Residential location choice in Lyon urban area

Location choice model is relatively old and has its roots in Von Thünen's theory of land use. In order to better explain residential choice behavior, this theory has been subject to different modifications to take household characteristics and urban environment into account.

The renewed interest in modeling household residential location choice in the last years is partly due to the development of several Land Use Transport Interaction models (Anas, 1994, 1995 ; Waddell, 2000, 2002). To tackle issues related to suburbanization, segregation, and traffic congestion, residential location choice is at the heart of LUTI (Land Use and Transport Interaction) models prospective approach. It is in this context that Nicolas (2010) has developed SIMBAD model which aims to simulate the impacts of different policy scenarios on land use and transport within Lyon urban area, taking the year 1999 as a reference. This is the date that the different models are estimated including the residential location choice.

Empirical studies addressing the determinants of residential location choice are poorly developed, especially in France, despite attempts to use sampling strategies (McFadden, 1978 ; Guevara and Ben-Akiva, 2013) to reduce the large number of alternatives used in the estimation of location choice at fine spatial scale. The multitude data sources, the difficulty to collect information describing alternative or households and to assemble it into a single data set, as well as the limits of software to deal with such complicated model are the main reasons. This article aims to address these constraints appropriately and enrich the empirical literature by providing a summary review of the determinants of residential and then use the data of Lyon urban area to test the influences of these determinants.

In this work, we chose to begin by estimating a multinomial logit model with sampling of alternatives to provide a better understanding of residential location choice process in Lyon Urban area. Being aware that the application of such a model in a spatial context may provide undesirable substitutions among the locations, due to the Independence of Irrelevant Alternatives (IIA) property, we remained cautious and we choose to use homogeneous geographical entities. These zoning (macro-zones) formed by INSEE from the combination of contiguous districts (IRIS). Such grouping meets some criterions of homogeneity between IRIS and consequently satisfied partially the IIA property.

The variables included in the model are made from several databases: i) the 1999 census data (households that recently moved: 1998, 1999), ii) official national business register (SIRENE) to describe amenities, iii) accessibility (transport model)

In terms of results, accessibility and social environment are the most important determinants of residential location choice. If market constraints and trade-offs between real estate prices and accessibility still weigh on household choice decision, socioeconomic

characteristics and preferences for social amenities are also important in location choice process. Indeed, rich households choose to locate in neighbors dominated by their own group and avoid neighbor with social problem (high concentration of social housing). From this point of view, we deduce that urban segregation is not only a consequence but also cause of the choice of residential location and that public policies have an important role to play in guiding local policies, such as social housing, as consequence these policies have an indirect effect on location choice decision.

To improve the predictive power of this first model and take into account the spatial autocorrelation, we work on the estimation of a nested logit model. The modeling process is split into two levels, first the household choose between macro-zone, before determining his district into the macro-zone.

The difficulty of estimating such a hierarchical structure is twofold. On the one hand we need to define the nests; a spatial hierarchical classification can be used to overcome this problem. On the other hand, we must address the problem of applying the recent methodology developed by Guevara and Ben-Akiva (2013) into larger database to correct sampling bias in nested logit models.