PLUREL

Land Use Relationships
in Rural-Urban Regions

Module 2

March 2010

PERI-URBAN LAND USE RELATIONSHIPS – STRATEGIES AND SUSTAINABILITY ASSESSMENT TOOLS FOR URBAN-RURAL LINKAGES, INTEGRATED PROJECT, CONTRACT NO. 036921

Response functions for housing demand

D2.3.5

Nathalie Bertrand*, Tina Rambonilaza, Mihai Tivadar, Sandrine Lyser
Cemagref

*Responsible partner and corresponding author
Tel: +33 4 76 76 27 42; Email: nathalie.bertrand@cemagref.fr

Document status:

Draft:	completed
Submitted for internal review:	completed
Revised based on comments given by internal reviewers:	completed
Final, submitted to EC:	completed

SIXTH FRAMEWORK PROGRAMME

Contents

Abstract

The spatial distribution of the various household categories is addressed here as a residential location choice that are related to demographic dynamics, socio-economic context, spatial planning in terms of services and infrastructures, and households preferences.

We adopt a standard approach in urban economics, by nesting a model of household location choice with a simplified life-cycle model. The quantification of response function proceed in two steps: describe the housing demand of each social class (retirees, middle-aged and young) according to 6 plurel RUR-subtypes then validate regional maps with regression analysis of the residential location choice of each social class according to each municipality typology for a RUR-typology, the Montpellier case study NUTS-3 region, which is a polycentric RUR configuration. We use the 1999 Census Data for 6 French NUTS-3 of 6 RUR-types.

Our empirical analysis showed that location patterns as the factors influencing housing demand vary strongly with the life-cycle stage of the household. Young households are massively attracted by the urban pole of the area for several reasons: proximity to employment centre, urban amenities and public services. Segregations indices confirmed that young households represent the most segregated and centralized social group in the area. Also there is almost no spatial autocorrelation for young distribution, as result of their high concentration in one spatial unit (Montpellier City). Thus, for young households, location characteristics (urban amenities) are primordial in their location choice and more important than their housing needs.

Middle-aged households present completely different location behaviour: they prefer periurban and rural zones, but in proximity of urban centres. There are several reasons: they need proximity to employment but in the same time, middle-aged households are usually bigger families, with higher housing demand, which couldn't be satisfied in urban centres. This location pattern will influence segregation indices, middle-aged households being the less segregated social group and with a quite correlated spatial distribution. In the same time, there centralisation stays relatively important. According to housing demand equation, for middle-aged households, urbanisation degree has a low impact on their location choice, but on contrary they have important needs for urban services (education and social and health facilities) and natural amenities, here mainly agricultural amenities than forest amenities due to the geographical configuration of the region. We can state that for the middle-aged households, contrary to young, even if they have strong preferences for urban amenities, the housing needs are more important in their trade-off for location choice decision.

Retirees have different behaviour and location patterns. There are no more employment reasons to choose urban areas and other factors emerge such as public services and natural amenities. Thus retirees have very decentralized

locations, in periurban and rural zones. We distinguished “young” retirees such an apart social group, because we can find some working households. This characteristic will make some differences between two categories of retirees. The households from first class of retirees are less segregated and prefer relatively more centralized location than the second class.

Classification of results/outputs:
For the purpose of integrating the results of this deliverable into the PLUREL Explorer dissemination platform as fact sheets and associated documentation please classify the results in relation to spatial scale; DPSIR framework; land use issues; output indicators and knowledge type.

Spatial scale for results: Regional, national, European	Regional
DPSIR framework: Driver, Pressure, State, Impact, Response	Response
Land use issues covered: Housing, Traffic, Agriculture, Natural area, Water, Tourism/recreation	Housing
Scenario sensitivity: Are the products/outputs sensitive to Module 1 scenarios?	Outputs sensitive to Module 1 scenarios
Output indicators: Socio-economic & environmental external constraints; Land Use structure; RUR Metabolism; ECO-system integrity; Ecosystem Services; Socio-economic assessment Criteria; Decisions	Socio-economic assessment Criteria
Knowledge type: Narrative storylines; Response functions; GIS-based maps; Tables or charts; Handbooks	Response functions

How many fact sheets will be derived from this deliverable:	**One**

Introduction

Objectives of WP and links to other WPs

Beyond the comprehension of mutations in progress, Plurel project aims to provide to the European commission a tool to assess sustainability of rural-urban relationships and a new framework to support the implementation of European integrated spatial policies. The project is organised in six modules. CEMAGREF is involved in two modules (Module 2 and module 3). Module 2 proposes to analyse the functional interaction between rural, peri-urban and urban areas. These interrelationships are tackled in terms of response functions of demands competition for resources (infrastructures, energy, water, biodiversity) and functions (residential, transportation, recreation, food supply according to main driving forces of land-use and function change such as : demographic trends, economic and technological development, climate change, ... Then the methodology that was collectively chosen is to use quantitative approach of the relationships between function variable such housing in our case and land-use.

In the case of housing and commuting the RF are of generic nature, derived from regression functions from single cases with detailed data availability, whereas the other RFs are regression functions derived from European databases at NUTS2 /NUTS3.

The parameterization of the housing response function is based on the residential location choice of three categories of the population (ageing population and retired, households with children, young couples and students) for each type of area (urban, rural, peri-urban) at the municipality level. In the difference with the works led in WP2.3.1 which addresses population and household structure response function, the analysis of the spatial distribution of the various household categories, through the housing perspective is connected with the capacity of the model to address the spatial differences in residential behaviours that are related to demographic dynamics, socio-economic context, spatial planning in terms of services and infrastructures, and households preferences.

Moreover specific products are indicators, databases, maps, theoretical framework, graphs and results of statistical regression. These outputs will complete the MOLAND results. Addressees are the scientific community and European policy makers that aim to assess the sustainability of urban growth through the rural-urban relationships and to implement integrated spatial planning policies. A direct connection is considered with SENSOR IP.

Contributions to PLURELs end-products

- PLUREL XPLORER: Fact sheets for response function and impact assessment

Objectives of the deliverable

CEMAGREF team had worked out the conceptual framework to better understand the relations between housing and urban shape, which means for the Plurel project RUR-typologies. Thereafter, they develop an econometric model, being able to quantify the functional relation between land-use as municipality type (urban, peri-urban, and rural) and housing demand.

We adopt a standard approach in urban economics, by nesting a model of household location choice with a simplified life-cycle model. The quantification of response function proceed in two steps: describe the housing demand of each social class (retirees, middle-aged and young) according to 6 plurel RUR-subtypes then validate regional maps with regression analysis of the residential location choice of each social class according to each municipality typology for a RUR-typology, the Montpellier case study NUTS-3 region, which is a polycentric RUR configuration. We use the 1999 Census Data for 6 French NUTS-3 of 6 RUR-types.

Housing demand (census data): the share of housing inhabited by each social class at the municipality level for each RUR-typology

Residential location choice modelling (census data of Montpellier Nuts3 region): Household choice of residential zone**;** municipality sub-types (urban, peri-urban, rural) demographic variables, infrastructures and services availability at municipality level.

Approach of the response function

Population dynamics is central to the development of the residential demand. A comparative analysis over a long period (50 years) of the relationship between population growth and urban growth of fifteen European cities, suggest that population growth over the last 50 years have favoured urban sprawl. Indeed, 90% of new residential areas were located in peripheral areas. Moreover, there is a narrow relationship between the housing demand and the household numbers. From the projections supplied by the World Population prospect, 2007, http://www.un.org/esa/population), because of its relatively weak increase or decline rate, the population of all the developed countries should remain practically unchanged between 2007 and 2050. In European countries, 20 % of the population is already 60 years old or more and this proportion should reach 33 % of the population in 2050. Demographic change will be then one of the key challenges many industrialized countries will face in the future. Globally, at least until 2050, the older population is expected to continue growing more rapidly than the population in other age groups. Such rapid growth will require far-reaching economic and social adjustments in most countries.

The issue of access to housing in conjunction with the new configurations of population structures is one aspect of this adjustment. When the household's size decreases, the number of households increases more quickly than the population. The satisfaction of the demographic demand in housing will require an increase of housing supply. Extending the housing stock will lead to two solutions of land consumption: consume not yet urbanized areas or increase density of areas already built. There is here a real dilemma for public policy because the changing preference of households in terms of housing and residential location has been very space consuming so far. Indeed, in a country like France, there was a preference for individual housing and if possible in areas of low density. That has resulted in urban sprawl and process of peri-urban development. This is not an isolated case. For most European countries, the functional interactions established between urban centres and the surrounding rural areas occur and raise questions concerning the new urban/rural relations and the underlying dynamics.

The main objective of this work is to identify and describe the interdependencies of different factors (variables) which are influencing the interactions between housing and land-use issues. Response functions should be useful in predicting the future development, given that main driving forces are identified and quantified (M1) for that purpose. One important limitation for the use of the response functions as a generic model for all configurations of urban shapes being able to be observed in Europe, the RUR-typologies provided in M2.1 is the availability of detailed data on European level. Therefore we are not able to quantify housing response function for the whole EU region. Instead we provide analytic housing response function, which is based on data at NUTs5 from the Montpellier case study delimited by its Nuts3 RUR, then the "*département*" of Hérault. However, descriptive schemes are provided for five Nuts3 French regions corresponding to each Plurel RUR typology.

We adopt here a standard approach in urban economics, by nesting a model of household location choice with a simplified life-cycle model for the analytic response function. The parameterization of the housing response function is based on residential demand of three categories of the population (ageing population and retired, households with children, young couples and students) for each type of area (urban, rural, peri-urban) at the municipality level. Moreover, we describe relationship between housing demands for each type of RUR with French database obtained from 6 French NUTS-3 areas which correspond to the 6 RUR-type. Main explanatory variables of residential demand are: the six PLUREL municipality typologies (from urban dense to rural with low density), the spatial effect and the time effect.

The report is divided into three parts

I. A theoretical framework to address housing demand and urban shape through a location choice model

II. Spatial distribution of housing within each RUR-typology

III. Housing demand and land-use type at Nuts3 region : location choice modelling in the Montpellier NUTs-3 region.

I. A THEORITICAL FRAMEWORK TO ADDRESS HOUSING DEMAND AND URBAN SHAPE THROUGH A LOCATION CHOICE MODEL

Many empirical studies argue that household migration and location choice differ by age (Nijkamp, 1993; Clark 1987). This is explained by changes in household composition, tastes, size and income caused by the household life cycle phenomenon. Thus, as individuals move through various stages of their life cycle, they constantly adjust their housing needs and make choices in incremental steps (Michelson; 1977, 1980).

Life-cycle position may explain location choices, with the features of different areas meeting the requirements of different categories of population (Clark, Hunter, 1992). In the early years of an individual's working life, professional motives far outweigh residential motives (i.e., attributes of housing and its environment). In an early paper, Beckmann (1973) was interested in the impact of household structure on urban location. He differentiated the workers and non-workers members of a household. This has a particular importance because only the workers are commuting to CBD for work, then labour market conditions was a main driving forces for residential choice location.

In the second stage of the life-cycle, the latter are likely to gain sway as the individual grows older and has a larger family. Labour market variables have less influence on the location decisions of the middle-aged, for whom residential motivations appear to be predominant. In this group, the increase in family size strongly influences location decisions, reflecting the consequences of increasing housing demand on migration toward suburban areas and rural areas under urban influence. This result is in line with Nivalainen's (2002) observation regarding the effect of household size increase on migration towards rural areas in Finland. But we can also find a situation where professional reasons play an important role in the location and, as a result, the urban-rural/urban-rural migration is not known a priori (Détang-Dessendre, 2002). The migration decisions of the 45-64 age-group are clearly residentially motivated (Cribier, Kych, 1992). Some migration is predominantly directed toward rural areas; such is the case for people whose children have left the family nest, for retired individuals. But we can also find households with high enough resources that return to urban centre because of the need for services.

By retiring age, one can suppose that residential location preferences are the only motives that remain. However, for western countries, with the strong desire for mobility among middle-age and older age groups, the spatial distribution of the different social category will end with young people highly concentrated in the core city, and a large percentage of middle and older age groups prefer to move to the intermediate area between the core and the periphery.

If the empirical studies on this topic are relatively abundant, there is a lack of theoretical work in the field. Then, as a first step of our work we suggest a theoretical model that analyzes the location behavior and spatial structure of urban areas, taking account of the life cycle heterogeneity of the population.

In this theoretical framework, one can prove that the ratio between workers and non workers in a household influences the choice of distance from the centre: the lower the weight of commuting costs compared to housing needs, the farther away households locate from the centre. In terms of location, this means for instance that working households experiencing a rise in the number of their children are likely to move to the suburbs. More generally, a decrease in household size should encourage people to move to the centre and vice versa.

I.1. Modeling assumptions framework

The starting point is the standard Alonso-Muth model (Alonso, 1964; Mills, 1967; Muth, 1969). We consider a closed monocentric urban region, with a predetermined centre where all firms are located and around which consumers distribute themselves. In a closed city framework, the population is fixed and the utility level of residents is endogenous.. This framework corresponds to intra-urban migration analysis which, according to many studies, is explained mainly by the life cycle.

As in the standard microeconomic models of urban areas, we postulate that residents commute to the Central Business District (CBD), where they engage in a production process. The system of roads is radial and dense, and the CBD is sufficiently concentrated to consider it as a point. Thus, transportation cost to the centre is proportional to the distance.

The urban region is located in a homogeneous plain and all land in the urban region is owned by absentee landlords. With no topographic features, location choice is reduced to a trade-off between commuting and land consumption. Because the only location variable is the distance from the CBD, we can use a linear representation of the urban area.

For each stage of its life cycle, a household is characterized by a certain level of resources and a different family structure (number of independent people and of children of the household). We consider that all household are composed by $n_{ind} = 2$ independent people (who have an income) and by a variable number n_i of dependant members (children). To take into account the social heterogeneity, we characterize each type of household *i* by two variables: the income y_i and the number of children of the household n_i. Thus, the household size is $n_{ind} + n_i$.

We assume that all members within a household have the same utility function. Thus; type *i* households choose their location so as to maximize the utility of the "representative" member of the family under their budget constraint. The utility of one person is a function of consumption of a monetary good (we assume that the price index for consumption goods is one), which equals the consumption of the household divided by the number of people in the household $z/(n_{ind} + n_i)$, and the part of house consumption corresponding to each member of the family $s/(n_{ind} + n_i)$. We make the standard choice of a Cobb-Douglas form, where, without loss of generality, $\alpha + \beta = 1$:

$$\max_{z,s,x} U_i\left(\frac{z}{n_{ind}+n_i}, \frac{s}{n_{ind}+n_i}\right) = \left(\frac{z}{n_{ind}+n_i}\right)^{\alpha}\left(\frac{s}{n_{ind}+n_i}\right)^{\beta} = \frac{z^{\alpha}s^{\beta}}{n_{ind}+n_i}$$

$$n_{ind}\left(y_i - c_i x\right) = z + R(x)s \qquad (1)$$

where y_i is the income of an independent member of household i, $c_i x$ is his/her commuting cost to the CBD (transportation cost is proportional to distance), and $R(x)$ is the land rent at x. The term $n_{ind}\left(y_i - c_i x\right)$ represents net income of households, which is equal to the total income of the two independent members reduced by the commuting cost. The unitary transportation cost c_i could be different for each household type, due to different levels of income (and thus of time cost) and number of children. The maximization program (1) will give the same solution as the maximization of the household global utility $z^{\alpha} s^{\beta}$ because of a linear transformation of $\dfrac{z^{\alpha} s^{\beta}}{n_{ind} + n_i}$.

I.2. Location behaviour

We define the bid rent function $\psi_i(x)$ of a household at x as the maximum rent per unit land this household is willing to pay for residing there while enjoying the level of utility u_i. Formally

$$\psi_i(x) = \max_{s,z}\left\{ \frac{2\left(y - c_i x\right) - z}{s} \,\middle|\, U\left(\frac{z}{n_{ind} + n_i}, \frac{s}{n_{ind} + n_i}\right) = u_i \right\} \quad (2)$$

A simple calculation gives:

$$\psi_i\left(x\right) = \beta \alpha^{\frac{\alpha}{\beta}} \left(\frac{n_{ind}}{n_{ind} + n_i}\right)^{\frac{1}{\beta}} \left(y_i - c_i x\right)^{\frac{1}{\beta}} u_i^{-\frac{1}{\beta}} \quad (3)$$

This willingness to pay depends on household characteristics (preferences, structure, income, satisfaction) and on location characteristics (accessibility to CBD). One can easily show that the bid rent function of a household is an increasing function of the income of each independent member and of the number of worker members of the family (in our model set to 2). If every independent member earns more or the number of independent members increases, total household income increases. Consequently, with the same level of utility, the household is willing to pay more per surface unit of a house. Bid functions are decreasing functions of unit transportation costs, of utility levels and number of dependents members. If commuting costs increase, the net income left for consumption and housing declines. Therefore, the household will be willing to pay less for the same dwelling. While the utility level of each household member increases, all things equal, this implies a reduction in housing spending. Finally, if a household has more children, with the same level of resources, it must pay less for housing in order to achieve the same level of utility.

Also, the bid function is decreasing and convex compared to the distance to the centre x:

$$\frac{\partial \psi_i\left(x\right)}{\partial x} = -\alpha^{\frac{\alpha}{\beta}} c \left(\frac{n_{ind}}{n_{ind} + n_i}\right)^{\frac{1}{\beta}} u_i^{-\frac{1}{\beta}} \left(y_i - c_i x\right)^{\frac{\alpha}{\beta}} < 0 \quad (4)$$

The bid rent function falls with distance to the CBD, compensating suburban consumers for their high commuting costs.

If a household has no children ($n_i = 0$), one can easily prove that:

$$\frac{\mathrm{d}\psi_{i=0}(x)}{\mathrm{d}n_{ind}} = 0 \;=>\; \psi_{n_{ind}=1}(x) = \psi_{n_{ind}=2}(x), \quad \forall x$$

One can interpret this relationship as follows: if the household is composed only by independent people, the bid rent function is homogenous of degree zero relative to the household size. As a result, singles and couples without children have the same location behavior. For this reason, we will take both households types as a single social class.

Suppose that households are distinguishable only in terms of number of children and all other characteristics (preferences, income, and utility level) are identical. Take two households, *i* and *j*, with n_i and n_j children, such as $n_i < n_j$. As a result of bid functions properties, it is easy to show that $\forall x,\ \psi_i(x) > \psi_j(x)$. Thus, in all locations, households with many children will be out-bided by households with fewer children. Therefore, if a household has more children, it must meet at least one of the following conditions to be able to locate somewhere in the urban area: a higher income, less individual satisfaction, or a change in preferences.

Each rentier is a local monopolist over a unique location. As such, he auctions his land only to those users offering the highest return per unit of land. This process of rent maximization is not unrestricted, however, and must be constrained by several equilibrium conditions that we will see later. The first of these is that all households who possess the same structure, tastes and income, must enjoy the same level of utility regardless of where they locate. Among different types of land use, we consider only residential and agricultural uses. According to standard urban model principles, a given social group will occupy the area of the city where it outbids the other groups for housing, and the equilibrium land rent curve *R(x)* is the upper envelope of the equilibrium bid rent curves of all household types and the agricultural rent line $R(x) = \max_i \{\psi_i(x), RA\}$

As a result, the urban area will be segregated into homogeneous population zones. The segregation boundary between two contiguous areas, $x_s^{i,j}$, is the point where the bid functions of highest bidder in theses zones (*i* and *j*) are equal. The boundary of the city (x_f) is the point where higher urban bids equal rural rent *RA*. Because of decreasing bid-rent functions, clearly no city with an unique CBD can expand beyond x_f.

The bid-max lot size function corresponds to the optimal lot size that the household would choose at *x* if the land rent at *x* were $\psi_i(x)$. This is obtained as solution of the maximization (2):

$$s_i(x) = \beta^{-\frac{\alpha}{\beta}} (n_{ind} + n_i)^{\frac{1}{\beta}} n_{ind}^{-\frac{\alpha}{\beta}} (y_i - c_i x)^{-\frac{\alpha}{\beta}} u_i^{\frac{1}{\beta}} \quad (5)$$

Knowing that at the equilibrium a social group will occupy the zone where the urban rent is equal to their bid function, the lot size function represents also the equilibrium housing demand, expressed in terms of housing size.

As the bid rent, the housing demand of a household depends on its characteristics (preferences, income, structure) and the location characteristics (accessibility). The housing demand function increases with the distance to CBD, transportation costs, the utility level and the number of dependant members of the family.

As we saw previously, within a homogenous population zone, housing prices must decline with distance to ensure that utility is the same for all households of the same group. Since the price of the non-housing consumer good is implicitly assumed to be constant over space, relatively more housing per household is consumed at greater distances from the CBD. It is not surprising that the larger the family, the higher the housing demand. The relation between housing demand and the number of independent members of the household is ambiguous $\operatorname{sgn}\left(\mathrm{d}s_i\left(x\right)/\mathrm{d}n_{ind}\right)=\operatorname{sgn}\left(\beta n_{ind}-\alpha n_i\right)$ and it depends on household preferences and structure.

If the household comprises only independent people, the housing demand function is homogenous of first degree relative to the size of the household (if one multiplies the number of people in the household with a positive constant, housing demand will be multiplied by the same constant). As a result, a couple without children has a housing size twice of a single.

4. The household typology

We will differentiate households depending on the stage of their life cycle. As we saw earlier, singles and couples without children have the same location behavior; consequently, we will only take into account couple households. We distinguish three types of households, based on the life cycle:

- *Young households*, without children $n_1=0$ and $n_{ind}=2$, with:

$$\psi_1\left(x\right)=\beta\alpha^{\frac{\alpha}{\beta}}\left(y_1-c_1x\right)^{\frac{1}{\beta}}u_1^{-\frac{1}{\beta}}$$

$$s_1\left(x\right)=2\beta^{-\frac{\alpha}{\beta}}\left(y_1-c_1x\right)^{-\frac{\alpha}{\beta}}u_1^{\frac{1}{\beta}}$$

- *Middle-aged households with children* : we suppose that, on average, each household has two children $n_2=2$ and $n_{ind}=2$. The bid function and housing demand of this category are:

$$\psi_2\left(x\right)=2^{-\frac{1}{\beta}}\beta\alpha^{\frac{\alpha}{\beta}}\left(y_2-c_2x\right)^{\frac{1}{\beta}}u_2^{-\frac{1}{\beta}}$$

$$s_2\left(x\right)=2^{\frac{1+\beta}{\beta}}\beta^{-\frac{\alpha}{\beta}}\left(y_2-c_2x\right)^{-\frac{\alpha}{\beta}}u_2^{\frac{1}{\beta}}$$

- *Retirees without children* to provide for, with the particularity that they should not commute to CBD to work (for them, transportation costs would be null):

$$\psi_3\left(x\right)=\beta\alpha^{\frac{\alpha}{\beta}}y_3^{\frac{1}{\beta}}u_3^{-\frac{1}{\beta}}$$

$$s_3\left(x\right)=2\beta^{-\frac{\alpha}{\beta}}y_3^{-\frac{\alpha}{\beta}}u_3^{\frac{1}{\beta}}$$

I.4. Social distribution equilibrium conditions

The first implication is that retirees' bid rent function $\psi_3(x)$ is constant over space, the result of null commuting costs. Thereafter, the first condition for the retirees to be located somewhere in the urban area is $\psi_3(x) > RA$. This implies:

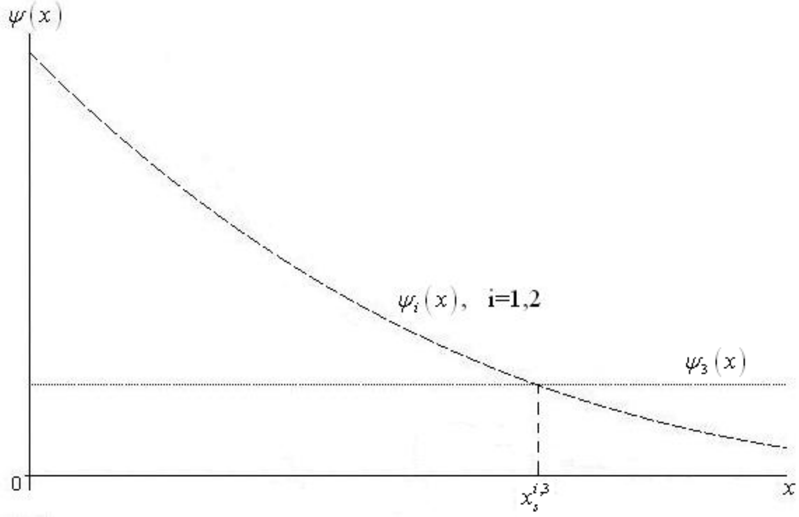
$$\frac{y_3}{u_3} > \alpha^{-\alpha}\beta^{-\beta}RA^{\beta} \qquad (6)$$

Condition (6) provides that retiree income should be sufficiently large so that their bid function (which is a constant) is higher than the agricultural rent. The fact that retirees have a constant bid rent function imposes a limit on the model: if we consider an "open-city" (with no migration costs), there is no spatial limit to the urban area because in all locations retirees outbid the rural land use. Therefore, in reality there are some mechanisms that make the retirees' bid function to start decreasing at a certain distance from the centre (for example, the reduction of public service provision).

As shown by Fujita (1989), if we have a set of continuous, non-increasing bid rent functions, the location of different social groups can be totally ordered according to the relative steepness of their bid functions: the steeper the bid function, the closer to the CBD the location of a group. That is, each type of household will locate in a concentric ring around the CBD. A consequence of constant retirees bid function is that they will locate at the periphery, where they outbid the other two social classes, with decreasing bid functions, and necessarily steeper in absolute value. If we consider $x_s^{1,3}$ and $x_s^{2,3}$ the unique points where retirees bid rent functions equalize, respectively, the young and middle-aged households bid functions, for *i*=1,2 we have the next property:

$$\psi_i(x) > \psi_3(x), \quad \text{for } x < x_s^{i,3}$$
$$\psi_i(x) = \psi_3(x), \quad \text{for } x = x_s^{i,3}$$
$$\psi_i(x) < \psi_3(x), \quad \text{for } x > x_s^{i,3}$$

Figure 1. Retirees location



Now we will take a look at the other two social groups: the young and middle-aged households. Let $x_s^{i,j}$ be the distance where two pairs of bid functions *i* and *j* are equal:

$$x_s^{1,2} = \arg\{\psi_1(x) = \psi_2(x)\} = \frac{2u_2y_1 - u_1y_2}{2u_2c_1 - u_1c_2}$$

$$x_s^{1,3} = \arg\{\psi_1(x) = \psi_3(x)\} = \frac{u_3y_1 - u_1y_3}{u_3c_1}$$

$$x_s^{2,3} = \arg\{\psi_2(x) = \psi_3(x)\} = \frac{u_3y_2 - 2u_2y_3}{u_3c_2}$$

To order the location of “worker” households, we will use the procedure of comparing the relative steepness of the bid rent functions at segregation point $x_s^{1,2}$. If the curve for young households is steeper (more negatively sloped) at $x_s^{1,2}$, then young households outbid middle-aged households for central locations while middle-aged households outbid young ones for more suburban locations. If the "middle-age" curve is steeper, then the reverse pattern is obtained. Thus, for ***i,j***=1,2, we have:

$$\psi_i(x) > \psi_j(x), \quad \text{for } x < x_s^{1,2}$$

$$\psi_i(x) = \psi_j(x), \quad \text{for } x = x_s^{1,2}$$

$$\psi_i(x) < \psi_j(x), \quad \text{for } x > x_s^{1,2}$$

The steepness of the bid rent function is given by the absolute slope in $x_s^{1,2}$:

$$\left|\frac{\partial \psi_1(x)}{\partial x}\right|_{x=x_s^{1,2}} = \frac{c_1\psi_1\left(x_s^{1,2}\right)}{\beta\left(y_1 - c_1x_s^{1,2}\right)}$$

$$\left|\frac{\partial \psi_2(x)}{\partial x}\right|_{x=x_s^{1,2}} = \frac{c_2\psi_2\left(x_s^{1,2}\right)}{\beta\left(y_2 - c_2x_s^{1,2}\right)}$$

Because $\psi_1\left(x_s^{1,2}\right)=\psi_2\left(x_s^{1,2}\right)$, we only need to compare $\frac{c_1}{y_1-c_1x_s^{1,2}}$ and $\frac{c_2}{y_2-c_2x_s^{1,2}}$.

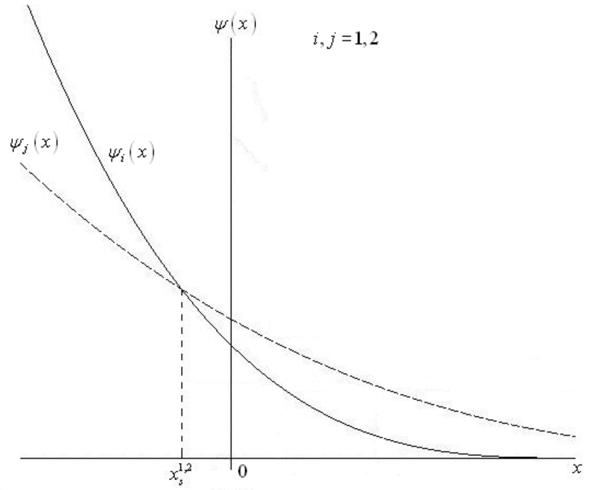
Thus, if $\frac{y_2}{c_2}>\frac{y_1}{c_1}$, the "young" curve is steeper and consequently they will locate in the city centre. We call this spatial structure a Y-A-R (Young-Adult-Retirees) pattern: the central part of the urban area with young households without children, surrounded by an area with middle-aged households with children, and finally the periphery inhabited by retirees. If $\frac{y_2}{c_2}<\frac{y_1}{c_1}$ the location pattern is different, with middle-aged households in the centre, surrounded by young households and then retirees (what we call an A-Y-R structure).

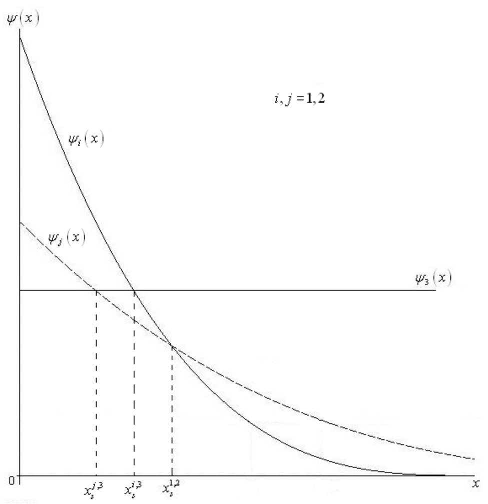
We find a classic result of residential choice that depends on income and commuting cost. Retirees do not commute to work, $c_3=0$, and the ratio y_3/c_3 is infinity, thus they surely locate in the most peripheral part of the area. For the other two social groups, we have two opposite effects. First, there is an income effect: a higher income of each worker member increases total household resources. As result, because of the assumption of normality of housing, high housing consumption by the rich implies an attraction to the suburbs, where housing prices are lower. There is a second effect, however: rich households also have a high opportunity cost of time and thus a high commuting cost per mile. Therefore, the rich value accessibility to the central business district (CBD) more than the poor. The net effect of these forces depends on the ratio of income and commuting cost per mile y_i/c_i.

In our case, this usual result is nuanced by the fact that for a household with many children, even if each worker member earns more, resources per person may be lower compared to households with fewer children. Thus, the first households will behave as being richer than the second ones, even if, in reality, their average income is lower.
To be sure that all three social groups are localized somewhere in the area, there is one necessary and sufficient condition: the segregation border between the young and middle-aged households $x_s^{1,2}$ (regardless of the order of location), must be positive and closer to the CBD than the segregation points with retirees $x_s^{1,3}$ and $x_s^{2,3}$.

If $x_s^{1,2}<0$ or $x_s^{1,2}>x_s^{i,3}$, for $i=1,2$ one of the worker groups will be out bided in all possible (positive) locations:

Figure 2. Equilibrium with only two instead of three social groups





We will now detail those conditions, starting with the constraint of a positive segregation distance between the two worker classes: $x_s^{1,2} = \dfrac{2u_2y_1 - u_1y_2}{2u_2c_1 - u_1c_2} > 0$. We have two situations only: both the nominator and the denominator are either positive or negative. The first possibility is that the two elements are positive:

$$\begin{cases} 2u_2y_1 - u_1y_2 > 0 \\ 2u_2c_1 - u_1c_2 > 0 \end{cases} \Rightarrow \begin{cases} \dfrac{u_2}{u_1} > \dfrac{y_2}{2y_1} \\ \dfrac{u_2}{u_1} > \dfrac{c_2}{2c_1} \end{cases}$$

If the location pattern follows a Y-A-R structure, we saw that $\dfrac{y_2}{c_2} > \dfrac{y_1}{c_1}$ and consequently $\dfrac{y_2}{2y_1} > \dfrac{c_2}{2c_1}$. Thus the necessary and sufficient condition for $x_s^{1,2} > 0$ is:

$$\frac{u_2}{u_1} > \frac{y_2}{2y_1} \qquad (7)$$

If the urban area is characterized by an A-Y-R spatial structure, these inequalities are met $\frac{y_2}{c_2} < \frac{y_1}{c_1} \Rightarrow \frac{y_2}{2y_1} < \frac{c_2}{2c_1}$, which implies as a necessary and sufficient condition:

$$\frac{u_2}{u_1} > \frac{c_2}{2c_1} \qquad (8)$$

A second possibility is when the two terms of $x_s^{1,2}$ are negative:

$$\begin{cases} 2u_2y_1 - u_1y_2 < 0 \\ 2u_2c_1 - u_1c_2 < 0 \end{cases} \Rightarrow \begin{cases} \frac{u_2}{u_1} < \frac{y_2}{2y_1} \\ \frac{u_2}{u_1} < \frac{c_2}{2c_1} \end{cases}$$

For an Y-A-R location pattern, we have $\frac{y_2}{2y_1} > \frac{c_2}{2c_1}$, and the necessary and sufficient condition for $x_s^{1,2} > 0$ is

$$\frac{u_2}{u_1} < \frac{c_2}{2c_1} \qquad (9)$$

The final situation for a A-Y-R configuration appears when $\frac{y_2}{2y_1} < \frac{c_2}{2c_1}$ and the condition for a positive $x_s^{1,2}$ will be:

$$\frac{u_2}{u_1} < \frac{y_2}{2y_1} \qquad (10)$$

We note that the conditions of a positive segregation distance $x_s^{1,2}$ are just reversed, depending on the spatial structure of the city.

Conditions $x_s^{1,2} < x_s^{1,3}$, $x_s^{1,2} < x_s^{2,3}$ and $x_s^{1,2} > 0$, ensure by transitivity the positive sign of $x_s^{1,3}$ and $x_s^{2,3}$. The distance $x_s^{1,3}$ is positive when $\frac{y_1}{y_3} > \frac{u_1}{u_3}$: the income ratio is higher than the utility ratio between young and retirees. Thus, young households must earn more than retirees to obtain the same satisfaction, a result of the assumption of null commuting cost for retirees. In order to have $x_s^{2,3} > 0$, we need $\frac{y_2}{y_3} > \frac{2u_2}{u_3}$, which is even more restrictive than the previous condition, because middle-aged households also have children to provide for and thus must earn twice what retirees make, in order to obtain the same level of satisfaction. This is result is compatible with the fact that increasing age is generally associated with increasing satisfaction (Campbell, 1976).

The last conditions for the presence of all social groups in the city are $x_s^{1,2} < x_s^{1,3}$ and $x_s^{1,2} < x_s^{2,3}$. These conditions are equivalent to $\psi_1\left(x_s^{1,2}\right) = \psi_2\left(x_s^{1,2}\right) > \psi_3\left(x_s^{1,2}\right)$ and if one is met, the other is, too. After calculating and arranging, we get:

$$\frac{c_1 y_2 - c_2 y_1}{2c_1 u_2 - c_2 u_1} > \frac{y_3}{u_3} \qquad (11)$$

One necessary condition is that the left term is positive. Again, we have two possible situations: the nominator and the denominator are both either positive or negative.

If $c_1 y_2 - c_2 y_1 > 0$, then $\frac{y_2}{c_2} > \frac{y_1}{c_1}$, which corresponds to a Y-A-R location scheme. Thereafter, $2c_1 u_2 - c_2 u_1 > 0 \Rightarrow \frac{u_2}{u_1} > \frac{2c_2}{c_1}$. This condition is contradictory with expression (9), and then the only possibility for a positive segregation point $x_s^{1,2}$ in a Y-A-R pattern is when condition (7) is met. Symmetrically, if $c_1 y_2 - c_2 y_1 < 0$, the city is structured following an A-Y-R location pattern, and the denominator of (11) must also be negative. We get $\frac{u_2}{u_1} < \frac{c_2}{2c_1}$, which excludes relationship (8). Thus, we can summarize that a positive segregation $x_s^{1,2} > 0$ is possible when one of next two conditions is met, depending on the location pattern:

- if young households are located in the centre then: $\frac{c_2}{c_1} < \frac{y_2}{y_1} < \frac{2u_2}{u_1}$ (12)
- if middle-aged households are located in the centre then: $\frac{2u_2}{u_1} < \frac{y_2}{y_1} < \frac{c_2}{c_1}$ (13)

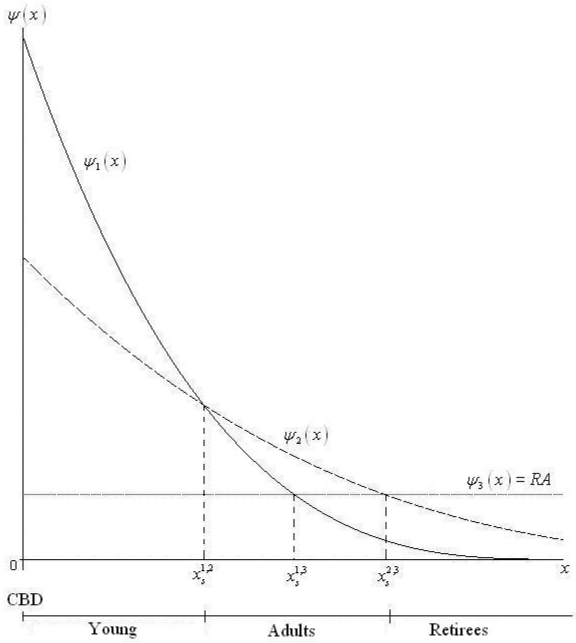
We can now write all these conditions (6), (12) and (13) in one inequality that ensures the presence of all social groups in the area, no matter the location pattern:

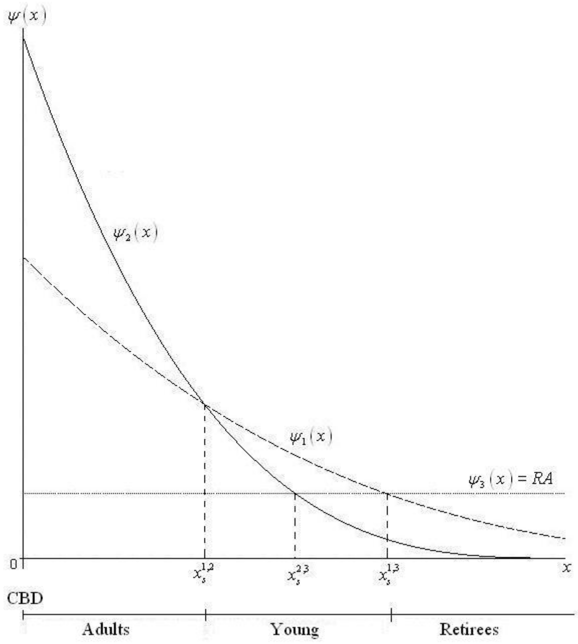
$$\alpha^{-\alpha} \beta^{-\beta} R A^{\beta} < \frac{y_3}{u_3} < \frac{c_1 y_2 - y_1 c_2}{2u_2 c_1 - u_1 c_2} \qquad (14)$$

I.5. Spatial distribution equilibrium conditions

As we have seen previously in this model, with two types of worker households and retirees, there are only two equilibrium spatial configurations: Y-A-R (from centre to the periphery: young households – middle-aged households - retirees) and A-Y-R (middle-aged households – young households - retirees). Once condition (14) is met, we know that there is an equilibrium, with two possible location patterns, depending on the resources- commuting costs rates y_1/c_1 and y_2/c_2.

<u>Figure 4. Location pattern: Y-A-R and A-Y-R structures</u>





Retirees represent a special category because they do not have to commute to the CBD for work. Thus, their bid rent function is constant over space and they are located on the periphery of the urban area. Consequently, the only constraint for retirees is given by the agricultural rent. Since their objective is to maximize their utility, they will minimize their bid function to the lowest possible price, which will be equal to the agricultural rent. Thus, at equilibrium, we have:

$$\psi_3^* = RA \Rightarrow u_3^* = \alpha^{\alpha} \beta^{\beta} RA^{-\beta} y_3 \qquad (15)$$

The housing demand will also be a constant at equilibrium:

$$s_3^* = 2\alpha^{\frac{\alpha}{\beta}} \beta^{\frac{\beta-\alpha}{\beta}} \frac{y_3}{RA} \qquad (16)$$

This situation represents one of the model limits. In the real world, retirees' bid function is influenced by other variables that are not taken into account: the level of public services, urban and natural amenities, etc. The consequence is that they have no spatial interactions with other social groups and their location will always be in the periphery. Because $\psi_3^* = RA$, the segregation points between worker households and retirees will be:

$$x_s^{1,3} = \arg\left\{\psi_2(x) = RA\right\} = \frac{1}{c_1}\left(y_1 - \alpha^{-\alpha}\beta^{-\beta}RA^{\beta}u_1\right) \qquad (17)$$

$$x_s^{2,3} = \arg\left\{\psi_2(x) = RA\right\} = \frac{1}{c_2}\left(y_2 - 2\alpha^{-\alpha}\beta^{-\beta}RA^{\beta}u_2\right) \qquad (18)$$

We suppose that, within the city boundaries x_f, the land is entirely allocated to residential use. Our city being in a perfectly plane area, a straightforward calculation gives the total population, $N = N_1 + N_2 + N_3$, where N_i is the number of type i households:

$$N_i = \int_0^{x_f} n_i(x)\,dx$$

and for all $x \in \left[0, x_f\right]$,

$$n_i(x) = \begin{cases} \dfrac{2\pi x}{s_i(x)}, & \text{if } \psi_i(x) \geq \psi_j(x),\ i, j = 1..3 \\ 0, & \text{else} \end{cases}$$

is the number of households located at distance x from the city centre as the ratio between the amount of land available for residential use at each distance x ($2\pi x$) and the surface occupied by each household.

Usually the equilibrium city's boundary is given by the point where the bid-rent function of the category located in the peripheral area equals the agricultural rent. In our case, retirees have a constant bid-rent function, and thus x_f will be influenced by the surface occupied by the other two social groups and by the number of retirees in the area.

For a given structure of population $\overline{N_1}$, $\overline{N_2}$ and $\overline{N_3}$ the closed city equilibrium is obtained by solving the system:

$$\left\{u_1^*, u_2^*, u_3^*\right\} = sol\left\{N_1 = \overline{N_1},\ N_2 = \overline{N_2}, N_3 = \overline{N_3}\right\}$$

This system of three equations and three unknown variables could be reduced using relation (15) to a system of two equations and two unknown variables $\left\{u_1^*, u_2^*\right\}$:

$$\begin{cases} \int_0^{x_f} n_1(x)\,dx = \overline{N_1} \\ \int_0^{x_f} n_2(x)\,dx = \overline{N_2} \end{cases} \qquad (19)$$

Once u_1^* and u_2^* are known, we can find $x_s^{1,3}$ or $x_s^{2,3}$ from (17) or (18) and depending on the spatial structure we can easily calculate the city's border:

Y-A-R structure:

$$\overline{N_3} = \int_{x_s^{2,3}}^{x_f} n_3(x)\,dx \;\Rightarrow\; x_f = \sqrt{\left(\frac{y_2 - 2\alpha^{-\alpha}\beta^{-\beta}RA^{\beta}u_2^*}{c_2}\right)^2 + \frac{2\alpha^{\frac{\alpha}{\beta}}\beta^{\frac{\beta-\alpha}{\beta}}y_3\overline{N_3}}{\pi RA}}$$

A-Y-R structure:

$$\overline{N_3} = \int_{x_s^{1,3}}^{x_f} n_3(x)\,dx \;\Rightarrow\; x_f = \sqrt{\left(\frac{y_1 - \alpha^{-\alpha}\beta^{-\beta}RA^{\beta}u_1^*}{c_1}\right)^2 + \frac{2\alpha^{\frac{\alpha}{\beta}}\beta^{\frac{\beta-\alpha}{\beta}}y_3\overline{N_3}}{\pi RA}}$$

I.6. A numerical illustration within a monocentric urban region

The system (19) cannot be calculated analytically and we will use numerical simulations to analyze the impact of demographic changes on the city. We follow two scenarios corresponding to each equilibrium spatial structure. The parameters are identical, except the commuting cost for middle-aged households, which is the key feature of each location pattern.

We set the number of households of each type at 50,000 and then we let the exogenous population levels vary from 10,000 to 100,000, only for young and middle-aged households. Our model does not allow interactions between retirees and worker households. Consequently, the only result of an increase in the retiree population will be an external enlargement of the peripheral zone, with no effects on other social groups nor on the housing market.

Table 1 : Values of the parameters used for the simulations

Parameter	Value
Share of non land expenditures (α)	0.6
Share of land expenditures (β)	0.4
Income of young households (y_1)	100
Income of middle-aged households (y_2)	200
Income of retiree households (y_3)	50
Unit transportation cost for young (c_1)	1
Unit transportation cost for middle-aged (c_2): YAR/AYR pattern	1,5/2,5
Unit transportation cost for retirees (c_3)	0
Agricultural rent (RA)	50

First we look at the impact of a change in the social composition on the spatial distribution of households for the two equilibrium structures (Appendix 1.1). In the case of an increase in the youth population, if they live in the centre, the central zone will enlarge, thus creating more competition for land and an expansion of all zones to the periphery. For retirees, the only impact will be a location even farther from the CBD, but with no effect on their satisfaction or on housing demand.

If young households live in the first ring, an increase in the young population will be followed by an enlargement toward the periphery but also to the city centre. Thus, middle-aged households suffer from this pressure, which will have impacts on prices, housing demand, consumption and consequently on their wealth. Once again, retirees are just changing the spatial position with no other effect.

In the case of a variation in the middle-aged population, the shifts are identical, but with a bigger impact on the spatial structure of the urban area. This is explained by the fact that middle-aged households, with more resources but also with dependent members, have a higher demand for housing than the young households.

We saw previously that an increase in social group size implies spatial interactions with the surrounding areas, which will be followed by pressures on the housing market. The competition for land becomes higher, which will increase land prices (Appendix 1.2) and consequently reduce housing demand (Appendix 1.3). An increase in the number of middle-aged households has more important consequences on equilibrium housing prices than the variation in young households. Prices are increasing everywhere in the two central zones (not only in the zone occupied by the population that suffers from demographical changes) but with less intensity. In Appendix 3, one can see that the impact of a social group increase is higher when it is localized in the residential ring, where housing demand is more important. As in the previous figures, the amplitude effect of middle-aged population variation is higher than for young population change. All these changes on the land market will have an impact on household welfare, except for retirees, who in our model do not interact spatially with other social groups.

The "general rule" is that when the number of households of a given type increases, then utility decreases for all households, but with a higher effect on the same given social group type (see Appendix). Middle-aged households are more impacted than young ones by the effect of their own demographic growth (see right-hand side of figures a) and b) from Appendix 1.4): the variation of middle-aged households' utility is less symmetrical than in left-hand side.

II. SPATIAL DISTRIBUTION OF HOUSING WITHIN EACH RUR-TYPOLOGY

Land-use typology at municipality level

One of the main tasks of this work is to learn about how the urbanism degree of a location (urban, periurban and rural) influences the residential choice of households. We will use the typology of PLUREL project (Peri-urban Land Use Relationships: Strategies and Sustainability Assessment Tools for Urban-Rural Linkages), proposed by Loibl et al. (2008) and Steinnocher et al. (2008). The authors established a typology of European Rural-Urban Regions – RUR (clusters of NUTS-3 regions) which are subdivided into urban, peri-urban and rural subregions, each of this type being divided on population density (high and low density) The typology is obtained using a high resolution population density map and CORINE land cover grid map. Their main output is a map with seven types of land use (combinations of three urbanism degree and two population density, plus areas without population, such as rocks, lakes, glaciers, etc…) at a 100x100 m level.

Because we are interested only about the urbanism degree of a location, we will use the "simplified" typology (urban, peri-urban, rural and areas without population) as input in our model. The available French data are at a municipality level, so first we need to aggregate this classification. Knowing that in the same municipality territory we can find several land use types, we exclude the non-habitable areas and then we define "dominant" codes as land use types that cover at least 1/3[1] of the municipality surface. If a municipality is characterized by more than one dominant code, we give priority to more urban types. We need also to take into a count some specific factors such as: municipality population (if more than 20.000, is considered automatically urban), spatial contiguity (the case of some very little rural municipalities, which are surrounded by urban or peri-urban zones).

Inputs (PLUREL sub-types)

- A data set based on the CORINE Land Cover 100x100 m grid

- 6 (populated) land use types at CLC level (PLUREL Typology)

Output: 6 types at municipality level, using the following algorithm

- municipalities with population > 20.000 and with CCL111 are directly considered of type 1 (urban high density)
- municipalities with population > 20.000 and without CCL111 are directly considered of type 2 (urban low density)
- we assume that a municipality could be characterized by several dominant codes

[1] This number is choose by the fact that in each municipality we can have maximum 3 land use types

- a code is dominant if the corresponding surface inside a municipality is bigger than the proportional part of the total surface

- if a municipality is characterized by several dominant codes, we give priority to the most urban (the lower density code)

– we assume that a municipality could be characterized by several dominant codes

– a code is dominant if the corresponding surface inside a municipality is bigger than the proportional part of the total surface

– if a municipality is characterized by several dominant codes, we give priority to the most urban (the lower code)

After these manipulations, we obtain an urban-periurban-rural typology at municipality level. The comparison between the land use typology at low scale spatial level and municipality level is reported in appendix for each RUR-Typology:

RUR 1.0 : Monocentric Very large ; RUR 1.0 : Monocentric large; RUR 1.0 : Monocentric medium

RUR 2 : Urban Polycentric; RUR 3 : Dispersed Polycentric; RUR 4: Rural

II.2 Spatial distribution of the housing inhabited by each social category

One of the main objectives of this task is to explain the location choice of households at different steps of their life cycle. Thus, we will consider the age of the households as differentiable variable of social groups. We classified households in 4 groups: young (age of household's head between 15 and 29 years), middle-age (age 30-59 years), "young" retirees or retirees1 (age of 60-74 years) and finally retirees2 which are older than 75 years. As we will see later, Retirees 1 have a very similar behaviour as middle-aged class. This is explained by low age limit of this category (60 years) while in France, the "normal" age of retiring is 65 years. For convenience, we called them Retirees 1, but we will keep in our mind the fact that a part of this population is still employed. The distinction of retirees in two classes is necessary, because the location behaviour and housing, amenities and public services needs of these households could change with age.

Then to have some address relationship between housing demands of each social class within each RUR-typology, we establish a statistical regression equation between the share of housing inhabited by each social class (1999 Census Data) and the municipality type (6 PLUREL subtypes, from urban dense to rural with low density). These regressions do not provide exact causality relationship between the two variables as predicted by household location choice. They just show us covariation relationships between the two variables. Do we find more housing inhabited by young households in the city? The results for each social category are provided in the following tables (2,3,4) provides a summary of the covariation relationships based on statistical regressions provided in the appendix).

Table 2 : Share of housing inhabited by young households

	RUR 1.0	RUR 1.1	RUR 1.2	RUR 2	RUR 3	RUR 4
Subtype 1 : urban high density	++	+++	+++	+++	+++	
Subtype 2 : urban low density			+++	++		
Subtype 3 : periurban high density		+		+		
Subtype 4 : periurban low density		+			+	
Subtype 5 : rural high density					+	+
Subtype 6 : rural low density						

Note : +++ : very high positive covariation; ++: high positive covariation;

+ : positive covariation

Table 3: Share of housing inhabited by middle-aged households

	RUR 1.0	RUR 1.1	RUR 1.2	RUR 2	RUR 3	RUR 4
Subtype 1 : urban high density						
Subtype 2 : urban low density						
Subtype 3 : periurban high density						
Subtype 4 : periurban low density	++			++	+	
Subtype 5 : rural high density	++		++	+	+	
Subtype 6 : rural low density		+++	++		+++	+

Note : +++ : very high positive covariation; ++: high positive covariation;

+ : positive covariation

Table 4 : Share of housing inhabited by retired households

	RUR 1.0	RUR 1.1	RUR 1.2	RUR 2	RUR 3	RUR 4
Subtype 1 : urban high density	++			++		
Subtype 2 : urban low density	++	+++	++	++		
Subtype 3 : periurban			+++	+++		

high density						
Subtype 4 : periurban low density	+		++	+		
Subtype 5 : rural high density	+			+	+	
Subtype 6 : rural low density						

Note : +++ : very high positive covariation; ++: high positive covariation;

+ : positive covariation

III. HOUSING DEMAND AND LAND-USE TYPE AT NUTS3 REGION : LOCATION CHOICE MODELLING

III.1. DATA AND DESCRIPTION OF STUDY AREA

The ambiguity of the notion of urban region at Nuts3 level for French case

These various considerations are to be taken into account in the choice of the spatial perimeter of the urban region to choose. Montpellier is the capital city of the administrative region Languedoc-Roussillon (Nuts 2) and Hérault *département* (Nuts 3).

The urban area of Montpellier recorded the strongest urban demographic progression between 1990 and 1999 with a growth rate of 8,4 % against 2,6 % for the average of the 15 greater agglomerations of France. Since the 60s the population of Montpellier more than doubled. Up to 2020 the region is facing a prognosis of a growth of 6.000 inhabitants a year. Employment in the urban area of Montpellier is largely concentrated in the urban core, the city of Montpellier. In 1999, on a total of almost 172,000 employments, 65% are in Montpellier.

The population increase materialises through the horizontal process of urban sprawl: with 22% of its buildings built after 1990, it presents the highest rate of recent buildings for cities of more than 450,000 inhabitants. The extension of the built surface is done by an expansion of its limits and also by a densification of the space already dense.

However, we have to extend our spatial frontier of our analysis to the Nuts3 level which is the Département de l'Hérault. This Nuts3 region is a very contrasted area which could be divided geographically in two parts: mountains in north, plains and see in south. This particular topology implied also strong economic and demographic differences, with the north with low population density and weak economic activity and in contrast the south where we find several urban agglomerations, namely Montpellier (about 250.000 in the core city and 600.000 inhabitants in the agglomeration) Béziers (125.000 inhabitants) and Sète (with more than 60.000 inhabitants)[2]. Then we cans consider here that le department

[2] 1999 French Census Data

de l'Hérault is a polycentric urban configuration and a monocentric large configuration in reference to Steinnocher et al. (2008). As they mention in their report, the application of RUR-typology to French regions is cautioned.

Figure 1. Land Use Typology at Corine Land Cover Level

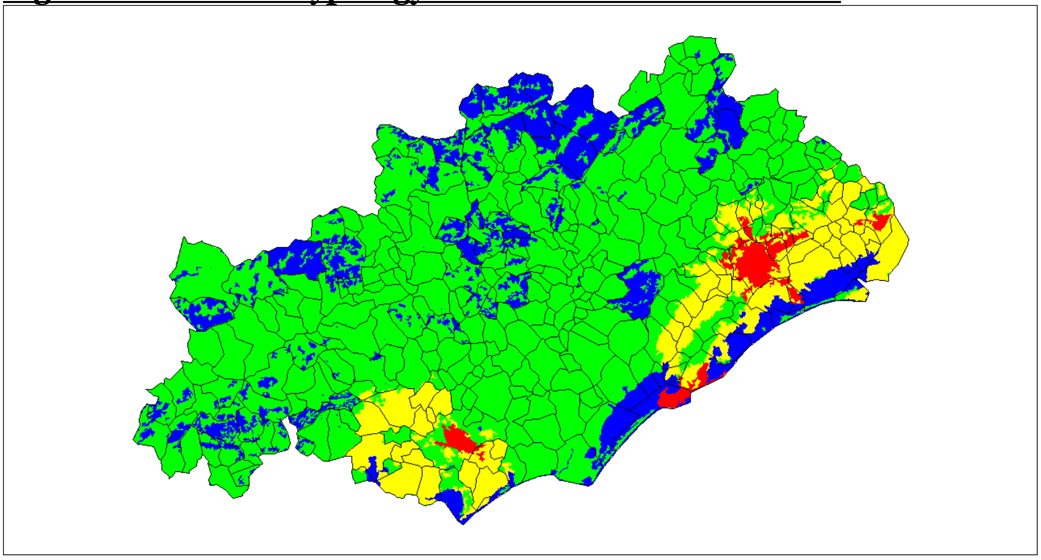
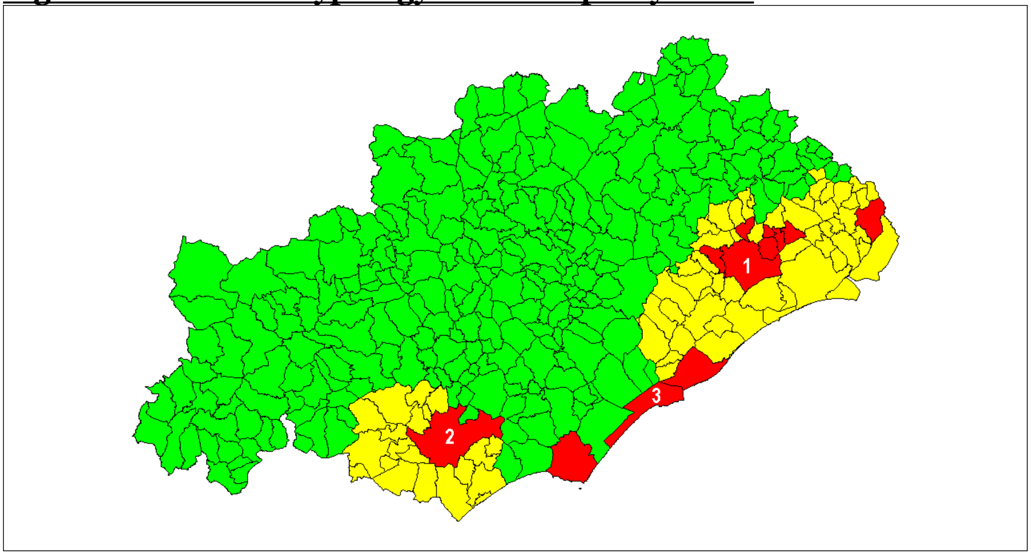
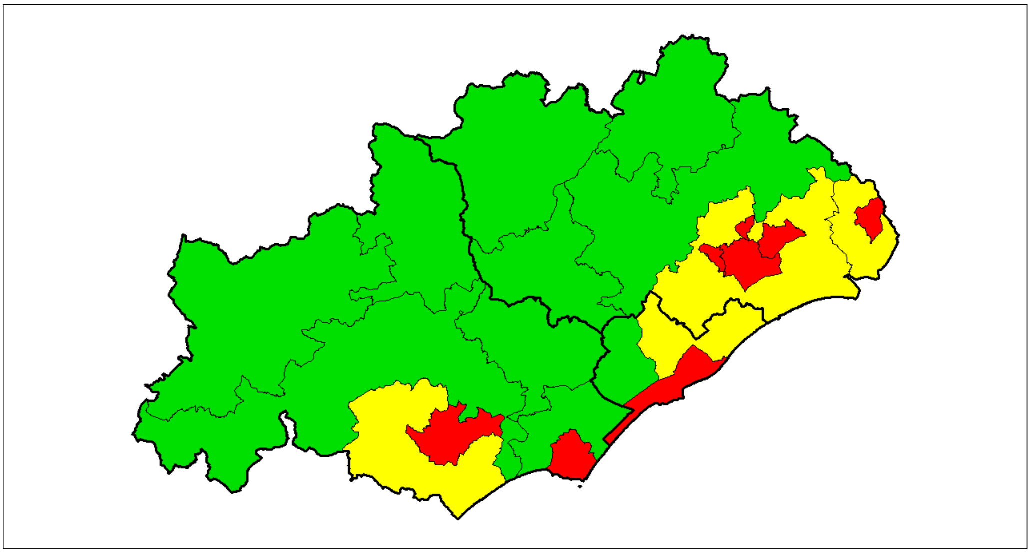


Figure 2. Land Use Typology at Municipality Level



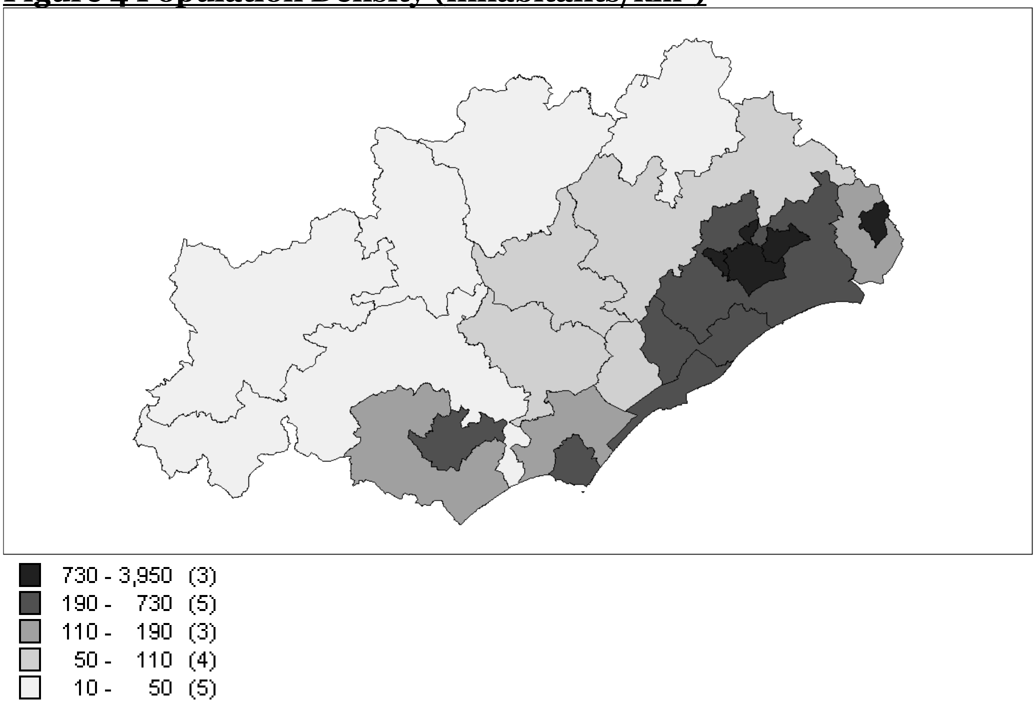
We consider that households choose their location in a two-stage process. They are choosing first one of the residential destination between three urban regions corresponding to the three centre of the Nuts3: Montpellier, Béziers and Sète. Then, inside each macro-zone, they choose a specific municipality destination according to their urban degree.

Figure 3: Spatial units after aggregation



III.2. Spatial distribution of the households

Figure 4 Population Density (inhabitants/km²)



In Table 1, we present the "top 5" of more populated zones of the area, which have more than 20.000 households.

Table 1: Most populated zones of the area

Zone	Households	% of total
Montpellier Core	112008	29.21

City		
Montpellier Peri-urban	54330	14.17
Béziers	31560	8.23
Sète-Frontignan	25675	6.70
Béziers Peri-urban	20039	5.23
Total Area	**383405**	**100**

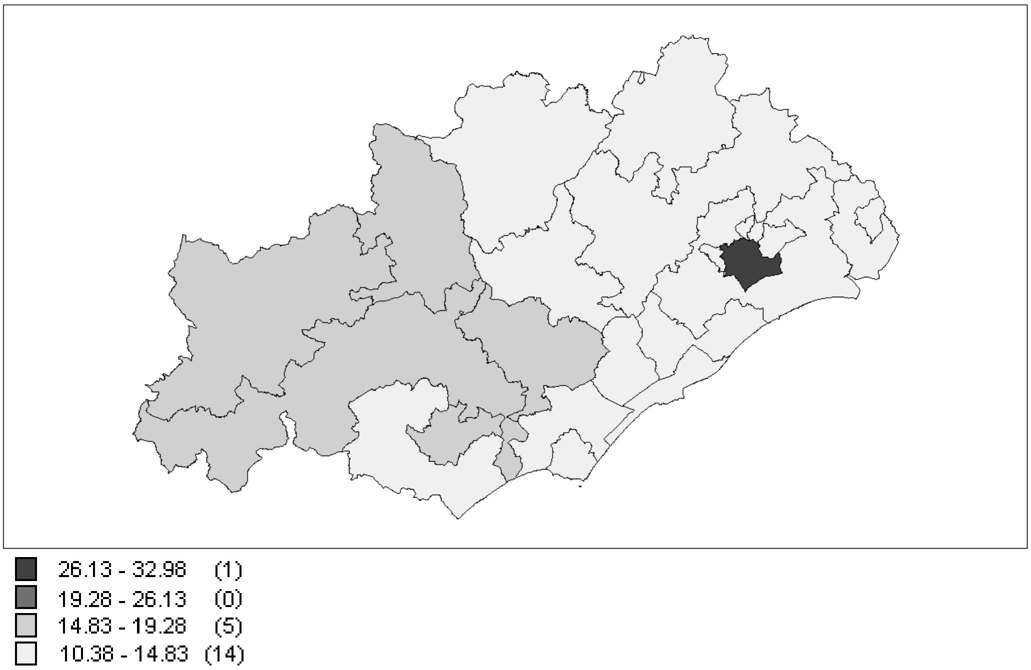
In Table 2, we are presenting some basic statistics of relative distribution of households by age. Excepting the middle-age households, which represent 43.45% of total households, the other three social groups are almost equally represented in the region. But, looking forward, we can see that the households have different location behaviour in their life cycle.

Table 2: Spatial distribution of households by age groups

	Young	**Middle-Age**	**Retirees1**	**Retirees2**
Area average	19.28	43.45	18.03	19.25
Max by zone	32.98	53.21	21.76	26.26
Min by zone	10.38	38.53	13.31	13.41
Standard Deviation	4.68	4.62	1.67	3.48

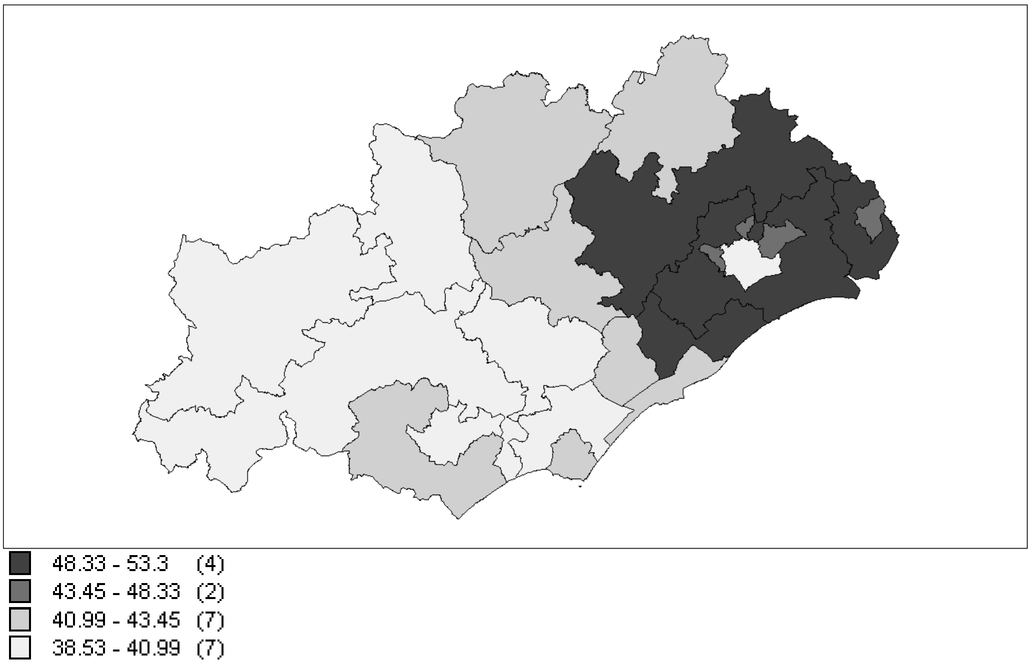
The young households are more unequally distributed than the other classes and in Montpellier Core City, almost a third of total households are young, compared with the area average which is less than 20%. This fact makes that in all other spatial units, the proportion of young households is less than the area average (see Figure 5). This confirm the fact the young prefer central locations, in proximity of employment but also urban amenities. In the same time, usually the young households are singles or they don't have children, and thus, their housing needs are weak, which corresponds with the housing offer in core cities.

Figure 5 Spatial Distribution of Young Households



Middle-age households present a complete different location pattern, with a strong preference on preferences on periurban and rural locations around the regional city centre (Figure 6). In these zones, more than half of households are middle aged. This fact is once again “natural” because, this social category still needs proximity to employment centre and urban services, but very often they are bigger families with higher housing needs, which may be satisfied in periurban and rural locations.

Figure 6 Spatial Distribution of Middle-Aged Households



In the case of retirees, we can see in Figure 7 and Figure 8 that the minimum share of this social group in total households is in Montpellier, which confirms the fact that they are no more linked to the employment centre and thus other variables represent the keys factors in residential choices, such as public services (health, social, cultural) and natural amenities. The retirees represent a little the opposite case of the young households: because of lowest representation in core city, in almost all the other spatial units the share of retirees is higher than the regional average (even completely for retirees1). If the “young” retirees are still massively located in the periurban zones of Montpellier, the most aged households are avoiding strongly the proximity of the most important urban centre.

<u>Figure 7 Spatial Distribution of Retirees1</u>

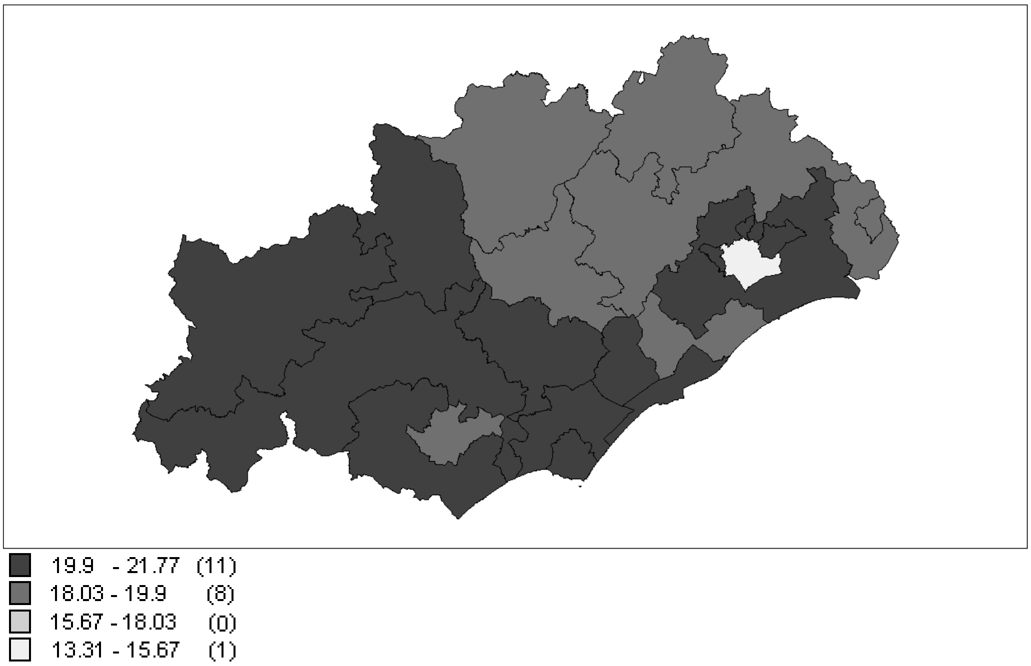
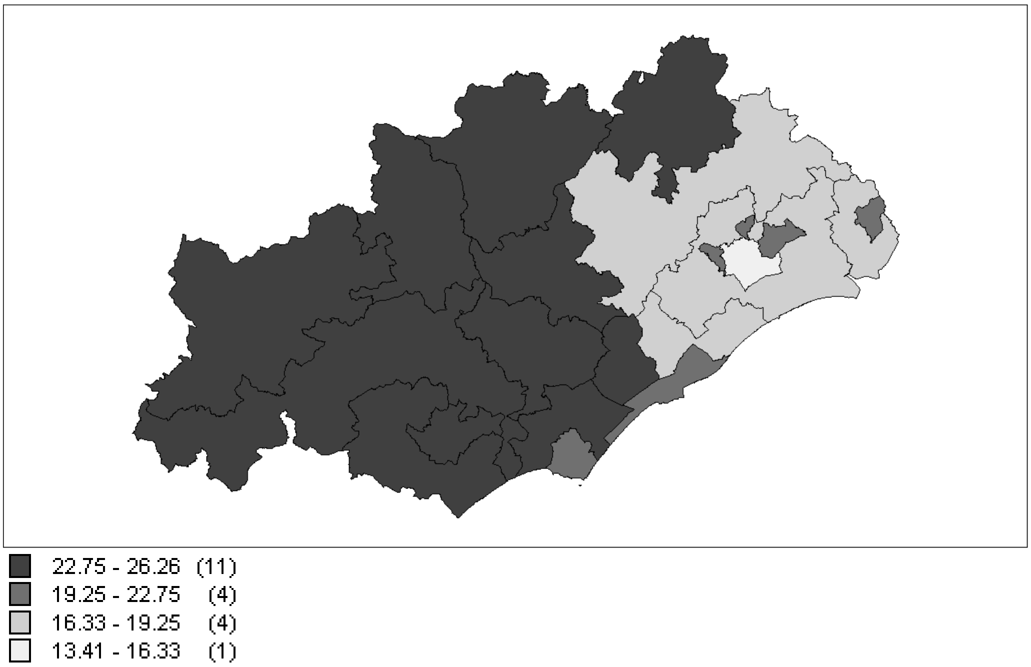


Figure 8 Spatial Distribution of Retirees2



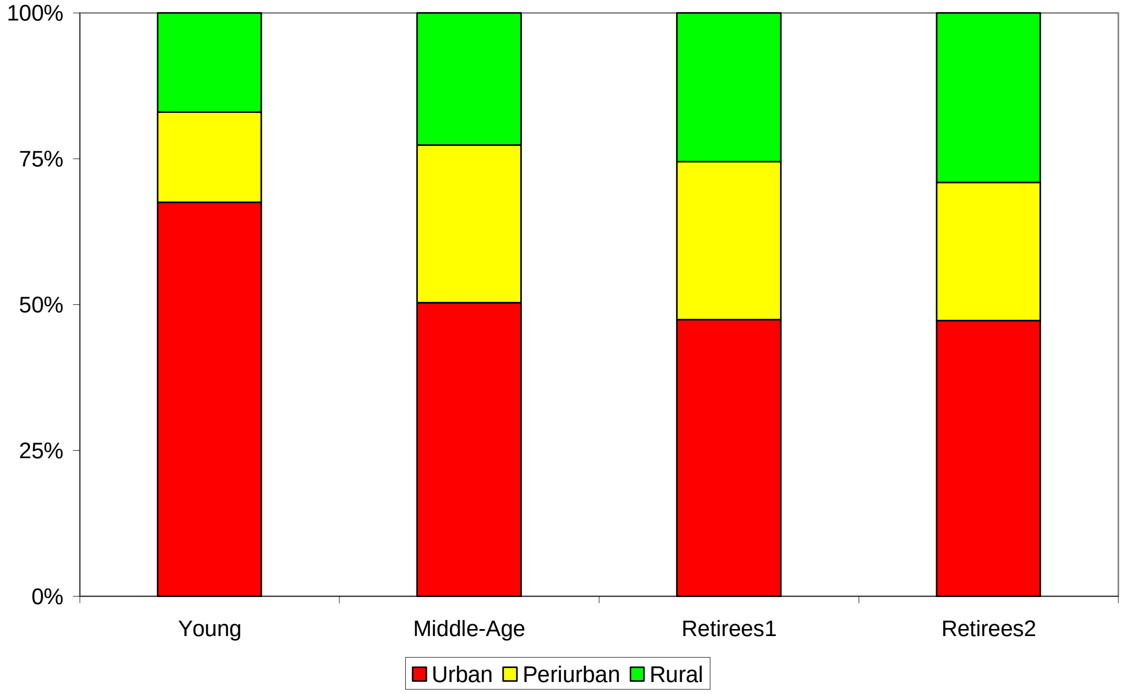
A complementary image is given by table 3, where we show the proportion of each social class which who lives in the most inhabited zones of the area. As we described previously, the young are very concentrated in Montpellier City: 50% of young households are located in this city, compared with 27% of middle-aged and about 20% of retirees. Then, lower age as the households, faster decreases the rate.

Table 3: Households representation in most inhabited spatial units

Part of the group located in	Young	Middle-Age	Retirees1	Retirees2
- Montpellier	49.98	27.10	21.57	20.36
- Montpellier Peri-urban	9.29	16.24	16.10	12.57
- Béziers	7.32	7.76	8.65	9.81
- Sète-Frontignan	4.78	6.59	7.71	7.91
- Béziers Peri-urban	3.46	5.14	5.82	6.64

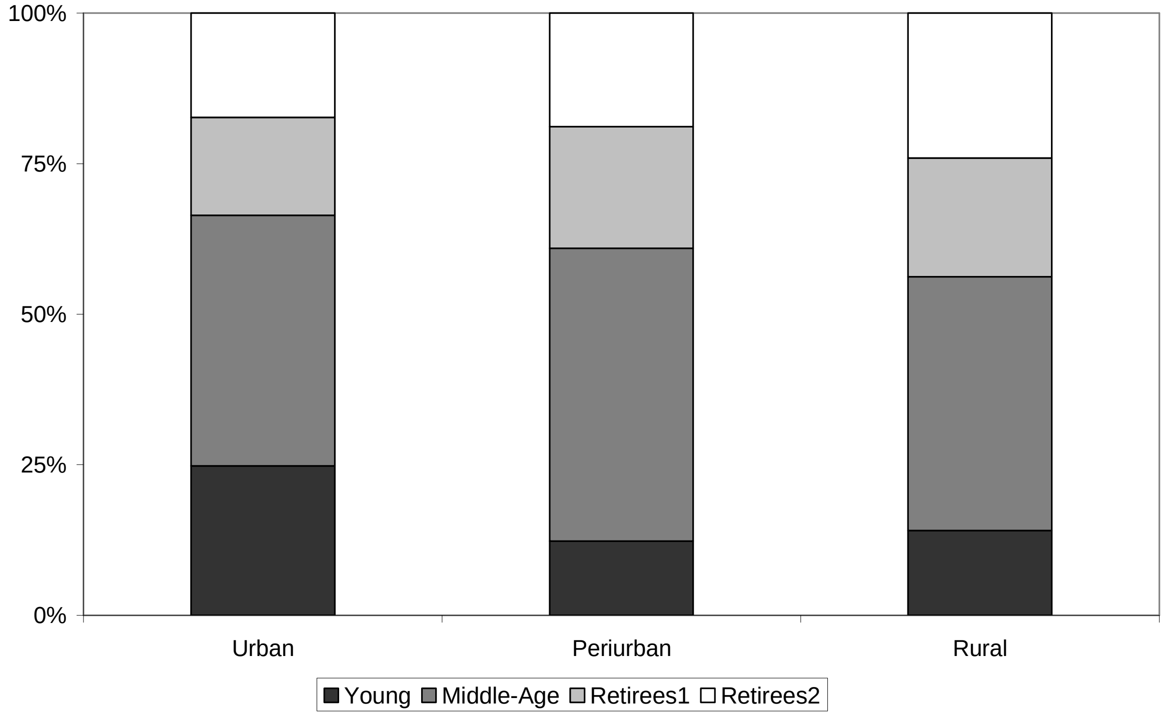
In Figure 9 we present households distribution between urban, periurban and rural zones. Almost 70% of young households live in urban zones, while the rest is equally partitioned between rural and peri-urban zones. For the other categories, half of the population (even less for retirees) lives in urban areas. If middle-aged households and first retirees are more often located in periurban than rural zones, for older retirees we have an opposite situation.

Figure 9. Household distribution between urban, periurban and rural zones



A symmetrical image is given by Figure 10, where we show social composition of urban, periurban and rural zones. Thus, young represent a quarter of urban population and less then 15% in other areas. Middle aged household proportion is highest in periurban zones (almost a half of population), while in urban and rural, slightly above 40%. All retirees are most underrepresented in urban zones, while first retirees represent maximum proportion in periurban zones (as middle-aged households) and oldest category in rural areas (almost a quarter of population of these zones).

Figure 10. Social composition of urban, periurban and rural zones



III.3. Spatial cohabitation of social categories

An interesting perspective on households' spatial distribution is given by residential segregation indices. These indices were developed by sociologists in the United States to measure the spatial segregation of ethnic and racial groups. Usually segregation indices are grouped according to five dimensions: evenness, exposure, concentration, clustering, and centralization. For each dimension, three types of indices are generally identified: one-group indices that measure a group's distribution compared to the entire population; intergroup indices that compare a group's distribution with that of another group; and multigroup indices that compare the spatial distribution of several groups at once. We will not present multigroup indices are used especially to compare social segregation between different areas.

Authors such as Massey and Denton (1988), Massey et al. (1996), Apparicio (2000), Hutchens (2001), Reardon and Firebaugh (2002), Reardon and O'Sullivan (2004) and Apparicio et al. (2008) have reviewed the literature on segregation indices, including the formulas and properties of each. It is not our goal here to make a detailed analyze of these statistical measures, but just to give an alternative perspective on households spatial distribution. This is the reason that we will make just on overview of the indices formulas (which are reported in Appendix 1) and of the results.

a) Evenness indices

Evenness indices refer to the distribution of one or more population groups and measure a group's over or under-representation in the spatial units. The more unevenly a population group is distributed across these spatial units, the more segregated it is. Generally, the indices have a value between 0 (perfect spatial distribution) and 1 (maximal segregation).

Table 4: One-group evenness indices of households' distribution

	SI **Duncan**	***SI(adj)*** **Morill**	***SI(w)*** **Wong**	***SI(s)*** **Wong**
Young	0.2572	0.2313	0.2436	0.2488
Middle age	0.0751	0.0366	0.0589	0.0667
Retirees1	0.0933	0.0819	0.0883	0.0903
Retirees2	0.1355	0.1118	0.1249	0.1297

All evenness indices are convergent, even if we take into account spatial interactions: segregation index adjusted for tract contiguity *SI(adj)*, for contiguous tract boundary lengths *SI*(*w*) and contiguous tract boundary lengths and perimeter/area ratio *SI*(*s*). They show that the young households are the most segregated social group. This is the result of the fact that they are very concentrated in Montpellier City. In the opposite situation, we find the middle-aged households, because of two reasons: they represent the majority in the area and also they have a uniform spatial location.

Table 5: Intergroup evenness indices of households' distribution

		Young	**Middle Age**	**Retirees1**	**Retirees2**
DI **Duncan**	**Young**	0	0.2288	0.2841	0.2962
	Middle Age	0.2288	0	0.0694	0.1263
	Retirees1	0.2841	0.0694	0	0.0605
	Retirees2	0.2962	0.1263	0.0605	0
DI (adj) **Morill**	**Young**	0	0.2029	0.2582	0.2703
	Middle Age	0.2029	0	0.0309	0.0878
	Retirees1	0.2582	0.0309	0	0.0492
	Retirees2	0.2703	0.0878	0.0492	0
DI (w) **Wong**	**Young**	0	0.2152	0.2704	0.2825
	Middle Age	0.2152	0	0.0533	0.1102
	Retirees1	0.2704	0.0533	0	0.0555
	Retirees2	0.2825	0.1102	0.0555	0
DI (s) **Wong**	**Young**	0	0.2204	0.2756	0.2877
	Middle Age	0.2204	0	0.061	0.1179
	Retirees1	0.2756	0.061	0	0.0575
	Retirees2	0.2877	0.1179	0.0575	0

If one dimensional evenness indices take into account only the spatial distribution of each group independently, the intergroup indices have the advantage that can measure also the interactions between social categories.

The dissimilarity index of Duncan & Duncan confirms the fact the young are the most segregated class, but have interesting additional information: more aged are other social groups, more spatial dissimilarity (segregation) appears. This means that when a household get older, the locations that is looking for and the reasons of the choice are changing. On contrary the two retirees groups have very similar location pattern. Middle-age households present also an important location similitude with young retirees but not with the most aged households.

It happens often that the results of dissimilarity index to be nuanced or even completely changed when we take into account spatial units aspects such as tract contiguity (Morrill *DI*(*adj*)), boundary lengths (Wong *DI*(*w*)) and tract boundary lengths and perimeter/area ratio (Wong *DI*(*s*)). In our case, even if absolute values are not the same, the relative "position" remains the same, which confirms the previous results.

b) Exposure indices

Exposure is the degree of potential contact between members of the same group or between members of two groups inside spatial units and measures the probability that members of one group will encounter members of their own group (isolation, *xPx*) or another group (interaction, *xPy*) in their spatial unit. The isolation index adjusted (*Eta2*) modifies the isolation index with the proportion of the group in the area to avoid the effects on the index of the population composition. The difference between these two indices appears very clear in Table 6. In the isolation index case, the middle-age households seems most isolated (highest probability to share same spatial units with the same social category), but this is just the effect of the fact that this group represents the majority in the area. Once the effect of the population composition removed, we see that the isolation index adjusted confirm the fact that the young households are the most segregated, contrary to middle-age households.

Table 6: Households one-group exposure indices

	xPx	***Eta2***
Young	0.2342	0.0513
Middle age	0.4388	0.0077
Retirees1	0.1856	0.0065
Retirees2	0.2035	0.0136

The interaction index (Table 7) shows that all households have high chances to share same spatial units with middle-aged households (second column), which represent majority in the area. For the other social groups, wee see that probability to share the same spatial unit varies very little.

Table 7: Households interaction index

		Young	**Middle**	**Retirees1**	**Retirees2**

		Age			
xPy	**Young**	0.2342	0.4227	0.1664	0.1767
	Middle Age	0.1876	0.4388	0.1817	0.192
	Retirees1	0.1779	0.4377	0.1856	0.1987
	Retirees2	0.177	0.4334	0.1862	0.2035

c) Concentration, clustering and centralisation indices[3]

Concentration refers to the physical space occupied by a group. Social groups of the same relative size occupying less space would be considered more concentrated and consequently more segregated. Thus, concentration indices, compared to evenness indices present the advantage that they take into account also the area of spatial units, which improves the quality of measure. The most known uni-dimensional concentration indices are reported in Table 8.

Table 8: Households one-group concentration, clustering and centralisation indices

	Delta	*ACO*	*ACL*	*ACE*
Young	0.637	0.755	0.1447	0.6030
Middle age	0.5459	0.5301	0.2649	0.4919
Retirees1	0.512	0.6164	0.0769	0.4230
Retirees2	0.4809	0.6084	0.0702	0.3616

The results show that the young households are not just spatially the most segregated social category but also the most concentrated. Once again, this is the result of the fact that an important part of these households lives in Montpellier

According to the clustering dimension of segregation, more contiguous spatial units a group occupies, forming an enclave within the area, more clustered and therefore segregated it is. Absolute clustering index (ACL from Table 8) shows a complementary image of previous indices. Thus, the middle-aged class seems more clustered, because of the fact that the spatial units where they are concentrated (the periurban and rural periphery of Montpellier) are contiguous.

Finally, centralization indices measure the degree to which a group is located near and in the centre of the metropolitan area. It is not a surprise that absolute centralisation index has highest value for young households (ACE in Table 8). Another usual result emerges: oldest is a household, the more chances there are to choose a distant location from urban centre. The importance of Montpellier

[3] We will analyze only one-group indices for those dimensions of spatial segregation, but for all indices, there are also two-group versions, but they just confirm the one-dimensional indicators and it is not our goal to make a detailed analysis of segregation indices.

city is highlighted by this index because it has only positive values for all age groups (tendency for centralisation location).

III. 4. HOUSING DEMAND RESPONSE FUNCTION

4.1. Methodology

At this step, the objective of our work is to describe a quantified relationship between housing demand and land-use for a specific RUR-type, the polycentric one with the Montpellier study case Nuts3 level.

For that, we analyze residential choice in using a nested logit model. In the first stage, they are choosing one of the three macro-zones of the area, which corresponds to the zone of influence of the most important municipalities: Montpellier, Béziers and Sète. Then, inside each macro-zone, the households are choosing one of three land-use type according to their urban degree.

As we discuss in first section, housing preferences are strongly related to life-cycle stage. We then differentiate housing function social class: young, middle and aged households split into two categories of retirees). Young households, often singles or couples without children, have more career and consumption oriented life style and thus may have a greater desire to central locations, to be close to work, shopping, entertaining, etc…Middle-aged households are often families with children, so with higher needs for living space. In the same time, this category is linked to employment, because the adult members of the households are working. These characteristics push middle-aged households to locate in periurban or rural zones in proximity of urban centres. Finally, retirees represent a different situation, being no longer in activity and the reasons for choosing a location are very often linked to life quality, natural amenities and public services such as social and health facilities.

Housing units are more complex than most of other commodities, because they are durable and they are linked to the location, providing a multitude of services and local public goods. Thus, the location choice and housing demand represent finally the same process: choosing a house, a household makes also a residential choice, and vice versa.

This is one of the reasons why there is a substantial and rich body of literature that developed and used discrete choice models to analyze demand for housing and housing associated goods.

Regarding our Nuts3 region, we consider the following choice sequence: in the first step the households choose on of three urban areas destination in the region (Montpellier, Bézier and Sète) and then in a second step they are choosing as residential location one of the RUR sub-type.

The two-stage optimization models are frequently used in the housing demand literature (e.g. Pollakowski 1982, Huff 1986, Chahill 1994, Tu and Goldfinch 1996) the most common are being nested multinomial logit models. Then, we assume that all households have the same full choice set which contains all combinations of location choices and each household possesses perfect information. A household n will choose one and only one residence; as a rationally seeking to optimize the utility U_n^i received from choosing alternative *i*. Let the set of alternatives *i* be partitioned into *K* nonoverlapping subsets denoted B_1, B_2,..., B_K and called nests. The utility that a household *n* obtains from alternative *i* in nest B_k is denoted in the usual manner $U_n^i = V_n^i + \varepsilon_n^i$. The nested logit model is obtained by assuming that the vector of unobserved utility $\varepsilon_n = \left(\varepsilon_n^1, \varepsilon_n^2, \ldots \varepsilon_n^I\right)$ has the following cumulative distribution:

$$\exp\left(-\sum_{k=1}^{K}\left(\sum_{i\in B_k} e^{-\varepsilon_n^i/\lambda_k}\right)^{\lambda_k}\right)$$

This distribution is a type of GEV (Generalized Extreme Value) distribution. For a logit model, each ε_n^i is independent with a univariate extreme value distribution. However, the ε_n^i's are correlated within the nests. For any two alternatives *i* and *j* in nest B_k, ε_n^i is correlated with ε_n^j. For any two alternatives in different nests, the unobserved portion of utility is still uncorrelated $\mathrm{cov}\left(\varepsilon_n^i, \varepsilon_n^j\right) = 0$ for any $i \in B_k$ and $j \in B_l$ with $k \neq l$.

The parameter λ_k is a measure of the degree of independence in unobserved utility among the alternatives in nest B_k and is called dissimilarity parameter. A high λ_k means greater independence and less correlation i.e. the alternatives in the nest are less similar for unobserved reasons. A value of $\lambda_k = 1$ means complete independence in nest B_k. Obviously, if $\lambda_k = 1$ for all nests, then the nested logit reduces to the standard logit model.

With this distribution, the probability that household *n* chooses alternative *i* from the choice set is:

$$P_n^i = \frac{\exp\left(\left(V_n^i/\lambda_k\right)\left(\sum_{i\in B_k} e^{V_n^i/\lambda_k}\right)^{\lambda_k - 1}\right)}{\sum_{l=1}^{K}\left(\sum_{i\in B_l} e^{V_n^i/\lambda_l}\right)^{\lambda_l}}$$

This equation it is relativity easy to show that IIA holds within nests but not across nests which is called also as IIN: independence from irrelevant nests (Train, 2007). Thus, in a nested logit model we do have a relaxation of the IIA assumption compared to a normal logit model but you still have IIA holding over alternatives in each nest and IIN holding over alternatives in different nests.

3.2. Econometric estimations

In our model, we have grouped spatial alternatives in two nests, the first nest corresponding to the influence zone of three urban poles of the area, but we will not use variables to describe each nest, just the attributes of each alternative. Also, the effects of demographic variable are typically handled in an ad hoc way in most housing demand analyses, either added to demand equations in a linear additive way or used as stratifying variables. Because these demographic variables, especially age, are well known to enter in the demand relationship rather in a nonlinear fashion, we partition the population into homogenous strata and regress the choice pattern of each stratum separately on the other variables[4].

We use a log-linear form of the model for quantitative variables, because it presents better results that the classic linear form. Maximum likelihood estimates of the parameters are obtained by an iterative process with a convergence criterion of .01. In all cases, less than twenty iterations were required. Another precision concerns the results which are not necessarily consistent with the notion of equilibrium in the housing market, because there is the possibility of demand or supply excess.

In appendix the table 13 we present the results of housing demand estimations for four age classes. The explanatory variables are those presented previously: Urbanism as the degree of urbanism from PLUREL typology (1 rural, 2 periurban and 3 urban), Education as the log of education institutions in the spatial unit, Social-Health as the log of social and health facilities, Sanitation as the percent of dwellings connected to the sanitation network, the percent of individual dwellings in housing stock and finally two natural amenities variables; the part of forest respectively agricultural land in each spatial unit. In the last three rows we reported dissimilarity parameter values for the three nests.

We will analyse each node equation separately, to highlight the impact of each variable on housing demand for each age category. In the young equation (the first column in the table) all variables are significant and most of them have the expected sign. Thus, urbanism degree is a positive factor to attract young households, which is not surprising, seeing their spatial distribution in the area. Also education has a positive and very high parameter, which involves that for this class the proximity to educational institutions is important. This could be also the result of the fact that in this category we find also students which are still linked to education. A surprising parameter is for Social Health, being negative, which means that young households are not attracted by the presence of these facilities. Knowing that the young are attracted by most urbanized zones, it is quit natural to find negative effects of share individual dwellings and natural amenities on the probability that a young household choose a certain location. We see that dissimilarity parameters for this equation are superior to one, which means that the model is consistent with utility-maximizing behaviour for only some range of the independent variables.

For middle-aged households, all variables are still significant but as we can see, the urbanism degree is still positive, but less important than for the young households. For public services, as expected, now both variable (education and social-health) are significant and positive, meaning that the presence of those services increase the probability that a middle-aged household choose a certain

[4] See for example Quigley (1973),), Li (1977), Börsch-Supan and Pitkin (1988), Tu and Goldfinch (1996)

location. Sanitation has a surprising sign, being negative, but we have to see it in pair with individual housing, which is very significant and with a high parameter. Compared to the other classes, only for middle-aged, individual housing is positive and significant, which confirms the fact that for this class, the housing demand in terms of dwelling size is highest. Knowing that forest areas are in the north side, with a low level of population and economic activity, its sign is negative. On contrary, for this time agricultural amenities are positive, because this category is less attracted by most urbanized zones, being very often located in periurban and rural zones at proximity of urban centres. The high level of the dissimilarity parameters means that there is an important independence of alternatives in each nest.

For the first group of retirees, we see that urbanism degree is significant and negative, showing their preferences for rural and periurban zones. If education is not significant, all other variables are significant with expected signs: positive impact of social and health facilities, but also of individual houses, sanitation network and natural amenities. The parameters values show that this class is in some kind a passage from middle-aged households to oldest category (retirees 2), having a combination of their preferences. For retirees1 equation, dissimilarity parameters are lower than previous cases. Finally, for the most aged households all variables are significant, with normal signs: positive impact of public services and natural amenities, but negative of the urbanism degree and individual housing.

3.3. Tests

We start with prediction success tables, where we compare the model's predictions and real data. Thus, 79% of the model's predictions are correct, but it is interesting to analyze deeper. A comparison between the four equations shows that the predictions are more accurate for young housing demand equation (97.5%) and less for other social groups: 77.5% for middle-aged households, 72.2 for retirees1 and 70.98 for retirees2. This could be the result of the fact that young households are massively concentrated in Montpellier core-city, so much easier to predict.

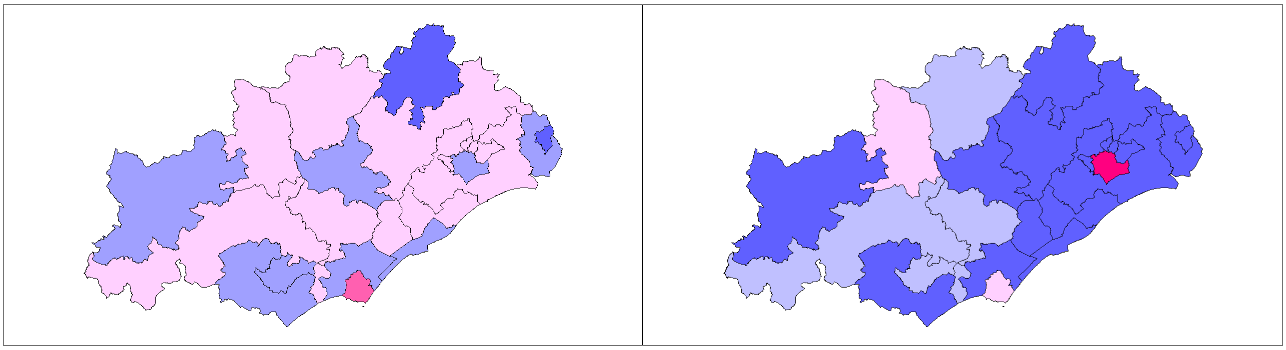
In figure 20 we present several maps of predictions for each age class, to see where housing demand predictions are similar or not with real spatial distribution of households. In blue areas the model predicts a housing demand inferior to real number of households in the spatial unit and in red areas we have an opposite situation. In dark areas, the error is more than 25%.

For young households, generally there is under prediction in urban zones of their housing demand, which determines a slight over prediction in the others. An exception is Agde Municipality, where the predicted housing demand is much higher, but this is explained by the particular case of this municipality (tourism). For the other social classes there is an important over prediction of housing demand in Montpellier city, the core of the area, which necessarily implies under predictions in other zones.

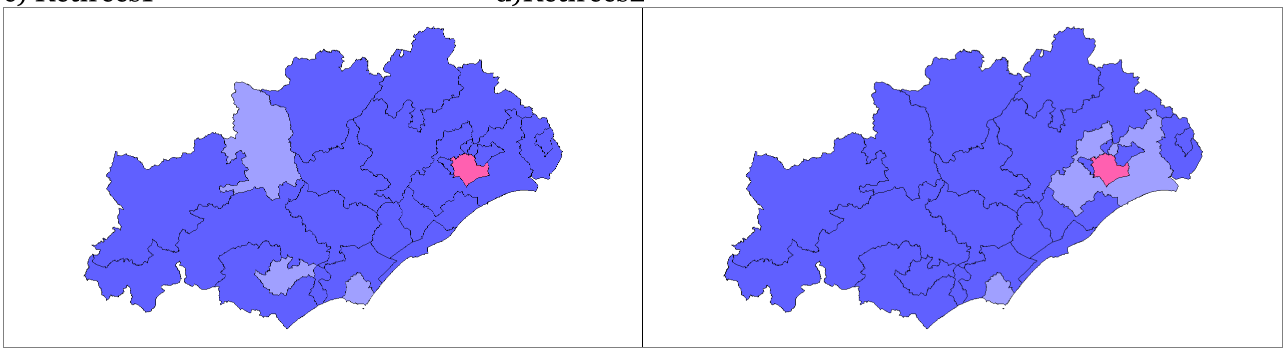
This result confirms that discrete choice models are not equilibrium models. An easy comparison shows that in Montpellier city housing demand obtained from the model exceed housing offer with almost 50%. This result is confirmed by the evolution of population in Montpellier urban region which increased from 1999 to 2006[5] of more than 13%. This could be explained by the fact that there is a very important gap between Montpellier and the rest of the region, in terms of public services and infrastructure.

<u>Figure 20: Model predictions compared to real distribution of households</u>

a) Young households b) Middle-aged households



c) Retirees1 d)Retirees2



[5] Insee Census Data

References

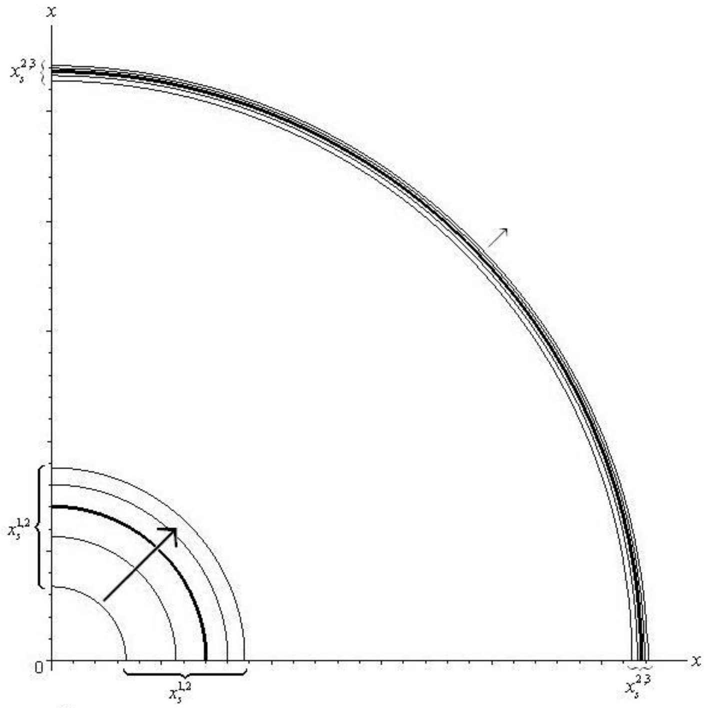
Loibl W. and Köstl M. (2008): "Report on a methodology to delineate RUR sub-regions", D2.1.4. Module 2. PLUREL report

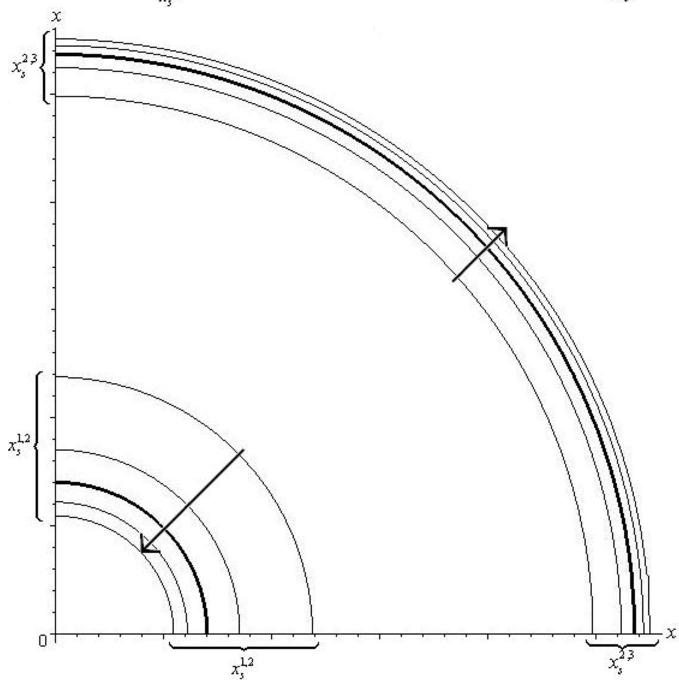
Steinnocher K., Köstl M., Loibl W. (2008): "Typology of rural-urban regions", Presentation at the PLUREL M2-meeting, Vienna

Appendix 1.1 Impact of demographic changes on location

a) Y-A-R location pattern

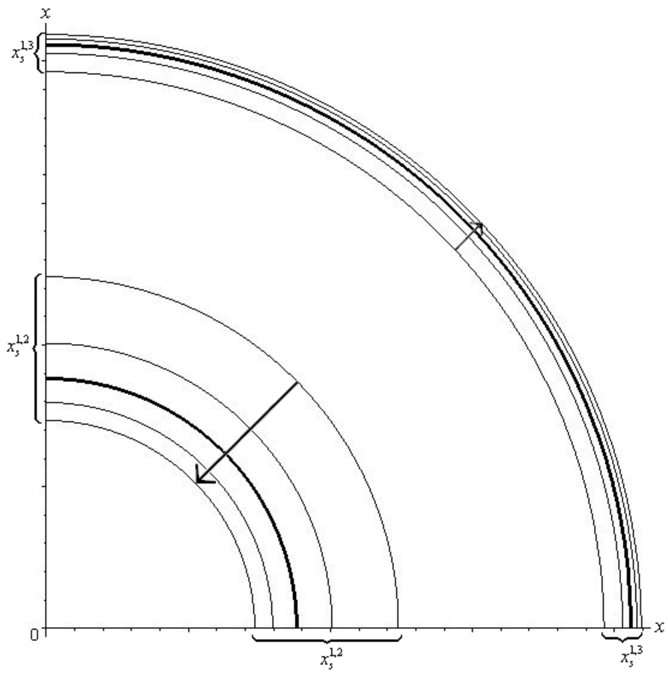
Young Population Variation Middle-aged Population Variation



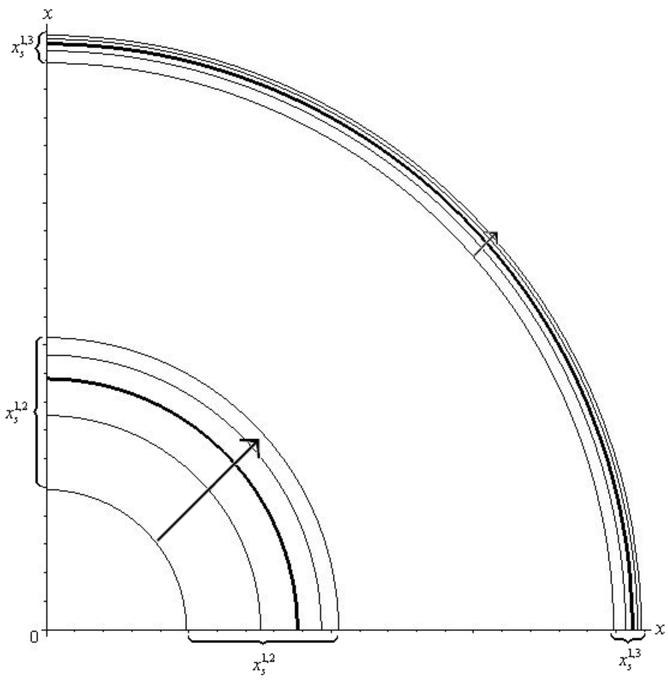


b) A-Y-R location pattern

Young Population Variation Middle-aged Population Variation



x
$x_s^{1,3}$
$x_s^{1,2}$
0
$x_s^{1,2}$
$x_s^{1,3}$
x



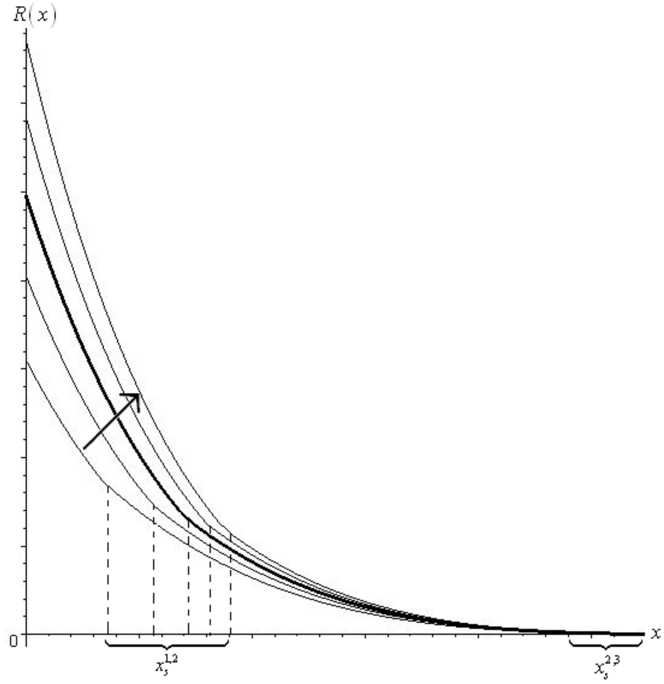
x
$x_s^{1,3}$
$x_s^{1,2}$
0
$x_s^{1,2}$
$x_s^{1,3}$
x

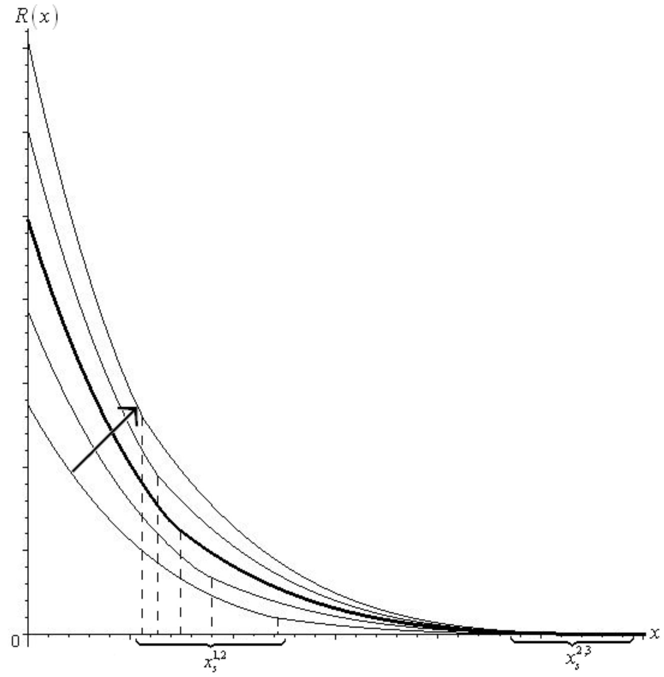
Appendix 1.2. Impact of demographic changes on housing prices

a) Y-A-R location pattern

Young Population Variation

Middle-aged Population Variation

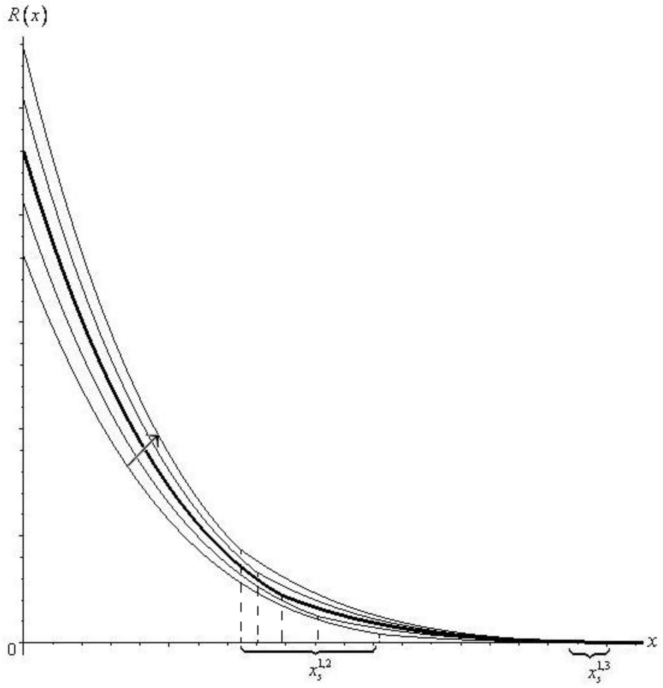


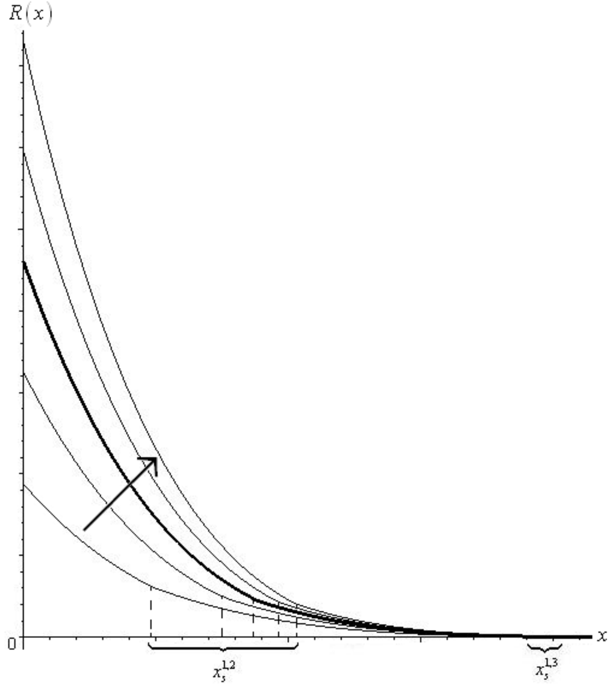


b) A-Y-R location pattern

Young Population Variation

Middle-aged Population Variation

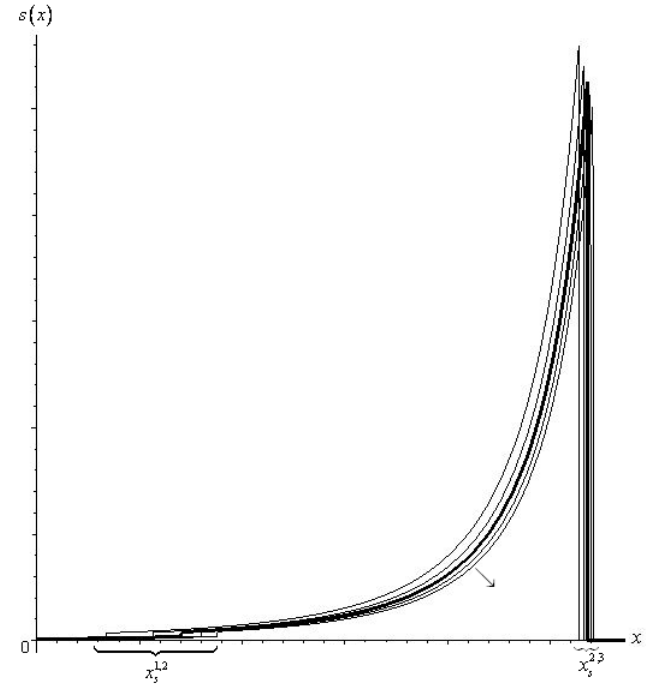

R(x)
0
x
$x_s^{1,2}$
$x_s^{1,3}$

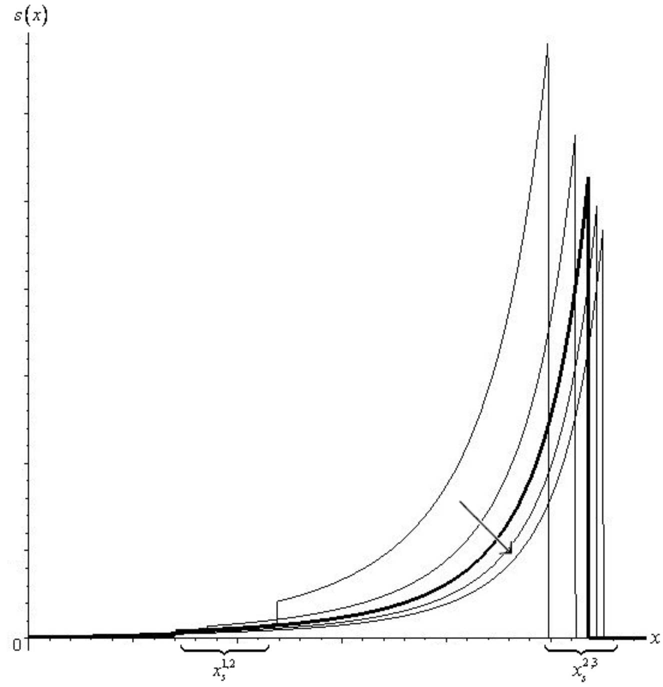

R(x)
0
x
$x_s^{1,2}$
$x_s^{1,3}$

Appendix 3. Impact of demographic changes on housing demand

a) Y-A-R location pattern

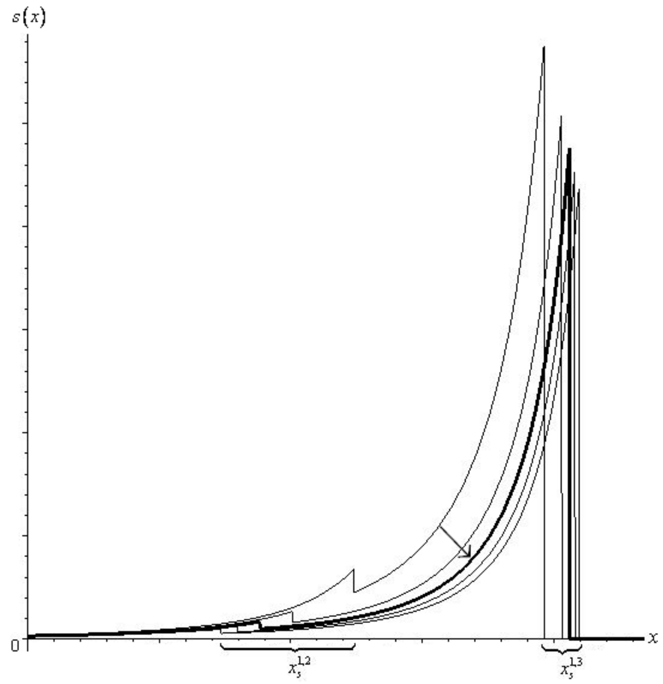
Young Population Variation Middle-aged Population Variation

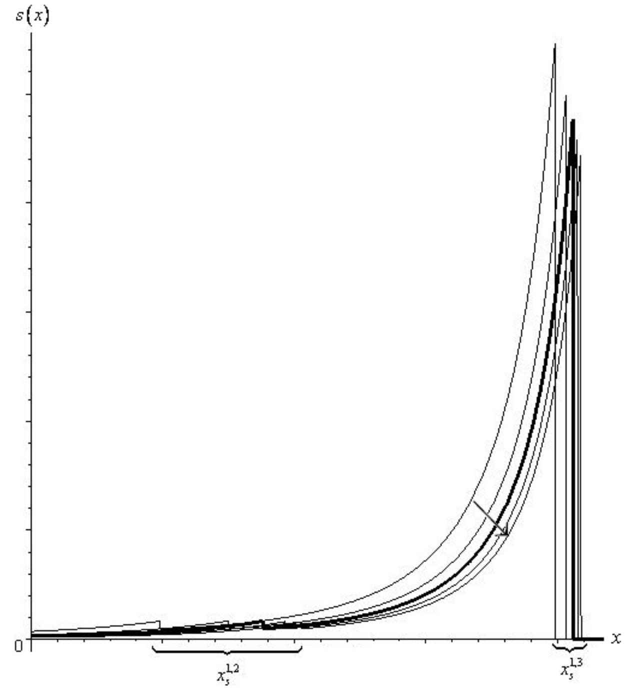




b) A-Y-R location pattern

Young Population Variation Middle-aged Population Variation

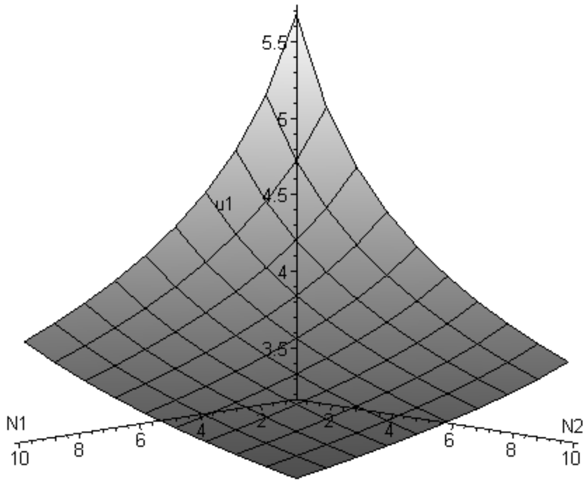

s(x)
0
x
$x_s^{1,2}$
$x_s^{1,3}$

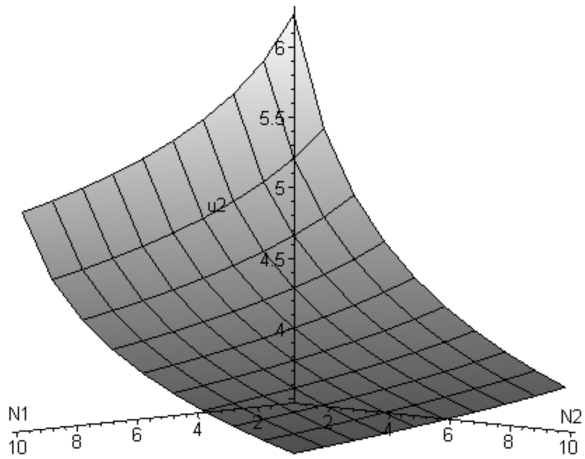

s(x)
0
x
$x_s^{1,2}$
$x_s^{1,3}$

Appendix 1.4. Demographic change impacts on welfare

a) Y-A-R location pattern

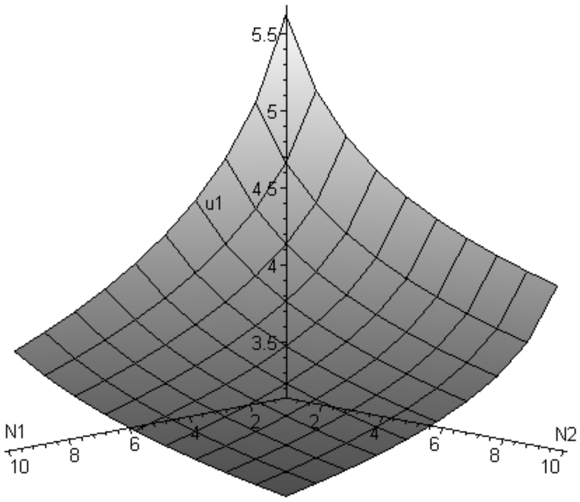
Young Household Utility Middle-aged Household Utility

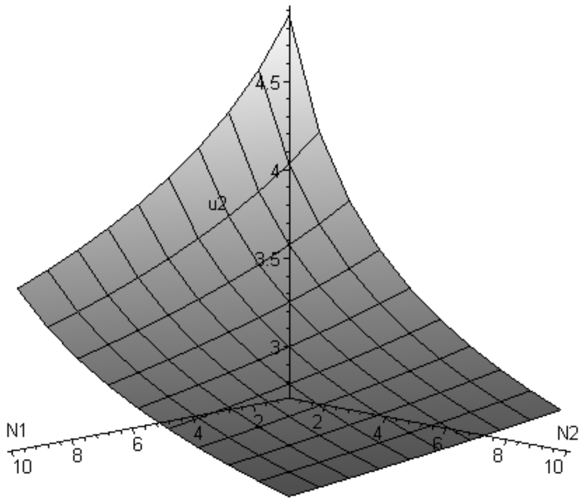




b) A-Y-R location pattern

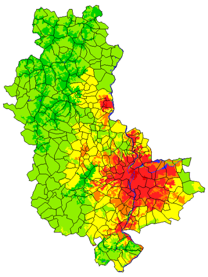
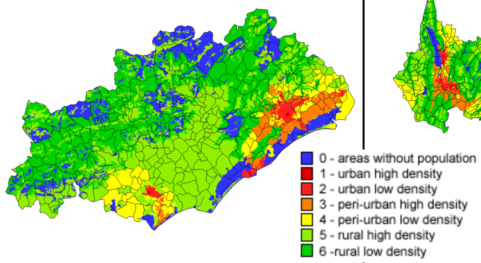
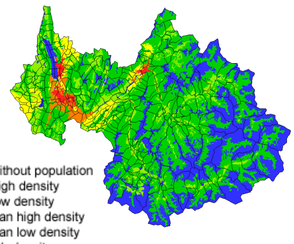
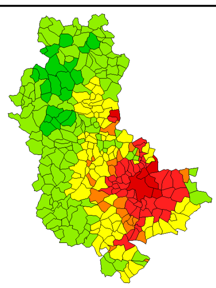
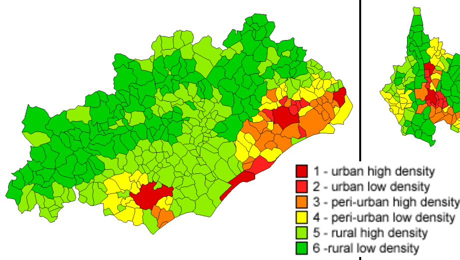
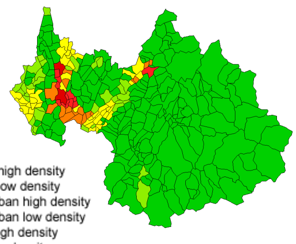
Young Household Utility Middle-aged Household Utility

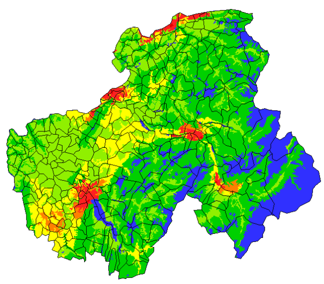
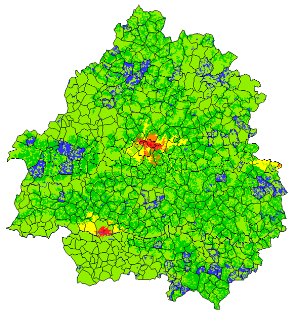
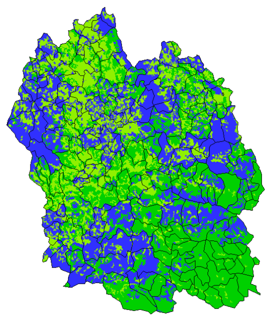
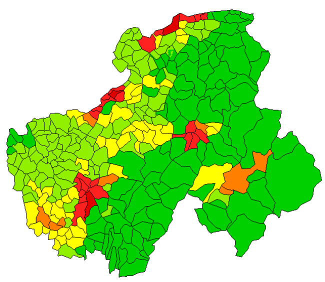
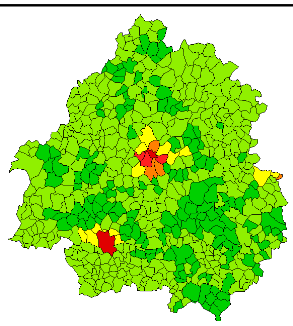
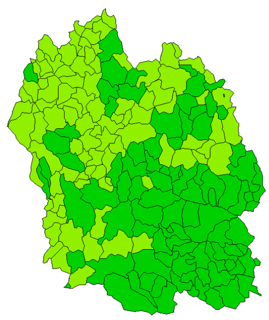




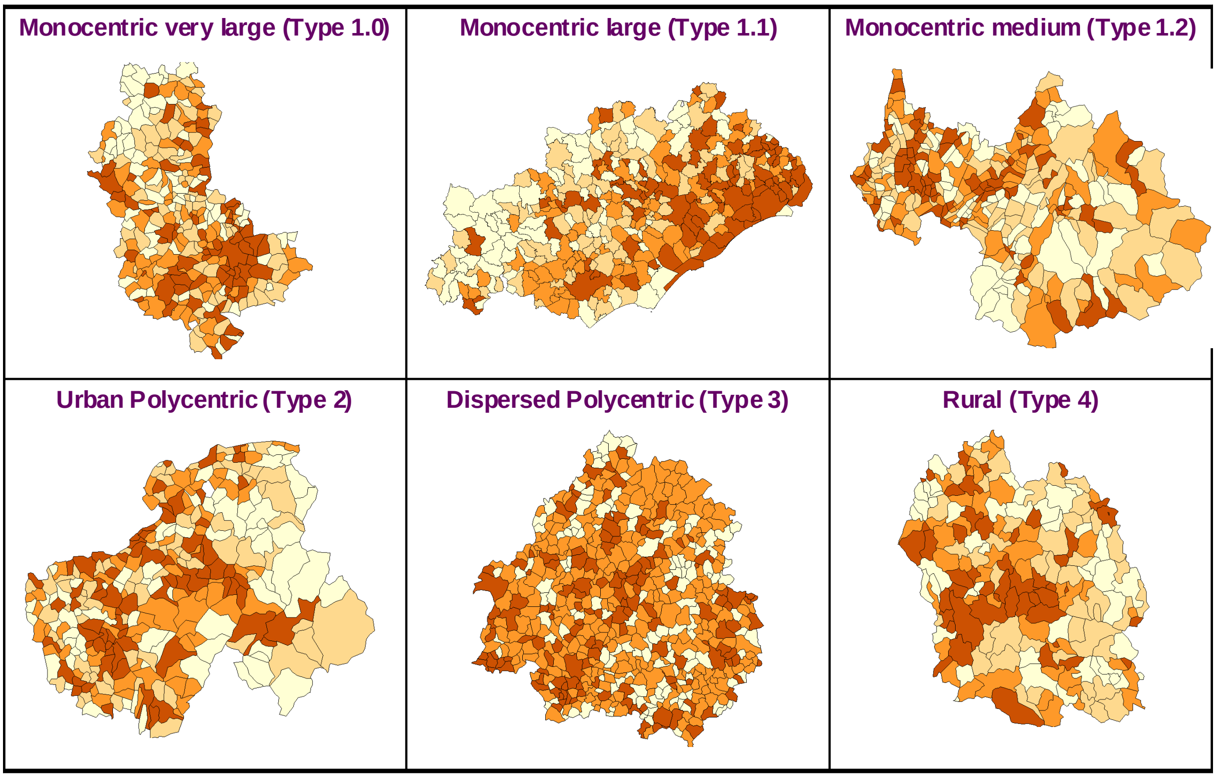
PLUREL

Appendix 2 : Municipality Typology Plurel/Cemagref methodology for French regions

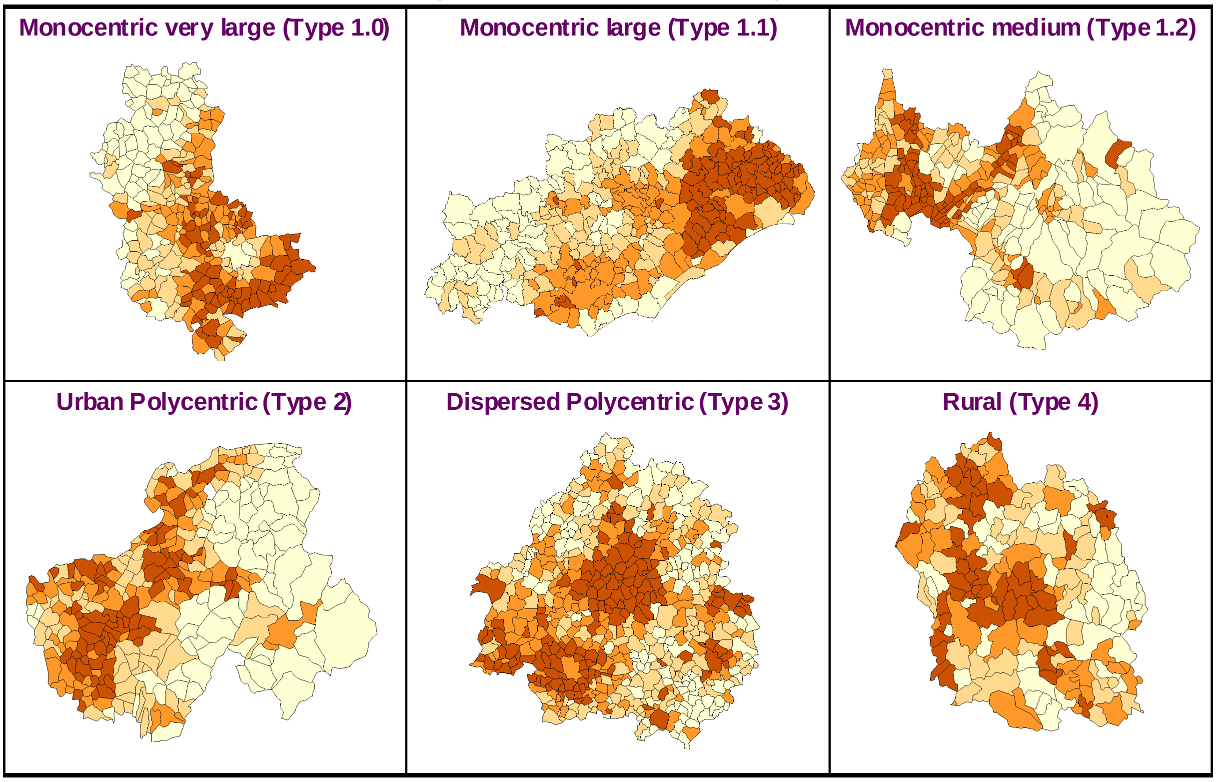
Rhône - Monocentric very large RUR TYPE 1.0	Hérault - Monocentric large RUR TYPE 1.1	Savoie - Monocentric medium RUR TYPE 1.2
		
		

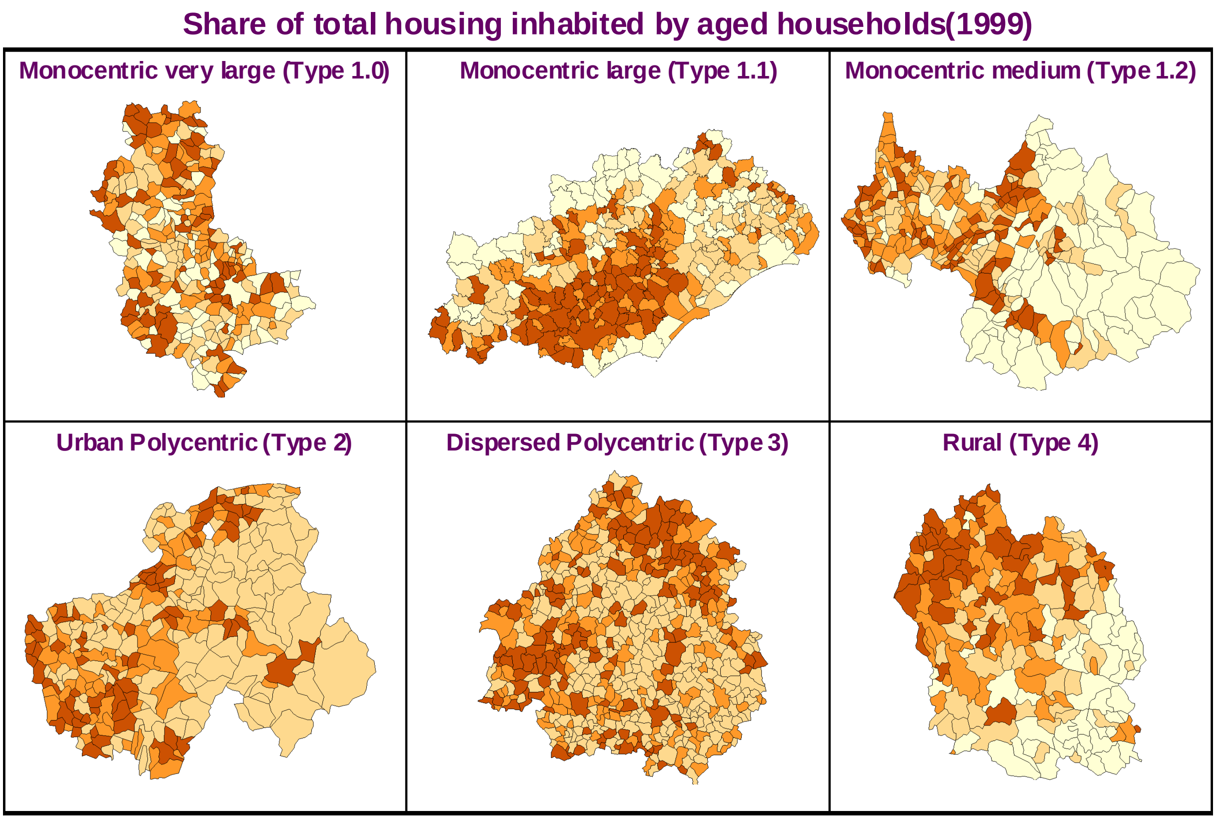
Haute Savoie – Urban Polycentric RUR TYPE 2	Dordogne - Dispersed Polycentric RUR TYPE 3	Lozère - Rural RUR TYPE 4
		
		

Share of total housing inhabited by young households(1999)



Share of total housing inhabited by middle-aged households(1999)




Share of total housing inhabited by aged households(1999)
Monocentric very large (Type 1.0)
Monocentric large (Type 1.1)
Monocentric medium (Type 1.2)
Urban Polycentric (Type 2)
Dispersed Polycentric (Type 3)
Rural (Type 4)

PLUREL

Table 2 Share of housing inhabited by middle-aged households

	Type 1.0	Type 1.1	Type 1.2	Type 2	Type 3	Type 4
Intercept	2.422	2.040	0.969	1.604	2.259	1.671
Spatial Lag	0.148	0.286	0.149	0.172	0.193	0.257
Time Lag	0.343	0.082	0.433	0.304	0.077	0.214
Subtype 1	3.770	7.346	5.255	4.546	8.322	
Subtype 2	0.177	1.257	4.168	2.283	1.471	
Subtype 3	0.008	1.038	0.536	1.602	1.057	
Subtype 4	-0.201	1.854	0.015	0.263	1.160	
Subtype 5	-0.036	0.332	0.479	0.123	0.285	0.056
Subtype 6	(Ref)	(Ref)	(Ref)	(Ref)	(Ref)	(Ref)
R-squared	0.475	0.197	0.458	0.495	0.148	0.149
Log likelihood	-581.465	-826.951	-733.623	-605.293	-1123.8	-418.913

Table 3 Share of housing inhabited by middle-aged households

	Type 1.0	Type 1.1	Type 1.2	Type 2	Type 3	Type 4
Intercept	4.393	2.664	0.901	2.292	2.027	1.337
Spatial Lag	0.397	0.451	0.283	0.349	0.605	0.412
Time Lag	0.517	0.542	0.749	0.632	0.344	0.609
Subtype 1	-2.576	-6.339	-5.295	-4.224	-7.827	
Subtype 2	0.033	-3.593	-3.250	-0.415	1.009	
Subtype 3	1.863	0.805	-2.021	0.055	4.514	
Subtype 4	3.082	-0.598	1.573	3.840	1.312	
Subtype 5	2.779	-0.291	2.841	1.923	-0.147	-1.519
Subtype 6	(Ref)	(Ref)	(Ref)	(Ref)	(Ref)	(Ref)
R-squared	0.844	0.823	0.871	0.887	0.660	0.695
Log likelihood	-855.314	-1135.41	-960.072	-900.184	-1756.77	-586.392

Table 4 Share of housing inhabited by retired households

	Type 1.0	Type 1.1	Type 1.2	Type 2	Type 3	Type 4
Intercept	14.987	5.634	2.087	2.536	12.254	4.022
Spatial Lag	0.083	0.058	0.051	0.350	0.327	0.072
Time Lag	0.338	0.338	0.648	0.471	0.295	0.424
Subtype 1	2.663	2.156	2.433	3.692	1.592	
Subtype 2	3.488	4.377	2.416	2.534	2.532	
Subtype 3	1.689	0.238	4.039	4.156	0.891	
Subtype 4	1.375	0.525	2.396	1.653	1.756	
Subtype 5	1.756	0.673	1.168	1.399	0.853	0.800
Subtype 6	(Ref)	(Ref)	(Ref)	(Ref)	(Ref)	(Ref)
R-squared	0.281	0.498	0.727	0.684	0.282	0.613
Log likelihood	-774.087	-1068.41	-910.171	-773.836	-1682.12	-531.708

Appendix 3 : Segregation Indices Formulas

1. Evenness Indices

a) One-group Indices

Duncan et Duncan (1955) proposed the segregation index which measures the spatial distribution of a social group and varies between 0 (perfectly equal distribution) and 1 (maximal segregate distribution). The index value expresses the percentage of the group which should move to obtain a perfect spatial distribution:

$$SI = \sum_{i=1}^{n}\left[\frac{t_i\left|p_i - P\right|}{2TP(1-P)}\right]$$

where x = social group
t_i, t_j = total population of spatial unit i (j)
p_i, p_j = proportion of group X in total population of spatial unit i (j)
T = total population of the area
P = proportion of group X de in total population T
n = number of spatial units

Morrill (1991) proposes a modified version of the index, which includes a spatial interaction component in it: the original dissimilarity index less the amount of potential interaction between different groups across spatial unit boundaries[6]:

$$SI(adj) = SI - \frac{\sum_i \sum_j \left|c_{ij}\left(p_i - p_j\right)\right|}{\sum_i \sum_j c_{ij}}$$

where c_{ij} is the i-th row and j-th column of a binary connectivity matrix

Wong (1993) further enhances this index by including more spatial information in the index formulation. He proposes that not just adjacency, but also the length of common boundary can affect the intensity of spatial interaction among population groups across zonal boundary, and hence affect the level of segregation[7]:

$$SI(w) = SI - \frac{1}{2}\sum_i \sum_j w_{ij}\left|p_i - p_j\right| \text{ with } w_{ij} = \frac{d_{ij}}{\sum_i d_{ij}}$$

[6] The original version was proposed only for two groups and this formula was adapted by Apparicio et al. (2008) to work for one group
[7] The original version was proposed only for two groups and this formula was adapted by Apparicio et al. (2008) to work for one group

$$SI(s) = SI - \frac{1}{2}\sum_i \sum_j \left(\frac{w_{ij}\left|p_i - p_j\right|}{\sum_i \sum_j w_{ij}} \frac{\frac{Per_i}{A_i} + \frac{Per_j}{A_j}}{2\max\left(\frac{Per}{A}\right)} \right)$$

where d_{ij}= the length of the common boundary between spatial units i and j

Per_i, Per_j = perimeter of spatial units i and j

A_i, A_j = area of spatial units i and j

max(Per/A) = maximum ratio between the perimeter and area in the area

b) Intergroup Indices

While one group indices can be used to determine whether a group is unevenly distributed across a set of spatial units, intergroup indices are used to verify if two different social groups have or not similar spatial distributions. As in the case of one-dimensional indices, the most used measure is the dissimilarity index (*DI*), proposed by Duncan and Duncan (1955) which was adjusted to take into account spatial units aspects such as tract contiguity (*DI*(*adj*), Morill 1991), boundary lengths (*DI*(*w*), Wong 1993) and tract boundary lengths and perimeter/area ratio (*DI*(*s*) , Wong 1993).

$$DI = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{x_i}{X} - \frac{y_i}{Y}\right|$$

where X,Y = population of group X (Y) in the area
x_i, y_i = population of group X (Y) in spatial unit i

$$DI(adj) = DI - \frac{\sum_i \sum_j \left|c_{ij}\left(p_i - p_j\right)\right|}{\sum_i \sum_j c_{ij}}$$

$$DI(w) = DI - \frac{1}{2}\sum_i \sum_j w_{ij}\left|p_i - p_j\right| \text{ with } w_{ij} = \frac{d_{ij}}{\sum_i d_{ij}}$$

$$DI(s) = DI - \frac{1}{2}\sum_i \sum_j \left(\frac{w_{ij}\left|p_i - p_j\right|}{\sum_i \sum_j w_{ij}} \frac{\frac{Per_i}{A_i} + \frac{Per_j}{A_j}}{2\max\left(\frac{Per}{A}\right)} \right)$$

2. Exposure Indices

Isolation index, proposed by Bell (1954), measures the probability that a member of a group shares same spatial unit with another member of his own group.

$$xPx = \sum_{i=1}^{n} \left(\frac{x_i}{X} \cdot \frac{x_i}{t_i} \right)$$

Adjusted isolation index takes into account the population composition which distorts *xPx* (Bell; 1954; White, 1986):

$$Eta^2 = \frac{xPx - P}{1 - P}$$

Interaction index (Bell, 1954) is the intergroup form of exposure and measure the probability that a member of a group shares same spatial unit with a member of another social group:

$$xPy = \sum_{i=1}^{n} \left(\frac{x_i}{X} \frac{y_i}{t_i} \right)$$

3. Concentration Indices

Delta index of Duncan (Duncan et al., 1961) is a one-group measure which can be derived from an application of the dissimilarity index defined above. With values ranging from 0 (no residential segregation) and 1 (complete segregation), the index is interpreted as the proportion of the group that should move to obtain a uniform density:

$$\Delta = \frac{1}{2} \sum_{i=1}^{n} \left| \frac{x_i}{X} - \frac{A_i}{A} \right|$$

Massey and Denton (1988) propose two more complex indices: absolute concentration index (one-group measure) and relative concentration index (two-group index). Absolute concentration index compares the total area inhabited by a certain category with the minimum and maximum spatial units which may reside members of the class in cases of maximum and minimum concentration. The index has values between zero and one, which correspond to a minimum concentration and maximum concentration:

$$ACO = 1 - \frac{\sum_{i=1}^{n} (x_i A_i / X) - \sum_{i=1}^{n1} (t_i A_i / T_1)}{\sum_{i=n2}^{n} (t_i A_i / T_2) - \sum_{i=1}^{n1} (t_i A_i / T_1)}$$

where spatial units are sorted by land area in ascending order
$n1$ = rank of spatial unit where the sum of all t_i equals X (from 1 to $n1$)

PLUREL

$n2$ = rank of spatial unit where the sum of all t_i equals X (from n to $n2$)
T_1 = sum of all t_i in spatial unit 1 to spatial unit $n1$
T_2 = sum of all t_i in spatial unit $n2$ to spatial unit n

4. Clustering Indices

The most known measure of clustering is absolute clustering index of Massey and Denton (1988) adapted from Geary (1954) and Dacey (1968). This index expresses the average number of group X members in nearby (areal units) as a proportion of the total population in those nearby (areal units). Absolute clustering index varies from 0 to 1.

$$ACL = \frac{\left[\sum_{i=1}^{n} \frac{x_i}{X} \sum_{j=1}^{n} (c_{ij} x_j)\right] - \left[\frac{X}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}\right]}{\left[\sum_{i=1}^{n} \frac{x_i}{X} \sum_{j=1}^{n} (c_{ij} x_j)\right] - \left[\frac{X}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}\right]}$$

5. Centralisation Indices

Absolute centralisation index (Massey & Denton, 1988) examines only the distribution of the minority group around the centre and varies between -1 and 1. Positive values indicate a tendency for the group members to reside close to the city centre, while negative values indicate a tendency to live in outlying areas. A zero score means that a group has a uniform distribution throughout the metropolitan area.

$$ACE = \left(\sum_i X_{i-1} S_i\right) - \left(\sum_i X_i S_{i-1}\right)$$

where spatial units are sorted by distance from the centre in ascending order

X_i = cumulative proportion of group X in spatial unit i (from 1 to i)

S_i = cumulative proportion of area of spatial i (from 1 to i)

Table 13: Housing demand estimations

	Young	Middle-aged	Retirees1	Retirees2
Urbanism	0.289*** (-0.025)	0.0166** (-0.00748)	-0.0725*** (-0.00893)	-0.0296*** (-0.00825)
Education Log	2.161*** (-0.0717)	0.347*** (-0.0185)	0.0106 (-0.023)	0.273*** (-0.0213)
Social Health Log	-0.550*** (-0.0527)	0.685*** (-0.0176)	0.753*** (-0.0234)	0.381*** (-0.0204)
Sanitation	1.665*** (-0.225)	-0.683*** (-0.0622)	0.446*** (-0.0825)	0.750*** (-0.0783)
Individual Housing	-1.001*** (-0.0772)	0.259*** (-0.0245)	0.0472 (-0.0297)	-0.00141 (-0.0267)
Forest	-1.096*** (-0.149)	-0.798*** (-0.0483)	0.116* (-0.0596)	0.157*** (-0.0547)
Agricultural	-0.774*** (-0.0739)	0.0368* (-0.0221)	0.153*** (-0.0283)	0.198*** (-0.0276)
λ_1 Montpellier urban region	1.481*** (-0.0349)	0.963*** (-0.00951)	0.815*** (-0.0108)	0.762*** (-0.0101)
λ_2 Béziers urban region	1.947*** (-0.0584)	0.978*** (-0.0139)	0.810*** (-0.0149)	0.799*** (-0.0137)
λ_3 Sète urban region	2.374*** (-0.106)	0.978*** (-0.0246)	0.665*** (-0.0224)	0.586*** (-0.0186)
Tests				
Log Likelihood	-150470.25	-420766.37	-181666.99	-197849.28
Wald chi2	2457.59	11566.29	6599.17	7381.10
LR Test chi2	1125.56	72.44	224.89	527.21

*** p<0.01, ** p<0.05, * p<0.1 Standard errors in parentheses