Neighborhood effects, public housing and unemployment in France*

Claire Dujardin
Université catholique de Louvain - FNRS †

Florence Goffette-Nagot
GATE - CNRS ‡

April 2006

Abstract

This paper is aimed at examining how individual unemployment is influenced both by location in a deprived neighborhood and public housing. Our identification strategy is twofold. First, we estimate a simultaneous probit model of public housing accommodation, type of neighborhood, and unemployment, thus accounting explicitely for correlation of unobservables between the three behaviors. Second, we take advantage of the situation of the public housing sector in France, which allows us to use public housing accommodation as a powerful determinant of neighborhood choices and to use household's demographic characteristics as exclusion restrictions. Our results show that public housing does not have any direct effect on unemployment. However, living within the 35% more deprived neighborhoods does increase the unemployment probability significantly. As expected, the effect of neighborhood substantially decreases when dealing with the endogeneity of neighborhood and when using public housing as a determinant of neighborhood choice.

Keywords: Neighborhood effects, public housing, unemployment, simultaneous probit models, simulated maximum likelihood.

JEL code: R2, J64.

*Claire Dujardin thanks the FNRS for financing a research stay at GATE. Florence Goffette-Nagot acknowledges funding from the CNRS APN-program that allowed initiating this work.

†CORE and Département de Géographie, Université catholique de Louvain, Place Pasteur 3, 1348 Louvain-la-Neuve, Belgium e-mail: dujardin@geog.ucl.ac.be.

‡GATE, CNRS - Université de Lyon 2 - ENS-LSH, 93 Chemin des Mouilles, 69131 Ecully, Tél: 33 +4.72.86.60.78. Fax: 33 +4.72.86.60.90. France email: goffette-nagot@gate.cnrs.fr. Corresponding author.

1 Introduction

A rapidly growing stream of research in the social interactions literature focuses on neighborhood effects, that is, the impact of neighbors' characteristics and behaviors on individual socio-economic outcomes.[1] Indeed, theoretical and empirical evidence suggests that interactions among neighbors are likely to affect individual labor-market outcomes through peer effects and role models in the human capital acquisition process, attitudes towards work, and dissemination of information on job opportunities. Arnott and Rowse (1987) show that less-able learners exerts negative externalities on the learning process of other students. Bénabou (1993) argues that the cost of education acquisition may be influenced by education decisions of neighbors. Wilson (1987) explains that the lack of successful role models among older adults in deprived neighborhoods may influence youths' motivations and attitudes. The role of social networks on information about job opportunities has also been highlighted, especially for low-skilled workers who often resort to informal search modes such as personal contacts. As a consequence, the percentage of employed individuals in the neighborhood may influence other residents' access to job opportunities (Topa, 2001; Bayer *et al.*, 2005). Finally, the stigmatization of deprived neighborhoods may lead employers to discriminate workers on the basis of their residential location (Zenou and Boccard, 2000).

Measuring neighborhood effects raises the issue of location choice endogeneity, which generates correlated effects (Moffitt, 2001; Durlauf, 2004). Indeed, urban economics has recognized for long that individuals with similar socio-economic characteristics, labor-market outcomes, and unobservable traits tend to sort themselves into certain areas of the urban space. Therefore, studies that do not control for the endogeneity of neighborhood choice will yield biased results. The inadequate correction for this bias has been put forward to explain the great divergence of results obtained by empirical studies and is one of the major focuses of recent research on neighborhood effects.

This paper aims to test for the existence of neighborhood effects on unemployment. Our identification strategy is twofold. First, it consists in correcting for endogenous selection into neighborhoods by estimating simultaneously two non-linear models of unemployment and neighborhood choice. More precisely, in a preliminary data analysis step, we classify neighborhoods as deprived or not deprived and then estimate a simultaneous probit model of unemployment and type of neighborhood. Second, the large share of public housing units in France and their

[1]See Durlauf and Young, 2001 for a review of the social interactions literature and Durlauf, 2004 for neighborhood effects.

concentration in poor neighborhoods, where they may represent as much as two thirds of housing units, allow us to use public housing accommodation as a powerful determinant of location in these neighborhoods. Criteria used in the public housing assignment process do not only rely on the household's economic situation, but depends also strongly on its demographic characteristics such as age of the spouse and composition. Those demographic criteria are used as exclusion restrictions in our system. In order to deal with the endogeneity of tenure choice, we add to our system a third probit equation of public housing accommodation. This strategy, that involves considering both public housing tenants and other households, also allows us to test for potential damaging effects of public housing accommodation, which is known to reduce residential mobility and may thus affect job search. Estimations of this simultaneous probit model are performed on a sample of approximately 10,000 individuals, taken from the 1999 French Census and representing about five percents of households' heads participating in the labor-market in Lyon, the third largest city in France.

The main contributions of this work are the treatment of neighborhood choice in a non-linear model of neighborhood effects on unemployment and the test for a negative influence of public housing accommodation on unemployment on European data. Our results show that public housing does not have any detrimental effect on unemployment, thus complementing Jacob's (2004) results concerning public housing and educational outcomes in the U.S. Further, living in a neighborhood displaying a combination of low-skilled population, high unemployment rate, and high proportion of foreigners increases the unemployment probability significantly. Our estimate is comparable to that Topa (2001) obtained for Chicago. These results also shed light on the potential effects of a recent French law aimed at achieving a more even spatial distribution of public housing units within cities. Indeed, our model enables us to simulate the impact of a change in the location of public housing tenants on unemployment.

The paper is structured as follows. Section 2 presents our identification strategy, the empirical model and the econometric method. Section 3 describes the database and gives a brief description of the spatial structure of Lyon. Section 4 presents the main results and section 5 concludes.

2 Model specification

2.1 Identification strategy

In his widely cited article, Manski (1993) considers two effects by which the social group may impact an individual's behavior. Individual behavior can be influenced either by the average behavior of his/her reference group, or by average characteristics of the members of this group. The first effect is referred to as an endogenous effect, while the latter is called a contextual effect. Moreover, similar behaviors in a group can be the consequence of exposure to common unobserved factors giving rise to correlated effects. Correlated effects may be caused by simultaneity in behaviors, common shocks or non random group selection. The goal of contemporaneous work on neighborhood effects is to disentangle these different kinds of mechanisms, in particular because endogenous and contextual effects, if shown to exist, have different policy implications (Moffit, 2001; Glaeser and Scheinkman, 2001). Recent empirical studies highlight the reduction of estimated neighborhood effects that stems from correcting for several biases (Ginther *et al.*, 2000; Krauth, 2005). The endogeneity of group membership in particular is likely to generate large biases, because individuals sort themselves into neighborhoods depending on their observable and unobservable characteristics.

The goal of this work is to estimate the intensity of neighborhood effects, focusing on the correction for selection into neighborhoods. Indeed, we do not try to disentangle endogenous and contextual effects, but we aim at providing an estimate of their global effect on unemployment probability. Our identification strategy consists in dealing with the endogeneity of neighborhood choice by estimating simultaneously two probits for unemployment and the choice of neighborhood.[2] Estimating simultaneously the two probits is a simple way to correct for endogeneity (Greene, 1998) that, to our knowledge, has not been used in the context of neighborhood effects. We treat neighborhood choice as a dummy variable indicating whether each neighborhood[3] of Lyon may be considered, on the basis of the social characteristics of its residents, as likely to generate negative spillovers in terms of unemployment. Specifically, the neighborhood type is

[2]Various strategies have been developed to correct for the endogeneity of neighborhood choice. Instrumental variables methods were often used, but Rivkin (2001) shows that using aggregate variables as instruments may actually increase the endogeneity bias. Quasi-experimental situations such as the Gautreaux Program and the Moving To Opportunity program provided more reliable estimates of neighborhood effects on labor-market outcomes (see Oreopoulos, 2003 for a review). However, we are not aware of any such possibility in the French case. A third strand of literature uses aggregate statistics and their variation in space to assess the importance of neighborhood effects (Glaeser *et al.*, 1996; Topa, 2001).

[3]See Section 3 for the definition of neighborhoods.

defined through a data analysis step, in which neighborhoods are classified as deprived or not according to characteristics likely to influence information on job opportunities, role models, peer effects in human capital acquisition or to generate statistical discrimination. This methodology is also motivated by the idea that individual outcomes are influenced by a wide variety of neighborhood characteristics. Introducing separately all of them is not desirable because of the high degree of correlation observed between such variables, which may cause instability in the parameters and significance levels (O'Regan and Quigley, 1998). The neighborhood type is then used to estimate simultaneously a probit model of unemployment and a probit model of location in a disadvantaged neighborhood. The simultaneous probit accounts for the correlation between unobservables by explicitly estimating the correlation matrix of residuals. Including neighborhood type in variables affecting unemployment allows to test for the presence of neighborhood effects.

Our identification strategy also takes advantage of the French process of assignment of households to public housing units. Indeed, although the identification of a simultaneous probit model does not formally require exclusion restrictions (Wilde, 2000), we have exclusion restrictions that are grounded on demographic criteria used by French public housing offices for giving access to public housing. In order to be eligible for public housing, French households must have an income below a certain threshold. Moreover, because demand largely exceeds supply, applications are ranked on a waiting list, subject to several criteria (for instance, households with a disabled person, or single-parent families are considered as having priority) and available housing units are proposed to households following their rank on the waiting list. They may then accept or refuse the proposal, and in the latter case may receive new proposals later. In 2002, one quarter of households housed in the public housing sector had rejected at least one offer before accepting one; half of these refusals were justified by the fact that *"the housing unit was in a neighborhood that did not fit household's preferences"* (Insee[4], 2002 French Housing Survey). Thus, the French public housing application process allows households to choose their neighborhood and forbids us to consider *a priori* the location of public housing renters as exogenous, as done by Oreopoulos (2003). Yet, as will be clear in our results, public housing accommodation is a strong determinant of the location in deprived neighborhoods and helps us identifying the effect of neighborhood on unemployment. Indeed, it is first worth noting that public housing units represent almost one half of the French renting sector (17% of the housing stock in 2002; Insee, 2003) and that a large part of those housing units belong to large projects located in the periphery of urban cores, thus providing a powerful source of income segregation. Consequently,

[4]French National Institute for Statistics.

a variable for public housing accommodation is included in the neighborhood equation. The potential endogeneity of public housing accommodation is dealt with by estimating a third probit model for housing tenure with the two former probits.

Finally, this strategy also permits us to test for potential detrimental effects of public housing accommodation on unemployment. Indeed, residing in a public housing unit may affect labor-market opportunities of individuals by constraining their residential location choices and subsequent residential mobility. In France, public housing renters are at risk of not obtaining another public housing unit if they move home. This may explain that annual mobility rates of public renters are at 10 percent against 16 percent in the private sector (Debrand and Taffin, 2005). Higher mobility costs of public renters may raise their reservation wage, thus increasing their unemployment probability. In order to test for such an effect, the public housing variable is also included in the unemployment probit equation.

2.2 Empirical model and econometric method

This study only deals with couple households, because the case of single adults suffers from a selection bias, young adults being less likely to form a separate household if they are unemployed.[5] Moreover, because dealing with women would imply to explain not only unemployment, but also labor-market participation, our study only concerns the household head.

Although the classical theory of job search ends up in the estimation of unemployment duration models, our dataset only allows us to estimate the probability of unemployment. This reduced form is assumed to gather both how neighborhood characteristics affect the arrival rate of job offers and how they impact reservation wages. Unemployment is then explained, in a classical manner, by individual characteristics relative to experience (that will be proxied by age and its square to allow for a non-linear effect), education and previous occupation. The individual's nationality is included in order to account for potential discrimination by employers. The spouse nationality is used as a proxy for the access to information on job opportunities through the network of relatives, as opposed to the social network provided by the neighborhood. Lastly, the two residential variables of neighborhood type and public housing accommodation are included as explanatory variables of unemployment in order to test our hypotheses.

As our model includes two endogenous observed discrete variables on the right hand side

[5]Note that because we deal explicitly with residential sorting, we did not find useful to work on young adults still living with their parents.

of the unemployment equation (and it is also the case of the public housing variable in the neighborhood choice equation), it amounts to a mixed model, which is consistent only if it has the form of a triangular system (Maddala, 1983). Therefore, the observed variable of unemployment can not be introduced in the other two equations of neighborhood type and public housing accommodation. Nonetheless, we may think that residential demand is influenced by the latent variable determining unemployment more than by the observed variable itself. Consequently, all the variables determining the latent variable of unemployment are included in the neighborhood and public housing equations. Moreover, as the simultaneous probit model enables us to deal with correlation between unobservables, the effect of unobservables determining both unemployment and residential choices are taken into account. Neighborhood type is also explained by spouse's educational level, that gives further information on permanent income of the household, and dummies for the number of children in the household, that determine housing floor space need and the propensity to settle in neighborhoods where housing rents are low. As a large share of deprived neighborhoods inhabitants are public housing tenants, public housing is also supposed to determine the probability to live in a deprived neighborhood.

As far as the public housing equation is concerned, one may think that neighborhood influences public housing choice: households that do not desire to locate in a deprived neighborhood might be deterred from applying to the public sector due to the location of public housing units. This means that the latent variable determining neighborhood choice may influence tenure choice and implies introducing all exogeneous variables influencing neighborhood choice in the public housing equation. Being housed in the public housing sector (which reflects both that the individual applied for and obtained a public housing unit) is then explained by the same variables as the neighborhood choice, with the addition of the spouse's age, because young households are given preferential attribution of public housing units. The complete list of variables and their descriptive statistics are given in Table 1.

In summary, the observed variables y_1, y_2 and y_3 referring respectively to unemployment, location in a disadvantaged neighborhood and public housing accommodation are defined by:

$$y_1 = \begin{cases} 1 & if \qquad y_1^* > 0, \\ 0 & otherwise \end{cases} \tag{2.1}$$

$$y_2 = \begin{cases} 1 & if \qquad y_2^* > 0, \\ 0 & otherwise \end{cases} \tag{2.2}$$

$$y_3 = \begin{cases} 1 & if \qquad y_3^* > 0, \\ 0 & otherwise \end{cases} \tag{2.3}$$

where y_1^*, y_2^* and y_3^* are latent variables influencing the probability of unemployment, the probablity to live in a deprived area, and the probability to be renter in the public sector respectively.

The system of latent variables is as follows:

$$\begin{cases} y_1^* &= \alpha_1 X_1 + \beta y_2 + \gamma y_3 + u_1 \\ y_2^* &= \alpha_2 X_2 + \delta y_3 + u_2 \\ y_3^* &= \alpha_3 X_3 + u_3 \end{cases} \tag{2.4}$$

where X_1 is a vector of exogenous variables including a constant, individual's age and its square, nationality, diploma and previous occupation as well as the spouse's nationality (each of them being a set of dummy variables), X_2 includes the same set of variables as X_1, the spouse's diploma and dummies for the number of children and X_3 includes the same set of variables as X_2 and the age of the spouse. β and γ test for the influence on unemployment probability of neighborhood type and public housing accommodation respectively.

As we assume that the sorting of households in deprived neighborhoods may be affected by unobserved characteristics influencing simultaneously unemployment and residential choice, the correlation terms between the residuals of the three probits (u_1, u_2 and u_3) are all supposed to be non-zero. The vector of residuals (u_1, u_2, u_3) follows thus a normal trivariate law with zero means and a covariance matrix that writes, after normalizations to 1 of the diagonal elements as usual in probit models:

$$Cov(u_1, u_2, u_3) = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix} \tag{2.5}$$

Such a system can be estimated by a maximum likelihood method. Endogeneity tests amount to test the significance of the correlation coefficients of residuals between two equations.[6] Note also that we use Huber adjusted standard errors, that is, we calculate a robust variance matrix which accounts for the potential dependence of residuals within neighborhoods. Indeed, the literature on neighborhood effects underlines the biases that stem from the possible existence of common random shocks affecting all the individuals in a neighborhood. Our sample having a large number of clusters and few individuals in each cluster, coefficients of cluster-level variables are consistently estimated, but the variance matrix must be corrected for within-cluster dependence (Wooldridge, 2003).

[6] Fabbri *et al.* (2004) show by means of a Monte-Carlo study that in a bivariate probit model, the likelihood ratio test performs well for testing this significance.

Individual contributions to the likelihood can be written as follows:

$$P(y_{i1}, y_{i2}, y_{i3}) = \Phi_3[q_{i1}(\alpha_1 X_{i1} + \beta y_{i2} + \gamma y_{i3}), q_{i2}(\alpha_2 X_{2i} + \delta y_{i3}), q_{i3}(\alpha_3 X_{3i}), q_{i1}q_{i2}\rho_{12}, q_{i1}q_{i3}\rho_{13}, q_{i2}q_{i3}\rho_{23}] \tag{2.6}$$

where $q_{ij} = 2y_{ij} - 1$ is equal to 1 whenever y_{ij} is 1 and to -1 whenever y_{ij} is 0, subscript i denotes individual i and $\Phi_3(.)$ is the trivariate normal cumulative distribution function. The log-likelihood function is then:

$$lnL = \sum_{i}^{N} lnP(y_{i1}, y_{i2}, y_{i3}) \tag{2.7}$$

The calculation of individual contributions requires to integrate over the distribution of the vector of three error terms, which means the calculation of a triple integral. Simulated maximum likelihood methods have been developed to circumvent this problem. One of the simulators commonly used is the GHK (for Geweke-Hajivassiliou-Keane) simulator.[7] The accuracy of the GHK simulator is good as soon as the number of random draws is equal to or higher than the square root of the sample size (Cappellari and Jenkins, 2003). With a sample of 10,473 individuals, we use 600 replications for each estimation, which is far above this threshold.

[7]The principle of this simulator is to use the lower triangular Cholesky decomposition of the covariance matrix of error terms to replace correlated random variables by uncorrelated ones, which are drawn from truncated normal density functions. Individual contributions to the likelihood are calculated as averages over several repeats of the random draw. See for example Bolduc, 1999 for a presentation of the GHK simulator and its use in a multinomial probit model.

3 Data and basic evidence

3.1 Data

This paper focuses on Lyon, the third largest city in France. Its agglomeration (defined here by its urban unit[8]) extends over a 958 km^2 area and hosts around 1.3 million inhabitants. As shown in the next subsection, Lyon is characterized by the existence of pockets of unemployment in the close periphery of its center and thus appears to be an adequate case study to test for the existence of neighborhood effects.

The empirical analysis conducted in this paper is based on two datasets extracted from the 1999 French Population Census. The neighborhoods are defined on the basis of *Iris*, the finest geographical level available in the French Census (they will be called neighborhoods in the rest of the paper, for the sake of simplicity). These neighborhoods are either municipalities, or subdivisions of municipalities if the latter have more than 10,000 inhabitants. They are created in order to represent homogenous entities in terms of housing and population. They are generally formed around well identified groups of buildings and respect frontiers such as main avenues, rivers or railways. Our study area has 540 neighorhoods[9] which have on average 2,428 inhabitants, a figure more or less comparable to the size of American Census tracts used in previous studies of neighborhood effects in the U.S.

Our first dataset gathers summary statistics at the neighborhood level and includes various indicators of the socioeconomic composition and average housing characteristics. This data is used to define the typology of neighborhoods (see next subsection). The second dataset corresponds to a sample of indiduals (1/20th of the total population), for whom detailed personal, household, and housing characteristics are provided (age, gender, education[10], employment status, household type, housing tenure[11],...) along with the characteristics of the other members of his/her household. This data allows to link each individual to the neighborhood in which s/he lives in. It is used to estimate our econometric model. As we already explained, our study deals

[8]The urban unit, *unité urbaine* in French, is a set of municipalities, the territory of which is covered by a built-up area of more than 2,000 inhabitants, and in which buildings are separated by no more than 200 meters. The urban unit of Lyon consists of 102 municipalities. For practical reasons, we added three municipalities which are enclosed within the urban unit of Lyon (Quincieux and Poleymieux-au-Mont-D'Or).

[9]A few *Iris* having less than 200 people had to be deleted for confidentiality reasons.

[10]In the whole paper, the following education levels will be used: No diploma, At most lower secondary school, Vocational training, High school final diploma, University degree. They correspond to the following French categories: no reported diploma, CEP or Brevet, CAP or BEP, *Baccalauréat*, DEUG or above, respectively.

[11]Note however that neither housing prices nor incomes are available in the French Census.

with heads of couple households, aged 19 to 64 and participating in the labor-market. Due to data availability on previous occupation, we deleted individuals who never worked, that is only 18 individuals. The final sample contains 10,473 individuals, all of them being males.

3.2 Neighborhood typology

The agglomeration of Lyon presents a well-marked spatial structure, with some parts of the city characterized by a concentration of disadvantaged communities. Figure 1 maps the percentage of unemployed workers among labor-force participants. In most American cities, central neighborhoods exhibit higher unemployment rates than peripheral neighborhoods. In Lyon also, the neighborhoods with the lowest unemployment rates are found in the far periphery, but Figure 1 shows that the highest unemployment rates are found in the close periphery of Lyon's municipality and not in the center.[12] As seen in Figure 2, in some of the neighborhoods displaying the highest unemployment rates, more than 50% of households (and even more than 70% for some of them) are housed in the public renting sector. This pattern is very typical of French cities and reveals the role that the public housing projects built in the 1970's had in spatially concentrating low-income households. The unemployment spatial structure is also quite related to the distribution of education levels and professional statuses as well as to the distribution of ethnic minorities. As a consequence, one can suspect the existence of neighborhood effects affecting labor-market outcomes of public housing tenants and of other individuals located in these neighborhoods.

Our typology of neighborhoods is aimed at reflecting for each neighborhood its social composition and the neighborhood effects that might potentially affect job search and unemployment. Therefore, it is built on the basis of the following variables: distribution of population by education levels, percentage of executives and blue-collars in labor force, percentage of unemployment, long-term unemployment and youth unemployment, percentage of household heads of foreign nationality, and percentage of lone-parent families. Each of these neighborhood characteristics is likely to affect individual unemployment propensity: low income levels (proxied by professional status distribution) may decrease the global investment in human capital and human capital spillovers; high unemployment rates as well as high rates of foreigners decrease information on job opportunities and may give rise to statistical discrimination; low education levels give low incentives for youths to invest in education and, together with high proportions of

[12] Preliminary tests of the impact of time distance on unemployment probability did not however reveal any empirical support for the spatial mismatch hypothesis, which was therefore left aside.

lone-parent families, provide few sucessful role models. This set of variables is treated by means of standard factorial ecology methods. We first ran a Principal Component Analysis to define a number of non-correlated factors summarizing the information carried by these variables (see Table A.1 in Appendix). Then, we gathered neighborhoods according to their respective coordinates on the factorial axes using a hierarchical ascending classification method (with the Ward method that minimizes intra-group variance). We obtained five[13] clusters of neighborhoods that are presented in Appendix (Table A.2).

In order to deal with a dummy variable, we grouped the least two favored neighborhood types as opposed to the rest of the city, thus defining the endogenous variable y_2. These neighborhoods, labelled "deprived" in the rest of the paper, represent 35% of the 540 neighborhoods and of the population of Lyon's city. They are spread in different parts of the city, still mostly concentrated in its eastern half (Figure 3). They are characterized by high unemployment rates (twice as high as the average unemployment rate of other neighborhoods), high percentages of foreigners and low educational levels and professional statuses (Table 2). Most of them have a large share of public housing, but 10% of them have less than 10% of public housing units.

3.3 Neighborhood, public housing and unemployment: descriptive statistics

Table 3 provides a few sample statistics by neighborhood type and by whether the individual is renter in the public sector or not. Deprived neighborhoods host almost one third of the individuals in our sample. Among deprived neighborhoods, 41% of individuals are renters in the public sector, against only 9% in other neighborhoods. Other residents in deprived neighborhoods are either renters in the private sector or homeowners (33% and 61% of them respectively). About one third of public housing renters in our sample are located in neighborhoods that are not classified as deprived. Thus, the diversity of situations regarding the combination of tenures and neighborhood types allows us to disentangle the effect of the two residential variables.

Compared with individuals having the same tenure (public housing versus others), individuals in deprived neighborhoods are less educated and have lower occupational statuses. Yet, they have similar demographic characteristics, except for public housing renters in deprived neighborhoods who have larger families than their counterparts in the rest of the city, owing to a large share of foreign families having more children than the average.

[13]This was the optimal number of clusters, according to a wide variety of criteria, including the Cubic Clustering Criterion, Pseudo-F and Pseudo-t values.

Unemployment rate varies markedly with respect to the residential situation.[14] Whatever their location, public housing renters are more often unemployed than others: one aim of the public housing sector is to provide individuals in a poor economic situation with affordable housing. Still, public housing renters in deprived neighborhoods are by 42% more often unemployed than other public housing renters. Individuals with other housing tenures display a similar picture: their unemployment rate in deprived neighborhoods is by about 50% higher than in other neighborhoods.

This differenciated unemployment rate of public housing renters depending on neighborhood type raises three interpretations. First, this can be the result of the variation in public housing rents depending on the location in the city. Although rents in the public sector are administrated, they vary in space and the most successful individuals on the labor market are likely to be able to afford the best-located public housing units. Second, this could account for peer effects that increase individual difficulties on the labor market when they live in a deprived neighborhood. Third, it could be the consequence of a self-selection effect, such that people less likely to find a job sort themselves in these neighborhoods. Our econometric analysis is intended to disentangle these different mechanisms.

4 Results

In this section, we present in turn results of simple probits, results of the simultaneous probit model, neighborhood and public housing predicted effects, and policy simulations.

4.1 Probit estimates

Table 4 contains marginal effects estimated from three simple probits: being a renter in the public sector, being located in a deprived neighborhood, and a probit of unemployment which is estimated in turn with and without the two residential variables. Demographic variables take a large part in determining the probability of being accommodated in a public housing. Younger households (as reflected by the spouse's age) and those with at least three children are more likely to rent a public housing unit, which is in line with assignment rules of public housing offices. For instance, having four children or more increases by 12 points the probability to be in a public housing unit. Individuals (or their spouse) of foreign nationality or, to a lower extent,

[14]Remark that the overall unemployment rate is different from the rate displayed in Table 2 due to the sample definition.

French people born abroad are more often housed in the public sector than French individuals. This observation might reflect an attempt by the public housing offices to compensate for discrimination on the private housing sector or the fact that foreign individuals are pushed toward the public housing sector due to this discrimination. As far as socioeconomic variables are concerned, occupational status along with education explain the propensity to live in a public housing unit. Blue-collar workers are more likely to rent a public housing unit than intermediate professions (the reference category) by 10 points, and office workers by 7 points. The lower the educational level is, the higher the probability of being renter in the public housing sector. Surprisingly, the spouse's educational level and not that of the household head is significant. This probably reflects the fact that it is the possibility to have or not a second wage in the household (low educated women having a weak incentive to take part in the labor-force) that determines income and is considered by public housing offices during the application process.

The second column gives marginal effects estimated from the neighborhood equation. A far as socioeconomic variables are concerned, marginal effects are very similar to marginal effects in the public housing equation. As expected, nationality, education and professional status determine the probability to live in a deprived neighborhood, even after conditioning for the accommodation in the public housing sector.[15] Only highly-educated individuals have significantly different behaviors regarding tenure and neighborhood choices: they do not differ from the reference category (high school final diploma) as far as tenure is concerned, whereas they are less likely than the reference to locate in a deprived neighborhood. We may think that skilled individuals are likely to apply for a public housing unit at the beginning of their career, but that in any case they avoid deprived neighborhoods. On the contrary, demographic variables do not explain the probability to live in a deprived neighborhood: neither the age of the household head[16], nor the number of children have significant coefficients. This accounts for the fact that the demographic situation of the household is among the criteria that are considered by public housing offices, whereas they are less relevant in determining residential location choice. Finally, the public housing variable is the more powerfull in explaining the neighborhood choice and it has the strongest marginal effect. The introduction of this variable significantly improves the likelihood of the model.[17] Being a renter in the public sector more than doubles (marginal effect +30 points) the probability to live in a deprived neighborhood, and as will be clear in next subsection from the simultaneous estimation of the three probits, this estimate does not

[15]Estimated coefficients do not change with the introduction of the public housing variable. Only the four-children variable looses it significance with the introduction of the public housing variable.

[16]Nor the age of the spouse introduced in a previous specification.

[17]The statistic of the likelihood ratio test is 570 for a $\chi^2_{0.5}$ critical value of 3.84.

suffer from any endogeneity bias.

The third and fourth columns of Table 4 give marginal effects for the unemployment equation. We find very conventional results regarding individual determinants. Young individuals are more often unemployed, and the probability to be unemployed declines until the age of 44, after which it increases again. Individuals without any diploma or with only a short vocational training are more likely not to find a job, whereas people who were previously independent workers or executives are less unemployed than others. Marginal effects do not change much with the introduction of the two residential variables (column 4 compared to column 3), with the exception of the blue-collars' marginal effect that looses its significancy. This result means that blue-collar workers seem more likely to be unemployed than technicians and supervisors, but that they in fact do not differ, when controlling for their tenure and location. Probit estimates show that unemployment probability increases both with location in a deprived neighborhood and with accommodation in the public sector, the latter being even more important than the former. However, these estimation results very likely suffer from an endogeneity bias.

4.2 Simultaneous probit model estimates

Table 5 presents the results of the simultaneous probit model. Coefficients of the public housing and neighborhood equations being very similar to the simple probit results, we do not comment them here. For the same reasons, we do not comment exogenous variables affecting unemployment propensity.

The correlation coefficient between the error terms of the neighborhood and the unemployment equations (ρ_{12}) is significantly different from zero at the 5% level, showing as expected that the neighborhood type is endogenous in the unemployment equation and that coefficients estimated from a simple probit are biased. The other two correlation coefficients are not significant, suggesting that the public housing variable is not endogenous in the unemployment equation, nor in the neighborhood equation. The latter result can be interpreted as showing that households are not deterred from applying to a public housing unit by the spatial distribution of the public housing sector and is coherent with the fact that the application process allows households to express spatial preferences.

While the coefficient of public housing in the simple probit of unemployment is highly significant (see Table 4), correcting for the endogeneity of neighborhood and taking into account its strong ties with public housing eliminate any effect of tenure on unemployment probability.

Flatau *et al.* (2003) find similar results for Australia, where public renters do not have higher unemployment probability once the endogeneity of tenure is accounted for. In our case, the apparent effect of public housing on unemployment is entirely due to its indirect influence through its positive effect on living in a deprived quarter, which itself raises unemployment.[18]

Indeed, the deprived neighborhood variable exerts a positive effect on unemployment probabilities. This variable is endogenous in the unemployment equation, that is, unobserved variables influencing unemployment are negatively correlated with unobserved characteristics affecting neighborhood choices. The negative sign of the correlation indicates that individuals having a higher propensity for unemployment than explained by their observed characteristics are less likely to live in a deprived neighborhood. While surprising at first sight, this result is in line with the observation of mixed neighborhoods in Lyon's urban core, that are not classified as deprived but that simultaneously have unemployment rates above the average and host younger individuals, with potentially less predictable paths.

As explained in the first part of the paper, estimating neighborhood effects requires to deal with correlated effects. Our identification strategy allows us to deal with the endogeneity of neighborhood choice. The estimated neighborhood effects could still be suspected to suffer from other biases due to random shocks common to all individuals in deprived neighborhoods. This is why our estimation method corrects the variance matrix of coefficients in order to account for dependencies within neighborhoods. This correction slightly changes the coefficient standard errors, but does not change the significativity with respect to conventional thresholds.[19] This is not surprising, since neighborhoods which are classified as deprived are spread in different parts of the city. There is no reason why each of the deprived neighborhoods would be concerned by a shock that would not affect the other types of neighborhoods in the same area. Note also that the estimations were performed for different initial values of correlation coefficients and all of them converged to the same correlation matrix and produced very similar coefficients. Other specifications differing with respect to exogenous explanatory variables were also estimated without changing the baseline results.

Because the public housing variable is not endogenous in the unemployment equation, nor in the neighborhood equation, and because it does not affect unemployment probability as soon as the endogeneity of neighborhood is properly dealt with, we base the assessment of neighborhood effects on the simultaneous estimation of two probits of unemployment and

[18]This result is confirmed by the estimation of a simultaneous model of two probits for unemployment and public housing accommodation, showing that public housing has no effect on unemployment probability.

[19]Detailed results available from the authors upon request.

neighborhood choice, including public housing in the neighborhood equation only.

4.3 Two probit estimates of neighborhood effects and public housing effects

A simultaneous model of two probits of unemployment and neighborhood choice is estimated by a classical likelihood maximization method, with the same exogenous variables as in the three probit model and gives very similar estimated coefficients. Given that the correlation term between the residuals of the two equations is significant, neighborhood effect on unemployment must be calculated as the difference in conditional probabilities, that themselves are calculated on the basis of joint probabilities. For instance, the effect on unemployment probability of living in a deprived quarter is:

$$P(y_{i1}=1|y_{i2}=1)-P(y_{i1}=1|y_{i2}=0)=\tfrac{P(y_{i1}=1,y_{i2}=1)}{P(y_{i2}=1)}-\tfrac{P(y_{i1}=1,y_{i2}=0)}{P(y_{i2}=0)} \tag{4.1}$$

This formula accounts both for the direct effect of neighborhood type and for the effect due to the correlation of unobservables between the two equations.

The public housing variable, as each exogenous variable affecting the probability to live in a deprived neighborhood, has an indirect effect on unemployment probability that may be calculated as:

$$\begin{aligned} P(y_{i1}=1|x=1)-P(y_{i1}=1|x=0) \quad = \quad & (P(y_{i1}=1,y_{i2}=1|x=1)+P(y_{i1}=1,y_{i2}=0|x=1)) \\ & -\ (P(y_{i1}=1,y_{i2}=1|x=0)+P(y_{i1}=1,y_{i2}=0|x=0)) \end{aligned} \tag{4.2}$$

Table 6 shows the effects of neighborhood and public housing accommodation on unemployment following several specifications that differ by the type of model (simple probit, seemingly unrelated probits and simultaneous bivariate probits) and by the presence of the public housing variable in the neighborhood equation. As suggested by Wooldridge (2001, p. 467), predicted effects are calculated for each individual and averaged over the sample. The standard errors of these effects are calculated by the delta method.[20]

As far as neighborhood type is concerned, column 1 displays the "naive" effect of +2.13 probability points that is calculated on the basis of the simple probit. In column 2, we take the correlation between unobservables into account by estimating a seemingly unrelated probit model; that is, we do not include neighborhood into the unemployment equation, but neighborhood type may still influence unemployment probability through the correlation between

[20] The delta-method allows to approximate the variance of a vector-valued function of a random vector X. It is based on the following general result: $Var(G(X))=(\partial G/\partial\bar{X})'Var(X)(\partial G/\partial\bar{X})$ where $\bar{X}$ is the mean of X, $Var(X)$ is the variance-covariance matrix of X, $G()$ is a vector function and $G'()$ its matrix of first derivatives.

unobservables. In this case, the estimated effect of living in a deprived neighborhood on unemployment is reduced by 4% compared with the simple probit estimate.[21] Then sorting on observable characteristics is accounted for by the estimation of the simultaneous model of two probits. The naive effect of living in a deprived neighborhood declines further to 1.94 (column 4). Finally, the comparison of neighborhood effects in columns 4 and 5 assesses the added value from explaining location in a deprived neighborhood by the public housing variable. This specification produces the strongest decrease in the estimated effect, that looses 40% as soon as the public housing variable is included among explanatory variables in the neighborhood equation (comparison of columns 4 and 5 or columns 2 and 3). In fact, when introducing the public housing variable in the neighborhood equation, we better account for the concentration of disadvantaged individuals in deprived neighborhoods. Therefore, we better control for self-selection effects and we obtain a much more reliable estimate of the neighborhood effect. This result shows that the particular situation of public housing renters in France provides a valuable opportunity to estimate the impact of neighborhood on socioeconomic outcomes.

As to public housing accommodation, the predicted effect on unemployment probability is 3.15 points in the baseline specification (Table 6, column 5). As we already explained, this effect is entirely due to the influence of public housing on neighborhood choice, and its intensity is due to the large impact of public housing accommodation on the probability to live in a deprived neighborhood.

As highlighted by Ginther *et al.* (2000), another potential concern in the estimation of neighborhood effects is an inadequate correction for unobserved heterogeneity. Although the estimation of the simultaneous probit system ensures that the correlation between unobservable characteristics is taken into account, it is worth performing an informal exercise in order to roughly evaluate the potential biases generated by unobservable traits. Therefore, we reestimate the model with two different specifications in which some known characteristics are assumed to be unobservables. The first specification consists in dropping the individual's occupational status in both equations, meaning that we neglect a characteristic which is quite important in determining the individual's behavior on the labor-market and on the housing market. The second specification eliminates the spouse nationality, a feature that has a weaker impact on unemployment (see Table 4). As expected, the correlation of residuals and the predicted marginal effect of neighborhood increase in both cases (Table 6, columns 6 and 7) compared with the

[21] The correlation between residuals is positive, because the fact that neighborhood type is explained by observable traits is not taken into account.

baseline specification.[22] However, the raise of neighborhood effect remains limited, with a maximum increase by only 0.18 points (that is, 15% of the baseline estimate) in the specification considering previous occupational status as unknown. Although this is only an informal way of assessing the effect of unobserved heterogeneity, these additional results suggest that our estimation method provides a reliable estimation of the neighborhood effect. This effect may be reasonably thought of as being a little higher than 1 point probability.

In summary, living in the 35% of neighborhoods that have been identified as having the worst combination of social characteristics in our data analysis step increases the probability of being unemployed by slightly more than 1%. The change of neighborhood type amounts to a decrease in neighbors' unemployment rate by 8.7%. By way of comparison, Topa (2001) found, in the case of Chicago in 1990, that an increase by 8% in the employment of neighboring tracts would increase employment rate by 1.3 %. As far as public housing is concerned, our results indicate that only an indirect effect exists, according to which being housed in the public sector increases unemployment probability by 3%. These effects can be compared with marginal effects of individual characteristics. For instance, the neighborhood effect is about as low as two-thirds that of spouse's foreign nationality, and it is twice as low as the effect of having the lower education level rather than having graduated from high school (Table 4, column 5). These effects can also be compared with differences in observed unemployment rates. On average, observed unemployment rate is by 4.9 points higher in deprived than in other neighborhoods (Table 7). According to our results, 1.18 probability points of this gap, that is a bit more than 20%, would be the consequence of neighborhood effects, the remaining part ensuing from spatial sorting. This results holds within each sub-category as defined by the two residential variables. For instance, the unemployment probability of individuals outside the public housing sector but living in deprived neighborhoods would decrease by 18% if they were located in another type of neighborhood.

4.4 Policy simulations

Our results give support to a law that was recently passed in France, aimed at achieving a more even distribution of public housing units in order to counter potentially harmful effects of public housing location (*Loi SRU "Solidarité et Renouvellement urbain"*, 2000). Our methodology allows us to go a step further and to assess the potential effect of a change in the spatial distribution of public housing units in French cities. Let us recall that for this purpose, it is

[22]Detailed results are available from the authors on request.

not necessary to distinguish endogenous and contextual effects, as both types of causalities are involved in the relocation of public housing units.

Table 7 displays predicted probabilities of unemployment that are issued from the baseline two probit model (results displayed in column 5 of Table 6). For each subsample, we give the observed unemployment rate and average predicted probabilities of unemployment if these individuals were to be located either in deprived or in other neighborhoods. The neighborhood effect is higher for public tenants and for individual in deprived neighborhoods. These figures show that relocating the public housing renters who live in deprived neighborhoods would reduce their individual unemployment probability from 14.2 to 11.0%.

To be more specific, assume that the location of other housing units remains identical. Then, given the initial distribution (Table 2), achieving an even distribution of the public housing tenants between both types of neighborhoods would imply transferring 65% of public housing units (that is, 43% of the public housing stock) from deprived to other neighborhoods. Supposing this new distribution to be implementable and assuming the unemployment rate of public housing renters in non-deprived neighborhood does not change, the overall unemployment rate of public housing tenants in Lyon's city would decrease from 12.5% to 11.2%.[23] This reduction is limited and in any case, such a change in the distribution of public housing units would be very costly.

Furthermore, it must be acknowledged that this simulation suffers from limitations. In particular, our simulation does not take into account the fact that with the change in the distribution of public housing, the percentage of public housing renters (characterized by low levels of education, higher unemployment rates, ...) in the rest of the city would be about twice as high as it is currently. Assessing more precisely the consequences of such a change would imply to estimate a continuous relationship between social composition and unemployment probability.

5 Conclusion

The objective of the present paper was to examine how unemployment probabilities are influenced both by accommodation in the public housing sector and location in a deprived neigh-

[23] This predicted average unemployment rate is calculated by applying the predicted unemployment probability conditional on living in a non-deprived neighborhood to 65% of public housing renters living in deprived neighborhoods, and the observed unemployment rates to the 35% remaining and to the public renters who are not in a deprived neighborhood.

borhood. Neighborhood types were defined through a data analysis step based on their social composition. We estimated simultaneously three probit equations relating respectively to unemployment, neighborhood type, and accommodation in the public housing sector, thus allowing to deal with endogeneity of the two residential variables with respect to unemployment. Potential dependencies within neighborhoods were accounted for by the estimation of a robust variance matrix. Demographic characteristics were used as exclusion restrictions. Estimation of this system by simulated maximum likelihood used the GHK simulator.

We observed that the endogeneity bias on coefficients of residential variables is relatively high. Our study also shows that the particular situation of public housing renters provides a valuable opportunity to estimate the impact of neighborhood on socioeconomic outcomes in France. Contrary to Oreopoulos (2003), this is not because the location of public housing tenants is exogenous, but because the tenure helps us to explain neighborhood choice by an exogenous characteristic.

Our results do not provide any support to the hypothesis according to which public housing accommodation would affect job search behavior and, in particular, would reduce residential mobility sufficiently so as to increase unemployment probability. As to residential location, we clearly observe a neighborhood effect on unemployment affecting, in particular, public housing renters. According to our results, living in one of the deprived neighborhoods (which represent 35% of Lyon's population) would increase the unemployment probability by 1.2 points. These results both add to the literature on neighborhood effects and give insight into a much debated policy issue in France and in other countries, that is, the effect of the location of public housing in cities on individual socioeconomic outcomes.

Of course, due to the chosen framework, this study does not allow us to estimate separately endogenous and contextual effects, because mean unemployment rate and neighbors' characteristics supposed to influence unemployment are used simultaneously in the classification of neighborhoods. Therefore, we are not able to test for the existence of a social multiplier, nor for specific mechanisms such as the role of social networks, stigma, or role models, but we keep these issues for future work. Further, we only estimate the change in unemployment occurring with a change of neighborhood type, and not a continuous effect. However, this strategy is relevant with respect to the fact that several correlated variables generate neighborhood effects and that some of them may have a non-continuous impact.

References

[1] ARNOTT, R. and ROWSE, J. (1987) Peer group effects and the educational attainment, *Journal of Public Economics*, 32(3), pp. 287-305.

[2] BAYER P., ROSS, S. and TOPA, G., (2005) Place of work and place of residence: informal hiring networks and labor market outcomes, *NBER Working Papers*, 11019.

[3] BENABOU, R. (1993) Workings of a city: location, education and production, *Quarterly Journal of Economics*, 108(3), pp. 619-652.

[4] BOLDUC, D. (1999) A practical technique to estimate multinomial probit models in transportation, *Transportation Research Part B: Methodological*, 33(1), pp. 63-79.

[5] CAPPELARI, L. and JENKINS, S. (2003) Multivariate Probit Regression Using Simulated Maximum Likelihood, *The Stata Journal*, 3(3).

[6] DEBRAND, T. and TAFFIN, C. (2005) La mobilité résidentielle depuis 20 ans: des facteurs structurels aux effets de la conjoncture, *Economie et Statistique*, forthcoming.

[7] DURLAUF, S.N. and YOUNG, H.P. (2001) *Social dynamics*, Mit Press.

[8] DURLAUF, S.N. (2004) Neighborhood effects, in Henderson, J.V. and Thisse, J.-F. (ed.) *Handbook of Regional and Urban Economics*, vol. 4, Elsevier, 2173-2242.

[9] FABBRI, D., MONFARDINI, C. and RADICE, R. (2004) Testing exogeneity in the bivariate probit model: Monte Carlo evidence and an application to health economics, Working Paper n° 514, Department of Economics, University of Bologna.

[10] FLATAU, P., FORBES, M., HENDERSHOTT, P.H. and WOOD, G. (2003) Homeownership and unemployment: the roles of leverage and public housing, *NBER Working Paper*, 10021.

[11] GINTHER, D., HAVERMAN R. and WOLFE B. (2000) Neighborhood attributes as determinants of children's outcomes: how robust are the relationships? *Journal of Human Resources*, 35(4), pp. 603-642.

[12] GLAESER, E.L., SACERDOTE, B. and SCHEINKMAN J.A. (1996) Crime and social interactions, *Quarterly Journal of Economics*, 111(2), pp. 507-548.

[13] GLAESER, E.L. and SCHEINKMAN J.A. (2001) Measuring social interactions, in: DURLAUF, S.N. and YOUNG, H.P. (ed.) *Social dynamics*, Mit Press., pp. 83-132.

[14] GREENE, W.H. (1998) Gender Economics Courses in Liberal Arts Colleges: Further Results, *Journal of Economic Education*, 29(4), pp. 291-300.

[15] INSEE (2003) *Les conditions de logement des ménages en 2002*, INSEE Résultats, n°20.

[16] JACOB, B.A. (2004) Public housing, housing vouchers, and student achievement: Evidence from public housing demolitions in Chicago, *American Economic Review*, 94(1), 233-258.

[17] KRAUTH, B. (2005) Peer effects and selection effects on smoking among Canadian youth, *Canadian Journal of Economics*, forthcoming.

[18] MADDALA, G.S. (1983) *Limited-dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

[19] MANSKI, C. (1993) Identification of endogenous social effects: the reflection problem, *Review of Economic Studies*, 60, pp. 531-542.

[20] MOFFITT, R.A. (2001) Policy intercention, low-level equilibria, and social interacions, in: DURLAUF, S.N. and YOUNG, H.P. (ed.) *Social dynamics*, Mit Press., pp. 45-82.

[21] O'REGAN, K.M. and QUIGLEY, J.M. (1998) Where youth live: economic effects of urban space on employment prospects, *Urban Studies*, 35(7), pp. 1187-1205.

[22] OREOPOULOS, P. (2003) The long-run consequences of living in a poor neighborhood, *Quarterly Journal of Economics*, 118(4), pp. 1533-1575.

[23] RIVKIN, S. (2001) Tiebout sorting, aggregation and the estimation of peer group effects, *Economic of Education Review*, 20, pp. 201-209.

[24] TOPA, G. (2001) Social interactions, local spillovers and unemployment, *Review of Economic Studies*, 68, pp. 261-295.

[25] WILDE, J. (2000) Identification of multiple equation probit models with endogenous dummy regressors, *Economic letters*, 69, pp. 309-312.

[26] WILSON, W.J. (1987) *The Truly Disadvantaged*, Chicago, University of Chicago Press.

[27] WOOLDRIDGE, J. (2001) *Econometric Analysis of Cross Section and Panel Data*, Cambridge: MIT Press.

[28] WOOLDRIDGE, J. (2003) Cluster-sample methods in applied econometrics, *American Economic Review*, 93, pp. 133-138.

[29] ZENOU, Y. and BOCCARD, N. (2000) Labor discrimination and redlining in cities, *Journal of Urban Economics*, 48(2), pp. 260-285.

Appendix: Building of the neighborhood typology

	Factor 1	Factor 2
Eigenvalue	4.19	4.61
Percent of variance explained	41.86%	46.15%
Loadings		
% families with foreign household head	**0.816**	-0.399
% monoparental households	**0.793**	-0.120
% pop. with at most lower secondary education	0.510	**-0.824**
% pop. with high school final diploma	-0.291	**0.949**
% pop. with a university degree	-0.212	**0.968**
% executives	-0.244	**0.931**
% blue-collars	0.486	**-0.820**
% unemployed workers	**0.921**	-0.308
% unemployed workers since more than one year	**0.908**	-0.314
% unemployed workers aged under 25	**0.730**	-0.437

Only factors with eigenvalues superior or equal to 1 were retained.

Table A.1: List of variables used in the principal component analysis and their contributions to factors

	Very well-off	Well-off	Mixed	Poor	Very poor	Total
Unemployment and tenure						
% unemployed workers	8.6	8.2	11.8	16.2	29.5	12.3
% public housing units	9.0	9.5	10.3	38.9	81.1	21.3
Demography						
% foreign household heads	5.5	5.9	8.5	17.4	34.3	10.9
% monoparental families	11.7	9.7	14.3	16.6	24.6	13.6
Education levels						
% at most lower secondary edu.	31.8	39.1	28.9	50.5	60.8	40.1
% university degrees	34.6	20.6	42.2	13.6	7.7	24.6
Occupational status						
% blue-collars	11.9	21.2	9.6	31.8	47.2	21.7
% executives	25.3	14.0	31.4	8.0	2.6	17.1

Table A.2: Mean population characteristics of the five types of neighborhoods defined by hierarchical ascending classification method

	Employed persons	**Unemployed persons**	**Total sample**
Number of observations	9,800	673	10,473
Residential characteristics			
Deprived neighborhood$^{(a)}$	2,980 (30.41)	323 (47.99)	3,303 (31.54)
Tenure			
Renter in the public sector	1,751 (17.87)	256 (38.04)	2,007 (19.16)
Renter in the private sector	2,517 (25.68)	217 (32.24)	2,734 (26.11)
Homeowner	5,155 (52.60)	178 (26.45)	5,333 (50.92)
Other tenures	377 (3.85)	22 (3.27)	399 (3.81)
Personal characteristics			
Age	41.84	41.55	41.83
Nationality			
French born in France	8,003 (81.66)	423 (62.85)	8,426 (80.45)
French born abroad	966 (9.86)	86 (12.78)	1,052 (10.04)
Foreign nationality	831 (8.48)	164 (24.37)	995 (9.50)
Education level			
No diploma	1,256 (12.82)	167 (24.81)	1,423 (13.59)
At most lower secondary edu.	1,223 (12.48)	112 (16.64)	1,335 (12.75)
Vocational training	2,796 (28.53)	183 (27.19)	2,979 (28.44)
High school final diploma	1,261 (12.87)	71 (10.55)	1,332 (12.72)
University degree	3,264 (33.31)	140 (20.80)	3,404 (32.50)
Occupational status			
Farmer or independent worker	1,041 (10.62)	37 (5.50)	1,078 (10.29)
Executive	2,468 (25.18)	89 (13.22)	2,557 (24.42)
Intermediate professions (b)	3,204 (30.59)	154 (22.88)	2,685 (25.64)
Office worker	957 (9.77)	64 (9.51)	1,021 (9.75)
Blue-collar	2,803 (28.60)	329 (48.89)	3,132 (29.91)
Characteristics of the spouse			
Age			
Nationality			
French born in France	8,133 (82.99)	451 (67.01)	8,584 (81.96)
French born abroad	871 (8.89)	78 (11.59)	949 (9.06)
Foreign nationality	796 (8.12)	144 (21.40)	940 (8.98)
Education level			
No diploma	1,171 (11.95)	159 (23.63)	1,330 (12.70)
At most lower secondary edu.	1,517 (15.48)	125 (18.57)	1,642 (15.68)
Vocational training	2,185 (22.30)	143 (21.25)	2,328 (22.23)
High school final diploma	1,626 (16.59)	80 (11.89)	1,706 (16.29)
University degree	3,301 (33.68)	166 (24.67)	3,467 (33.10)
Number of children			
None	2,749 (28.05)	220 (32.69)	2,969 (28.35)
One	2,488 (25.39)	167 (24.81)	2,655 (25.35)
Two	2,905 (29.64)	144 (21.40)	3,049 (29.11)
Three	1,187 (12.11)	81 (12.04)	1,268 (12.11)
Four or more	471 (4.81)	61 (9.06)	532 (5.08)

Figures give the mean value for continuous variables and frequency for discrete variables. Figures in brackets are % of the corresponding subsample.
(a) See definition in subsection 3.2. (b) Intermediate professions includes teachers and related, social and healthcare workers, clergy, civil service middle managers, sales and administrative middle managers, technicians, and supervisors.

Table 1: List of variables and summary statistics

	Deprived neighborhoods			Other neighborhoods			Total
	Mean	Min	Max	Mean	Min	Max	Mean
Public housing units (%)	44.0	0.0	98.5	8.9	0.0	50.4	21.3
Demography							
Foreign household heads (%)	19.5	4.8	56.9	6.3	0.0	30.9	11.0
Lone-parent families (%)	17.7	6.7	33.3	11.4	0.0	28.6	13.6
Education levels							
At most lower secondary edu. (%)	51.5	31.3	69.7	33.8	19.2	62.0	40.1
University degree (%)	13.2	4.0	28.4	30.8	6.0	54.3	24.6
Unemployment							
Unemployed workers (%)	17.9	8.2	37.3	9.2	4.0	21.5	12.3
Unemp. for more than 1 year (%)	9.7	3.2	22.0	4.5	1.3	12.4	6.3
Occupational status							
Blue-collars (%)	33.6	15.5	62.9	15.1	2.8	46.4	21.7
Executives (%)	7.4	0.3	23.5	22.3	0.0	47.6	17.1
Population	2,409	270	5,041	2,438	247	5,730	2,428
Total population		460,100			858,200		1,318,300
Number of neighborhoods		191			349		540

Table 2: Mean characteristics of neighborhoods by type

	Deprived neigh.		Other neighborhoods		Total
	Public housing	Other tenures	Public housing	Other tenures	
Number of individuals	1,348	1,955	659	6,511	10,473
% of total sample	12.9	18.7	6.3	62.2	100.0
Unemployment rate (%)	14.2	6.7	9.9	4.4	6.4
Tenure					
Homeowner	0.0	60.7	0.0	63.7	50.9
Renter in the private sector	0.0	33.1	0.0	32.0	26.1
Renter in the public sector	100.0	0.0	100.0	0.0	19.2
Other renter	0.0	6.1	0.0	4.3	3.8
Individual characteristics					
Age	40.1	42.2	38.9	42.4	41.8
Nationality					
French born in France	59.6	76.9	70.1	86.9	80.4
Fr. born abroad	14.5	11.2	12.6	8.5	10.0
Foreign nation.	25.8	11.9	17.3	4.6	9.5
Education					
No diploma	30.8	17.4	22.6	8.0	13.6
At most lower sec. edu.	16.5	14.7	14.6	11.2	12.7
Vocational training	33.9	32.33	35.8	25.4	28.4
High school final diploma	9.4	12.7	12.9	13.4	12.7
University degree	9.3	22.9	14.1	42.0	32.5
Occupational status					
Farmer or independent w.	2.7	10.5	4.6	12.4	10.3
Executive	3.3	15.4	8.3	33.1	24.4
Intermediate professions	14.8	27.6	22.1	27.6	25.6
Office worker	13.6	10.9	14.0	8.2	9.7
Blue-collar worker	65.5	35.6	51.0	18.7	29.9
Spouse characteristics					
Age	37.1	40.1	36.6	40.6	40.8
Nationality					
French born in France	62.3	78.2	74.6	87.9	81.9
Fr. born abroad	13.2	10.9	9.9	7.6	9.1
Foreign nation.	24.5	10.9	15.5	4.5	9.0
Education					
No diploma	32.5	15.6	22.5	6.7	12.7
At most lower sec. edu.	19.4	17.0	18.2	14.2	15.7
Vocational training	27.6	25.4	27.8	19.6	22.2
High school final diploma	10.2	15.1	15.8	17.9	16.3
University degree	10.3	26.8	15.8	41.4	33.1
Households characteristics					
Number of children					
None	21.7	29.6	24.7	29.7	28.3
One	24.3	25.8	21.8	25.8	25.3
Two	26.6	27.5	29.3	30.1	29.1
Three	14.8	11.8	15.0	11.3	12.1
Four of more	12.6	5.4	9.1	3.0	5.1

Figures give the mean value for continuous variables and frequency for discrete variables.

Table 3: Sample characteristics by residential situation

Dependent variable	**Public housing**		**Deprived neighborhood**		**Unemployment Model 1**		**Model 2**	
Residential variables								
Deprived neigh.							0.0122^{*}	(0.0053)
Public housing			0.3189^{***}	(0.0137)			0.0324^{***}	(0.0071)
Personal characteristics								
Age	-0.0023^{NS}	(0.0042)	0.0001^{NS}	(0.0049)	-0.0081^{***}	(0.0019)	-0.0080^{***}	(0.0019)
Squared age	$2*10\text{-}6^{NS}$	(5*10-5)	$-8*10\text{-}6^{NS}$	(5*10-5)	$9*10\text{-}5^{***}$	(2*10-5)	$9*10\text{-}5^{***}$	(2*10-5)
Nationality								
French nationality	Ref.		Ref.		Ref.		Ref.	
Fr. born abroad	0.0562^{***}	(0.0140)	0.0362^{**}	(0.0171)	0.0226^{***}	(0.0092)	0.0190^{**}	(0.0089)
Foreign nation.	0.0887^{***}	(0.0180)	0.0745^{***}	(0.0228)	0.0610^{***}	(0.0135)	0.0517^{***}	(0.0128)
Education								
No diploma	0.0295^{**}	(0.0153)	0.0554^{***}	(0.0209)	0.0249^{**}	(0.0110)	0.0203^{**}	(0.0106)
$\leq$ lower sec. edu.	0.0239^{NS}	(0.0152)	0.0281^{NS}	(0.0197)	0.0228^{**}	(0.0110)	0.0208^{**}	(0.0107)
Vocational training	0.0108^{NS}	(0.0124)	0.0166^{NS}	(0.0166)	0.0036^{NS}	(0.0082)	0.0029^{NS}	(0.0081)
High school final dip.	Ref.		Ref.		Ref.		Ref.	
University degree	-0.0213^{NS}	(0.0132)	-0.0355^{**}	(0.0170)	-0.00095^{NS}	(0.0083)	0.0012^{NS}	(0.0083)
Occupational status								
Independent w.	-0.0950^{***}	(0.0089)	-0.0579^{***}	(0.0164)	-0.0297^{***}	(0.0062)	-0.0270^{***}	(0.0064)
Executive	-0.0841^{***}	(0.0100)	-0.1050^{***}	(0.0139)	-0.0197^{***}	(0.0064)	-0.0169^{**}	(0.0065)
Intermediate prof.	Ref.		Ref.		Ref.		Ref.	
Office worker	0.0714^{***}	(0.0714)	0.0338^{*}	(0.0181)	-0.0062^{NS}	(0.0080)	-0.0098^{NS}	(0.0075)
Blue-collar worker	0.1023^{***}	(0.0119)	0.0658^{***}	(0.0146)	0.0139^{**}	(0.0071)	0.0062^{NS}	(0.0068)
Characteristics of the spouse								
Age	-0.0118^{***}	(0.0035)						
Squared age	0.0001^{**}	(0.00004)						
Nationality								
French nationality	Ref.		Ref.		Ref.		Ref.	
Fr. born abroad	0.0615^{***}	(0.0147)	0.0818^{***}	(0.0182)	0.0190^{**}	(0.0093)	0.0139^{*}	(0.0088)
Foreign nation.	0.0522^{***}	(0.0172)	0.0674^{***}	(0.0236)	0.0233^{**}	(0.0108)	0.0176^{*}	(0.0102)
Education								
No diploma	0.1428^{***}	(0.0187)	0.1010^{***}	(0.0208)				
$\leq$ lower sec. edu.	0.0917^{***}	(0.0162)	0.0456^{**}	(0.0182)				
Vocational training	0.0674^{***}	(0.0134)	0.0593^{***}	(0.0163)				
High school final dip.	Ref.		Ref.					
University degree	-0.0443^{***}	(0.0112)	0.0021^{NS}	0.0154)				
Number of children								
None	Ref.		Ref.					
One	0.0165^{NS}	(0.0105)	0.00024^{NS}	(0.0133)				
Two	0.0268^{**}	(0.0112)	-0.0209^{NS}	(0.0136)				
Three	0.0592^{***}	(0.0158)	-0.0091^{NS}	(0.0176)				
Four of more	0.1281^{***}	(0.0248)	0.0126^{NS}	(0.0252)				
Log likelihood	-4,033		-5,527		-2,359		-2,338	
Pseudo-R2	0.212		0.153		0.056		0.064	
# Observations	10.473		10.473		10.473		10.473	

Notes: ***, ** and * denote significance at the 1%, 5% and 10% level respectively. Each equation also includes a constant. Marginal effect are (a) for the age variables: $\beta\Phi(\beta\text{X})$ with $\Phi()$ the normal cumulative distribution function and β the vector of estimated coefficients and (b) for each dummy explanatory variable X_k: $\Phi(\beta X_{-k}+\beta_k)-\Phi(\beta X_{-k})$ with X_{-k} the vector of explanatory variables except X_k. X is taken at the sample mean.
Figures in brackets give standard errors of the marginal effects calculated by the delta method.

Table 4: Marginal effects from the three simple probits

	Public housing		Deprived neighborhood		Unemploy-ment	
Intercept	0.412^{NS}	(0.344)	-0.908***	(0.310)	-0.307^{NS}	(0.333)
Residential characteristics						
Public housing	-		0.980***	(0.258)	-0.168^{NS}	(0.249)
Deprived neighborhood	-		-		0.621**	(0.279)
Personal characteristics						
Age	-0.013^{NS}	(0.020)	0.005^{NS}	(0.014)	-0.072***	(0.016)
Squared-age	0.00005^{NS}	(0.0002)	-0.00007^{NS}	(0.0002)	0.0008***	(0.0002)
Nationality						
French born in France	Ref.		Ref.		Ref.	
French born abroad	0.310***	(0.051)	0.097*	(0.051)	0.155**	(0.067)
Foreign nationality	0.380***	(0.062)	0.199***	(0.069)	0.354***	(0.084)
Level of education						
No diploma	0.118*	(0.064)	0.152***	(0.058)	0.146*	(0.090)
At most lower sec. edu.	0.100^{NS}	(0.064)	0.077^{NS}	(0.057)	0.160*	(0.083)
Vocational training	0.041^{NS}	(0.055)	0.048^{NS}	(0.052)	0.020^{NS}	0.075)
High school final diploma	Ref.		Ref.		Ref.	
University diploma	-0.097*	(0.056)	-0.102*	(0.054)	0.014^{NS}	(0.080)
Ocucpational status						
Farmer or independent worker	-0.566***	(0.074)	-0.162***	(0.059)	-0.287***	(0.087)
Executive	-0.441***	(0.060)	-0.312***	(0.046)	-0.138**	(0.072)
Intermediate professions	Ref.		Ref.		Ref.	
Office worker	0.291***	(0.058)	0.084^{NS}	(0.057)	-0.084^{NS}	(0.080)
Blue-collar	0.430***	(0.047)	0.169***	(0.053)	0.061^{NS}	(0.070)
Characteristics of the spouse						
Age	-0.054***	(0.018)	-		-	-
Squared-age	0.0004**	(0.0002)	-		-	-
Nationality						
French born in France	Ref.		Ref.		Ref.	
French born abroad	0.253***	(0.061)	0.221***	(0.056)	0.095^{NS}	(0.069)
Foreign nationality	0.221***	(0.069)	0.177***	(0.068)	0.132^{NS}	(0.088)
Level of education						
No diploma	0.534***	(0.064)	0.266***	(0.067)	-	-
At most lower sec. edu.	0.367***	(0.061)	0.125**	(0.052)	-	-
Vocational training	0.284***	(0.054)	0.161***	(0.048)	-	-
High school final diploma	Ref.		Ref.		-	-
University diploma	-0.215***	(0.059)	0.016^{NS}	(0.048)	-	-
Household characteristics						
Number of children						
None	Ref.		Ref.		-	-
One	0.079^{NS}	(0.048)	-0.010^{NS}	(0.036)	-	-
Two	0.127**	(0.052)	-0,082*	(0.044)	-	-
Three	0.255***	(0.067)	$-0{,}047^{NS}$	(0.055)	-	-
Four or more	0.487***	(0.091)	0.010^{NS}	(0.087)	-	-
Correlation of residuals unemp./deprived neigh. ρ_{12}			-0.303**(0.154)			
Correlation of residuals unemp./public housing ρ_{13}			0.152^{NS}(0.119)			
Correlation of residuals deprived neigh./pub. housing ρ_{23}			-0.073^{NS}(0.144)			
Log likelihood			-11,897			
LR test ($\rho_{12} = \rho_{23} = \rho_{23} = 0$)			3.2112			
Pseudo-R^2			0.159			
Number of observations			10,473			

Notes: ***, ** and * denote significance at the 1%, 5% and 10% level respectively.
Estimation by simulated maximum likelihood with 600 draws.
Figures in brackets give robust standard errors corrected for dependencies within neighborhood.

Table 5: Results of the three probits system

	Simple probit	Seemingly unrelated probits[(a)]		Simultaneous probits[(b)]			
Public housing in neighborhood eq.	-	No	Yes	No	Yes	Yes	Yes
Dropped variable in unemployment eq.	None (1)	None (2)	None (3)	None (4)	None (5)	Occupational status (6)	Spouse's nationality (7)
Deprived neighborhood							
Coefficient	0.173***	-	-	0.631**	0.788***	0.869	0.805
	(0.042)			(0.270)	(0.116)	(0.107)	(0.114)
Correlation	-	0.0992***	0.0604**	-0.274*	-0.390***	-0.426***	-0.396***
		(0.0255)	(0.0257)	(0.155)	(0.065)	(0.060)	(0.064)
Marginal effect	0.0213	0.0205	0.0125	0.0194	0.0118	0.0136	0.0125
	(0.00593)	(0.00602)	(0.00587)	(0.00700)	(0.00626)	(0.00630)	(0.00624)
	[0.00591]	[0.00596]	[0.00563]	[0.00705]	[0.00611]	[0.00627]	[0.00612]
Public housing							
Marginal effect	-	-	-	-	0.0315	0.0386	0.0327
					(0.00632)	(0.00658)	(0.00628)
					[0.00677]	[0.00709]	[0.00676]
Log likelihood	-2351.34	-8164.32	-7883.84	-8162.24	-7865.29	-7932.85	-7881.14
Pseudo-R2	5.88	9.55	12.66	9.57	12.86	12.11	12.68

Notes: Control variables include, unless otherwise mentioned, individual's age, nationality, education level, occupational status, spouse' age, nationality, education level and number of children. (a) Neighborhood variable is not in the unemployment equation. (b) Neighborhood variable is in the unemployment equation. Mean marginal effects are calculated as $1/N\sum(P(y_{i1}=1|y_{i2}=1)-P(y_{i1}=1|y_{i2}=0))$ for neighborhood type and $1/N\sum(P(y_{i1}=1,y_{i2}=1|y_{i3})+P(y_{i1}=1,y_{i2}=0|y_{i3})-P(y_{i1}=0,y_{i2}=1|y_{i3})-P(y_{i1}=0,y_{i2}=0|y_{i3}))$ for public housing. Predicted probabilities are calculated for each individual vector of characteristics. Standard errors of marginal effects are calculated by the delta method. Standard errors between square brackets are corrected for dependencies within neighborhoods.

Table 6: Effect of public housing and neighborhood on unemployment probability

	Deprived neighborhoods			Other neighborhoods			Total		
	Public housing	Other tenures	All tenures	Public housing	Other tenures	All tenures	Public housing	Other tenures	All tenures
Number of observations	1348	1955	3303	659	6511	7170	2007	8466	10473
Observed unemployment rate	14.17	6.75	9.78	9.86	4.38	4.88	12.76	4.93	6.43
Predicted unemp. prob. conditional on location									
in a deprived neighborhood	14.12	6.85	9.82	12.04	5.02	5.67	13.44	5.44	6.98
in another type of neigh.	11.04	5.62	7.83	9.21	4.42	4.86	10.44	4.70	5.80

Table 7: Observed and predicted unemployment rates

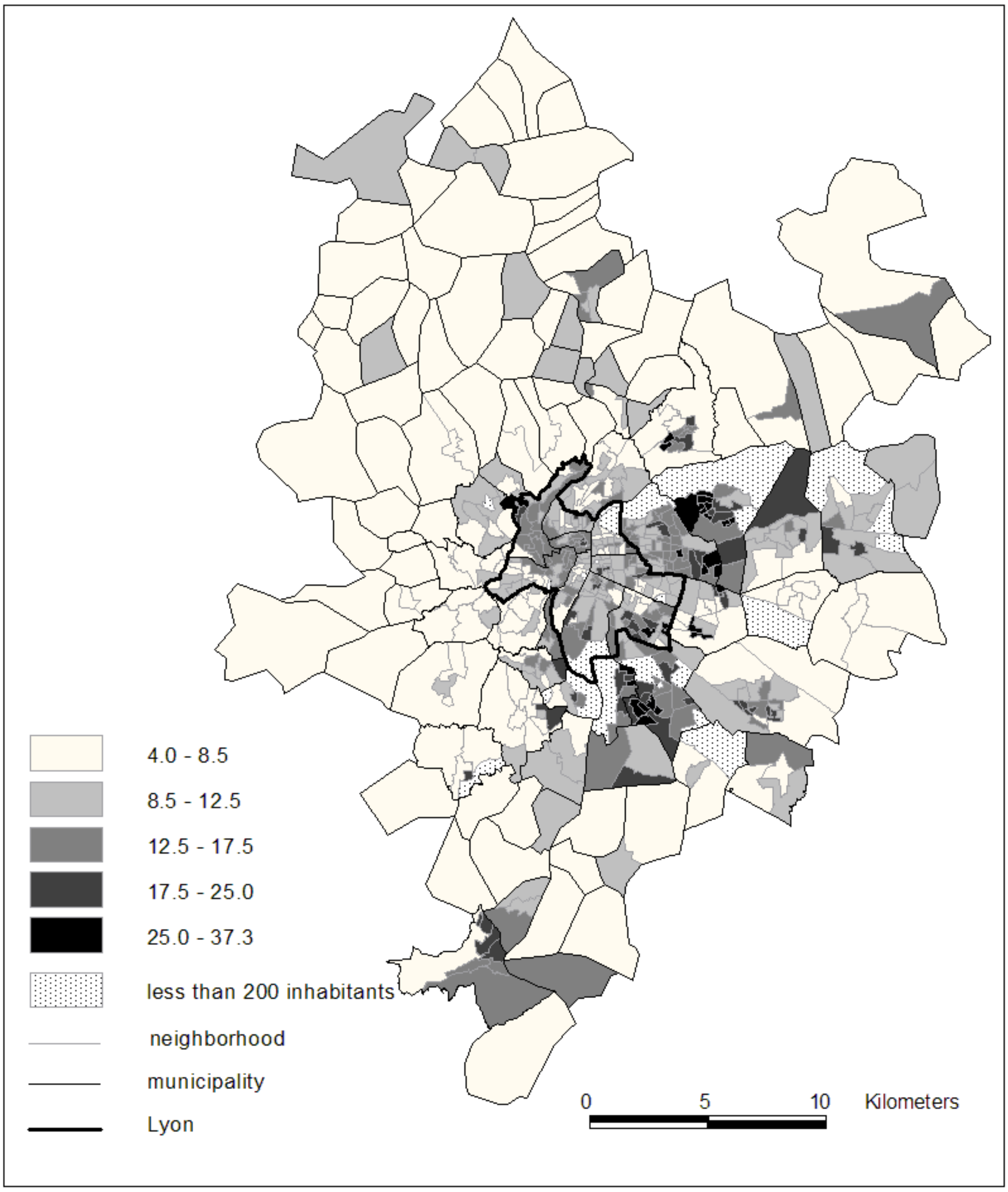

Figure 1: Percentage of unemployed workers within labor-force participants

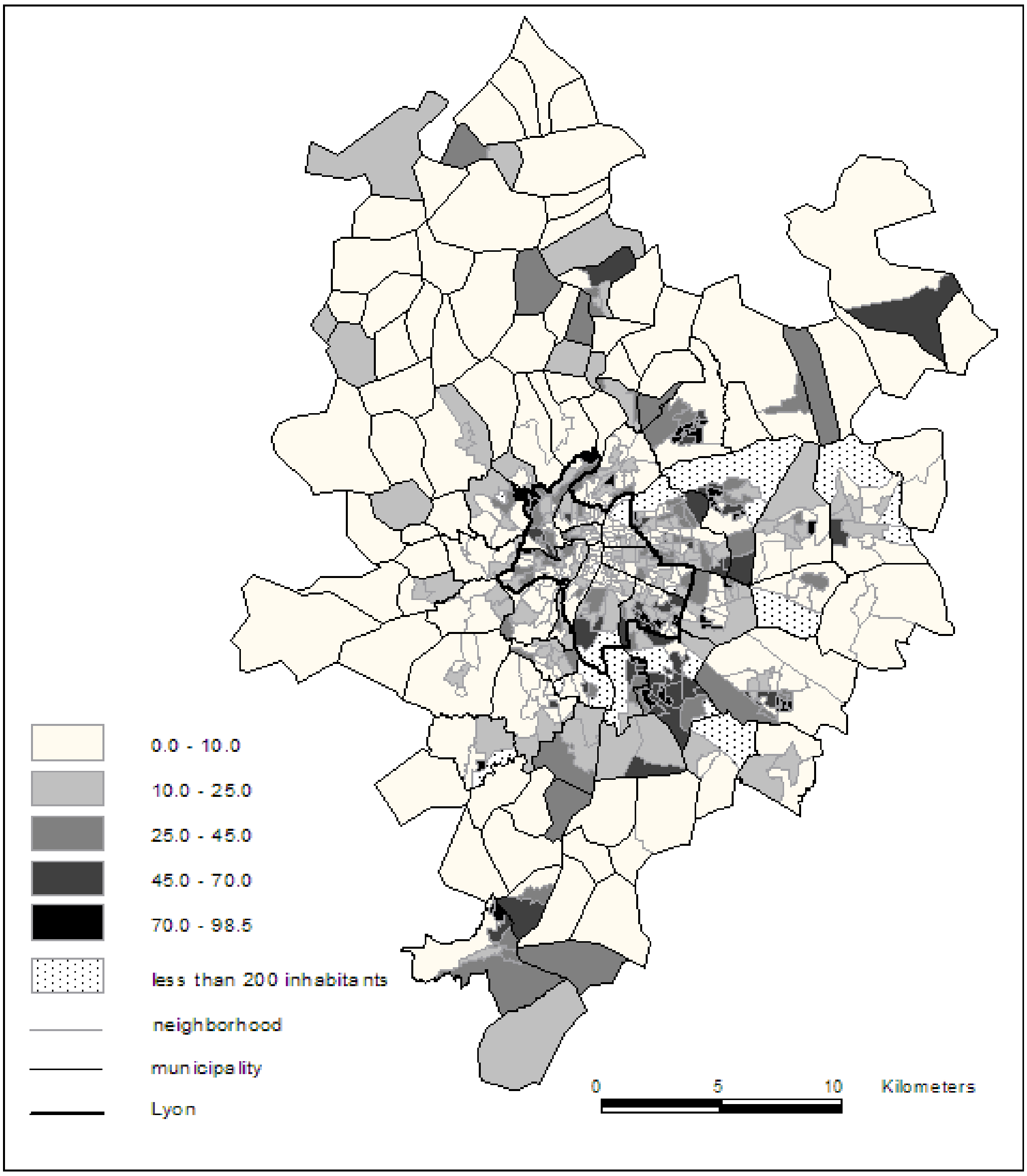

Figure 2: Percentage of housing units in the public renting sector

Figure 3: Location of deprived neighborhoods