Income Segregation and Suburbanization in France : a discrete choice approach

Florence Goffette-Nagot, Yves Schaeffer

GROUPE D'ANALYSE ET DE THÉORIE ÉCONOMIQUE LYON - ST ÉTIENNE

W P 1112

Income Segregation and Suburbanization in France: a discrete choice approach

Florence Goffette-Nagot, Yves Schaeffer

Mars 2011

Documents de travail | Working Papers

GATE Groupe d'Analyse et de Théorie Économique Lyon-St Étienne

93, chemin des Mouilles 69130 Ecully – France
Tel. +33 (0)4 72 86 60 60
Fax +33 (0)4 72 86 60 90

6, rue Basse des Rives 42023 Saint-Etienne cedex 02 – France
Tel. +33 (0)4 77 42 19 60
Fax. +33 (0)4 77 42 19 50

Messagerie électronique / Email : gate@gate.cnrs.fr
Téléchargement / Download : http://www.gate.cnrs.fr – Publications / Working Papers

Income Segregation and Suburbanization in France: a discrete choice approach.

Florence Goffette-Nagot* [+],
Yves Schaeffer**

Abstract:

This paper focuses on residential sorting by social and ethnic status in large French urban areas. Our objective is to assess the relative importance of two major determinants of segregation stressed by the economic literature (Bartolome and Ross, 2003; Brueckner *et al.*, 1999): (i) "Alonso sorting over space", due to the trade-off between land consumption and accessibility to the central city and (ii) "Tiebout sorting over jurisdictions", due to the taste for local public goods and by extension for all kinds of local public amenities (e.g. neighborhood externalities). Our methodology draws on Schmidheiny (2006). First, a conditional logit model is estimated for each urban area, in which moving households are assumed to sort based on jurisdiction distance to the central city and jurisdiction mean of households' incomes (as a proxy for the level of public amenities). Second, our estimation results are used to simulate the counterfactual residential patterns that would prevail if, alternatively, one or the other of these mechanisms were inactive (setting the coefficients of the corresponding variables to zero). The contribution of each mechanism to the observed social and ethnic segregation is finally appreciated by comparing the values of dissimilarity indexes computed on the basis of the counterfactual households distributions and on the observed households distribution. "Tiebout-sorting" emerges as the primary cause of social segregation among wage-earning households. On the contrary, "Alonso-sorting" appears to be the main driver of segregation between economically active and inactive households, as well as between French-citizen and Foreign-citizen households.

JEL classification: R21, R23, R1.
Keywords: Income segregation, Ethnic segregation, Suburbanization, Local amenities, Migrations, Conditional logit, French urban areas.

* University of Lyon, CNRS, GATE
** CEMAGREF de Grenoble, Saint-Martin-d'Hères cedex, F-38402.
[+] Corresponding author, GATE-CNRS, UMR 5824, 93 chemin des Mouilles, F-69130 Ecully, France; tel: 33+4 72 86 60 78; fax: 33+4 72 86 60 90; mail: goffette-nagot@gate.cnrs.fr
We thank Sylvie Charlot for her help at various stages of this work.

1. Introduction

Residential segregation by income and ethnicity is a major feature of contemporary Western cities (see e.g. Card et al., 2008; Wheeler, LaJeunesse, 2008) and is generally considered to be undesirable. It may be the source of poverty traps due for instance to peer effects in education and lack of role models. It may weaken social cohesion and redistribution mechanisms (Bjorvatn and Capellen, 2003). In short, segregation may be the source of short-term as well as long-term inequalities and social tensions.

In the United States, at least for fifty years, economists have been trying to explain residential sorting by income and race in American metropolitan areas. Two main strands of literature are competing to this purpose. The local public finance literature and its numerous extensions, rooted in Tiebout's (1956) model of fiscal competition, suggests a "sorting across jurisdictions" that indirectly contributes to the understanding of urban configurations. The urban economic literature, originally based on the Alonso (1964), Mills (1967) and Muth (1969) model of a monocentric city, suggests a "sorting over space". To improve the understanding of income sorting, an integrated modelling approach is clearly required (Nechyba and Walsh, 2004) and has only been intended recently (Brueckner, Thisse and Zenou, 1999; Bartolome and Ross, 2003, 2004, 2007).

The aim of this paper is to provide a contribution to this integated approch on the basis of an empirical analysis of residential choices and their impact on urban segregation. Our objective is to assess empirically the relative contribution of Tiebout-sorting and Alonso-sorting mechanisms to social and ethnic segregation in 37 large French urban areas. Similar to what is observed in other industrialized countries, French urban areas, that host in 1999 77% of the French population, have been continuously sprawling since the sixties and residential segregation by income and ethnicity is commonly acknowledged as a striking issue (Pan Ké Shon, 2009; Gobillon and Selod, 2007; Préteceille, 2006). In this context, improving our understanding of the causes and consequences of households location decisions within French urban areas is obviously of great political relevance, in particular because political answers to segregation are not the same whether it is due to Alonso-like or Tiebout-like mechanisms.

This inquiry requires the empirical analysis of residential location choices, in terms of distance to the central city and local economic conditions, as a function of households' income. The literature interested in this question mainly relies on housing price hedonic estimations aimed at estimating willingness to pay for accessibility and local amenities. Nevertheless, some recent works use discrete choice modeling. For instance, Ioannides and Zanella (2008) analyze the demand for neighborhood "quality" in contrasting location choices of households with and without children. Schmidheiny (2006) focuses on the impact of local progressive income taxes on location choices in Switzerland.

We follow the methodology proposed by Schmidheiny (2006). The approach consists in two steps. In a first step, we estimate on each of the 37 urban areas a conditional logit model of residential location choice, in which households are supposed to select a community depending on their income interacted with distance to the central city and mean income of potential destinations. The specific advantage associated with relocating close to one's former community is also taken into account through a specific former location dummy. This advantage is due to better information about housing supply, lower moving costs, access to previously built local social networks, etc. Because considering all the communities of an urban area as potential destinations in the location choice set would give far too many alternatives in the logit model, we built a smaller location choice set by grouping jurisdictions according to their mean household income and distance to the central city in 1990. As households' incomes are not directly available in our data, occupational status is used as a proxy. Ethnicity is also accounted for. Our sample encompasses all migrations that occurred between 1990 and 1999 within 37 urban areas of more than 200,000 inhabitants and contains 210,611 households.

In a second step, social and ethnic dissimilarity indexes of migrants are calculated for each urban area either with the observed household distribution or with counterfactual household distributions predicted based on the estimated conditional logit model. The relative importance of the two "segregation channels" stressed above – i.e. the choice of distance to the central city and of neighborhood economic conditions – can be disentangled by predicting distributions of households between types using only the coefficients of our model associated with either the choice of distance to the central city or the choice of local economic conditions. As, according to the third segregation channel in our model, moving households are *ceteris paribus* more likely to choose to relocate nearby their former location rather than elsewhere, the segregated pattern prevailing in 1990 is partly translated to 1999.

Our main results are the following. Going beyond the standard urban model to take into account income-related amenities is strongly justified by our results: segregation among the economically active social groups appears to be mainly driven by the income channel, especially for the most affluent social group. Nevertheless, segregation between economically active and inactive households is mainly explained by distance, as predicted by the standard urban economic model. Regarding ethnic segregation, the distance channel dominates the income one, but this is likely to be due to a non-market spatial structure, i.e. the centralization of the public housing supply. More important, both of them are dominated by the inertia of the previously prevailing segregated pattern.

The remaining of the paper is organized as follows. Section 2 presents the background and the hypotheses of our empirical analysis. Section 3 presents the empirical model. Section 4 describes the data, the sample definition and a few descriptive statistics. Results are presented in the last section.

2. Background, theoretical model and overview of the empirical approach

2.1. Background

In the local public finance literature, sorting across jurisdictions is driven by households' tastes for public amenities. In the original Tiebout's model, public amenities are the level of public goods produced by jurisdictions. The level of public good is a normal good, which means that its demand increases with income. Households with different income demand different public service levels and thus choose different jurisdictions, each household choosing the jurisdiction which provides his preferred public service level. Assuming that jurisdictions are formed on a featureless plain, jurisdictional boundaries may be freely adjusted, the public good is financed by a head-tax and a households' income does not depend on the jurisdiction in which it resides, income stratification across jurisdictions should be perfect at the long term equilibrium (Tiebout, 1956). This jurisdictional sorting is a potential source of suburbanization, as people move to suburban jurisdictions to get their desired level of public services.

Subsequently, the local public finance framework has been extended to other kinds of public amenities, produced by local governments or directly produced by the population, either in jurisdictions or in neighborhoods. Some extensions specifically aim at explaining suburbanization, income and racial sorting, i.e. the well-known US urban pattern displaying rich households in suburbs and poor households in central cities. The implied public amenities can be divided into factors that pull rich white households into the suburbs and factors that push them out of city centers (Nechyba and Walsh, 2004). From the pull-side, sprawl can be explained by the possibility of implementing zoning regulations in suburban jurisdictions to exclude those supposed to bring with them negative fiscal externalities (free riding on tax payments, etc.) or peer externalities (crime rates, school qualities, etc.). In turn, if zoning makes high density impossible in the suburban jurisdictions closer to the city center, this will increase the development of sprawl on the city edge. The push-side corresponds to the "flight from blight" hypothesis (Jackson, 1985; Mills and Lubuele, 1997): rich households move from the central city to suburban jurisdictions in order to flee the negative externalities produced by poor households (high crime rates, low school quality, general fiscal distress, etc). These push-factors has to be related to Schelling's (1971) model of segregation, which features externalities linked to the neighborhood population composition. This model shows that even mild preferences for having similar neighbors may lead to sharp segregation patterns. More generally, literature on neighborhood externalities is a particularly vivid field of research (Ioannides and Zabel, 2008; Ioannides and Zanella; 2008). It appears that the demand for the "quality" of the neighborhood social composition – in terms of providing positive externalities for access to employment and human capital accumulation for instance – is linked to the local income level and generates income segregation (Benabou, 1993; Durlauf, 2004).

In the urban economic literature, sorting over space is driven by households' tastes for land and accessibility to jobs. In the standard monocentric city model, jobs are all in the central business district (CBD), so that a location closer to the CBD has the advantage of lower commuting costs but – due to land market competition – the inconvenient of higher land prices. Both land and accessibility to jobs are normal goods. Rich households have a high land consumption and thus are more strongly attracted than the poor by low land prices. However, they also have a high opportunity cost of time, so that they value more accessibility to jobs than the poor. Sorting between rich and poor households depends on the value of the income elasticity of commuting costs relative to the value of the income elasticity of the demand for land. If the former is higher than the latter, then the accessibility effect dominates and the rich households outbid the poor households for locations closer to the CBD. If, on the contrary, the former is lower than the latter, then the land consumption effect dominates and the rich households tend to live further to the CDB than the poor households. This case is supposed to explain the commonly observed urban pattern in the United States. In both cases, the model displays a monotonic relationship between households' income and distance from the CBD (Wheaton, 1977) and produces income segregation.

In addition, a major strength of the monocentric model is that it allows to identify the primary causes for urban sprawl in the twentieth century. Firstly, households' incomes rose importantly during this period (Margo, 1992). If the income elasticity of demand for land is greater than the income elasticity of commuting costs, as it is traditionally assumed, rising income is associated with rising land consumption and thus with decreased density in the urban area. Secondly, transportation costs have dramatically decreased during this period, especially due to the development of car ownership (Glaeser and Kahn, 2003). The fall of commuting costs generates an additional income effect and above all weakens the centripetal force of the model: the amount of land consumed increases and the edge of the city expands.

A model integrating these two strands of the literature must take into account simultaneously households' tastes for public amenities, land consumption and accessibility to jobs. This is the case in the model of Brueckner, Thisse and Zenou (1999) (BTZ in the following). Income sorting is due to the conventional forces of the monocentric city model, but also to the households' tastes for urban amenities. Urban amenities are of three kinds: natural, historical and modern. "*While natural and historical amenities are largely exogenous,* modern amenities *are endogenous, with their levels depending on the current economic conditions in a neighborhood, especially the local income levels. Such amenities might include restaurants, theaters, and modern public facilities such as swimming pools and tennis courts*." (Brueckner, Thisse and Zenou, 1999, p.94). The authors make the traditional assumption that the conventional location forces drive the rich to the suburbs and the poor to the city center (i.e. the income elasticity of commuting costs is lower than the income elasticity of land demand). They show that if the center's exogenous amenity advantage is sufficiently large, the equilibrium outcome can be reversed: the rich households outbid the poor household for locations in the city center. This additional location force could be at the origin of the differences between US and European cities. Europe's longer history and differences in government investment in central city infrastructures is likely to explain the differences in the spatial pattern of exogenous amenities. In addition, modern endogenous amenities introduce the possibility of multiple equilibria: they make the existing location of the rich attractive to them wherever it might be. This possibility may help explain the variety of location patterns by income observed in reality.

The BTZ model constitutes a first step toward the integration of Tiebout-sorting and Alonso-sorting models but it unfortunately does not solve their main caveats. As stressed by Bartolome and Ross (2003), their predictions are hardly supported by facts: income sorting between jurisdictions is far from perfect; empirical estimations suggest that the income elasticity of commuting costs is greater than the income elasticity of land demand (Wheaton, 1977; Glaeser, Kahn and Rappaport, 2000); and the relationship between households' income and distance from the CBD does not appear to be monotonic in general (Glaeser, Kahn and Rappaport, 2000). Another integrated model that better matches empirical facts was developed by Bartolome and Ross (2003, 2004, 2007), hence named the BR model. This is a model of a monocentric city with jurisdictions providing public services. Households sort over jurisdictions based on conventional location forces and public service levels. The level of public services in a jurisdiction is decided by majority voting and rich households must be the majority to be able to implement their preferred level of local public goods. When it is not the case, they tend to vote with their feet for jurisdictions with high public service levels. Income elasticity of commuting costs is now assumed to be greater than the income elasticity of land demand, but rich households may nevertheless settle in the suburb because of the higher public good level there. Therefore, the model is able to predict complete as well as partial income sorting between jurisdictions and across space, and the relationship between households' income and distance from the CBD may be, or not, monotonic, depending on households tastes for local public goods and of the unitary commuting time.

2.2. Segregation mechanisms in French urban areas

In this article, our objective is to assess empirically the relative contribution of Tiebout-sorting and Alonso-sorting mechanisms to social and ethnic segregation in 37 large French urban areas. Our empirical framework relies upon a few important assumptions. Put in a nutshell, we assume that within urban areas, households sort across jurisdictions primarily on the basis on jurisdictions' distances to the central city and jurisdictions' means of households' incomes. Due to data constraints, we however do not focus on households sorting by income but household sorting by "social status".

More precisely, we first assume that these urban areas are essentially monocentric in terms of employment. Considering French urban areas in 1999, we know that central cities contain 27.2% of the total employed population but 41.5% of total jobs (Julien, 2001). Secondly, as in the BTZ and BR models, we assume that households sort between jurisdictions both on the basis on their preferred accessibility to the central city (resulting from their tastes for land consumption and accessibility to jobs) and their preferred level of public amenities. The public amenities supposed here to be relevant regarding households' location decisions are the jurisdiction's *endogenous modern amenities*, including notably the public service level and the external effects of population composition. Thirdly, we make important assumptions about proxies. Modern amenities are assumed to be well proxied by the jurisdictions' mean of households' incomes: high income jurisdictions offer high quality public and private services, positive peer externalities, etc. Accessibility to the central city is assumed to be well proxied by the jurisdictions' bird's eye distances to the central city.

2.3. Methodology overview

Our methodology draws on Schmidheiny (2006), who studies the impact of local progressive income taxes on households' location choices and income sorting in the city of Basel, Switzerland. This methodology is in two steps. The first step consists in the estimation of a conditional logit model of households' location decisions, one for each the 37 urban areas considered in our study. For each urban area, we consider a 1/20 sample of households that *moved within* the urban area *between 1990 and 1999* (all urban areas together, there are 210,611 households in our sample). The choice set of each moving household is defined on the basis on all jurisdictions included in the urban area in 1999. However, in urban areas of more than 200,000 inhabitants, considering all jurisdictions as potential destinations would give far too many alternatives for estimating a conditional logit model and hence we group jurisdictions of each urban area according to their distances to the central city and to their means of households' incomes in 1990. The choice set finally contains 17 classes of jurisdictions.

The main explanatory variables of households' choices between these classes are interactions between households' characteristics and characteristics of the jurisdiction classes. On one hand, we interact either the social or the ethnic status of households with the classes' weighted mean of jurisdictions' distances to the central city (the weight corresponds to the number of houses of each jurisdiction). On the other hand, we interact either the social or the ethnic status of households with the location class mean of households' incomes. Importantly, the households' incomes taken into account are measured in 1990, i.e. before migrations took place. Indeed, as stressed by the BZT model, modern amenities are endogenous: they are both a cause and consequence of the location patterns of different income groups. There is a simultaneity issue that we try to limit in our econometric model by taking the past value of location income.

The second step relies on predicted location choices and segregation measures. For each urban area, we can easily measure the *observed level of segregation* by social and ethnic status across classes of jurisdictions of moving households, based on the computation of dissimilarity indexes. Then, for each urban area, we can use the coefficients obtained in our first step estimation to predict the counterfactual location pattern of these households and measure the *predicted level of segregation* by social and ethnic status across classes. By comparing the observed and predicted segregation levels, we are able to assess the prediction power of our model, thus the joint explanatory power of Tiebout-sorting and Alonso-sorting mechanisms. Furthermore, the relative importance of these two "segregation channels" can be disentangled. Indeed, for each urban area, we can select the coefficients associated with explanatory variables corresponding to one segregation channels only (interactions with classes' distances only or with classes' mean incomes only), setting all the other coefficients to zero, and predict the corresponding counterfactual location patterns of households across classes. We can thus measure and compare the "*Tiebout-sorting*" *predicted level of segregation* and the "*Alonso-sorting*" *predicted level of segregation*. A third "segregation channel" is also included in our model, namely, the specific advantages of relocating nearby one's former location. These advantages may be due to better information about housing supply, lower moving costs, access to previously built local social networks, etc. We expect that moving households are more likely to choose to relocate nearby their former location rather than elsewhere, and that the segregated pattern prevailing in 1990 is partly translated to 1999 due to this additional effect.

3. Data and sample definition

Before to expose the econometric model in detail in section 5, we present in this section our data and the variables used in the estimations. We also give a few descriptive statistics concerning the location classes that constitute the explained variable of our model, as well as the sample statistics.

3.1. Data

Our empirical investigation is essentially based on the 1999 French Population Census (produced by INSEE, the French National Institute for Statistics and Economic Study). A 1/20th sample is drawn from the Census, in which detailed characteristics of households and their members are available. Residential location of these households at the Commune level is known in 1999 and 1990, allowing to trace households' moves. Unfortunately, households' incomes are not. We will therefore make use of the occupational status of the household's reference person as a proxy for household's income. We will however give statistics in the following in order to show that using occupational status categories is indeed useful to measure income segregation.

As to location characteristics, our analysis requires information on jurisdiction average household income as well as jurisdiction's distances to the employment center of the urban area. The mean households' incomes of each Commune (i.e. French jurisdiction) comes from the French Tax Authorities (INSEE/DGI). The French National Geographic Institute (IGN) provides the geographic coordinates of each Commune's town hall, allowing to compute the straight-line distance between the Commune's townhall and the center of the urban area. Other characteristics of locations are taken from the 1999 Population Census aggregated at the Commune level.

In the descriptive statistics, we also make use of 1993 local housing tax rates provided by the French Authority for Local Governments (INSEE/DGCL). We also own data from the French and the Parisian notary societies (PERVAL, Chambre des notaires de Paris) giving mean housing prices in 2002.

3.2. Choice of urban areas and definition of the location choice set

Our study is aimed at explaining location choices *within* urban areas, that is, in labor-market areas within which households are assumed to choose their residential location considering their workplace as given. Urban areas are defined in France based on commuting flows as measured in 1999 from Census data. An urban area comprises a city center and inner suburbs divided into several jurisdictions and a ring of outer suburbs composed of jurisdictions that do not belong to the urban agglomeration but which are tightly tied to it by commuting flows.[1] There were 354 urban areas in France in 1999 with a total of 45 millions inhabitants representing 77% of the French population. Because sorting is a more striking issue the bigger the urban area, we focus on urban areas of more than 200,000 inhabitants in 1999.

Within each urban area, considering each jurisdiction separately as a potential destination would give far too many alternatives in the location choice model.[2] Consequently, within each urban area, we form groups of jurisdictions that will be considered as alternatives in the choice model. These groups are aimed at being as homogenous as possible with respect to the two main characteristics of our analysis: distance to the city center and household mean income. Therefore, in each urban area, we classify jurisdictions: (i) first of all, according to their position in the urban area: city center, inner suburbs, outer suburbs; (ii) then, in each of these preliminary groups - -except for the city center that has only one jurisdiction-, according to their position relatively to the group median distance to the city center: close and distant jurisdictions; (iii) and eventually, according to quartiles of the jurisdictions' average household income. Therefore, we obtain 17 groups of jurisdictions that are considered by moving households as potential destinations.

Note that this classification imposes to have at least eight communes in the inner suburbs and eight in the outer suburbs to be able to define a choice set of 17 types. Applying this criteria to urban areas of more than 200,000 inhabitants yields 37 urban areas (out of 41 French urban areas of this size).[3]

We now describe the 17 types of location thus defined. In particular, we want to stress the correlation between the types' average household income and other characteristics, namely: the percentage of foreign population and local taxes. For the sake of clarity, we present the descriptive statistics after making an additional pooling of jurisdictions: at the last step of the

[1] In France, an urban unit (*unité urbaine* in French), is a set of communities, 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. Each urban area is built around an urban unit having at least 5 000 jobs.

[2] For instance, Lyon urban area has as many as 296 communities. Toulouse urban area encompasses 342 communities.

[3] The excluded urban areas are those of Brest, Reims, Limoges and Nîmes.

classification, we group the jurisdictions which belong to the second and third quartiles of average household income. Thus, we present statistics only for 13 types in Table 1. As urban areas can have different average household income and different spatial ranges, several statistics in Table 1 are given relatively to the urban area average characteristics.

Table 1: Characteristics of the 17 location classes relatively to the urban area average

			Central jurisdiction	
City center	income [a]		93.2	
	distance [b]		0	
	% foreign		8.0	
	housing tax rate [b]		18.20	
	housing prices		102.3	
		Low-income jurisdictions	Medium-income jurisdictions	High-income jurisdictions
Close inner suburbs	income [a]	84.4	104.3	139.1
	distance [a]	42.2	41.7	42.8
	% foreign [b]	7.5	4.1	3.0
	housing tax rate [b]	14.16	12.95	10.84
	housing prices [a]	90.0	107.8	124.5
Distant Inner suburbs	income [a]	85.0	101.9	129.8
	distance [a]	97.6	94.9	84.8
	% foreign [b]	6.6	4.4	3.2
	housing tax rate [b]	13.16	12.16	9.98
	housing prices [a]	90.3	105.6	122.6
Close Outer Suburbs	income [a]	82.0	99.4	127.4
	distance [a]	124.2	110.2	92.8
	% foreign [b]	3.2	2.4	2.2
	housing tax rate [b]	10.49	9.88	9.70
	housing prices [a]	85.8	101.3	113.8
Distant outer suburbs	income [a]	69.1	83.0	101.3
	distance [a]	204.5	184.3	179.9
	% foreign [b]	2.5	2.4	2.6
	housing tax rate [b]	9.03	9.51	9.26
	housing prices [a]	74.2	84.0	97.7

[a] These figures read as follows: on average over the 37 urban areas of our sample, the mean of households' incomes in the close inner suburbs of an urban area equals 84.4% of the average value of the same statistic computed over all types of the same urban area.
[b] These are to be read more directly: on average over the 37 urban areas of our sample, 7.5% of all households are headed by a foreign person in close inner suburbs and the housing tax rate equals 14.16%.

Broadly speaking, except in the close inner suburbs, high-income jurisdictions appear to be closer to the city center than medium-income jurisdictions, which in turn are closer than low-income jurisdictions. Within distant inner-suburbs for instance, low-income jurisdictions are at 0.98 from the average distance, medium-income jurisdictions at 0.95 and high-income jurisdictions at 0.85. Housing prices are decreasing with distance from the city center and increasing with the jurisdiction's average household income, ranging from 74% of the urban area mean in low-income jurisdictions of distant outer suburbs to 124% in high-income jurisdictions of close inner suburbs. The proportion of foreign citizen decreases with distance and with the jurisdiction's average household income. Housing tax rates follow the same pattern, with higher tax rates for lower income jurisdictions. Note that prices used in these statistics are those of 2002, so that we cannot use them in the estimations due to simultaneity biases. Tax rates are given in 1993 and the percentage of foreign citizen in 1990, but both create multicollinearity issues when included simultaneously with distances and mean incomes of types.

3.3. Sample definition and descriptive statistics

Following Schmidheiny (2006), we focus on the behavior of households that *moved within* an urban area, ignoring immobile households and other moving households. Indeed, we suppose that local migrations (i.e. *within* an urban area) are mainly driven by motivations related to housing (broadly understood as including access to employment and amenities), whereas long distant migrations are essentially linked with the search of study and employment opportunities, that may be associated with different location behaviors. Furthermore, people newly arrived in an urban area may not know well the characteristics of locations available in this urban area. Defined on these criteria, the total estimation sample (summing the 37 urban areas) contains 210,611 households that moved between 1990 and 1999 within urban areas. Some commentaries and simple statistics are given to justify our choices and present the sample.

Considering the 37 urban areas of our study, 51.5% of households moved between 1990 and 1999. Among the movers, 67.7% were living in the same urban area in 1990. Descriptive statistics seem to corroborate the idea that short distance moves are mainly driven by residential considerations whereas moves to a new urban area are employment or study related. Indeed, households that moved, but *not* within an urban area, have five times more often a student as reference person of the household (and logically display a lower mean age and mean size) (see Table 2). The reference person is also more often a foreign citizen, what may be partly due to the arrival of new immigrants looking for economic opportunities in French urban areas between 1990 and 1999.

Table 2: Characteristics of the population and of mobile households

	Household with student as RP[1] (%)	Househ. with foreign citizen as RP (%)	Mean age of the RP	Mean household size	Household in a new location class in 1999 (%)
Whole population	5.33	7.83	50	2.39	
Migrants *within* urban areas	4.03	8.39	42	2.55	45.79
Others migrants	20.23	10.21	36	2.15	73.65[2]

[1] RP: Reference Person
[2] Calculated for households that were living in another urban area of our selection in 1990.

We define three "ethnic groups" and six "social groups" based on the characteristics of the reference person of the household.[4] The former are foreign citizens, French citizens born

[4] The reference person of the household is always the man in households where a man and a woman are in couple and either a man or a woman in all other cases.

abroad and French citizens born in France. The latter are built on the basis of occupational status as follows: (i) executives and high intellectual professions; (ii) white-collars in mid management positions; (iii) white-collar subordinates; (iv) blue-collar workers; (v) independent workers; (vi) economically inactive people, i.e. retirees and people who never worked. The sizes of these groups are given in Table 3. As we would like to be able to interpret the spatial segregation of these groups in terms of income segregation, Table 3 also displays the results of a simple linear regression aimed at predicting the average income of each category. This regression is performed on data taken from the French Housing Survey 1996 (see Appendix 1). Table 3 shows that the classification into these four groups translates in a hierarchy in terms of mean annual income by consumption unit. Therefore, measuring segregation on the basis of groups (i) to (iv) provides us information about income segregation. On the contrary, we know that in groups (v) and (vi), incomes and professional situations are much dispersed (for instance, independent workers can be farmers, craftsmen or big entrepreneurs).

Table 3: Sample distribution by occupational status and nationality of the reference person and estimated mean annual income by consumption unit (in French Francs)

	French		Foreign nationality		All	
	N[1]	Mean[2]	N	Mean	N	Mean
Executives	31,046	152,527	1,041	127,639	32,087	151,720
Intermediate professions	40,779	102,776	1,687	81,454	42,466	101,929
White-collar workers	34,673	73,546	2,613	57,042	37,286	72,390
Blue-collar workers	40,616	69,768	7,621	56,116	48,237	67,611
All	147,114	97,273	12,962	65,345	160,076	94,688

[1] The number of observations corresponds to the estimation sample.
[2] The mean income is estimated based on French Housing Survey data.

4. Empirical model of location choice and measure of residential segregation

4.1. A conditional logit model of location choice

We consider a random utility model, according to which utility of a household *i* in a location *j* is the sum of a deterministic and a random part:

$$V^*_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

where V_{ij} is the deterministic part representing the influence of observed household and jurisdiction characteristics and ε_{ij} is the idiosyncratic random term specific to household *i* and jurisdiction *j*.

In order to test our hypotheses, utility is supposed to take the following form:

$$\begin{aligned} V^*_{ij} = {} & \alpha_j + \beta_1 \ln(y_j)O_i + \beta_2 \ln(y_j)E_i + \beta_3 \ln(y_j)S_i \\ & + \gamma_{11} d_j O_i + \gamma_{12} d_j^2 O_i + \gamma_{21} d_j E_i + \gamma_{22} d_j^2 E_i + \gamma_{31} d_j S_i + \gamma_{32} d_j^2 S_i \\ & + \delta\, F_{ij} + \varepsilon_{ij} \end{aligned} \tag{2}$$

where $\ln(y_j)$ is the average population income in location j, d_j distance between location j and the city center, O_i and E_i are two vectors of dummy variables relating respectively to the occupational status and ethnic category of the household, S_i represents its size and F_{ij} is a dummy variable indicating whether location i was the former location of the household. α_j, β_1 to β_3 , γ_{11} to γ_{32} and δ are vectors of coefficients to be estimated.

By doing so, we suppose that households differ in terms of locational choices based on their income, ethnic origin and size. We also assume that they make their choice based on average income in the jurisdiction and size to the center. All the choice determinants that do not differ with household are in the location classes fixed effects α_j. In other words, the impact of the location characteristics that are constant across households are left in the location dummy, together with unobserved variables.

A household chooses among potential locations by comparing its utility level in the different location types and select location j which maximizes his utility:

$$V^*_{ij} \geq V^*_{ik} \quad \forall k \in C = (1,...,K) \tag{3}$$

where C is the choice set of K alternative locations.

We assume that the error terms are identically and independently distributed following an extreme value distribution, of which cumulative distribution function is given by:

$$F(\epsilon_{ij}) = e^{-e^{-\epsilon_{ij}}} \tag{4}$$

As a result, the probability for a household i to choose location j is:

$$P_{ij} = \frac{e^{V_{ij}}}{\sum_{k=1}^{J} e^{V_{ik}}} \tag{5}$$

This conditional logit model is estimated by the maximum likelihood method.

Our empirical model then focuses on interaction variables between the location characteristics and household's characteristics. As stated in the theoretical model, we expect that households sort by income relatively to the distance to the central city. We introduce interactions between dummies for household's social categories (white-collars in mid management positions being the reference) and location distance to the city center (and this variable squared). The same is done with a dummy for household's ethnicity (French born in France being the reference). To control for household's size, we also introduce an interaction between household size and type distance to the center (and this variable squared). We also expect that migrants take account of social context and related endogenous amenities. The same interaction variables are built with the location mean income instead of distance to the central city.

To sum up, household variables interacted with location variables are the following:
- dummies for the reference person occupational status,
- dummy for the reference person not being a French citizen,
- dummy for the reference person being a French citizen born abroad,
- number of persons living in the household.

Location variables are:
- mean fiscal income in 1990[5] divided by its average over the 17 types,
- distance to the central city in km[6] less the mean distance to the city center over the 17 types in the urban area, so as to avoid colinearity with their squared counterparts.
- distance less mean distance squared.

A variable is both a household and location variable:
- dummy for the location being the former location of the household.

4.2. Choice probabilities and measures of social and ethnic segregation

Our aim is not only to test for the different conjectured factors of location choices, but also to assess the importance of each of them in segregation levels observed in French urban areas. We can do so by comparing segregation levels in different counterfactual cities that are predicted by the model estimated coefficients.

Segregation levels are classically measured by different spatial concentration indexes. Among them, the dissimilarity index is the most commonly used. Based on socio-occupational categories and ethnic origins that we use in the location model, we compute multi-group dissimilarity indexes (Readon and Firebaugh, 2002) of which the general expression is the following:

$$D = \frac{1}{2I}\sum_{m=1}^{M}\pi_m \sum_{j=1}^{J} w_j \left|\pi_{jm} - 1\right| \quad (6)$$

where m indexes the different groups of population and j the different locations. π_m is the proportion of group m in the population, w_j is the weight of location j in the total population and π_{jm} is the share in group m for the population in location j. I is equal to $\sum_{m=1}^{M}\pi_m(1-\pi_m)$ and measures the diversity of groups among the population.

Estimation of the conditional logit model in a given urban area provides for each household its probabilities to choose each available location alternatives. Of course we also know the true distribution of moving households among locations. Hence, we can compute measures of segregation in both cases and compare observed and predicted patterns, so as to assess how well our estimated model accounts for the observed residential segregation.

Then, using only the coefficients attached to one dimension of location choice (e.g. distance to the central city) and setting all other coefficients except fixed effects to zero, a new set of choice probabilities can be predicted. Measuring segregation with the resulting counterfactual household distribution provides information as to the contribution of this causal channel to the production of segregation.

More specifically, we will compute three types of predicted location choice probability.
The general form of this probability is:

---

[5] For each type, we sum the total fiscal income of communities belonging to the type and divide by the total number of households of these communities.
[6] For each type, this distance is the average over all communities belonging to the type of the community level distances to the city center, weighted by the number of housing offered by each community.

$$P_{ij} = \frac{e^{V_{ij}}}{\sum_{k=1}^{J} e^{V_{ik}}} \tag{7}$$

where V_{ij} takes different forms as follows:

(i) Probabilities predicted by the full model

$$\begin{aligned}\widetilde{V}_{ij} = \alpha_j &+ \beta_1 \ln(y_j)O_i + \beta_2 \ln(y_j)E_i + \beta_3 \ln(y_j)S_i \\ &+ \gamma_{11} d_j O_i + \gamma_{12} d_j^2 O_i + \gamma_{21} d_j E_i + \gamma_{22} d_j^2 E_i + \gamma_{31} d_j S_i + \gamma_{32} d_j^2 S_i \\ &+ \delta\, F_{ij}\end{aligned} \tag{8}$$

This is simply the prediction of the full model and allows us to assess the explanatory power of the location model.

(ii) Probabilities predicted by a given segregation channel

Our conditional logit model includes only three locational characteristics: household mean income, distance to the central city and, for each household, if it is her former location. The first two location characteristics are interacted with household characteristics, allowing location behaviors to differ along the line of the latter. The third is *per se* both a location and a household characteristic. Taking our estimation results and setting all the coefficients corresponding to these variables to zero, except those associated with one of these segregation channels provides us the choice probabilities that would be relevant if only this channel produced social and ethnic segregation or in other words if households considered only this determinant in their location choice. Computing the corresponding dissimilarity indexes, we can assess the relative contribution to segregation of each of these channels.

For instance, the predicted probability based on the distance segregation channel would be:

$$\widetilde{V}_{ij} = \alpha_j + \gamma_{11} d_j O_i + \gamma_{12} d_j^2 O_i + \gamma_{21} d_j E_i + \gamma_{22} d_j^2 E_i + \gamma_{31} d_j S_i + \gamma_{32} d_j^2 S_i \tag{9}$$

We refer in particular to the inertia model, which gives the following predicted probabilities:

$$\widetilde{V}_{ij} = \alpha_j + \delta\, F_{ij} \tag{10}$$

(iii) Probabilities predicted by behaviors differing with occupational or ethnic group

In our conditional logit model, households differ by their belonging to a social group, their ethnic status, their size and their former urban location. Let us focus on social and ethnic status. Taking our estimation results and setting the coefficients of all interactions implying ethnic status (resp. social status) to zero provide the choice probabilities that would be relevant if behaviors did not differ between ethnic groups (resp. social groups) and was the same as the behavior of the reference category in the econometric model. Thereby, computing the corresponding dissimilarity indexes, we can check if social segregation is mainly explained by the differing behaviors of households belonging to differing social groups or by the combination of the differing behaviors of households belonging to different ethnic groups and the social composition of ethnic groups. In the same manner, we can check if ethnic segregation is mainly explained by the differing behaviors of households belonging to differing ethnic groups or by the combination of the differing behaviors of households belonging to different social groups and the ethnic composition of social groups.

The corresponding predicted probabilities are for instance, considering that households only differ in terms of social categories:

$$\widetilde{V}_{ij} = \alpha_j + \beta_1 \ln(y_j) O_i + \gamma_{11} d_j O_i + \gamma_{12} d_j^2 O_i \quad (8)$$

5. Results

5.1 Estimation results.

The estimated coefficients for the five largest urban areas, namely Paris, Lyon, Marseille, Lille and Toulouse, are presented in Table 4 as examples. Joint significativity tests aimed at assessing the significance of each of the interactions in the model are presented for the 37 urban areas in table 5. The signs of estimated significant coefficients are presented in Table 6. First of all, fixed effects are always jointly significant (Table 5). Estimated fixed effects for the 17 location classes are all negative: the central city is always more attractive, what obviously reflects differences in housing stocks, which are also differences in housing opportunities. More densely populated locations mechanically receive more migrants. As expected, it is also observed that outer suburbs – which exhibit a very low density compared to others – generally display the lowest coefficients. These fixed effects thus control, among other effects, for the size differences between locations.

We comment only briefly the coefficients corresponding to the effect of household size and previous location. Household size significantly influence distance choice in the vast majority of the urban areas. The effect is the one predicted by urban economic models: large households locate further away from the city center, although the effect is not linear, as shown by the coefficient of quadratic distance. The interaction between household size and average location income is significant and positive in 13 urban areas, which can be interpreted in particular as the fact that the presence of children increases the preference for wealthy locations. Finally, there is a very significant effect of previous location on location choices: locations where the household was initially are very more likely to be chosen as the new destination when moving.

The interactions of social category dummies with location average income are jointly significant in 29 out of 37 urban areas. This result points out the power of the search for income-related amenities in the production of residential social segregation. When significant, the coefficient of the "executive" category is positive, whereas it is always negative for white-collar subordinates and blue-collars (Table 6). This is true in particular for the five urban areas in Table 4. As expected, the most affluent social groups are more attracted by the most affluent locations than the less favored social groups. However, in most urban areas, the coefficient of the interaction with location income is higher for blue-collar than for white-collar subordinates: although the income differential between these two categories is very small (see Table 3 in Section 3.3), they behave somewhat differently regarding income-related amenities.

The effect of social status on the choice of distance is slightly less often significant. Nevertheless, these interactions are jointly significant in 27 out of 37 urban areas (Table 5).

On average, white-collar subordinates have the most negative coefficients: they are the least attracted by distant locations, compared to intermediate categories (Table 6). Apart from Paris, it is the case in the large urban areas considered in Table 4. In Paris urban area, the executives are the category that is the most reluctant to settle in distant locations, *ceteris paribus*. This can be explained by the level of traffic congestion, that makes wealthy households locate close to the center. Blue-collar workers behave differently from white-collar subordinates: they behave either like the intermediate category (reference) or are more attracted by distant locations (in only three urban areas, among which Paris). This result is in line with what we already know from urban configurations in France: it has been observed that blue-collar workers are more prone to locate in outer suburbs than white-collars (Goffette-Nagot, 2000). What we show here is that it remains true after controlling for preferences regarding income-related amenities.

The effect of ethnic origin, after controlling for social status, is slightly weaker than that of occupational status. Still, the coefficients of the interactions between ethnic origin and distance from the central city are significant and negative in the majority of the urban areas (21 urban areas). This centralization of foreigners could be the consequence of a strong concentration of public housing in the inner suburbs and the fact that foreign households are often housed in public housing. Further, only in half of the urban areas (18) do foreign citizens behave differently than French citizen born in France regarding location income, by locating in less affluent municipalities. French citizens born abroad behave more often as their fellow-citizens, except in the four biggest urban areas (Paris, Lyon, Marseille and Lille) and in three others of medium size. This result also can be the consequence of public housing accommodation, as the average income in the municipality is likely to be correlated with the percentage of public housing in the housing stock. The significantly different behavior of foreign households can also be the consequence of network effects, as immigrants often choose to settle near individual belonging to the same ethnic group.

5.2. Analysis of social and ethnic segregation.

In the three following sections, we present and analyze simple statistics for the dissimilarity indexes computed for each of the 37 urban areas. In the first section, we look at how well our full model can predict the observed socially and ethnically segregated patterns. Based on predictions obtained with partial models, we then try to assess the relative contributions to segregation of the choices of "neighborhood income" and "distance to the central city". Lastly, we try to disentangle the social and ethnic determinants of the observed segregated pattern.

5.2.1. Observed segregation patterns and predictions of the full model

Table 7 and 8 present the observed and predicted dissimilarity indexes for the five largest urban areas taken as examples, for income and ethnic segregation respectively. The following tables display indexes averaged over the 37 urban areas: observed indexes in Table 9 and the ratio of predicted to observed indexes in Table 10.

Broadly speaking, we can first note that the mean values of the dissimilarity indexes computed here at a supra-municipality geographic level are low compared to what is generally obtained in studies working with municipality or infra-municipality levels. This is expected as social and ethnic segregation is likely to be stronger the smaller the spatial scale considered. Note also that individuals considered here are migrants and it is not clear *a priori* whether their spatial segregation is likely to be stronger or weaker than the segregation level of the whole population.

The household charateristics considered in the estimated model allows to consider separately households of which the reference person is out of labor force and those of which the reference person is in labor force. Among the latter, five categories are considered: executives, intermediate professions, white-collars, blue-collars and independent workers. We first discuss segregation between these five categories and then oppose them to the category of inactive individuals. We base our discussion on indexes averaged over the 37 urban areas.

The mean value of the multi-group dissimilarity index corresponding to the observed spatial distribution of categories in labor force is equal to 0.14, considering either four or five social groups (Table 9).[7] Regarding binary oppositions between the four income-ordered social categories, we observe that the highest index is obtained for the opposition between executives and blue-collars (0.25) and the second highest for the one between executives and white-collar subordinates (0.19). A lower value is obtained for the opposition between executives and white-collars in mid-management positions (0.15). A value of 0.14 is then obtained for all other oppositions implying the latter. Thereby, *social segregation seems to obey to income hierarchy between social groups: the higher the income differential, the higher the value of the dissimilarity index*. However, less expected in this respect is the strong segregation between white-collar subordinates and blue-collars (0.18), two categories that display a nearly equal mean income level. Finally, the mean dissimilarity index increases to 0.16 when the category of economically inactive households is added to the five-group index. This result may be surprising given the low value of the dissimilarity index corresponding to the binary opposition between economically active and inactive households (0.13).

Regarding now ethnic segregation, the mean value of the multi-group dissimilarity index opposing foreign citizen, French citizen born abroad and French citizen born in France is 0.18. As expected, the highest value of binary indexes opposing these groups is obtained for the opposition between foreign citizen and French citizen born in France (0.24). More interesting is the higher value of the index corresponding to the opposition between French born abroad and foreign citizen (0.19) than that between French born abroad and French born in France (0.15). "Spatial integration" among French citizens is far from perfect but seems to be at work.

Yet, the question we are interested in is the following: to what extent is our parsimonious location choice model able to predict this observed segregation pattern? Recall that our model includes only three location characteristics: location mean income, location distance to the central city and, for each household, if it is its former location. Thereby, this model features only three "segregation channels": differing choices of "income" and "distance" between households of differing characteristics, as well as the inertia of the previously prevailing segregated pattern due to the advantages associated with relocating nearby one's former location (better information about housing supply, lower moving costs, access to previously built social networks, etc.). Knowing the predictive power of our full model provides information on the importance of these factors in the formation of social and ethnic segregation. To support this point, dissimilarity indexes are now computed considering the distribution of households corresponding to the choice probabilities predicted by our

[7] The four group index opposes the four categories that can be meaningfully ordered by mean income by consumption unit, i.e. executives, white-collars in mid-management positions, white-collar subordinates and blue-collars (see section 3.3). Independent workers are added to build the five group index.

conditional logit estimation results (see equation 8 section 4.2). For each of these indexes, we comment the average value of the predicted index relatively to the observed index.

The model predicts, on average over the 37 urban areas, 72% of the value of the observed multi-group dissimilarity indexes built with economically active households (either with four or five groups) and 80% of the value of the multi-group index built with both active and inactive households (Table 9). It predicts on average 71% of the value of the observed binary index opposing active to inactive households, and from 65% to 76% of the values of the ones opposing occupational categories. Regarding ethnic segregation, the model predicts on average 64% of the value of the observed ethnic multi-group dissimilarity index. It predicts 68% of the value of the observed binary index opposing foreign citizen to French citizen born in France, 73% of the one opposing foreign citizen to French citizen born abroad and 54% of the one opposing French citizen born in France to French citizen born abroad.

This set of results shows that the estimated model reproduces quite well the observed segregation patterns. Even in the urban areas where it is the less explanatory, it still predicts more than 30% of the observed social segregation level. *This indicates clearly that the three segregation channels included in our model indeed contribute strongly to the formation of social and ethnic segregation.*

5.2.2. "Income driven" vs. "Distance driven" ethnic and social segregation.

As explained in details in section 4.2, it is possible, by computing predicted location choices considering only one characteristics of location (income, distance or being the former household's location), to assess how strong the segregation would be if households considered only this characteristic when choosing their location. By doing so, we can determine whether the observed social segregation is mainly explained by the Alonso model or by income-related amenities.

Table 11 shows that the proportion of the value of the observed five-group dissimilarity index explained by the "income" channel partial model is 40% vs. 37% for the "distance" channel partial model and 32% for the "inertia" channel partial model.[8] *Segregation among the economically active social groups thus appears to be mainly driven by choices of location average income, but choices of distance to the central city and the location inertia also contribute significantly.*

Adding the economically inactive population in the analysis reverses this conclusion: segregation between social groups is now mainly (but only slightly) driven by distance, the "income" model explaining 56% vs. 58% for the "distance" model and 52% for the "inertia" model. Looking at binary oppositions allows to refine the analysis. *Segregation between economically active and inactive households is clearly driven by the "distance" channel*: the "income" model explains only 29% vs. 57% for the "distance" model and 25% for the "inertia" model. As predicted by the standard urban economic model, estimation results for most of the urban areas show that inactive households settle significantly further away from the central city.

Regarding economically active households, on the one hand, all the oppositions implying executives show a much stronger contribution to segregation of the "income" channel

[8] The same proportions are respectively 42%, 36% and 32% for the four-group index.

compared to the "distance" channel. The proportion of the value of the observed index opposing executives to blue-collars explained by the "income" model is 51% vs. only 16% for the "distance" model (and 32% for the "inertia" model). The same proportions are respectively 45% vs. 21% (and 46%) for the opposition between executives and white-collars in mid-management positions and 59% vs. 50% (and 38%) for the opposition between executives and white-collar subordinates. *Executives presumably settle in affluent neighborhoods hardly accessible for others groups. Amenity considerations seem more important than the standard urban economic trade-off between proximity to the central business district and land consumption to explain the segregation between this group and poorer households.*

The same prevails for the opposition between white-collars in mid management positions and blue-collars (proportions are respectively 42%, 19% and 35%). On the other hand, the "distance" channel is clearly dominant in the explanation of the segregation between white-collars in mid management positions and white-collar subordinates (35% for the "income" model vs. 54% for the "distance" model and 24% for the "inertia" model), and it is even more the case regarding the segregation between the latter and blue-collars (23% vs. 55% and 27%). The specific behavior of white-collar subordinates with respect to the choice of distance to the central city is at the source of the surprisingly strong segregation level observed between this group and blue-collars and the surprisingly low segregation level observed between this group and executives already stressed in section 5.2.1. Estimation results show that in most of the urban areas, *white-collar subordinates – although displaying a similar income compared to blue-collars – do not follow them in their migration toward peripheral urban locations. They tend to settle in central locations, so that they do not segregate too much from the more economically favored groups (white-collars in mid management positions and executives)*, although they choose more central locations than the former and less affluent locations than the latter. This point would be worth further investigations.

Regarding ethnic segregation, the proportion of the value of the observed multi-group dissimilarity index explained by the "income" model is 24% vs. 35% for the "distance" model and 43% for the "inertia" model (Table 12). Ethnic groups were segregated in 1990 and as moving households preferentially chose to relocate close to their initial location, they are still segregated in 1999: *this segregation inertia appears to be the first segregation force*. One reason may be that foreign citizens are less able to get information about housing supply in other parts of the urban area and less able to support large moving costs. Another reason may be that they are more dependent on local social networks, for instance due to their larger participation to informal economic activities.

Differing choices of distance to the central city between ethnic groups is the second segregation force. This could be explained by the over-representation of foreign citizen and French citizen born abroad in social housing and the historical clustering of social housing in close inner suburbs. Binary oppositions provide additional insights. Segregation between foreign citizens and French citizens born in France is also mainly driven by the "distance" channel, but the "income" channel appear more important than in the multi-group case: the "income" model explains 32% vs. 39% for the "distance" model and 41% for the "inertia" model. Regarding segregation between foreign citizens and French citizens born abroad, the "distance" channel is more important than both the "income" and the "inertia" channels (37% vs. 52% and 39%). This may be explained by the exit of social housing and consecutive decentralization of a large number of French citizens born abroad being former foreign citizens having simultaneously acquired the French nationality and improved their economic conditions of living. Eventually, regarding segregation between French citizens born abroad and French citizens born in France, the "income" channel is still dominated by the "distance" channel, but most of all, the "inertia" channel strongly dominates both other channels (11% vs. 29% and 44%). Again, we can think that a large number of French citizens born abroad are former foreign citizens having acquired the nationality and improved their conditions of living: they were thus likely to be strongly segregated from French citizen born in France in 1990. Now, although their behaviors regarding location characteristics came closer to the behaviors of French citizens born in France, the former level of segregation is partly reproduced due to the advantages procured by the relocation nearby one's former location. However, the thrust of this last comment is limited by the relatively poor performance of our model in explaining this binary index (only 54% of the observed index is explained by the full model, which is the lowest explanation power displayed in our study).

5.2.3. Ethnic vs. social determinants of ethnic and social segregation.

As explained in details in section 4.2, it is possible, by computing predicted location choices considering only occupational status (respectively origin), to assess how strong the segregation would be if households located in the same way, irrespective of their origin (respectively occupational status). By doing so, we can determine whether the observed social segregation is mainly explained by differences in location behaviors according to occupational status or by the differences in behaviors of ethnic groups and the social composition of these groups.

Results are unambiguous. The proportion of the value of the observed multi-group dissimilarity indexes built with economically active households (indexes with four or five groups) explained by the partial model that does not account for ethnic status is 73% *versus* 32% for the one that does not account for social status (Table 11). The same proportions are respectively 80% *versus* 51% regarding the multi-group index including inactive households (six groups index). Finially, the same differences are observed when considering the binary indexes opposing social groups. *These results show that both social status and ethnicity contribute to the production of social segregation, differing behaviors between social groups being however clearly the main determinant of social segregation.*

Regarding ethnic segregation, the proportion of the value of the observed ethnic multi-group dissimilarity index explained by the partial model that do not account for ethnic status is 38% vs. 64% for the one that do not account for social status (Table 12). The same proportions are respectively 38% vs. 66% regarding the binary index opposing foreign citizen to French citizen born in France, 48% vs. 59% regarding the one opposing foreign citizen to French citizen born abroad, and 38% vs. 57% regarding the one opposing French citizen born in France to French citizen born abroad. *These results show that both social status and ethnicity contribute to the production of ethnic segregation, differing behaviors between ethnic groups being clearly the main determinant of ethnic segregation.*

As could be expected, household social characteristics appear more important to predict ethnic segregation than household ethnic characteristics to explain social segregation. In other words, social segregation is not explained by a social composition effect of ethnic groups and conversely, ethnic segregation is not the result of the ethnic composition of social groups.

6. Conclusion

This paper provides a framework aimed at analyzing the determinants of location choices and social and ethnic segregation within urban areas. Our objective is to assess empirically the relative contribution of Tiebout-sorting and Alonso-sorting mechanisms to social and ethnic segregation in 37 large French urban areas. We propose a conditional logit model of urban location choic in which moving households are assumed to sort based on jurisdiction distance to the central city and jurisdiction average household income (as a proxy for the level of public amenities). Estimation of this model provides for each household her probabilities to choose each one the available location alternatives and allow the comparison of various predicted segregation patterns with the observed segregation pattern.

Our main results are the following. Going beyond the standard urban model to take into account income-related amenities is strongly justified by our results: segregation among economically active social groups appears to be mainly driven by the income channel, especially for the most affluent social group. Nevertheless, segregation between economically active and inactive households is mainly explained by distance, as predicted by the standard urban economic model. Regarding ethnic segregation, the distance channel dominates the income one, but this is likely to be due to a non-market effect, i.e. the location of the public housing supply in close inner-suburbs. More important, both of them are dominated by the inertia of the previously prevailing segregated pattern.

Our analysis thus confirms the importance the choices of distance to the central city and neighborhood income in the formation of residential segregation. It also shed light on a third “segregation channel” that may be worth further investigation: the tendency to relocate nearby one’s former location, which is presumably linked with the question of moving costs and access to social networks.

Several limits should now be overrun. First, our analyses are mainly based on mean tendencies among the 37 urban areas considered. We should try to explain the heterogeneity displayed in our results. Second, our econometric model assumes the Independence of Irrelevant Alternatives hypothesis (IAA). We should test for it and if necessary turn to another model.

References

Alonso W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*, Harvard Univ. Press, Cambridge, 1964.

de Bartolome, CAM; Ross, SL. 2007. Community income distributions in a metropolitan area. *Journal of Urban Economics* 61 (3): 496-518.

de Bartolome, CAM; Ross, SL. 2004. Who's in charge of the central city? The conflict between efficiency and equity in the design of a metropolitan area. *Journal of Urban Economics* 56 (3): 458-483.

de Bartolome, CAM; Ross, SL. 2003. Equilibria with local governments and commuting: income sorting vs income mixing. *Journal of Urban Economics* 54 (1): 1-20.

Benabou, R. (1993), 'Workings of a City: Location, Education, and Production', *Quarterly Journal of Economics* 108, 619-652.

Bjorvatn, K., Capelen, A. W. (2003), 'Inequality, segregation and redistribution', *Journal of Public Economics*, 87: 1657-1679.

Brueckner, J.K., Thisse, J., Zenou, Y. (1999), 'Why is central Paris rich and downtown Detroit poor?: An amenity-based theory', *European Economic Review* 43, 91-107.

Card, D., Mas, A., Rothstein, J. (2008), 'Tipping and the Dynamics of Segregation', *Quarterly Journal of Economics*, 123, 177-218.

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.

Glaeser E.L., Kahn M. (2003). Sprawl and urban growth. in Henderson V., Thisse J.-F. *Handbook of Urban and Regional Economics*, Elsevier.

Glaeser E. L., Kahn M.E. and Rappaport J. (2000) “Why Do the Poor Live in Cities?” NBER Working Paper #7636.

Gobillon L. and H. Selod (2007), Les déterminants locaux du chômage en région parisienne, *Economie et Prévision*, 180-181, pp. 19-38.

Goffette-Nagot, F. (2000), 'Urban spread beyond the city edge', in Huriot J.-M., Thisse J.-F. (ed.) *Economics of Cities : Theoretical Perspectives*, New York, Cambridge University Press, 318-40.

Ioannides, Y., Zabel, J.E. (2008), 'Interactions, Neighborhood Selection and Housing Demand', *Journal of Urban Economics*, 63, 229-252.

Ioannides, Y., Zanella, G. (2008), '*Searching for the Best Neighborhood: Mobility and Social Interactions*', Working paper 533, University of Siena, Department of Economics, available from http://www.econ-pol.unisi.it/dipartimento/it/node/920.

Jackson K. (1985) *Crabgrass Frontier. The Suburbanization of the United States.* Oxford University Press.

Julien P. (2001), Les grandes villes françaises étendent leur influence, *INSEE Première* N°766.

Margo R. (1992). Explaining the Postwar Suburbanization of the Population in the United States; the Role of Income. *Journal of Urban Economics* 31: 301-310.

Mills E.S. (1967). An aggregative model of resource allocation in a metropolitan area, *Papers and Proceedings of the American Economic Association* 57: 197–210.

Mills E.S., Lubuele L.S. (1997). Inner Cities, *Journal of Economic Literature*, 35: 727-756.

Muth, R.F. (1969), *Cities and Housing*, University of Chicago Press.

Nechyba, TJ; Walsh, RP. 2004. Urban sprawl. *Journal of Economic Perspectives* 18 (4): 177-200.

Pan Ke Shon J.L. (2009). Ségrégation ethnique et ségrégation sociale en quartiers sensibles. *Revue française de Sociologie*, 50: 451-487

Préteceille E., 2006, "La ségrégation sociale a-t-elle augmenté ? La métropole parisienne entre polarisation et mixité", *Sociétés Contemporaines*, 62, 69-93

Reardon S.F., Firebaugh G. (2002), 'Measures of multigroup segregation', *Sociological Methodology* 32, 33-67.

Schelling T. (1971), Dynamic models of segregation, *Journal of Mathematical Sociology*, 1, 143-186.

Schmidheiny, K. (2006), 'Income segregation and local progressive taxation: Empirical evidence from Switzerland', *Journal of Public Economics*, 90, 429 - 458.

Tiebout, C. (1956). A pure theory of local expenditures. *Journal of Political Economy* 64: 416-424.

Wheaton W.C. (1977). Income and Urban Residence: An Analysis of Consumer Demand for Location, *American Economic Review*, 620-632.

Wheeler, C.H. (2006), '*Urban Decentralization and Income Inequality: Is Sprawl Associated with Rising Income Segregation Across Neighborhoods*?', Federal Reserve Bank of St. Louis Working Paper No. 2006-037A.

Wheeler, C.H., La Jeunesse, E. (2008), 'Trends in Neighborhood Income Inequality in the U.S.: 1980-2000', *Journal of Regional Science,* 48, pp. 879-891.

Appendix 1

The French Population Census does not contain information regarding the household's income. However, it contains detailed information concerning the households' members, that allows to predict its income. Coefficients used in this prediction are based on a regression performed on a sample taken from the French Housing Survey (hereafter FHS) 1996.

On this secondary sample, the natural logarithm of the household's income is regressed on the main determinants of wages and income subsidies available in the FHS. The following variables are included in the estimation:
- Age (and squared age): age of the “reference person of the household”
- Gender: “reference person of the household” is a woman,
- Marr: “reference person of the household” is married,
- Divo: “reference person of the household” is divorced,
- Wido: “reference person of the household” is widowed,
- Foreign: “reference person of the household” is not a French citizen,
- Educ_1-6: level of education of the “reference person of the household” from 1 to 6,
- Status_1-8: socio-occupational status of the “reference person of the household” from 1 to 6,
- Nbwork: number of persons having a job in the household,
- Nbchild: number of children living in the household.

The R^2 of the estimation is equal to 0.62. All the estimated coefficients are statistically significant at the 5% level and the Ficher test indicates significance of the overall estimation. Detailed results of this estimation are available on request.

Table 4: Estimated coefficients of the conditional logit model for the five largest urban areas.

	Paris		Lyon		Aix-Marseille		Lille		Toulouse	
	coeff.	std	coeff.	std	coeff.	std	coeff.	std	coeff.	std
Interactions with location av. income										
av. income x executive	1.25***	0.05	1.37***	0.19	1.81***	0.39	1.3***	0.21	1.34***	0.38
av. income x white-collar	-0.88***	0.05	-0.54***	0.18	-0.82**	0.37	-0.89***	0.21	-1.05**	0.43
av. income x blue collar	-1.36***	0.06	-0.85***	0.16	-1.87***	0.35	-1.8***	0.19	-1.24***	0.39
av. income x indep. work.	0.66***	0.08	0.75***	0.24	0.37	0.47	0.67**	0.30	0.27	0.49
av. income x out of lab. force	-0.2***	0.06	-0.27	0.18	-0.01	0.35	-0.73***	0.21	-0.81*	0.43
av. income x F. born abroad	-0.61***	0.06	-0.63***	0.20	-0.68**	0.32	-0.81**	0.34	0.3	0.38
av. income x foreigner	-1.21***	0.06	-1.82***	0.22	-1.73***	0.59	-2.22***	0.36	-1.02	0.80
av. income x hsld size	-0.06***	0.01	0.27***	0.04	0.01	0.09	0.2***	0.05	0.37***	0.10
Interactions with location av. distance										
distance x executive	-1.01***	0.11	-0.87	0.58	1.25*	0.70	-0.52	1.05	-0.48	0.89
dist.2** x executive	2.48***	0.41	7.93*	4.30	7.43**	3.37	17.75	11.88	9.06	6.37
distance x white-collar	-0.76***	0.11	-1.46***	0.54	-1.59**	0.66	-1.26	1.01	-1.71**	0.81
dist.2** x white-collar	-0.3	0.43	5.97	4.02	-3.16	3.03	6.41	11.11	-0.53	5.96
distance x blue collar	0.86***	0.10	0.99**	0.43	-0.05	0.58	-1.92**	0.93	0.17	0.71
dist.2** x blue collar	0.29	0.41	-4.71	3.51	-5.49**	2.73	-34.53***	10.35	-4.39	5.25
distance x indep. work.	0.62***	0.14	1.1	0.68	0.81	0.81	2.73*	1.42	2.54***	0.93
dist.2** x indep. work.	4.06***	0.58	-0.01	5.24	1.53	3.87	2.88	16.05	5.5	6.91
distance x out of lab. force	0.48***	0.11	-0.69	0.53	0.22	0.59	0.76	1.02	0.68	0.77
dist.2** x out of lab. force	4.05***	0.43	-0.96	4.01	1.85	2.83	10.93	10.99	8.95	5.91
distance x F. born abroad	-2.52***	0.14	-1.76***	0.57	-0.21	0.52	-3.22	2.61	-0.77	0.75
dist.2** x F. born abroad	-1.5***	0.50	-2.31	4.29	-0.63	2.43	-8.89	25.72	-2	5.57
distance x foreigner	-3.59***	0.14	-4.62***	0.65	-1.78*	1.01	-5.84**	2.89	-5.77***	1.55
dist.2** x foreigner	-0.95*	0.51	0.17	4.77	3.42	5.14	-5.43	27.79	1.68	11.20
distance x hsld size	0.58***	0.02	0.91***	0.11	0.07	0.14	1.85***	0.23	1.41***	0.18
distance2** x hsld size	-2.69***	0.09	-11.4***	0.90	-2.91***	0.68	-7.1***	2.51	-9.74***	1.36
Previous residence in same location	2.15***	0.01	2.11***	0.02	2.86***	0.02	2.32***	0.02	2.11***	0.03
Averaged fixed effects (all fixed effects are signif. at 1% level)										
Close inner suburbs	-0.8		-1.73		-1.91		-0.84		-2.15	
Distant inner suburbs	-2.03		-3.03		-1.93		-1.42		-3.01	
Close outer suburbs	-2.64		-3.23		-2.58		-2.64		-3.21	
Distant outer suburbs	-3.19		-3.1		-2.72		-3.08		-3.49	
Log likelihood	-163,352		-20,636		-12,608		-13,335		-10,731	
Number of observations	89,823		11,895		10,799		7,846		6,242	

***, **, *: significant at the 1%, 5% and 10% level respectively

Table 5: Joint significativity tests of estimated coefficients of the conditional logit – 37 urban areas.

	Paris	Lyon	Aix-Marseille	Lille	Toulouse	Nice	Bordeaux	Nantes	Strasbourg	Toulon	Douai-Lens	Rennes	Rouen
Likelihood	-16,352	-20,636	-12,608	-13,335	-10,731	-9,161	-11,074	-7,905	-6,264	-5,622	-4,650	-5,478	-7,348
Interactions w/ occupation	4282.3***	293.8***	142.0***	268.8***	146.5***	72.3***	216.6***	117.8***	93.6***	30.85**	45.06***	131.56***	168.8***
av. income x occ.	2695***	167.52***	103.5***	239.7***	529***	38.08***	89.57***	18.64***	23.94***	3.48	39.86***	6.31	106.6***
distance x occ.	890.6***	103.9***	50.9***	39.5***	56.4***	22.57**	81.35***	68.46***	58.03***	12.67	11.49	80.07***	64.43***
Interactions w/ origin	1856.3***	146.7***	54.6***	73.0***	51.0***	22.93***	40.05***	29.6***	133.09***	14.14**	5.43	49.16***	48.99***
av. income x origin	483.6***	68.29***	13.31***	45.0***	-0.56	1.1	4.36	-0.3	-0.19	-0.57	2.76	-1.62	10.99***
distance x origin	1640.9***	104.35***	26.95***	29.8***	47.3***	21.03***	37.41***	29.46***	126.03***	8.66	2.79	45.3***	35.31***
Fixed effects	92511***	10304***	13052***	8688***	4580***	8263	4357***	3642***	3318***	5073***	4064***	2335***	3342***
Whole model	182271***	26131***	35975***	17789***	13909***	21059***	12123***	10489***	9436***	10248***	7886***	7578***	6452***

	Grenoble	Mont-pellier	Metz	Nancy	Clermont-Ferrand	Valenciennes	Tours	Caen	Orléans	Angers	Dijon	Saint-Étienne	Le Havre
Likelihood	-6,090	-3990	-4721	-4813	-5023	-3914	-4044	-4530	-3845	-3343	-3440	-2542	-2309
Interactions w/ occupation	139.42***	102.69***	92.03***	90.83***	87.11***	63.85***	65.3***	82.79***	99.19***	39.23***	86.63***	23.44*	90.65***
av. income x occ.	108.94***	41.56***	52.4***	49.6***	37.1***	41.8***	18.8***	40.5***	48.6***	20.9***	47.1***	8.5	53.2***
distance x occ.	40.95***	47.85***	19.01**	32.88***	54.85***	1.63	38.81***	40.65***	52.31***	14.82	43.62***	12.81	15.61
Interactions w/ origin	28.85***	61.44***	44.74***	47.2***	27.06***	-0.13	12.43*	8.53**	16.49**	27.45***	32.05***	26.54***	15.54**
av. income x origin	5.73	3.18	40.99***	25.88***	2.1	-3.89	-3.5	2.32	-1.83	5.23*	11.37***	9.36**	-3.73
distance x origin	19.45***	46.3***	20.79***	38.41***	24.07***	-0.41	10.78**	2.74*	15.6***	21.73***	19.79***	15.79***	15.08***
Fixed effects	3008***	1812***	2760***	2049***	2074***	2823***	1516***	1584***	1388***	1194***	1190***	2196***	945***
Whole model	7408***	7586***	5461***	5185***	5458***	4735***	4792***	4534***	3982***	4918***	5184***	7927***	7894***

	Le Mans	Avignon	Mulhouse	Amiens	Béthune	Dunkerque	Perpignan	Besançon	Pau	Bayonne	Genève(CH)-Annemasse
Likelihood	-2799	-2687	-2770	-2295	-2140	-3027	-2377	-1901	-2086	-2245	-2462
Interactions w/ occupation	57.24***	36.72***	43.47***	57.24***	28.65**	52.8***	19.28	58.74***	28.47**	20.5	29.9**
av. income x occ.	14.62**	18.09***	18.06***	8.98	23.09***	38.61***	0.24	12.3**	0.61	7.15	6.54
distance x occ.	26.32***	18.34**	20.44**	23.24**	5.96	10.35	15.65	35.01**	22.58**	9.69	21.66**
Interactions w/ origin	18.29**	20.3***	23.31***	10.38	0.67	8.07	14.25**	25.77**	7.44	9.07	4.08
av. income x origin	2.45	14.09***	7.75**	-3.2	-3.03	-4.16	-0.41	-0.35	0.07	1.42	0.53
distance x origin	15.17***	5.15	7.62	9.86**	-2.45	7.21	12.54**	25.49***	5.93	0.71	3.33
Fixed effects	1127***	2335***	1115***	769***	1690***	1971***	1674***	614***	561***	1211***	1727***
Whole model	5180	4388***	4586***	4340***	3840***	3585***	4289***	4006***	2645***	2820***	2210***

***, **, *: significant at the 1%, 5% and 10% level respectively

Table 6: Signs of significant coefficients – 37 urban areas

	Paris	Lyon	Aix-Mar-seille	Lille	Tou-louse	Nice	Bor-deaux	Nantes	Stras-bourg	Toulon	Douai-Lens	Ren-nes	Rouen	Greno-ble	Mont-pellier	Metz	Nancy	Cler-mont-Ferrand	Valen-ciennes
inc x ex	+	+	+	+	+	+			+				+	+	+	+	+	+	+
inc x bl	-	-	-	-	-		-	-	-		-	-	-	-		-	-	-	-
inc x wh	-	-	-	-	-	-	-	-			-	-	-			-	-		-
inc x ind	+	+		+		+	-								+				
inc x ina	-	-		-	-		-				-					-			-
inc ffor	-	-	-	-													-		
inc x for	-	-	-	-			-					-	-	-	-	-	-	-	
inc x siz	-	+		+	+	+		+	+	-		+	+	+	+	+	+		
dis x ex	-		+					-											
dis2 x ex	+	+	+										+	+					
dis x bl	+	+		-			+												
dis2 x bl			-	-										-				-	
dis x wh	-	-	-		-		-	-	-			-	-	-			-	-	
dis2 x wh																			
dis x ind	+			+	+	+	+	+				+	+		+	+	+	+	
dis2 x ind	+					+													
dis x ina	+								-			+							
dis2 x ina	+							+	+		+								
dis ffor	-	-				-	-	-	-				-	-			-		
dis2 x ffor	-					-													
dis x for	-	-	-	-	-		-	-	-			-	-	-	-	-	-	-	
dis2 x for	-					+													
dis x siz	+	+		+	+		+	+	+			+	+	+	+	+	+	+	
dis2 x siz	-	-	-	-	-		-	-	-	-	-	-	-	-	-	-	-	-	-
prev res.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Table 6: Signs of significant coefficients – 37 urban areas (continued)

	Tours	Caen	Or-léans	An-gers	Dijon	Saint-Étien-ne	Le Havre	Le Mans	Avi-gnon	Mul-house	A-miens	Béthu ne	Dunker-que	Perpi-gnan	Besan-çon	Pau	Bayon -ne	Genève-Anne-masse	Total +	Total -
inc x ex	+	+	+		+				+			+						+	21	
inc x bl	-	-	-	-	-	-	-	-			-		-		-					27
inc x wh		-	-		-		-			-	-		-		-					22
inc x ind								-	+										6	2
inc x ina	-				-		-					-	-		-					14
inc ffor								-	-											7
inc x for		-		-	-	-			-	-										18
inc x siz		+																	13	2
dis x ex		-																	1	3
dis2 x ex			+		+														7	
dis x bl																			3	1
dis2 x bl																		+	1	4
dis x wh		-	-		-		-	-		-	-				-					20
dis2 x wh																+			1	
dis x ind	+	+	+		+								-		+			+	18	1
dis2 x ind										+									3	
dis x ina		-								-									2	3
dis2 x ina		+							+									+	7	
dis ffor		-		-	-						-									13
dis2 x ffor																				2
dis x for			-	-	-			-		-					-					21
dis2 x for																			1	1
dis x siz	+	+	+	+	+		+	+		+	+		+		+				25	
dis2 x siz	-	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-		35
prev res.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		37

Table 7: Income segregation indexes of the five largest urban areas

	Paris	Lyon	Marseille	Lille	Toulouse
Four groups					
Observed	0.18630	0.15553	0.09821	0.17279	0.10778
Whole model	0.17881	0.13167	0.10218	0.14914	0.08644
Distance only	0.06219	0.05096	0.07443	0.03592	0.05497
Income only	0.12292	0.06211	0.08159	0.11830	0.07325
Previous location only	0.07201	0.05820	0.04165	0.06860	0.02701
Five groups					
Observed	0.17710	0.15139	0.09398	0.17009	0.11091
Whole model	0.17037	0.12527	0.09241	0.14871	0.09324
Distance only	0.05900	0.04692	0.06682	0.03833	0.05626
Income only	0.11619	0.06122	0.07373	0.11701	0.06615
Previous location only	0.06917	0.05389	0.03811	0.06742	0.02907
In/out labor force					
Observed	0.07774	0.09185	0.08969	0.11147	0.14391
Whole model	0.06795	0.06839	0.06110	0.07603	0.10411
Distance only	0.07007	0.07075	0.02411	0.06723	0.08083
Income only	0.01251	0.01783	0.01574	0.03995	0.05487
Previous location only	0.02417	0.01723	0.04789	0.03922	0.02023
Executives/blue-collars					
Observed	0.42443	0.33294	0.16615	0.36385	0.20225
Whole model	0.41523	0.29237	0.16213	0.31904	0.13947
Distance only	0.12259	0.03914	0.11887	0.04776	0.01776
Income only	0.27565	0.15010	0.22771	0.28244	0.16632
Previous location only	0.16370	0.15299	0.07087	0.15337	0.07380
Executives/intermediate category					
Observed	0.23412	0.16674	0.09519	0.15833	0.10911
Whole model	0.22340	0.15274	0.09324	0.12183	0.08536
Distance only	0.07078	0.01664	0.04918	0.02116	0.02867
Income only	0.11697	0.09960	0.12012	0.12902	0.10267
Previous location only	0.10884	0.08466	0.03944	0.06338	0.03551
Executives/white-collars					
Observed	0.29837	0.18975	0.15868	0.23487	0.13885
Whole model	0.28396	0.17606	0.17019	0.23017	0.15475
Distance only	0.01632	0.09703	0.10607	0.03949	0.13310
Income only	0.20839	0.13725	0.16935	0.21222	0.16984
Previous location only	0.12416	0.10002	0.06575	0.10208	0.04252
Intermediate category/blue-collars					
Observed	0.20267	0.17502	0.10495	0.22826	0.11524
Whole model	0.19793	0.14805	0.10243	0.20746	0.08562
Distance only	0.05182	0.04387	0.06969	0.05231	0.01302
Income only	0.15989	0.06402	0.10830	0.15450	0.06425
Previous location only	0.06185	0.07166	0.03879	0.09590	0.03866
Intermediate category/white-collars					
Observed	0.10477	0.12263	0.08894	0.13906	0.10348
Whole model	0.09431	0.09954	0.09747	0.10961	0.09286
Distance only	0.06667	0.09051	0.09256	0.01833	0.10444
Income only	0.09157	0.04537	0.04928	0.08423	0.06777
Previous location only	0.02615	0.02437	0.04167	0.04341	0.00838
White-collars/blue-collars					
Observed	0.14617	0.18004	0.13582	0.15549	0.17186
Whole model	0.14694	0.14603	0.13326	0.13000	0.13040
Distance only	0.11849	0.13439	0.15092	0.05981	0.11535
Income only	0.06895	0.01865	0.05902	0.07052	0.00352
Previous location only	0.04087	0.05303	0.05351	0.06680	0.03264

Table 8: Ethnic segregation indexes of the five largest urban areas

	Paris	Lyon	Marseille	Lille	Toulouse
Multigroup ethnic segregation					
Observed	0.14330	0.14217	0.09389	0.22632	0.08820
Whole model	0.14036	0.12330	0.08008	0.18648	0.06527
Distance only	0.09096	0.06378	0.02013	0.04057	0.04602
Income only	0.11391	0.07457	0.05522	0.12344	0.00331
Previous location only	0.06518	0.06544	0.06145	0.10850	0.06183
Foreign citizen/French citizen					
Observed	0.17102	0.18351	0.17504	0.30040	0.22004
Whole model	0.16759	0.17023	0.14130	0.25067	0.19473
Distance only	0.09567	0.09149	0.05894	0.07068	0.15070
Income only	0.16486	0.12213	0.10550	0.18528	0.01096
Previous location only	0.07452	0.07462	0.07982	0.13696	0.11018
French born abroad/French born in France					
Observed	0.11836	0.11046	0.06582	0.14055	0.05172
Whole model	0.11493	0.08409	0.05941	0.10917	0.01840
Distance only	0.08738	0.04045	0.01452	0.02671	0.02933
Income only	0.06182	0.03424	0.03817	0.04979	0.00203
Previous location only	0.05628	0.05853	0.05723	0.07449	0.04484
Foreigners/French born abroad					
Observed	0.11055	0.11562	0.12364	0.17990	0.20150
Whole model	0.10010	0.09682	0.09285	0.16276	0.17873
Distance only	0.01828	0.05633	0.06823	0.09603	0.17637
Income only	0.11084	0.09140	0.07404	0.13839	0.01273
Previous location only	0.02789	0.02563	0.03889	0.06973	0.07151
Foreigners/French born in France					
Observed	0.18387	0.19271	0.18646	0.30700	0.22270
Whole model	0.18027	0.17861	0.15162	0.25624	0.19701
Distance only	0.10567	0.09564	0.05696	0.06932	0.14704
Income only	0.17184	0.12564	0.11221	0.18781	0.01071
Previous location only	0.08096	0.08052	0.08858	0.14077	0.11577

Table 9: Observed dissimilarity indexes – averages for 37 urban areas

	Mean	Std Dev	Min.	Max.
Social segregation				
Multi-group indexes				
Four groups[1]	0.1414	0.0277	0.0725	0.1906
Five groups[2]	0.1426	0.0262	0.0801	0.1873
Six groups[3]	0.1552	0.0290	0.0852	0.2284
Two-groups indexes				
Executives/blue collars	0.2547	0.0807	0.1269	0.4244
Executives/intermediate categ.	0.1503	0.0502	0.0673	0.3008
Executives/white-collars	0.1893	0.0509	0.1050	0.2984
Intermediate/blue collars	0.1548	0.0443	0.0605	0.2292
Intermediate/white collars	0.1407	0.0335	0.0783	0.2360
White collars/blue collars	0.1764	0.0486	0.0654	0.2956
In labor force/out LF	0.1254	0.0359	0.0687	0.1969
Ethnic segregation				
Multi-group index				
Three origin groups	0.1766	0.0559	0.0882	0.3172
Binary indexes*				
French born abroad/born in Fr.	0.1560	0.0597	0.0517	0.2688
Foreign/French born abroad	0.1868	0.0565	0.0799	0.2763
Foreign/French born in France	0.2410	0.0641	0.1138	0.4098

[1] The four groups are the following: executives, intermediate category, white-collars, blue-collars.
[2] The five groups are the previous ones, plus the independent workers.
[3] The six groups are the previous ones, plus the out-of-labor force category.

Table 10: Ratio of predicted over observed dissimilarity indexes – averages for 37 urban areas

	Mean	Std Dev	Minimum	Maximum
Social segregation				
Multi-group indexes				
Four groups[1]	0.7160	0.1674	0.3097	1.0405
Five groups[2]	0.7209	0.1499	0.3310	0.9834
Six groups[3]	0.7952	0.1336	0.3337	0.9855
Two-group indexes				
Executives/blue collars	0.7016	0.1993	0.2465	0.9783
Executives/intermediate categ.	0.6441	0.2575	0.1344	1.0736
Executives/white-collars	0.7636	0.2369	0.2152	1.1145
Intermediate/blue collars	0.6881	0.2720	0.1932	1.2046
Intermediate/white collars	0.7011	0.2062	0.2677	0.9766
White collars/blue collars	0.7094	0.2595	0.3056	1.6982
In labor force/out LF	0.7062	0.2260	0.1400	1.0717
Ethnic segregation				
Multi-group index				
Three origin groups	0.6356	0.2078	0.1845	0.9890
Two-group indexes*				
French born abroad/born in Fr.	0.5413	0.2341	0.2089	0.9710
Foreign/French born abroad	0.7284	0.2581	0.2839	1.3347
Foreign/French born in France	0.6788	0.2372	0.2033	1.0574

[1] The four groups are the following: executives, intermediate category, white-collars, blue-collars.
[2] The five groups are the previous ones, plus the independent workers.
[3] The six groups are the previous ones, plus the out-of-labor force category.

Table 11: Proportion of observed social segregation predicted by the partial models

	Mean	Std. Dev.	Min	Max
Four groups[1]				
Occupational status only	0.7268	0.1701	0.3097	1.0382
Origin only	0.3189	0.0994	0.1347	0.5977
Distance only	0.3639	0.2197	0.0682	0.9132
Income only	0.4150	0.2595	2.14 E-15	0.9045
Previous location only	0.3229	0.0951	0.1347	0.5941
Five groups[2]				
Occupational status only	0.7323	0.1545	0.3310	0.9824
Origin only	0.3185	0.0994	0.1448	0.5722
Distance only	0.3663	0.2048	0.0631	0.7884
Income only	0.3966	0.2476	2.35 E-15	0.8945
Previous location only	0.3227	0.0947	0.1574	0.5698
Six groups[3]				
Occupational status only	0.8010	0.1357	0.3337	0.9924
Origin only	0.5180	0.1202	0.2755	0.7716
Distance only	0.5813	0.1604	0.1177	0.9085
Income only	0.5620	0.1996	0.0133	0.8884
Previous location only	0.5202	0.1187	0.2921	0.7716
Executives/blue collars				
Occupational status only	0.7092	0.2004	0.2502	0.9844
Origin only	0.3934	0.1129	0.1425	0.7069
Distance only	0.1590	0.1797	0.0027	0.7154
Income only	0.5068	0.3587	3.43 E-15	1.3705
Previous location only	0.3989	0.1104	0.1656	0.7027
Executives/white-collars				
Occupation status only	0.7672	0.2342	0.2373	1.1121
Origin only	0.3686	0.1448	0.1227	0.6460
Distance only	0.4951	0.4573	0.0157	1.8095
Income only	0.5879	0.4026	2.64E-15	1.2837
Previous location only	0.3678	0.1456	0.1226	0.6484
Executives/intermediate categ.				
Occupational status only	0.6447	0.2587	0.1322	1.0829
Origin only	0.4119	0.1671	0.1617	0.8566
Distance only	0.2124	0.2285	0.0007	1.1112
Income only	0.4580	0.4600	3.30 E-15	1.5689
Previous location only	0.4106	0.1669	0.1676	0.8566
Intermediate categ./white-collars				
Occupational status only	0.6856	0.2709	0.1932	1.2117
Origin only	0.2408	0.1170	0.0692	0.4808
Distance only	0.5410	0.3971	0.0104	1.1425
Income only	0.3515	0.3431	9.44 E-16	1.4135
Previous location only	0.2388	0.1161	0.0692	0.4684
Blue-collars/white-collars				
Occupational status only	0.7349	0.2732	0.3056	1.7863
Origin only	0.2628	0.1353	0.0679	0.6472
Distance only	0.5501	0.3377	0.1117	1.6159
Income only	0.2276	0.2222	1.85E-15	0.7979
Previous location only	0.2711	0.1328	0.0679	0.6547
Intermediate categ./blue-collars				
Occupational status only	0.7197	0.2082	0.2677	0.9850
Origin only	0.3398	0.1069	0.1444	0.5716
Distance only	0.1889	0.1693	0.0140	0.7902
Income only	0.4207	0.3351	2.37E-15	1.4252
Previous location only	0.3481	0.1046	0.1455	0.6554
In labor force/out of labor force				
Occupational status only	0.7010	0.2280	0.1471	1.0717
Origin only	0.2580	0.1235	0.0775	0.5751
Distance only	0.5686	0.2950	0.0517	1.2200
Income only	0.2925	0.2871	9.42E-16	1.1941
Previous location only	0.2532	0.1216	0.0700	0.5751

[1] The four groups are the following: executives, intermediate category, white-collars, blue-collars.
[2] The five groups are the previous ones, plus the independent workers.
[3] The six groups are the previous ones, plus the out-of-labor force category.

Table 12: Proportion of observed ethnic segregation predicted by the partial models

	Mean	Std. Dev.	Min	Max
Foreigners/French born abroad/French born in France				
Occupation status only	0.3767	0.1142	0.1490	0.6164
Origin only	0.6419	0.1981	0.2278	1.0276
Distance only	0.3481	0.2205	0.0355	0.8289
Income only	0.2409	0.2689	3.80 E-14	0.8628
Previous location only	0.4313	0.1153	0.1733	0.7010
French born abroad/French born in France				
Occupation status only	0.3824	0.1449	0.1583	0.7341
Origin only	0.5669	0.2258	0.1947	1.0580
Distance only	0.2915	0.2536	0.0095	0.8417
Income only	0.1148	0.1744	3.86E-14	0.5965
Previous location only	0.4379	0.1585	0.1877	0.8694
Foreign/French born abroad				
Occupation status only	0.4766	0.3125	0.0850	2.0262
Origin only	0.5854	0.1944	0.2551	1.0211
Distance only	0.5213	0.5573	0.0251	2.0379
Income only	0.3716	0.3925	0	1.3362
Previous location only	0.3853	0.1697	0.0894	0.8063
Foreign/French born in France				
Occupation status only	0.3779	0.1260	0.1255	0.6753
Origin only	0.6636	0.2064	0.2996	1.0021
Distance only	0.3899	0.2595	0.0319	0.9831
Income only	0.3157	0.3366	3.38E-14	1.1697
Previous location only	0.4098	0.1132	0.1641	0.6426