), "Daily moves: Ile-de-France population more mobile" (in French), 2 pages.

[13] Gobillon L. and H. Selod (2007), "The effect of segregation and spatial mismatch: evidence from France", CEPR Working Paper 6198.

[14] Gobillon L., Selod H. and Y. Zenou (2007), "The mechanisms of spatial mismatch", Urban Studies, 44 (12), 2401-27 .

[15] Gouriéroux C., A. Monfort, E. Renault and A. Trognon (1985), "Asymptotic Least Squares" (in French), Annales de l'INSEE, 58, 91-122.

[16] Holzer H. and K. Ihlanfeldt (1998) "Customer discrimination and employment outcomes for minority workers", Quarterly Journal of Economics, 113, 835-867.

[17] Holzer H., Ihlanfeldt K. and D. Sjoquist (1994), "Work, search and travel among white and black youth", Journal of Urban Economics, 35, 320-345.

[18] Horowitz J., (1998), "Bootstrap Methods for Covariance Structures", Journal of Human Resources, 33, 39-61.

[19] Ihlanfeldt K. (1993), "Intra-urban job accessibility and Hispanic youth employment rate", Journal of Urban Economics, 33, 254-271.

[20] Ihlanfeldt K. (1997), "Information on the spatial distribution of job opportunities within Metropolitan Areas", Journal of Urban Economics, 41, 218-242.

[21] Ihlanfeldt K. and D. Sjoquist (1998), "The spatial mismatch hypothesis: a review of recent studies and their implications for welfare reform", Housing Policy Debate, 9, 849-892.

[22] Johnson R. (2006) "Landing a job in urban space: the extent and effects of spatial mismatch", Regional Science and Urban Economics, 36, 331-72.

[23] Lancaster T. (1990), The Econometric Analysis of Transition Data, Cambridge University Press: Cambridge.

[24] Manski C. F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem", Review of Economic Studies, 60(3), 531-42.

[25] Maurin E. (2004), The French Ghetto (in French), ed. Seuil, 95p.

[26] Moulton B.R. (1990), "An Illustration of Pitfalls in estimating the Effects of Aggregate variables on Micro Units", Review of Economics and Statistics, 72, 334-8.

[27] Mortensen D. et T. Vishwanath (1994), "Personal contacts and earnings. It is who you know!", Labour Economics, 1, 187-201.

[28] Patacchini E. Y. Zenou (2006), "Search activities, cost of living, and local labor markets", Regional Science and Urban Economics, 36(2), 227-48.

[29] Ridder G. and I. Tunali (1999), "Stratified partial likelihood estimation", Journal of Econometrics, 92(2), 193-232.

[30] Rogers C. (1997), "Job Search and Unemployment Duration: Implications for the Spatial Mismatch Hypothesis", Journal of Urban Economics, 42, 109-32.

[31] Selod H. and Y. Zenou (2006), "City structure, job search, and labor discrimination. Theory and policy implications", Economic Journal, 116, 1057-87.

[32] Smith T. and Y. Zenou (2003), "Spatial mismatch, search effort and urban spatial structure", Journal of Urban Economics, 54, 129-156.

[33] Stoll M. (1999), "Spatial job search, spatial mismatch, and the employment and wages of racial and ethnic groups in Los Angeles", Journal of Urban Economics, 46, 129-55.

[34] Stoll M. and S. Raphael (2000), "Racial differences in spatial job search patterns: Exploring the causes and consequences", Economic Geography, 201-223.

[35] Weinberg B. (2000), "Black residential centralization and the spatial mismatch hypothesis", Journal of Urban Economics, 48, 110-134.

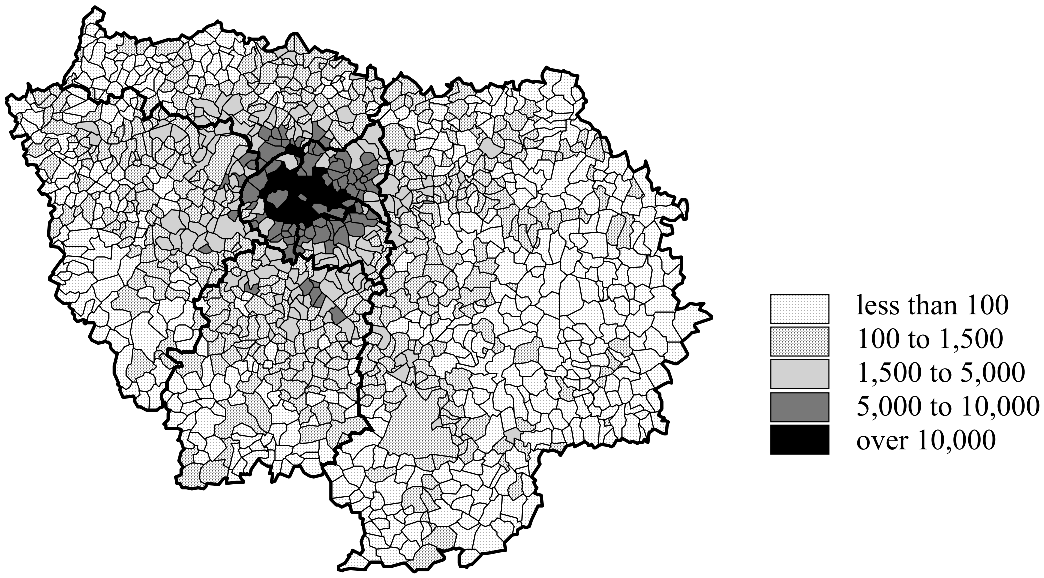
[36] Weinberg B. (2004), "Testing the spatial mismatch hypothesis using inter-city variations in industrial composition", Regional Science and Urban Economics, 34, 505-32.

[37] Wilson,J. (1996), When Work Disappears: The World of the New Urban Poor, New York: Alfred A. Knopf.

[38] Zax J. and J. Kain (1996), "Moving to the suburbs: Do relocating companies leave their Black employees behind?", Journal of Labor Economics, 14, 472-504.

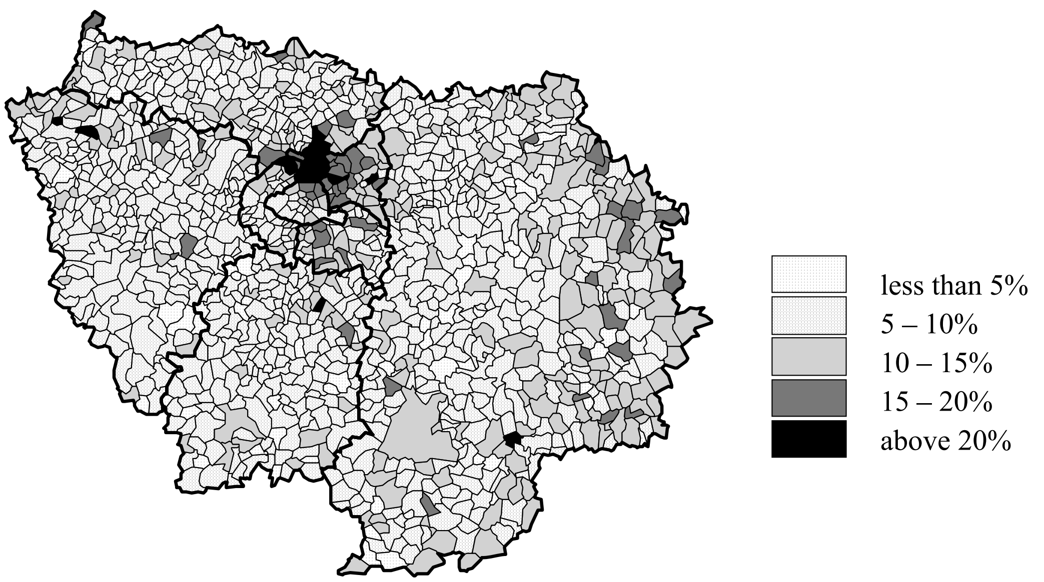
[39] Zenou Y. (2002), "How do firms redline workers?", Journal of Public Economics, 52(3), pages 391-408.

Graph 1: Population density (per sq km) in the Paris region in 1999



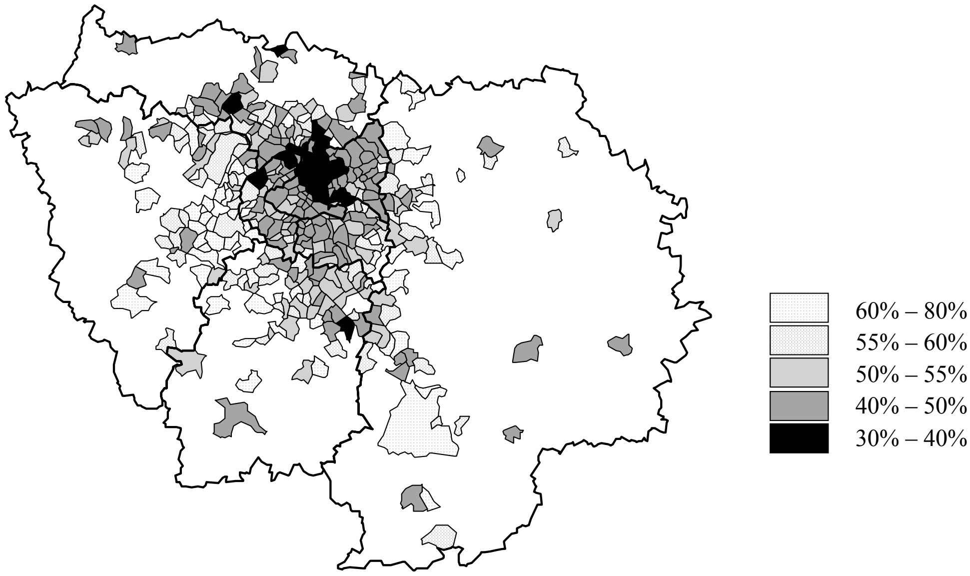
Source: constructed from the 1999 Population Census, INSEE. The geographical unit is the subdistrict for the city of Paris or the municipality for the rest of the region. Bold lines represent the boundaries of the city of Paris (the turtle-shaped area in the middle of the map) and of the seven surrounding subregional administrative districts (*départements*).

Graph 2: Unemployment rates in the Paris region in 1999



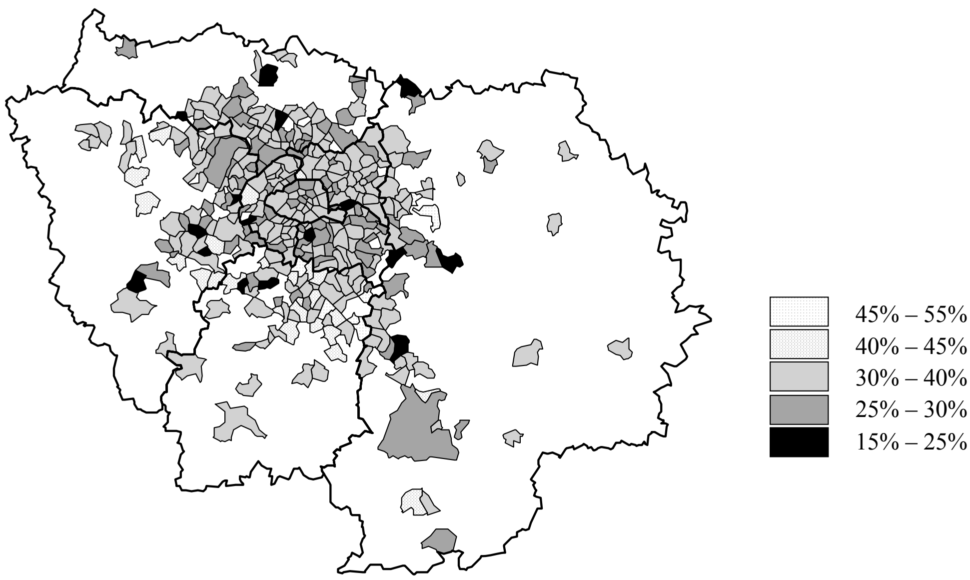
Source: constructed from the 1999 Population Census, INSEE.

Graph 3: probability of finding a job before 24 months (Kaplan-Meier)
for municipalities with at least 5,000 inhabitants



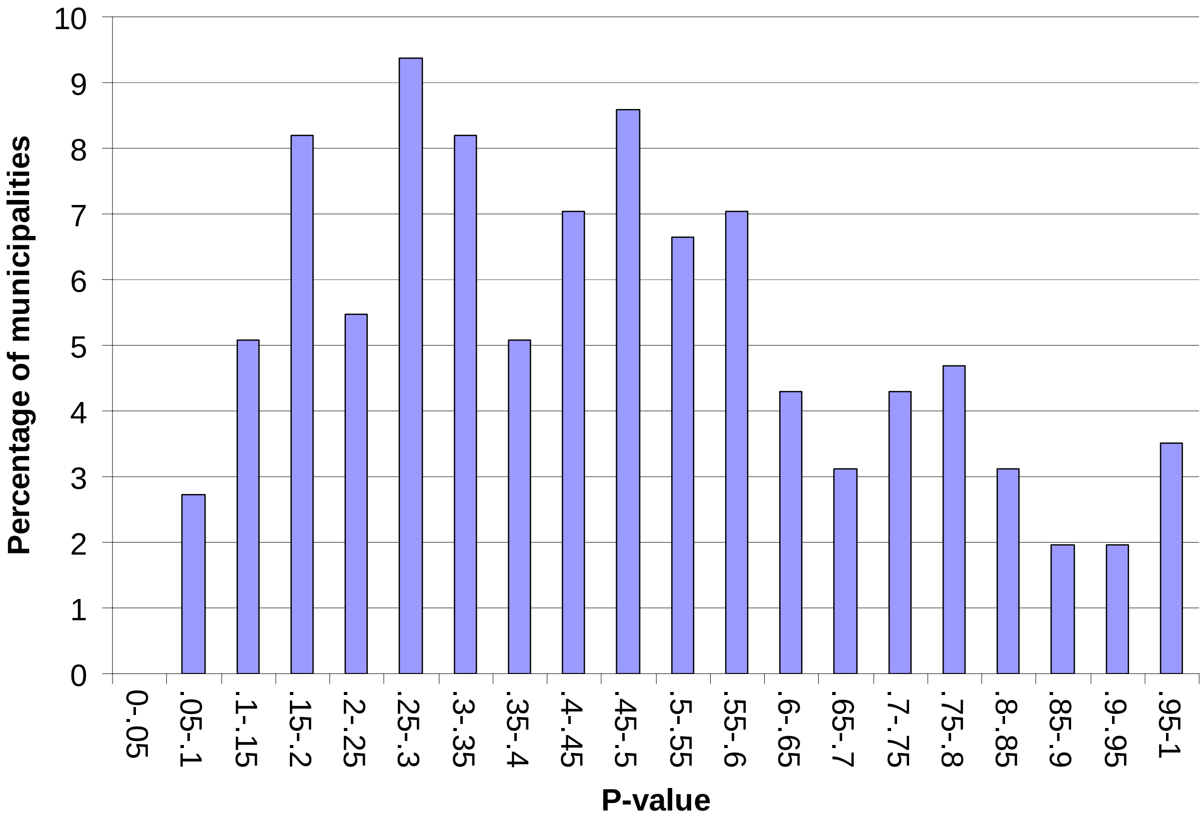
Source: constructed from the ANPE file.

Graph 4: probability of leaving for non-employment before 24 months
(Kaplan-Meier) for municipalities with at least 5,000 inhabitants



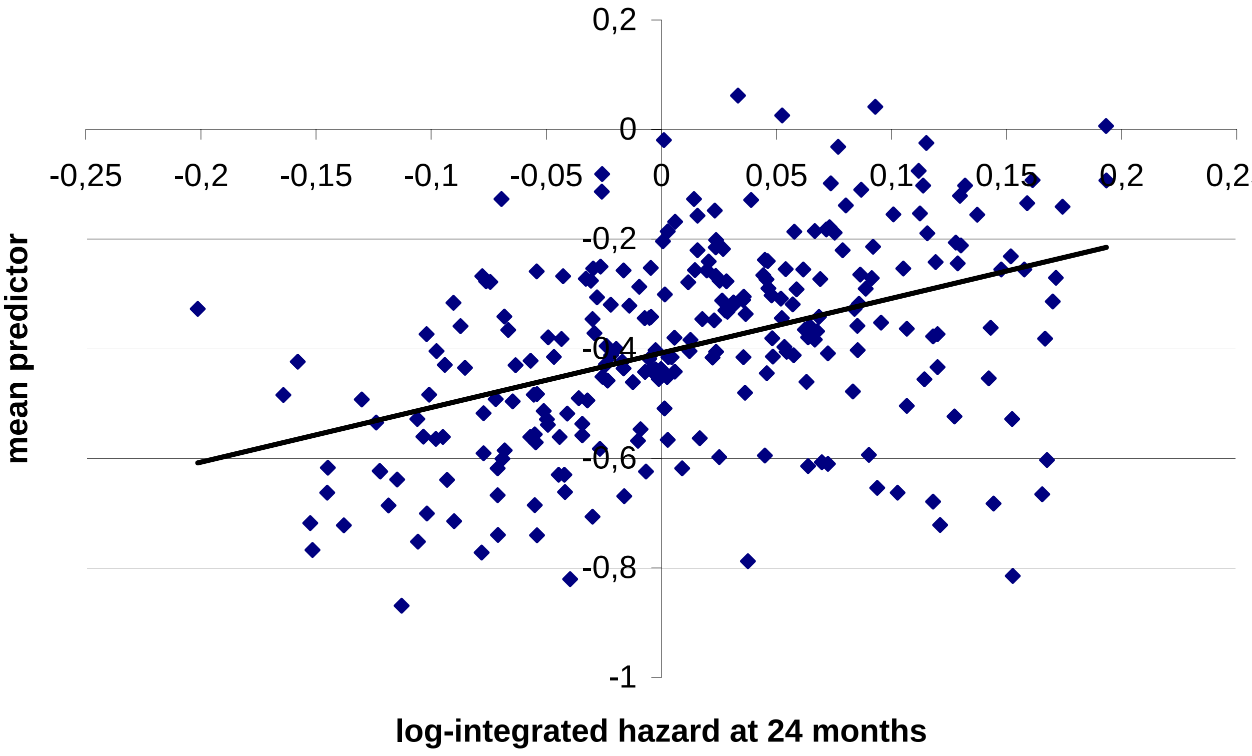
Source: constructed from the ANPE file.

Graph 5: The spatial distribution of p-values of the Kolmogorov test statistic



Source: constructed from the ANPE file.

Graph 6: Municipality average of individual effects $X^j\beta$
as a function of log-integrated hazard at 24 months for exit to job



Source: constructed from the ANPE file.
Municipalities with more than 10,000 inhabitants.

Table 1: Spatial inequality indices for measures of segregation and job accessibility calculated using census data

Variables	Mean	q90/q10	q90-q10	Gini	Coeff. of variation	Number of obs.
Segregation variables						
Unemployment rate	.116	2.408	.100	.196	.355	1300
% French	.813	1.277	.198	.054	.095	1300
% European (other)	.070	2.080	.050	.160	.294	1300
% North African	.054	9.464	.097	.396	.730	1300
% Sub-Saharan African	.026	9.371	.046	.389	.700	1300
% Other Nationality	.037	6.522	.057	.361	.688	1300
% Secondary School Diploma	.399	1.895	.248	.130	.227	1300
% Technical diploma	.192	2.526	.156	.175	.311	1300
% High school Diploma	.148	1.453	.054	.076	.137	1300
% College Diploma	.261	3.883	.333	.270	.479	1300
Job-accessibility variables						
45mn job density by public transport	1.062	2.995	1.009	.211	.459	1300
45mn job density by car	.856	1.615	.400	.104	.181	1300

Source: constructed from the 1999 Population Census and the 2000 General Transport Survey (*Enquête Globale de Transport*). The unemployment rate is weighted by the labor force. Nationality rates are weighted by the population. Diploma rates are computed for the population over 15 and are weighted by the population over 15. Job-accessibility variables are weighted by the labor force.

Table 2: Spatial inequality indices calculated using the ANPE file

Variables	Mean	q90/q10	q90-q10	Gini	Coeff. of variation	Number of obs.
Exit types and unemployment spells						
% Exit to job	.280	1.734	.152	.121	.224	1289
% Exit to non-employment	.167	1.370	.052	.070	.148	1289
% Right-censoring	.553	1.322	.152	.218	.426	1289
Duration if exit to job	276	1.374	87	.070	.138	1254
Duration if exit to non-employment	369	1.433	131	.083	.179	1156
Duration if right-censoring	334	1.753	179	.130	.281	849
Characteristics of unemployed workers						
% Age	32.610	1.080	2.499	.017	.032	1289
% Male	.518	1.164	.078	.033	.068	1289
% Female	.482	1.177	.078	.035	.073	1289
% Single	.606	1.327	.174	.062	.113	1289
% Couple	.394	1.599	.175	.095	.174	1289
% 0 child	.613	1.352	.185	.065	.116	1289
% 1 child	.163	1.458	.061	.085	.174	1289
% 2 children	.124	1.875	.074	.135	.264	1289
% 3 children	.057	2.814	.056	.212	.404	1289
% 4 children	.023	5.281	.032	.306	.569	1289
% 5 children and more	.019	6.938	.032	.378	.703	1289
% French	.782	1.315	.214	.060	.107	1289
% European (other)	.064	2.636	.061	.209	.402	1289
% North African	.077	5.444	.110	.305	.541	1289
% Sub-Saharan African	.045	5.665	.063	.303	.531	1289
% Other Nationality	.032	1.671	.051	.371	.695	1289
% College diploma	.239	3.942	.315	.285	.513	1289
% High School (excluding final year)	.165	1.636	.079	.106	.205	1289
% High school (final year and diploma) and technical diploma	.327	2.199	.231	.152	.272	1289
% Secondary school	.269	2.455	.226	.179	.314	1289
% Disabled	.033	2.631	.030	.195	.399	1289

Source: constructed from the ANPE file, sample of workers whose unemployment spell started between January 1996 and June 1996. All indices are weighted by the number of unemployed workers.

Table 3: Estimation results of the first-stage equation (SPLE)

Variables	Job	Non-employment
Age /100	-2.9289***	-9.0729***
	(.2801)	(.3253)
(Age/100) squared	1.210***	11.330***
	(.387)	(.442)
Male	*<ref>*	*<ref>*
Female	-.1819***	.3486***
	(.0060)	(.0079)
Single	*<ref>*	*<ref>*
Couple	.1089***	.0710***
	(.0077)	(.0094)
No child	*<ref>*	*<ref>*
1 child	-.0815***	.0834***
	(.0093)	(.0110)
2 children	-.0266**	.0375***
	(.0106)	(.0130)
3 children	-.1312***	.0352**
	(.0149)	(.0174)
4 children	-.1823***	.0428*
	(.0245)	(.0260)
5 children and more	-.2425***	.0852***
	(.0299)	(.0281)
French	*<ref>*	*<ref>*
European (other)	-.0510***	-.1732***
	(.0124)	(.0168)
North African	-.4455***	-.0810***
	(.0143)	(.0154)
Sub-Saharan African	-.6638***	-.0244
	(.0209)	(.0198)
Other Nationality	-.5629***	.0248
	(.0231)	(.0224)
College diploma	*<ref>*	*<ref>*
High School (first grade)	-.2296***	-.0970***
	(.0089)	(.0118)
High school (other grade) and technical diploma	-.3349***	-.2176***
	(.0078)	(.0107)
Secondary school	-.5872***	-.4252***
	(.0095)	(.0119)
Not disabled	*<ref>*	*<ref>*
Disabled	-03837***	.4653***
	(.0197)	(.0168)
Number of observations	430,695	

Source: constructed from the ANPE file.
***: significant at 1% level; **: significant at 5% level; *: significant at 10% level.
Monthly dummy variables were also included to control for seasonality but are not reported in the table.

Table 4: Disparity indices at the municipality level

Statistics on durations	Mean	P90 / P10	P90 - P10	Gini	Coeff. of variation
Until exit to job					
Survival at 6 months					
Kaplan-Meier	.801	1.173	.127	.035	.065
Model	.811	1.132	.100	.028	.053
Multiplicative model	.811	1.135	.103	.027	.055
Survival at 24 months					
Kaplan-Meier	.533	1.503	.213	.088	.164
Model	.537	1.426	.188	.076	.143
Multiplicative model	.535	1.416	.185	.076	.142
Until exit to non employment					
Survival at 6 months					
Kaplan-Meier	.893	1.059	.051	.012	.027
Model	.896	1.049	.043	.011	.024
Multiplicative model	.894	1.068	.059	.011	.037
Survival at 24 months					
Kaplan-Meier	.681	1.200	.123	.039	.093
Model	.686	1.175	.109	.036	.075
Multiplicative model	.683	1.175	.110	.036	.087

Source: constructed from the ANPE file.
Fixed effects in the multiplicative model are computed using 8 intervals of 90 days and one interval covering the remaining days. Municipalities are weighted by the number of unemployed workers.

Table 5: Variance analysis at the municipality level

	Exit to Job		Exit to non-employment	
	Variance	Correlation With $X^j\beta$	Variance	Correlation With $X^j\beta$
$X^j\beta$	.0068	1	.0016	1
$\ln H_k 6$	.0750	.668	.0387	.127
$\ln H_m 6$	.0556	.468	.0381	-.081
$\ln H_m 6 + X^j\beta$	.0805	.679	.0385	.126
$\ln H_{mm} 6$	.0499	.437	.0309	-.030
$\ln H_{mm} 6 + X^j\beta$	.0728	.667	.0322	.197
$\ln H_k 12$	.0702	.693	.0358	.177
$\ln H_m 12$	.0508	.490	.0346	-.049
$\ln H_m 12 + X^j\beta$	.0757	.700	.0355	.167
$\ln H_{mm} 12$	.0449	.435	.0342	-.024
$\ln H_{mm} 12 + X^j\beta$	.0726	.666	.0355	.193
$\ln H_k 24$	.0602	.646	.0374	.149
$\ln H_m 24$	.0465	.387	.0372	-.087
$\ln H_m 24 + X^j\beta$	.0669	.639	.0375	.123
$\ln H_{mm} 24$	.0481	.435	.0340	.034
$\ln H_{mm} 24 + X^j\beta$	.0705	.668	.0352	.035

Source: constructed from the ANPE file.
Fixed effects in the multiplicative model are computed using 8 intervals of 90 days and one interval covering the remaining days.
$X^j\beta$: average effect of individual explanatory variables at the municipality level. $\ln H_k T$: log of integrated hazard at T days using the Kaplan-Meyer estimator. $\ln H_m T$: log of integrated hazard at T days for the model. $\ln H_{mm} T$: log of integrated hazard at T days for the model under the multiplicative assumption. Statistics are computed weighting municipalities by their number of unemployed workers.

Table 6: Regressions of town fixed effects (for exit to job) on municipality variables

	(1)	(2)	(3)
Constant	-6.717***	-6.147***	-6.644***
	(.120)	(.035)	(.121)
Proportion of technical diplomas	1.861***		1.966***
	(.337)		(.338)
Proportion of high school diplomas	-.078		-.260
	(.384)		(.386)
Proportion of college diplomas	.099		.354*
	(.178)		(.192)
Proportion of European (other)	-1.394***		-1.402***
	(.246)		(.245)
Proportion of North Africans	-1.756***		-1.344***
	(.220)		(.250)
Proportion of Sub-Saharan Africans	-3.775***		-3.872***
	(.513)		(.512)
Proportion of other nationalities	-.458*		-.491**
	(.246)		(.245)
Job density within 45mins by public		-.092***	.000
transport		(.016)	(.012)
Job density within 45mins by		-.517***	-.176***
private transport		(.047)	(.051)
Number of observation	1254	1254	1254
Weighted number of observations	430602	430602	430602
Error rate	.468	.246	.473
Pseudo-R^2	.724	.259	.730

Source: constructed from the ANPE file.
***: significant at 1% level; **: significant at 5% level; *: significant at 10% level.
Estimates are computed using 8 intervals of 90 days and one interval covering the remaining days. Municipalities are weighted by the number of unemployed workers.

Table A1: Descriptive statistics on variables used in the study

Variable	Number of obs.	Mean	Standard deviation	Minimum	Maximum
Exit types and unemployment spells					
Exit to job	430,695	.280	.449	.000	1.000
Exit to non-employment	430,695	.167	.373	.000	1.000
Right-censoring	430,695	.672	.470	.000	1.000
Duration if exit to job	120,502	273	337	1	2818
Duration if exit to non-employment	71,807	368	452	1	2813
Duration if right-censoring	238,386	287	378	1	2829
Characteristics of unemployed workers					
Age	430,695	32.610	9.222	16.000	54.000
Male	430,695	.518	.500	.000	1.000
Female	430,695	.482	.500	.000	1.000
Single	430,695	.606	.489	.000	1.000
Couple	430,695	.394	.489	.000	1.000
0 child	430,695	.613	.487	.000	1.000
1 child	430,695	.163	.369	.000	1.000
2 children	430,695	.124	.330	.000	1.000
3 children	430,695	.057	.233	.000	1.000
4 children	430,695	.023	.150	.000	1.000
5 children and more	430,695	.019	.137	.000	1.000
French	430,695	.782	.413	.000	1.000
European (other)	430,695	.064	.245	.000	1.000
North African	430,695	.077	.267	.000	1.000
Sub-Saharan African	430,695	.045	.207	.000	1.000
Other Nationality	430,695	.032	.175	.000	1.000
College diploma	430,695	.239	.427	.000	1.000
High School (excluding final year)	430,695	.165	.371	.000	1.000
High school (final year and diploma) and technical diploma	430,695	.327	.469	.000	1.000
Secondary school	430,695	.269	.443	.000	1.000
Disabled	430,695	.033	.178	.000	1.000
Segregation variables					
Unemployment rate	430,695	.127	.043	.000	.246
% French	430,695	.794	.078	.569	1.000
% European (other)	430,695	.071	.020	.000	.265
% North African	430,695	.064	.041	.000	.218
% Sub-Saharan African	430,695	.030	.019	.000	.086
% Other Nationality	430,695	.041	.027	.000	.230
% No diploma	430,695	.412	.091	.202	.663
% Technical diploma	430,695	.191	.057	.040	.402
% High school	430,695	.145	.020	.000	.266
% University	430,695	.252	.123	.030	.571
Job-accessibility variables					
45mn job density by public transport	430,695	1.085	.436	.076	19.920
45mn job density by car	430,695	.860	.152	.152	1.200

Source: constructed from the ANPE file.