YAN KESTENS

# UTILISATION DU SOL, ACCESSIBILITÉ ET PROFIL DES MÉNAGES: EFFETS SUR LE CHOIX RÉSIDENTIEL ET LA VALEUR DES PROPRIÉTÉS

# LAND USE, ACCESSIBILITY AND HOUSEHOLD PROFILES: THEIR EFFECTS ON RESIDENTIAL CHOICE AND HOUSE VALUES

Thèse présentée
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de
doctorat en aménagement du territoire et développement régional
pour l'obtention du grade de Philosophiae Doctor (Ph.D.)

FACULTÉ D'ARCHITECTURE, D'AMÉNAGEMENT ET DES ARTS VISUELS
UNIVERSITÉ LAVAL
QUÉBEC

AVRIL 2004

# Résumé

**Résumé 150 mots :**

Cette thèse explore et développe différentes méthodes d'analyse afin de mieux comprendre les choix de localisation résidentielle et les impacts de certaines externalités sur la valeur des propriétés unifamiliales. En ayant recours à la modélisation hédonique, le premier chapitre analyse l'impact, variable selon la proximité au centre-ville, de l'utilisation du sol et de la végétation sur les valeurs résidentielles. Dans le deuxième chapitre, les raisons de déménager et les critères de choix de la propriété et du quartier de résidence sont étudiés en détail. Une analyse des correspondances met en lien ces critères avec les théories psychologiques et géographiques de « place-identity » et d'espaces de perception, tandis que des modèles de régression logistique mesurent la probabilité d'évoquer un critère en fonction du profil du ménage. Enfin, dans un troisième chapitre, le profil des ménages acheteurs est introduit dans deux types de modèles hédoniques. L'hétérogénéité des valeurs implicites est alors mesurée et comparée selon le recours à l'expansion spatiale ou aux Geographically Weighted Regressions.

**Résumé 350 mots :**

Cette thèse explore et développe différentes méthodes d'analyse afin de mieux comprendre les choix des ménages en terme de localisation résidentielle et les impacts de certaines externalités sur la valeur des propriétés unifamiliales. Le territoire d'étude couvre la ville de Québec, tandis que l'essentiel des analyses repose sur l'analyse de transactions effectuées pendant les périodes 1986-1987 et 1993-2001. De plus, une enquête téléphonique réalisée entre 2000 et 2002 a permis d'obtenir des informations complémentaires sur les critères de choix et le profil socio-démographique de quelque 800 ménages acheteurs de propriétés unifamiliales à Québec.

Dans un premier chapitre, l'impact de l'utilisation du sol et plus particulièrement de la végétation est analysée, en ayant recours à la modélisation hédonique. Les données d'utilisation du sol, extraites de photos aériennes et d'une image satellite, sont compilées au sein d'un système d'information géographique, et ce, à différentes échelles. L'impact de la végétation, variable selon la proximité au centre ville, est clairement démontré.

Dans un deuxième chapitre, les motivations liées au déménagement et les critères de choix de la résidence et du quartier par les ménages acheteurs sont étudiés. Une analyse des correspondances souligne le lien entre les critères de choix exprimés et les théories cognitive et géographique de « place-identity » et d'espaces de perception. Aussi, des régressions logistiques mesurent la probabilité d'exprimer un critère en fonction du profil du ménage et de la localisation. Le fait d'avoir ou non été préalablement propriétaire, l'âge, le type de ménage, le revenu, le niveau d'éducation ainsi que la localisation sont des facteurs significativement liés à divers critères de choix. Enfin, dans un troisième et dernier chapitre, les données décrivant le ménage sont introduites dans deux types de modèles hédoniques, les uns ayant recours à l'expansion spatiale et les autres utilisant les « Geographically Weighted Regressions ». L'hétérogénéité des valeurs implicites est alors analysée en considérant le profil des ménages. Il apparaît non seulement que la valeur marginale de plusieurs attributs varie en fonction du ménage acheteur, mais que le revenu et le statut de l'acheteur (ancien vs. nouveau propriétaire) ont un impact direct sur le prix d'achat de la propriété.

Cette thèse, s'appuyant sur des méthodes d'analyse des marchés résidentiels et ayant recours à divers outils d'analyse spatiale, parvient à établir des liens entre le statut socio-démographique des ménages, leurs critères de choix résidentiels, et la structure spatiale de la ville de Québec.

**Abstract (350 words) :**

This thesis explores and develops various analytical methods in order to better understand residential choice and the implicit prices of single-family property markets. The area of study is Quebec City, whereas most of the work relies on single-family property transactions that occurred during the 1986-1987 and 1993-2001 periods. A phone survey held between 2000 and 2002 gave additional information on the choice criteria and household profiles of 800 of these actual property buyers.

In a first chapter, the impact of the surrounding land use and vegetation is measured using hedonic modelling. Land-use data are extracted from both a mosaic of aerial photographs, and from a Landsat TM-5 image. Various measures of land use, at different spatial scales, are introduced within the hedonic models. More specifically, the heterogeneous impact of vegetation, depending on relative proximity to the Main Activity Centre, is shown.

In a second chapter, motivations for moving and residential and neighbourhood choice criteria are analysed. A Correspondence Analysis underscores the links between choice criteria and the psychological and geographical theories of Place-Identity and perception spaces. Also, logistic regressions measure the odds of mentioning a criteria depending on the household profile and location. Previous tenure status, age, income, household structure and location are significantly related to various residential choice criteria. Finally, in a third chapter, the household-level data are introduced within the hedonic framework, using Casetti's expansion method and Geographically Weighted Regressions. The heterogeneity of implicit prices is analysed regarding the buyer's household profile. Not only does the marginal value of certain attributes vary regarding the buyer's profile, but it appears that income and previous tenure status have a direct impact on property values.

This thesis, through the development of new methods aiming at analysing residential markets and residential choices, contributes to further understanding the complex links between the socio-demographic dimension of households, their residential choice criteria, and the spatial structure of Quebec City.

Candidat: Yan Kestens

Directeur: Marius Thériault

Co-directeur: François Des Rosiers

# Avant propos

*"Il ne fait aucun doute qu'il existe un*
*monde invisible,*
*Cependant il est permis de se demander*
*à quelle distance il se trouve du centre-ville*
*et jusqu'à quelle heure il est ouvert."*
Woody Allen, Dieu, Shakespeare…et moi, Opus 1 et 2, Édition Solar, p.58

Pour la rédaction d'une thèse, il existe aussi un monde invisible qu'il convient ici de remarquer, et par la même occasion, de remercier.

Mes premiers remerciements s'adressent à mon directeur, Marius Thériault, qui, à l'aube du troisième millénaire, a eu la bonne idée de me proposer de rejoindre l'équipe dynamique du Centre de Recherche en Aménagement et Développement pour entreprendre une thèse. Tout au long de cette aventure, ses conseils, ses suggestions et sa vivacité d'esprit m'ont beaucoup apportés. De même, je remercie mon co-directeur François Des Rosiers, qui a su me transmettre les joies et subtilités de la modélisation hédonique. Tous deux trouvent donc tout naturellement leur place comme coauteurs d'un article accepté (chapitre 1), d'un deuxième soumis (chapitre 2), et d'un troisième à soumettre (chapitre 3). Enfin, je remercie les autres membres de mon comité, Paul Villeneuve, Danielle Marceau et Geoffrey Edwards pour leurs judicieux commentaires et le temps qu'ils ont consacré à ma thèse.

Mes remerciements s'adressent également à tout le personnel du CRAD et du département d'aménagement, Line Béland, Josée Bouchard, Solange Gouygou, Suzanne Larue, Pierre Lemieux, François Théberge, ainsi qu'à mes compagnons d'(in)fortune Rémy Barbonne, Housseini Coulibali, Salomon Gonzalez, Daniel Lachance et Catherine Trudelle, avec qui les échanges ont été une continuelle source d'inspiration.

Je tiens aussi à remercier le ministère des affaires étrangères du gouvernement du Canada pour son support financier.

Finalement, mes remerciements s'adressent particulièrement à ma famille; mes parents qui m'ont depuis toujours encouragé dans mes études, et Pauline, qui m'a accompagné, encouragé (et supporté!) au quotidien dans cette aventure. Sa joie de vivre et sa présence ont grandement contribué à l'aboutissement de ce projet. Je la remercie tout particulièrement pour ces neufs derniers mois, qui annoncent le début d'une extraordinaire aventure…

# Table des matières

# Liste des tableaux et figures

# Introduction

Human beings organise their surroundings in order to adapt their environment to their needs. Since the protecting caves and the limited land occupation of the beginnings, societies have gone through major evolutions, namely in the development of their social, political, economic and cultural structures. Concomitantly, the spatial evolutions of land occupation have changed the visible marks of the human grip on the environment. Since the year 1000, the world population has multiplied by fifteen, and following the recent demographic transition and industrialisation of most countries, the major form of the spatial organisation of human societies has become predominantly urban, while less than one in ten people were living in cities at the beginning of the 20th century. However, the notion of city itself and its spatial and social organisation has also gone through important changes since the Greek origins of the « polis ». The development of social interaction and economic exchanges have dictated the location of housing and of production facilities. Globally, the spatial organisation of human societies, and more precisely of our cities, is therefore the result of complex interactions between social, cultural, political, economic and technological influences.

Understanding the numerous dimensions of the spatial structure of our cities is a challenging task. In fact, a global understanding requires the contribution of various disciplines, like the cognitive sciences, psychology, sociology, but also economics, and more generally, as the spatial dimension is a key concept to urban structure, geography. A multidisciplinary approach is required today in order to grasp the whole array of interacting effects resulting from the social and physical organisations of our cities. The consideration of these multiple dimensions opens great possibilities for further understanding, whereas new challenges emerge with the rising consciousness of the complexity of the spatial dimension, from cognition to land occupation. The understanding of the processes underlying spatial organisations is a key factor for a better planning of our surroundings, and is needed for future generations. In fact, among the major transformations we have witnessed during this past century, human societies have demonstrated a massive capacity to harm themselves and to impact dangerously on their environment. A growing awareness of this reality has led political spheres and research communities to focus mainly on two points in the last half of the 20th century : (i) to establish political stability between nations in order to avoid massive wars; and (ii) to develop sustainable ways of living in order to avoid starvation and to protect our resources and environment. Henceforth, it appears that the processes underlying the spatial use of our environment constitutes a central question, and that answering this may contribute (i) to the development of knowledge and, more importantly, (ii) to better life conditions for future generations. As most of the world's inhabitants now live in cities, and while the urbanisation rate of developing countries is at its peak, analysing and understanding the spatial organisation of housing, services, production and consumption facilities and the interaction between this spatial setting and people's lifestyles and choices is a challenging but useful task.

In the past decades, the developed countries have gone through major societal changes, with direct effects on the way space has been perceived and occupied. In Canada, the concomitant post-industrialisation, investment in car-oriented infrastructures, women's access to the labour market, changes in family structures as well as important immigration rates in certain areas, have led to major transformations in the spatial organisation of our cities. The declines in the densities of central locations as well as the extension of cities toward low-density suburban areas – eventually causing urban sprawl – represent two clear trends in North American cities. This phenomenon is both the result and the determinant of new ways of life mainly dictated by household needs as well as individualistically oriented types of behaviour. Although this situation has generally made it easier to access homeownership, new challenges have risen for today's inhabitants. In fact, the impact of their lifestyle choices continues to affect future generations and are not bound by the city limits. Social cohesion, chronic illnesses, pollution and degradation of the environment are among the major challenges we are confronted with today. In this context, it appears that the analysis of housing markets and location choices may contribute to a better understanding of (i) people's preference and (ii) the impact of certain urban externalities. In order to do so, appropriate space-sensitive tools need to be developed and validated through empirical studies. Understanding the links between residential location choice, preference, residential market values and externalities, may contribute to shape adapted policies concerning tomorrow's planning decisions.

Von Thünen, in his pioneering work « Der isolierte Staat », was among the first to develop a theory that linked economic activities (and rent values) to location (Von Thünen, 1850). Adjusted to agricultural-oriented land use, this model was further extended and adapted to urban settings by Alonso (1964). For the latter, and considering a monocentric conception of the city, the competition between businesses and households in order to maximise satisfaction leads to the development of a high-density Central Business District, with concentric declining bid rent curves reflecting the possible trade-off between space consumption and accessibility to the city centre. More specifically, the study of housing markets has known an important turn with Rosen's (1974) seminal contribution to the hedonic theory, as applied to the property market. The hedonic pricing method was first developed by Court (1939, see also Goodman, 1998), who applied it to the automobile market. The theoretical justification to this approach was provided several decades later by Lancaster (1966), whose consumer theory stated that (i) it is not the good itself, but rather the characteristics of the good, that procure its utility; furthermore, that (ii) a good is generally composed of several characteristics; and that, (iii) goods in combination may detain characteristics that are not the sole sum of their parts. Finally, the consumer may seek to find a good whose combination of characteristics maximises utility. The multiplicity of characteristics that define a housing good, the numerous possibilities of combinations of these characteristics, and yet the conception of housing as a common unique good, make residential properties perfect examples of differentiated goods. The market value is defined through the interaction between supply and demand. Using the envelop theorem, hedonic price functions decompose the result of bid and supply interactions, that is, the market values, in implicit prices, by regressing the price on a number of variables describing the characteristics of the good (Rosen, 1974). The marginal contribution of each characteristic of the property can therefore be measured. The major characteristics of a non-movable property being linked to its location, the estimation of the implicit or hedonic prices of spatial externalities has been the subject of much research in the area of urban economics and more generally urban studies. However, the hedonic framework had not originally focussed on the spatial dimension of urban markets, and various concomitant disciplines like geography could therefore bring additional expertise. The growing use of hedonic analysis methods has been attendant to the recent boom in the past two decades of both computing capacities and the growing development and availability of large databases. Simultaneously, these technical advances led to the development in the early eighties of geographic information sciences, that is, a group of disciplines that study geographic phenomena using Geographic Information Systems (GIS), remote sensing data, and quantitatively-oriented geostatistical tools. GIS makes it possible to store, handle, analyse, visualise and model spatial data and therefore may contribute to better understand geographical phenomena. These new technical possibilities have contributed to the observation and analysis of spatial data, and spatial statistics have been of growing interest, both for urban and regional studies.

The attendant development of these capacities in spatial data handling, production and analysis and the possibility and need for economic disciplines to better consider the spatial dimensions of market structures have led to growing interaction between the geographer's expertise in space, and the economist's expertise in markets. This multidisciplinary aspect is of course highly appropriate for urban studies, especially when observing location choices and urban externalities through the analysis of the residential market. However, as this task requires an analysis of the result of past choices from an array of actors ranging from the mayor's planning division to our neighbour's household who recently moved in, it is necessary to account for the psychological and social dimensions of people, and to account for the political structure of institutions. People's perceptions and choices in terms of location, activities, labour or leisure do contribute to the shaping of our cities. Hedonic modelling makes it possible to estimate the implicit price of any significant externality, thereby reflecting people's willingness to pay for locating nearby. Hedonic prices are therefore good approximations of people's preferences. However, in order to be measured properly, it is recommended that people's preferences be put in relation with existing psychological and cognition theories, like place-identity theory or cognitive [hierarchical] perception spaces.

Finally, it can be said that an array of concordant developments may contribute to the analysis of urban structures. In fact, large spatial databases are available today. These can be properly integrated using appropriate tools like GIS and may be further analysed using spatial statistics. Furthermore, when integrated in efficient econometric modelling frameworks, and considering the spatial cognitive and perception theories, it is possible to answers specific questions about urban structure.

The global objectives of this research are therefore:

- to further our understanding of the links between urban structure, residential choice and residential markets; and
- to develop appropriate methodological procedures to achieve this goal.

Accordingly, this thesis proposes to further explore and develop the combination of existing tools and concepts. The three main hypotheses are the following: through the combination and development of appropriate spatial-sensitive tools and modelling procedures, it is possible to better measure and understand:

- the impact of certain externalities on property values, with a special emphasis on the nature of land use and vegetation;
- the effects of housing profiles on location choice; and, finally
- the spatial variability of marginal prices considering the buyer's profile.

More specifically, we think that

1. land use and vegetation has a significant impact on property values,
2. this impact may vary through space, especially regarding relative centrality,
3. the eventual non-stationarity of the effect of a property attribute may be linked to the buyers' household profiles, this statement leading to the two following, that is,
4. the motivations for moving and the residential choice criteria are not homogeneous but may vary with the buyers' or sellers' socio-economic characteristics, and,
5. if (4) proves right, the differences in choice criteria may be partly internalised in the selling prices. Also, appropriate econometric modelling may measure the resulting heterogeneity of implicit prices regarding the socio-economic profile of the buyer or/and the seller.

Furthermore, it is assumed that in order to validate

- (1) and (2), the extraction of land use and vegetation data from both aerial photographs and remote sensing data and its subsequent integration within the hedonic framework can prove efficient. Furthermore, the integration of the property's surrounding land use should integrate the various spatial scales of the hierarchy of spatial perception and should be spatial-sensitive, using for example Casetti's spatial expansion method;
- (4), detailed survey data procuring information about the motivation for moving and residential choice criteria of actual property buyers is needed. This data can be properly analysed using Correspondence Analysis and Logistic Regressions.
- (5), by integrating the buyer's socio-economic profile within the hedonic framework, and by using and comparing Casetti's spatial expansion method and the Geographically Weighted Regressions, the socio-spatial dimension of the possible heterogeneity of implicit prices may be measured.

Although the methods presented in the following chapters can be used in various contexts, they are applied here to Quebec City, which is the territory used for this study. Several datasets of single-family property transactions, mainly occurring in the 1986-1988 and 1993-2001 periods, are used for modelling purposes. The results and interpretations presented here are therefore limited to this type of tenure. Important datasets were available at the Centre de Recherche en Aménagement et Développement at the beginning of this project. These are databases describing the property – originating from the valuation roll – and accessibility measures to various infrastructures and services. Additional datasets were collected and computed during the research. Two land-use maps were developed from a set of aerial photographs and from a Landsat TM-5 image. Furthermore, a phone survey, held between 2000 and 2002, yielded additional information on over 800 property buyers' choice criteria and household profile.

The first chapter focuses on the marginal value of a property's surrounding land use, paying special attention to the role of vegetation. Two single-family property transaction datasets are analysed using hedonic models. The first series of models integrates land-use data derived from a mosaic of aerial photographs, whereas the second series integrates remote sensing data. Using Casetti-type expansion variables, the spatial variation of the impact of an attribute is examined. The second chapter studies the residential and neighbourhood choice criteria of actual property buyers. Canonical correspondence analysis is used in order to sort out actual choice criteria and suggest certain links with place-identity theories. Also, logistic regressions make it possible to analyse the probability of mentioning a criteria depending on the household profile and on the house location. Finally, in the third and last chapter, household-level data is introduced in hedonic models, and two spatial-sensitive methods – Casetti's expansion method and Geographically Weighted Regression – are compared, in order to analyse the eventual implicit price variations related to the buyer's profile. A general conclusion summarises the main findings and opens further research possibilities.

# References

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

Court A T, 1939, "Hedonic Price Indexes with Automotive Examples", in *The Dynamics of Automobile Demand* Ed. G Motors New York) pp 98-119

Goodman A C, 1998, "Andrew Court and the Invention of Hedonic Price Analysis" *Journal of Urban Economics* **44** 291-298

Lancaster K J, 1966, "A New Approach to Consumer Theory" *Journal of Political Economics* **74** 132-157

Rosen S, 1974, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" *Journal of Political Economy* **82** 34-55

Von Thünen J H, 1850 *Der isolierte Staat in Beziehung auf Landwirschaft und Nationalökonomie* (Scientia Verlag, Aalen)

# Chapter 1: The Impact of Surrounding Land Use and Vegetation on Single-Family house prices

**Résumé: Cet article explore l'impact des externalités géographiques liées à la nature de l'utilisation du sol par la modélisation hédonique de propriétés unifamiliales résidentielles. Cette approche intègre à la fois la hiérarchie spatiale, en accord avec les théories cognitives, et le compromis distance-centralité. À partir de deux échantillons de transaction de vente de propriété unifamiliales transigées à Québec, et intégrant des données d'utilisation du sol et de végétation extraites d'images satellites et de photographies aériennes, deux séries de modèles hédoniques sont élaborés. Un modèle de base intègre des variables de propriété, de recensement, d'accessibilité et de localisation. Dans une deuxième étape, des variables décrivant l'utilisation du sol et la végétation à plusieurs échelles sont intégrées. Enfin, dans une dernière étape, l'interaction entre l'utilisation du sol et la localisation est intégrée, la localisation étant mesurée en terme de proximité aux principaux lieux d'activité de la ville. Ceci permet de mesurer la variation de l'impact de l'environnement à travers la ville sur les valeurs immobilières. L'intégration significative de variables environnementales considérant la localisation améliore notre compréhension locale des impacts de l'utilisation du sol et de la végétation. Ceci améliore également la performance du modèle en réduisant significativement l'autocorrélation spatiale des résidus. Ce type de modèle pourrait s'avérer un outil performant pour évaluer l'impact fiscal de différentes politiques d'aménagement du territoire.**

**Abstract: The aim of this paper is to assess the marginal effect of land-use locational externalities on the sale price of single-family houses, considering various spatial scales – in accordance with perception theories – and trade-off with accessibility to the city centre. Using land-use and vegetation data derived from aerial photographs and Landsat TM satellite images, two sets of hedonic models using OLS regression are built using two samples of single-family properties sold in Quebec City. A standard model integrates property-specifics, Census factors, accessibility and location attributes. In a second model, land-use and vegetation variables are considered on various spatial scales, whereas a third step introduces the interaction effect of the surrounding land use with location, using car time distance to the main activity centres as the main indicator. This allows for analysing the spatial variation of the environmental impact throughout the city considering relative proximity to the centre. The successful integration of environmental variables considering location enhances our understanding of the local land-use and vegetation effects. It also improves the overall performance of the model while virtually removing spatial autocorrelation among residuals. Such**

**models could be used in order to assess the fiscal impacts of various land zoning by-law policies, thereby providing planning administrations with a useful decision-making tool.**

## 1.1 Introduction : Focus and Objectives of Study

The ongoing development of GIS, combined with the increasing availability of spatial data, is opening vast opportunities for better analysing and understanding our world. Urban planning authorities have already largely benefited from these developments to improve their handling of space-related issues, especially for facilities management. These new technologies have also contributed to the emergence of new approaches in urban studies and in the field of real estate. Real estate markets can be analysed in greater detail, thanks to computer-assisted mass appraisal (CAMA) relying on the combination of large databases with effective statistical and spatial analysis methods. Hedonic modelling is part of the prevalent statistical analyses used for analysing housing market components. A hedonic model expresses the market value of some composite good as a function of its various intrinsic and environmental attributes and reflects the envelope function of both supply and demand sides (Can, 1990; Can, 1992; Dubin, 1998; Rosen, 1974). This approach is derived from the consumer theory that states that the characteristics of any commodity determine its utility (Lancaster, 1966). Applied to the housing market, the coefficients of the house-price function reflect the probability distribution of the combined buyers' and sellers' willingness to pay and be paid for the defined attributes, as an expression of their own utility level.

The main purpose of this paper is to model the marginal effect of neighbouring land cover on the market value of residential properties. Furthermore, special attention is given to the trade off between accessibility to jobs and services – using Car-Time Distance (CTD) to the Main Activity Centres (MACs) as a proxy[1] – and environmental locational externalities[2] – using land-use and vegetation data as a proxy. Our test case integrates large databases into a GIS (Census data, services and facilities, aerial

---

[1] In order to partly integrate the polycentricity of the city and based on previous work by Des Rosiers *et al.* (2000), we use the mean car time distance to both the old city core and Laval University as a proxy for accessibility to main work and shopping locations.

[2] The term "locational" externalities, proposed by Orford (1999), refers to various types of externalities related to the property's specific location.

photographs, remote sensing data), combining statistical (semi-log regression analysis) and spatial analysis (image classification, buffer functions, and autocorrelation). A straightforward methodology for city planners and fiscal authorities is described in this paper to improve coordination of their actions. In fact, assessing the effect of land zoning on consumer satisfaction could improve future city development and, by maximising the overall utility of residential owners, increase values and the tax base.

One of the main difficulties with this kind of application comes from the prevalence of systemic relationships among several closely related phenomena for competing land uses, which generates spatial patterns. In turn, this geographical structure involves spatial constraints leading to trade-offs among limited choice sets. Therefore, it is often difficult to distinguish from among specific effects when they show similar spatial patterns. For example, home buyers often have to choose between proximity to nature (or large lot size for their family) and access to urban amenities. Is the hedonic modelling approach suitable for distinguishing between those very different, yet related and intricate factors?

In order to integrate environmental proxies, both land-use maps and remote sensing images are used in this paper. A manually supervised classification method is applied to aerial colour photographs (shot in 1987) to build a standard land-use map. In addition, a semi-automatic classification procedure is applied to a Landsat TM 5 image (1999) to derive land-use categories and vegetation indices. The usefulness of the resulting classifications is tested by integrating the ensuing data into two distinct sets of residential hedonic models built on town cottage sales that occurred in Quebec City from 1986 to 1987 and from 1993 to 1996. Town cottages correspond to single-family detached houses with more than one above-ground floor.

In Section 0, previous work and theoretical bases justifying certain technical choices are reviewed. Section 0 presents the data bank and modelling procedure, whereas Section 0 reveals the empirical results of the analyses. A discussion follows in Section 0, and concluding remarks are given in 0.

## 1.2 Previous Work

One of the main difficulties – but also one of the major stakes – in the modelling of urban – more generally geographical – phenomena is to handle and integrate the spatial dimension. Furthermore, property values are considered by many authors to be the result of the complex and intricate combination of externalities and location rents (Can, 1992; Dubin, 1998; Hoch and Waddell, 1993; Krantz *et al.*, 1982; Strange, 1992). According to Tse (2002), the relationship between house values and location effects results from unobservable links across house attributes coupled with the heterogeneity of the market and of its players. That is why it is important to distinguish between structural spatial dependence through observations and spatial dependence in error terms. The latter is "generally due to omitted variables, which are themselves spatially correlated, or due to errors in measurement that are systematically related to location" (Tse, 2002:1168). Better integrating land-use and vegetation locational externalities should help to lower the spatial autocorrelation in error terms.

Moreover, homeowners and local communities are increasingly preoccupied with environmental issues and sensitive to the overall quality of their neighbourhood, as shown by the abundant and growing literature about NIMBYs (Not In My Backyard), LULUs (Locally Unwanted Land Use) or even NOPEs (Not On Planet Earth) (Davy, 1997; Foldvary, 1994; McAvoy, 1999).

Numerous preference studies have highlighted the positive impact of vegetation in urban scenes (Cooper-Marcus, 1982; Kaplan, 1983), while also showing that the relationship is not monotonous and that an excess of vegetation can affect preferences negatively (Buyhoff *et al.*, 1984; Payne, 1973). Following a recent survey on perception of the environmental quality of residential real estate held in Geneva, Lugano and Zurich, Switzerland, Bender *et al.* (2000) show that from among eight criteria, degree of quietness and distance to nature were the two most important factors rated in Geneva, whereas quality of view was considered most important for Zurich's respondents.

Certain hedonic models have integrated environmental quality measures, analysing the impact of noise (Freeman, 1979; Huang and Palmquist, 2001; Weinstein, 1976), air quality (Anderson and Crocker, 1971; Beron *et al.*, 2001; Graves *et al.*, 1988; Murdoch and Thayer, 1988; Smith and Huang, 1995), water quality (Des Rosiers *et al.*, 1999; Leggett and Bockstael, 2000; Michael *et al.*, 1996) or proximity to potential toxic sites (see Boyle and Kiel, 2001, for a survey of the models covering the three last themes). However, as pointed out by Nasar (1983), sight can be considered the most important sense in our immediate interaction with our surroundings. What potential buyers can see from and around a property has an impact on their estimation of its value. Vegetation attributes – principally trees – have often been studied, with a positive contribution to property values associated with tree presence generally ranging from +3% to +8% (Anderson and Cordell, 1985 & 1988; Luttik, 2000; Morales *et al.*, 1976; Seila and Anderson, 1982; Tyrvainen and Miettinen, 2000) and specific landscaping attributes deserving an additional premium (Des Rosiers *et al.*, 2002). The overall quality of the view is sometimes integrated, showing positive premiums (Do and Sirmans, 1994; Rodriguez and Sirmans, 1994). Specific elements of the view have also been studied, such as a view of a forest, mountain, lake, river, ocean or open space (Benson *et al.*, 2000; Benson *et al.*, 1998; Luttik, 2000; McLeod, 1984; Powe *et al.*, 1995).

It is important to note that in these hedonic models, the data relating to environmental features is collected either *in situ* by visual observations (Benson *et al.*, 2000; Benson *et al.*, 1998; Des Rosiers *et al.*, 2002; Luttik, 2000; Morales *et al.*, 1976), gathered manually from maps (Gallimore *et al.*, 1996; Garrod and Willis, 1992; Powe *et al.*, 1995; Tyrvainen and Miettinen, 2000), or from photographs (Anderson and Cordell, 1985 & 1988; Seila and Anderson, 1982). The fact that these methods are highly time-consuming and potentially subjective probably explains in part the relative scarcity of environmental consideration in hedonic modelling.

Some recent residential hedonic models integrate environmentally related factors computed within a GIS. Powe *et al.* (1997) calculate distance to and area of forests using a GIS in order to build an index for measuring access to woodland in the New Forest Area, Great Britain, England. Lake *et al.* (1998 & 2000) compute view extent and composition from a Digital Elevation Model (DEM) and land-use data in Glasgow, Scotland. View-obstructing buildings and distance-decay weightings are considered in order to build refined visibility variables. However, only three of the visual impact variables appear significant (amount of road and rail visible from the property) whereas some present unexpected coefficient signs. The authors point to the difficulty of producing a precise DEM as well as that of characterising vegetation land use. More recently, Patterson and Boyle (2002) applied a similar approach to properties of the Farmington River Valley in Connecticut, U.S.A. Viewshed extension and composition (developed area, agriculture, woodland and water) is calculated in a 1000-metre radius around each property. For each land-use category, both the percentage of area and the percentage of visible area within one kilometre are computed. Surprisingly, the coefficient of overall visibility extent is negative, as is the percentage of developed area and water surface in a one kilometre radius. When all environmental variables are taken into account, both visible developed area and visible forest proportions have a negative impact on property values.

It is interesting to note that in very few models only have researchers tried to integrate the environmental impact on various spatial scales, although the hierarchical approach provides an interesting framework for apprehending complex spatial systems (Wu *et al.*, 1997; Yuan, 2000). Perceptual regions can also be defined as hierarchical spatial units (Mesarovic *et al.*, 1970; Reginster and Edwards, 2001). In fact, in keeping with previous work concerning the differentiation of spatial perception according to spatial behaviour (Remy and Voyé, 1992) and to activity patterns (Walmsley and Lewis, 1993), Reginster and Edwards (2001) consider life spaces hierarchically ordered and directly related to location and activities. The perceptual region of a household can be defined by two main spaces: the vista space and the local displacement-reinforcement space (Reginster and Edwards, 2001). The first is the spatial region with perceptually similar characteristics apprehended from a single place, which corresponds to a sense of belonging to what is considered home. The second is the region of belonging to the immediate environment around the vista space, which can be apprehended by a memory reinforcement rating that is related to the use of locomotion inside reasonable temporal limits. We assume that the perceptual regions, hierarchically structured, should be better integrated into the assessment of the impact of locational externalities on house prices. For each scale of observation, one reality pattern can be seen (Hay *et al.*, 2001), and it is the intricate system of patterns that has to be understood. A recent multilevel model using transaction sales in the inner area of the Welsh capital of Cardiff, integrates locational externalities – measured within a GIS and derived from the Housing Condition Survey (HCS) – at various spatial scales, e.g. house level, street level, HCS-area level and community level (Orford, 2002). Our hypothesis is that the use of buffer functions measuring the land-use characteristics for various distances constitutes an efficient proxy for the previously defined hierarchical perceptual regions.

It is also important to consider the eventual variation *through* space of the impact of any locational externality. If the objective is strictly to predict property values, it is possible to integrate the spatial dimension using a "location factor" or a "location value response surface," as is often done in real estate assessment (Eichenbaum, 1989; Gallimore *et al.*, 1996; Shi *et al.*, 2000). However, in order to explain the spatial influence of externalities, additional well-defined variables and interaction terms can be integrated into existing hedonic models. Alonso's bid rent function theory (Alonso, 1964), extending Von Thünen's model to urban land use (Von Thünen, 1850), suggests that a number of economic and social patterns are the expression of a function of the distance to the CBD. These functions are the solution to an economic equilibrium for the market of space. In fact, the derivation of bid-price curves represents a set of combinations of land rents and distances from the CBD to which the potential buyer is indifferent. A specific bid-price curve corresponds to each utility level. The residential bid price curve is defined as the "set of prices for land the individual could pay at various distances while deriving a constant level of satisfaction" (Alonso, 1964: 59). Some extensions of the standard urban model include the environmental quality variations in the structure of urban equilibrium (Latham and Yeates, 1970; Newling, 1966; Papageorgiou and Pines, 1998). Specifically, in a recent paper an attempt was made to assess the impact of environmental amenities on the urban residential land-use structure (Cho, 2001). Following Alonso's theory and its more recent developments, our assumption is that the environmental marginal contribution to the housing market is uneven through space, and that its spatial variation can, in most cases, be expressed as a function of the distance to the MACs using Casetti's spatial expansion method (Can and Megbolugbe, 1997; Casetti, 1972). Geoghegan *et al.* (1997) considered these assumptions and built a residential hedonic model covering urban, suburban and rural areas in the Patuxent Watershed, Maryland, U.S.A. Interactive variables are used to analyse the spatial variation of derived landscape pattern indices (fragmentation and diversity) and proportion of land-use types within 100 metres and one kilometre around each property considering also the distance to Washington, DC. Derived landscape pattern indices proved particularly significant when combined with the distance to the city, showing that the marginal contribution of landscape characteristics vary, depending on the location (Geoghegan *et al.*, 1997). Unfortunately, in the absence of multicollinearity or heteroskedasticity measures, the obtained coefficients can be interpreted only with considerable caution. Furthermore, several non significant variables were kept in the final model specification, possibly inducing biases in the value of significant coefficients.

To the best of our knowledge in current research, there is a need to better analyse the impact of locational externalities on real estate markets. Combining the efficiency of GIS, remote sensing and multiple regression analysis, this paper proposes a straightforward and easy-to-use method in order to integrate land-use locational externalities in hedonic property models. Furthermore, by considering various measurement scales for locational externalities as well as the variable impact of distance to MACs, this study lays particular emphasis on the spatial variation (scale and location) of land-use related externalities.

## 1.3 Data Banks and Modelling Procedure

Two data sources were used to extract the land cover information: 126 aerial colour photographs (1/20 000 scale) shot in June 1987, and a Landsat TM 5 satellite image of 30 metres resolution obtained for August 23rd, 1999. Due to discrepancies in the dates, two sets of hedonic models are used. The first set analyses 724 houses sold in 1986 and 1987, sale prices ranging from \$40 000 to \$180 000. Mean (median) value is about \$86 000 (\$80 000). Some 42 cases were randomly selected for validation purposes, the remaining transactions (682) being used for modelling. The second set of transactions contains 2 278 houses sold between 1993 and 1996 for prices ranging from \$50 000 to

$250 000. Average (median) value is around $123 000 ($119 500). Again, a sample of 2 058 properties was used for modelling, with the remaining cases (220) being kept separate for validation purposes.

Each transacted house is described by 80 property-specifics derived from the city's assessment records. Census data, available for 786 enumeration areas, is used to reflect the social and demographic attributes of the neighbourhood. In order to avoid insidious multicollinearity and in line with previous work showing the efficiency of integrating Census PCA-factors as proxy for the neighbourhood socio-economic status (Des Rosiers *et al.*, 2000), factor scores of two principal component analyses held on 1986 and 1991 Census attributes (Table 1) were used. As for accessibility to the MACs, and in order to partially integrate the polycentricity of the city (Musterd and Van Zelm, 2001), the average CTD to the previously identified principal employment and shopping locations – e.g. the historical centre of Quebec City and Université Laval– were computed. CTD were computed for each property using the TransCAD GIS to carry out simulations using a city-wide road network (19 250 street intersections), and taking into account various impedance constraints and turn penalties (Nijkamp *et al.*, 1993; Thériault *et al.*, 1999). Proximity to power lines and freeways, causing eventual visual and noise externalities (Delaney and Timmons, 1992; Des Rosiers, 2002), were also measured within a GIS. The definition and statistical description of significant variables are shown in Table 2 and Table 3.

Table 1. PCA of socio-economic Census attributes (1986 and 1991 Census)

| | Rotated Component Matrix | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Census Factor 1** | | **Census Factor 2** | | **Census Factor 3** | | **Census Factor 4** | |
| **Census Attributes** | **1986** | **1991** | **1986** | **1991** | **1986** | **1991** | **1986** | **1991** |
| % 0-14 years | -0.766 | -0.811 | 0.580 | 0.505 | | | | |
| % 15-24 years | | | -0.378 | | | | 0.760 | 0.808 |
| % 25-44 years | | | 0.917 | 0.933 | | | | |
| % 45-64 years | | | -0.881 | -0.883 | | | | |
| % 65+ years | 0.662 | 0.648 | -0.564 | -0.548 | | | -0.329 | -0.370 |
| % women | 0.554 | 0.545 | -0.319 | -0.339 | | | | |
| Persons per household | -0.952 | -0.959 | | | | | | |
| % non-family households | 0.935 | 0.945 | | | | | | |
| % single-person households | 0.931 | 0.933 | | | | | | |
| Children per family | -0.843 | -0.574 | | | | | | |
| % single-parent families | 0.745 | 0.759 | -0.325 | | | -0.304 | | |
| % families with children | -0.833 | -0.820 | | | | | | |
| % families children 0-6 years | | | 0.860 | 0.850 | | | | |
| % families children 6-14 years | -0.613 | -0.585 | 0.611 | 0.557 | | | | |
| % detached dwellings | -0.614 | -0.932 | | | | | | |
| % dwellings in large buildings | 0.500 | 0.473 | | | 0.378 | | | |
| Persons per room | -0.330 | | 0.472 | 0.425 | -0.530 | -0.668 | | |
| % dwellings built before 1946 | 0.508 | 0.525 | | | | | -0.559 | -0.494 |
| % dwellings built 1946-60 | | 0.309 | -0.662 | -0.504 | | | | |
| % dwellings built 1961-70 | | | -0.371 | -0.575 | | | 0.800 | 0.636 |
| % of tenants | 0.905 | 0.910 | | | | | | |
| % households with housing costs > 30% of income | 0.827 | 0.635 | | | -0.336 | -0.306 | | |
| % secondary school diploma | | | | | 0.915 | 0.903 | | |
| % university degree | | | | | 0.912 | 0.944 | | |
| % men with college degree | | | | | 0.943 | 0.935 | | |
| % women with college degree | | | | | 0.923 | 0.919 | | |
| Household income $ | -0.643 | -0.699 | | | 0.688 | 0.629 | | |
| % moved during last 5 years | 0.757 | 0.605 | 0.381 | 0.608 | | | | |
| Population density persons/hectare | 0.735 | 0.760 | | | | | | |
| Dwelling density per hectare | 0.786 | 0.804 | | | | | | |
| **Percentage of explained variance** | **37.4** | **36.4** | **17.2** | **16.0** | **16.2** | **15.9** | **7.0** | **6.3** |
| **Interpretation** | **Urban Centrality** | | **Family Cycle** | | **Socio-Economic Status** | | **Replacement** | |
| Positive values | Small households in the city centre | | Young families with children living in new suburbs | | Well-educated persons with high income and large houses | | Young adults living with their parents in low density suburbs | |
| Negative values | Family households, homeowners in suburbs | | Empty-nesters and retirees living in older suburbs | | Low educated poor persons in overpopulated houses | | Retirees living in the old city core | |

Adapted from Des Rosiers, F., Thériault, M. and P. Villeneuve (2000) Sorting Out Access and Neighbourhood Factors in Hedonic Price Modelling. Journal of Property Valuation and Investment, Vol 18(3); 291-315.

Table 2. Set 1: Operational definition and statistical description of significant variables

| | Variable name | Description | Type* | Main Sample (N=682) Minimum | Maximum | Mean | Std. Deviation | Validation Sample (N=42) Mean | Std. Deviation |
|---|---|---|---|---|---|---|---|---|---|
| | SPRICE | Sale price of the property ($) | C | 40 000 | 180 000 | 86 155 | 31 837 | 86 595 | 34 067 |
| | LNSPRICE | Natural logarithm of the sale price ($) | C | 10.60 | 12.10 | 11.30 | 0.36 | 11.29 | 0.39 |
| Property Specifics | LTaxRate | Local tax rate ($/100$ of assessed value) | C | 1.6 | 4.2 | 0 | 0 | 0.03 | 0.01 |
| | AppAge | Apparent age (years) | C | 0 | 48 | 17.3 | 13.2 | 19.10 | 12.7 |
| | LnAppage | Natural logarithm of apparent age (years) | C | 0.00 | 3.89 | 3 | 1 | 2.73 | 0.82 |
| | LnAppage_Sq | Natural logarithm of apparent age squared (years) | C | 0.0 | 7.5 | 1.1 | 1.63 | 0.7 | 0.9 |
| | LivArea | Living area ($m^2$) | C | 81.0 | 273.7 | 144.2 | 34.7 | 150.16 | 33.74 |
| | LnLivarea | Natural logarithm of the living area ($m^2$) | C | 4.4 | 5.6 | 4.9 | 0.23 | 5.0 | 0.2 |
| | LnLivarea_Sq | Natural logarithm of the living area squared ($m^2$) | C | 0.00 | 0.39 | 0.05 | 0.07 | 0.04 | 0.05 |
| | LotSize | Lot size ($m^2$) | C | 211.4 | 2952.0 | 756.1 | 403.8 | 790.93 | 352.80 |
| | LnLotsiz | Natural logarithm of the lot size ($m^2$) | C | 5.4 | 8.0 | 6.5 | 0.44 | 6.59 | 0.40 |
| | Quality | House quality index | C | -2 | 2 | 0.00 | 0.42 | -0.05 | 0.31 |
| | InferiorFoundation | Inferior foundation quality | B | 0 | 1 | 0.16 | 0.37 | 0.21 | 0.42 |
| | FinishedBasement | Finished basement | B | 0 | 1 | 0.30 | 0.46 | 0.29 | 0.46 |
| | KitchenCab. | Kitchen cabinets made of hardwood | B | 0 | 1 | 0.21 | 0.41 | 0.24 | 0.43 |
| | FirePlace | Number of fireplaces | C | 0 | 4 | 0.40 | 0.54 | 0.48 | 0.55 |
| | Washrooms | Number of washrooms | C | 1 | 4.5 | 1.54 | 0.43 | 1.61 | 0.42 |
| | Dishwasher | Build-in dishwasher | B | 0 | 1 | 0.48 | 0.50 | 0.55 | 0.50 |
| | DetGarage | Detached garage | B | 0 | 1 | 0.21 | 0.41 | 0.26 | 0.45 |
| | InGroundPool | In-ground pool | B | 0 | 1 | 0.06 | 0.23 | 0.07 | 0.26 |
| Neighbourhood Quality and Accessibility | CSF1 Core/Periphery | Socio-economic component 1 (Centrality) | B | -1.78 | 2.34 | -0.48 | 0.70 | -0.53 | 0.74 |
| | CSF2 FamCycle | Socio-economic component 2 (Family Cyle) | B | -2.21 | 2.84 | -0.01 | 1.17 | 0.21 | 1.29 |
| | CSF3 Socio-EconomicStatus | Socio-economic component 3 (Socio-Economic status) | B | -2.37 | 3.49 | 0.57 | 1.15 | 0.85 | 1.24 |
| | CarTimeMACs | Car-time distance to main activity centres (min) | B | 4.63 | 27.02 | 12.42 | 4.13 | 12.02 | 4.44 |
| | CTDtoMACs_Sq | Car-time distance to main activity centres squared (centered) | B | 0.00 | 213.81 | 17.05 | 20.48 | 19.36 | 33.37 |
| Land-Use Locational Externalities | Ln%Mineral 100m | Natural logarithm of density of mineral surfaces within a 100-m radius | C | 0.00 | 4.09 | 1.59 | 1.09 | 1.54 | 1.17 |
| | Water 100m | Percentage of water surfaces within a 100-m radius | C | 0.00 | 35.00 | 0.31 | 2.37 | 0.14 | 0.89 |
| | Woodland 1km | Percentage of woodlands within a 1-km radius | C | 0.00 | 85.53 | 17.93 | 18.03 | 18.05 | 20.37 |
| | LnIndustrial 500m | Natural logarithm of percentage of industrial land-cover within a 500-m radius | C | 0.00 | 4.1 | 0.32 | 0.9 | 0.18 | 0.60 |
| Interactive Variables | Commercial500m * CTDtoMACs | Percentage of commercial land-use within a 500-m radius * Car-time distance to main activity centres | C | -299.27 | 48.26 | -20.82 | 40.16 | -25.51 | 49.93 |
| | Lawn300m * CTDtoMACs | Percentage of lawn within a 300 m radius * Car-time distance to main activity centres | C | -177.6 | 37.0 | -0.1 | 9.8 | -0.95 | 11.08 |
| | AgricCultPast300m * CTDtoMACs | Percentage of agricultural cultures and pastures within a 300 m radius * Car-time distance to main activity centres | C | -3.8 | 173.0 | 0.6 | 7.9 | 1.84 | 11.25 |

***Type: C=continuous; B=Binary**

Table 3. Set 2: Operational definition and statistical description of significant variables

| | | | | Main Sample (N=2 058) | | | | Validation Sample (N=220) | |
|---|---|---|---|---|---|---|---|---|---|
| | **Variable name** | **Description** | **Type*** | **Minimum** | **Maximum** | **Mean** | **Std. Deviation** | **Mean** | **Std. Deviation** |
| | SPRICE | Sale price of the property ($) | C | 50 000 | 250 000 | 123 657 | 41 352 | 122 703 | 39 435 |
| | LNSPRICE | Natural logarithm of the sale price ($) | C | 10.82 | 12.43 | 11.67 | 0.34 | 11.66 | 0.33 |
| Property Specifics | LTaxRate | Local tax rate ($/100$ of assessed value) | C | 1.19 | 2.73 | 2 | 0 | 2.22 | 0.40 |
| | AppAge | Apparent age (years) | C | 0 | 52 | 14.2 | 13.8 | 15.20 | 14.5 |
| | LnAppage | Natural logarithm of apparent age (years) | C | -0.69 | 3.95 | 2 | 1 | 2.11 | 1.28 |
| | LnAppage_Sq | Natural logarithm of apparent age squared (years) | C | 0.0 | 7.4 | 1.6 | 1.89 | 1.6 | 1.9 |
| | LivArea | Living area ($m^2$) | C | 76.2 | 287.4 | 146.7 | 33.2 | 148.51 | 31.44 |
| | LnLivarea | Natural logarithm of the living area ($m^2$) | C | 4.3 | 5.7 | 5.0 | 0.22 | 5.0 | 0.2 |
| | LnLivarea_Sq | Natural logarithm of the living area squared ($m^2$) | C | 0.00 | 0.48 | 0.05 | 0.07 | 0.04 | 0.06 |
| | LotSize | Lot size ($m^2$) | C | 178.6 | 4 666.4 | 675.2 | 377.5 | 745.12 | 461.67 |
| | LnLotsiz | Natural logarithm of the lot size ($m^2$) | C | 5.2 | 8.4 | 6.4 | 0.39 | 6.50 | 0.44 |
| | Quality | House quality index | C | -2 | 2 | -0.01 | 0.30 | -0.02 | 0.26 |
| | SuperiorRoofQual | Superior roof quality | B | 0 | 1 | 0.00 | 0.07 | 0.00 | 0.00 |
| | SuperiorFloorQual | Superior floor quality | B | 0 | 1 | 0.66 | 0.47 | 0.69 | 0.46 |
| | FacingStoneBrick51%+ | More than 50% of facing made of stone or brick | B | 0 | 1 | 0.37 | 0.48 | 0.35 | 0.48 |
| | FinishedBasement | Finished basement | B | 0 | 1 | 0.36 | 0.48 | 0.34 | 0.48 |
| | Storey | Number of storeys | C | 1.5 | 2 | 1.84 | 0.23 | 1.84 | 0.24 |
| | Stairs | Stairs made of hardwood | B | 0 | 1 | 0.49 | 0.50 | 0.45 | 0.50 |
| | Fireplace | Number of fireplaces | C | 0 | 2 | 0.39 | 0.51 | 0.35 | 0.50 |
| | Dishwasher | Build-in dishwasher | B | 0 | 1 | 0.70 | 0.46 | 0.74 | 0.44 |
| | AttGarage | Attached garage | B | 0 | 1 | 0.19 | 0.40 | 0.15 | 0.36 |
| | DetGarage | Detached garage | B | 0 | 1 | 0.19 | 0.39 | 0.23 | 0.42 |
| | InGroundPool | In-ground pool | B | 0 | 1 | 0.08 | 0.27 | 0.07 | 0.26 |
| | WaterSewer | Linkage to the municipal waterworks and sewer network | B | 0.00 | 1.00 | 0.98 | 0.13 | 0.99 | 0.12 |
| Neighbourhood Quality, Accessibility and Localisation | CSF2 FamCycle | Socio-economic component 2 (Family Cycle) | C | -2.46 | 2.78 | 0.68 | 1.18 | 0.60 | 1.18 |
| | CSF3 Socio-EconomicStatus | Socio-economic component 3 (Socio-Economic status) | C | -1.86 | 2.46 | 0.34 | 0.84 | 0.29 | 0.85 |
| | CSF4 Replacement | Socio-economic component 4 (Replacement) | C | -2.10 | 1.02 | 0.08 | 0.36 | 0.07 | 0.38 |
| | CTDtoMACs | Car-time distance to main activity centres (min) | C | 4.63 | 27.62 | 13.47 | 3.68 | 13.55 | 3.86 |
| | CTDtoMACs_Sq | Car-time distance to main activity centres squared (centered) | C | 0 | 199.8 | 13.57 | 19.2 | 14.81 | 19.35 |
| | Highway150m | Located within 150-m of nearest highway | C | 0 | 1.00 | 0.02 | 0.13 | 0.04 | 0.19 |
| Land-Use Locational Externalities | ResidMatureTrees 100m | Percentage of residential land-use with mature trees within a 100-m radius | C | 0 | 85.5 | 15.3 | 16.8 | 15.55 | 16.09 |
| | ResidMatureTrees 500m | Percentage of residential land-use with mature trees within a 500-m radius | C | 0 | 46.9 | 12.2 | 9.8 | 12.53 | 9.41 |
| | ResidLowTreeDensity 500m | Percentage of residential land-use with low tree density within a 500-m radius | C | 0.08 | 36.0 | 15.8 | 6.1 | 16.0 | 6.2 |
| | Woodlands 500m | Percentage of woodlands within a 500-m radius | C | 0 | 81.7 | 16.1 | 15.2 | 16.5 | 15.8 |
| | Agriculture/Disp.Trees 100m | Percentage of agricultural land with dispersed trees within a 100-m radius | C | 0 | 49.7 | 1.9 | 4.4 | 1.7 | 3.5 |
| | NDVI StdDev 1km | NDVI standard deviation within a 1-km radius (Heterogeneity of land-use pattern) | C | 0.18 | 0.58 | 0.33 | 0.07 | 0.33 | 0.07 |
| | Mean NDVI 40m | Mean NDVI value within a 40-m radius | C | -0.77 | 0.43 | -0.30 | 0.16 | -0.28 | 0.15 |
| Interactive Variables | ResidMatureTrees100m * CTDtoMACs | Percentage of residential land-use with mature trees within a 100-m radius * Car-time distance to main activity centres | C | -582.7 | 207.1 | -23.16 | 83.43 | -24.95 | 71.94 |
| | Woodlands500m * CTDtoMACs | Percentage of woodlands within a 500-m radius * Car-time distance to main activity centres | C | -104.13 | 692.52 | 20.73 | 58.59 | 25.23 | 68.15 |
| | Agric/DispTrees100m * CTDtoMACs | Percentage of agricultural land with dispersed trees within a 100-m radius * Car-time distance to main activity centres | C | -119.7 | 192.7 | 2.7 | 14.8 | 3.9 | 12.3 |
| | Agric/BarrenLand500m * CTDtoMACs | Percentage of agriculture/barren land within a 500-m radius * Car-time distance to main activity centres | C | -98.6 | 168.3 | 1.7 | 17.7 | 0.9 | 15.3 |

***Type: C=continuous; B=Binary**

Aerial photographs and a Landsat TM 5 image were used to provide land-use data around the location of each transaction. The GIS-integrated aerial photographs mosaic was categorised using a manual procedure. The territory was divided into roughly 8000 polygons, first using the road network to build city blocks, second manually dividing them when the variations in land use types were important. Two types of information were associated with each polygon: the main land-use type defined for three levels of categorisation (see Table 4 for land-use-type definitions) and an estimation of the density of trees, lawn, built and mineral cover – that is, all concrete surfaces except buildings – all expressed in percentage of the total area of each polygon.

The Landsat image was categorised using the semi-automated ISODATA (Iterative Self-Organising Data Analysis) technique, widely used and implemented in some GIS packages (Duda and Hart, 1973). The 16 initial categories – interpreted using the aerial photographs – were combined in nine final categories (Table 5). Furthermore, the Normalised Difference Vegetation Index (NDVI) was computed using the red wavelengths of the visible spectrum (RED: from 0.63 to 0.69 μm) and the near infrared (NIR: 0.76 to 0.90 μm) spectral wavelengths. NDVI is a widely used and sensitive indicator of the green biomass (Tucker, 1979; Tueller, 1989; Wu *et al.*, 1997). Chlorophyll strongly absorbs visible light while intensely reflecting NIR; this causes a greater difference between the reflections in those wavelength windows. The NDVI index, ranging from – 1 to +1 – higher values indicating higher density of vegetation – is defined as:

$$NDVI = (NIR - RED) / (NIR + RED) \qquad (1)$$

Table 4. Land-use/cover classification system for aerial photographs categorisation

| Level 1 | | Level 2 | | Level 3 | |
|---|---|---|---|---|---|
| 1 | Urban | 11 | Residential | 111 | Residential high-built density |
| | | | | 112 | Residential medium-built density |
| | | | | 113 | Residential low-built density |
| | | 12 | Commercial and services | 120 | Commercial and services |
| | | 13 | Industrial | 130 | Industrial |
| | | 14 | Transport, communication and facilities | 140 | Transport, communication and facilities |
| | | 15 | Vacant lots | 151 | Barren vacant lots |
| | | | | 152 | Vacant lots with vegetation |
| | | 16 | Landscaped green areas | 160 | Landscaped green areas |
| | | 17 | Recreational facilities | 170 | Recreational facilities |
| 2 | Agriculture | 21 | Pastures | 210 | Pastures |
| | | 22 | Cultures | 220 | Cultures |
| | | 23 | Pastures and cultures | 230 | Pastures and cultures |
| 3 | Low vegetation | 31 | Lawn | 310 | Lawn |
| | | 32 | Shrubs | 320 | Shrubs |
| | | 33 | Lawn with shrubs | 330 | Lawn with shrubs |
| 4 | Woods | 40 | Woods | 400 | Woods |
| 5 | Water surfaces | 50 | Water surfaces | 500 | Water surfaces |
| 6 | Wetland | 60 | Wetland | 600 | Wetland |
| 7 | Barren land | 70 | Barren land | 700 | Barren land |

Table 5. Final classification of Landsat TM 5 image and NDVI variables

| Level 1 | Level 2 |
|---|---|
| Water surfaces | Water surfaces |
| Woods | Woods |
| Agriculture / Lawn | Agriculture / Lawn |
| Agriculture / Dispersed trees | Agriculture / Dispersed trees |
| Agriculture / Barren Land | Agriculture / Barren Land |
| Residential with trees | Residential with mature trees<br>Residential with low-tree density |
| Residential high-built density | Residential high-built density |
| Industrial / Built infrastructures | Industrial / Built infrastructures |
| **NDVI variables** | **Interpretation** |
| Mean NDVI value | Greenness of area |
| NDVI Standard Deviation | Relative homogeneity of area in terms of vegetation |
| NDVI Range | Indication of difference between extreme values in terms of vegetation/built surfaces |

***Grey cells: NDVI variables**

Land-use information was subsequently computed for each property using buffer functions. The radii of the buffers range from 40 metres to one kilometre, a 40 metres radius buffer – representing roughly an area of 8 pixels of 25-metre resolution (8·625 square metres) – being the minimum significant computable area with this medium. The variety of radii used (40, 100, 500 and 1000 metres) is an attempt to consider the previously defined different hierarchical scales of environmental perception, the buffers representing a crude proxy for vista and local displacement-reinforcement spaces. Whereas the 40 and 100-metre radii are meant to approximate the vista space, 500 metres represent roughly the limits of frequent walking distance. The one-kilometre radius was added to measure the overall effect of the large neighbourhood.

Interactive variables are computed in order to estimate the potential variation of the land-use impact on house value considering location (CTD to the MACs). In order to avoid multicollinearity, interactive variables are built using previously centered variables, thereby reflecting the departure from the overall market's average values (Jaccard *et al.*, 1990: 31).

The models, based on OLS specification, are computed using a stepwise procedure. The logarithm of the sale price is used as the dependent variable, thus optimising the linear relationship with the input variables. Bearing in mind the objective of transferring the developed methodology to planning authorities, both "minimal mathematical complexity" and "ease of use" criteria were considered. Therefore, alternative methods – such as generalised least squares (GLS) (Fletcher *et al.*, 2000; Goodman and Thibodeau, 1995), geographically weighted regression (GWR) (Fotheringham *et al.*, 2002), spatial autoregressive specification (SAR) (Anselin, 1990; Pace and Gilley, 1997), SAR with Similarity components (SARS) (Besner, 2002), spatial filtering techniques (Cliff and Ord, 1981; Getis, 1990; Getis and Griffith, 2002; Griffith, 1996), artificial neural networks (ANN) (Din *et al.*, 2001; Nguyen and Cripps, 2001; Tay and Ho, 1992; Worzala *et al.*, 1995), the stochastic approach (Tse, 2002), or the multilevel approach (Orford, 2000 & 2002) – although presenting interesting avenues, were not used at this time. Some of them still lack conclusive results (Din *et al.*, 2001; Tse, 2002; Worzala *et al.*, 1995), and further investigating these approaches from a comparative perspective although undoubtedly potentially useful, is beyond the scope of this paper.

A three-step procedure is applied to each set of transactions. A first model is built integrating property-specifics, Census factors, location and accessibility attributes. A second model further integrates land-use locational externalities, whereas a third model adds interaction effects. Some of the variables have been expressed both in their linear and quadratic form, in order to check for an eventual non-linear effect. Squared variables are computed using the previously centered original variable, thus avoiding insidious multicollinearity effects. Eventual time shift effect was controlled using a temporal variable but did not prove significant, indicating a relative price stability of studied markets over time for both periods (1986-1987 and 1993-1996).

Tests are conducted throughout the modelling process in order to verify the eventual presence of multicollinearity, heteroskedasticity and spatial autocorrelation issues (Anselin and Can, 1986; Anselin and Rey, 1991; Des Rosiers and Thériault, 1999; Goodman and Thibodeau, 1995 & 1997). One of the advantages of OLS regression is the possibility of measuring multicollinearity using the Variance Inflation Factor (VIF), which indicates how strongly each explaining variable is correlated to the others. The higher the VIF value, the more multicollinearity. Knowing that the goal of hedonic modelling is to distinguish the marginal effects of various attributes, the eventual dilution of a specific effect on several multicollinear variables must be detectable. The presence of heteroskedasticity is verified both visually and using the Goldfeld-Quandt (1965) test, while spatial autocorrelation in the residuals structure is measured using Moran's index (Moran, 1950).

## 1.4 Summary of Results

### 1.4.1 Set 1 Models: Land-Use Data Extracted from Aerial Photographs

The first set of models is built with a sample of 682 single-family cottages sold in 1986 and 1987. The land-use locational externalities introduced in models $1_B$ and $1_C$ are extracted from the aerial photographs.

# Table 6. Set 1 hedonic models

| | Model 1A | | | | Model 1B | | | | Model 1C | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dependent Variable: LnSprice | B | (t-value) | Sig | VIF | Model B | (t-value) | Sig | VIF | B | (t-value) | Sig | VIF |
| (Constant) | 8.7995 | (51.04) | *** | | 8.8810 | (51.99) | *** | | 8.8335 | (51.94) | *** | |
| LtaxRate | -0.0653 | (-6.07) | *** | 1.46 | -0.6785 | (-6.9) | *** | 1.42 | -0.6714 | (-6.9) | *** | 1.43 |
| LnAppage | -0.2005 | (-14.66) | *** | 5.25 | -0.2040 | (-15.06) | *** | 5.33 | -0.2088 | (-15.47) | *** | 5.42 |
| LnAppage_Sq | -0.0441 | (-6.76) | *** | 3.15 | -0.0442 | (-6.77) | *** | 3.26 | -0.0462 | (-7.1) | *** | 3.32 |
| LnLivarea | 0.5129 | (15.06) | *** | 1.70 | 0.5059 | (15.08) | *** | 1.71 | 0.5134 | (15.44) | *** | 1.72 |
| LnLivarea_Sq | 0.3083 | (3.17) | *** | 1.15 | 0.3090 | (3.23) | *** | 1.14 | 0.2906 | (3.07) | *** | 1.15 |
| LnLotSize | 0.1086 | (6.24) | *** | 1.65 | 0.1189 | (6.8) | *** | 1.72 | 0.1230 | (7.11) | *** | 1.73 |
| InferiorFoundation | -0.0610 | (-3.47) | *** | 1.15 | -0.0660 | (-3.74) | *** | 1.20 | -0.0689 | (-3.92) | *** | 1.22 |
| Quality | 0.0415 | (2.47) | ** | 1.39 | 0.0438 | (2.64) | *** | 1.40 | 0.0418 | (2.55) | ** | 1.41 |
| FinishedBasement | 0.0428 | (2.85) | *** | 1.31 | 0.0393 | (2.65) | *** | 1.32 | 0.0381 | (2.6) | *** | 1.32 |
| FirePlace | 0.0665 | (5.11) | *** | 1.36 | 0.0620 | (4.81) | *** | 1.37 | 0.0653 | (5.09) | *** | 1.39 |
| Washrooms | 0.0443 | (2.66) | *** | 1.43 | 0.0437 | (2.66) | *** | 1.43 | 0.0424 | (2.61) | *** | 1.43 |
| KitchenCab. | 0.0641 | (3.98) | *** | 1.21 | 0.0574 | (3.63) | *** | 1.20 | 0.0554 | (3.54) | *** | 1.21 |
| Dishwasher | 0.0382 | (3) | *** | 1.13 | 0.0359 | (2.86) | *** | 1.14 | 0.0343 | (2.76) | *** | 1.14 |
| DetGarage | 0.0456 | (2.89) | *** | 1.15 | 0.0526 | (3.38) | *** | 1.16 | 0.0550 | (3.54) | *** | 1.17 |
| InGroundPool | 0.0714 | (2.57) | ** | 1.17 | 0.0741 | (2.7) | *** | 1.17 | 0.0700 | (2.58) | ** | 1.18 |
| CSF1 Core/Periphery | 0.0546 | (4.07) | *** | 2.48 | 0.0596 | (4.65) | *** | 2.35 | 0.0510 | (3.89) | *** | 2.51 |
| CSF2 FamCycle | 0.0447 | (4.72) | *** | 3.43 | 0.0429 | (4.89) | *** | 3.06 | 0.0389 | (4.43) | *** | 3.13 |
| CSF3 Socio-Economic-Status | 0.0977 | (12.43) | *** | 2.28 | 0.0907 | (11.29) | *** | 2.46 | 0.0864 | (10.78) | *** | 2.50 |
| CTD to MACs | -0.0174 | (-6.11) | *** | 3.88 | -0.0184 | (-6.29) | *** | 4.22 | -0.0190 | (-6.55) | *** | 4.25 |
| CTD to MACs_Sq | 0.0006 | (1.88) | * | 1.35 | - | - | - | - | - | - | - | - |
| Ln%Mineral 100m | | | | | -0.0191 | (-3.16) | *** | 1.25 | -0.0190 | (-3.16) | *** | 1.27 |
| Water 100m | | | | | 0.0068 | (2.65) | *** | 1.08 | 0.0066 | (2.57) | ** | 1.09 |
| Woodland 1km | | | | | -0.0010 | (-2.56) | ** | 1.51 | -0.0013 | (-3.11) | *** | 1.56 |
| LnIndustrial 500m | | | | | -0.0158 | (-2.04) | ** | 1.26 | -0.0143 | (-1.85) | * | 1.28 |
| Commercial500m * CTDtoMACs | | | | | | | | | -0.0004 | (-2.71) | *** | 1.21 |
| Lawn300m * CTDtoMACs | | | | | | | | | 0.0015 | (2.47) | ** | 1.09 |
| AgricCultPast300m * CTDtoMACs | | | | | | | | | -0.0016 | (-2.15) | ** | 1.05 |

*** significant at the 1 per cent level; ** significant at the 5 per cent level; * significant at the 10 per cent level.

| | | Model 1A | Model 1B | Model 1C |
|---|---|---|---|---|
| **Model Specification** | Nb of cases | 682 | 682 | 682 |
| | R-square | 0.821 | 0.828 | 0.839 |
| | Adj. R-Square | 0.815 | 0.821 | 0.825 |
| | SEE | 0.156 | 0.154 | 0.152 |
| | SEE in % | 16.9% | 16.6% | 16.4% |
| | F ratio | 151 | 137 | 125 |
| | Sig. | 0.0000 | 0.0000 | 0.0000 |
| | Df1/Df2 | 20/661 | 23/658 | 26/655 |
| | Ind. Variables | 20 | 23 | 26 |
| | Maximum VIF value | 5.25 | 5.33 | 5.42 |
| **Spatial Autocorrelation** | 1500m Moran's I | 0.026 | 0.020 | 0.003 |
| | Sig. | 0.421 | 0.436 | 0.486 |
| **Heteroskedasticity** | G-Q test Price (Sig.) | 0.023 | 0.024 | 0.022 |
| | G-Q test Appage (Sig.) | 0.005 | 0.001 | 0.000 |
| **Validation Sample** | Nb of cases | 42 | 42 | 42 |
| | R-square | 0.811 | 0.801 | 0.847 |
| | Adj. R-Square | 0.647 | 0.570 | 0.607 |
| | SEE | 0.194 | 0.177 | 0.176 |
| | SEE in % | 21.4% | 19.4% | 19.2% |
| | Max. Abs. Residual | 27 719 $ | 27 719 $ | 31 118 $ |
| | Mean Abs. Residual | 10 059 $ | 9 449 $ | 9 000 $ |
| | Stdd of Residuals | 6 992 $ | 6 632 $ | 7 118 $ |

**Model $1_A$.** The first model, integrating 13 property specifics, three Census factors, one accessibility factor and one taxation variable, expresses 81.5% of the price variance (see Table 6 for an overall summary of set 1 models). Standard Error of Estimation (SEE) amounts to 16.88%, with an *F*-ratio of 151. Some 17 out of the 20 variables are significant at the .01 level, whereas the squared form of CTD to the MACs is significant at the 10% level. Multicollinearity is well under control, with a maximum VIF value of 5.3 (Neter *et al.*, 1990). Regression coefficients are consistent in sign and magnitude and in accordance with expectations. The three most significant variables are the natural logarithm of the living area (positive effect, t-value 15.06), the natural logarithm of the apparent age (negative effect, t-value -14.66), and the third socio-economic component indicating the overall level of schooling and income in the neighbourhood (positive effect, t-value 12.42). In a log-linear functional form, when the independent variable is also expressed as the logarithm of the variable, the coefficients are expressed as elasticities. Therefore, for the apparent age, the living area and the lot size, the related coefficient values measure the ratio of the average percentage of change in the property value to a one percent change in the explanatory variable.

**Model $1_B$**. A first environmental model includes four additional variables relating to land-use characteristics, while all preceding variables remain significant at the .01 level. Explained variance rises to 82.1%, for a SEE of 16.6% and an *F*-value of 137. Significant land-use variables are as follows:

1. The logarithm of the mean density of mineral surfaces (concrete surfaces excluding buildings and houses, i.e. roads, sidewalks, parking lots) proves significant, showing that an increase in the mineral density in a close radius around the property – which can be assimilated to the vista space – impacts negatively on property value.
2. The presence of water within close proximity to the property has a positive impact, in accordance with several previous studies.
3. Industrial areas impact negatively in a 500 metres radius around the property, i.e. in the displacement-reinforcement space: a 10% increase in the industrial coverage results in a 16% value drop.
4. Woodland, when considered in a one kilometre radius, has a negative impact on property value.

**Model $1_C$.** The last step integrates three additional interactive variables, slightly raising the percentage of explained variance (from 82.1 to 82.5%), whereas SEE decreases from 16.6% to 16.4%. All other variables remain significant. Major findings are as follows:

1. The relative abundance of commercial land use within 500 metres adds a premium when located near the MACs, but has a negative impact in suburban areas. Commercial areas in suburbs are mainly shopping centres and big boxes, which are caracterised by the presence of important parking lots, road traffic and noise, associated to negative externalities. Also, proximity to commercial facilities may represent a premium in the central locations.
2. Proportions of lawn areas above mean value within 300 metres add value in suburban areas, but devalue the properties near the MACs.
3. The importance of agricultural land (cultures and pastures) adds a negative value when distance to MACs is above average (more than 12 minutes).

### 1.4.2 Set 2 Models: Environmental Data Extracted from Landsat TM 5 Images

This second set of models (Table 7), built with a main sample of 2 058 cottages sold between 1993 and 1996, uses environmental data semi-automatically extracted from a Landsat TM 5 image.

Table 7. Set 2 hedonic models

| Dependent Variable: LnSprice | Model $2_A$ B | (t-value) | Sig | VIF | Model $2_B$ B | (t-value) | Sig | VIF | Model $2_C$ B | (t-value) | Sig | VIF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (Constant) | 8.7226 | (91.52) | *** | | 8.7339 | (90.18) | *** | | 8.7545 | (93.54) | *** | |
| LtaxRate | -0.0573 | (-6.53) | *** | 1.63 | -0.0449 | (-4.93) | *** | 1.83 | -0.0407 | (-4.47) | *** | 1.92 |
| LnAppage | -0.1279 | (-29.09) | *** | 4.07 | -0.1417 | (-31.98) | *** | 4.30 | -0.1432 | (-33.11) | *** | 4.27 |
| LnAppage_Sq | -0.0429 | (-20.75) | *** | 1.96 | -0.0464 | (-22.78) | *** | 1.98 | -0.0488 | (-24.27) | *** | 2.02 |
| LnLivarea | 0.4909 | (29.66) | *** | 1.70 | 0.4892 | (29.91) | *** | 1.73 | 0.4786 | (30.01) | *** | 1.71 |
| LnLivarea_Sq | 0.1271 | (2.9) | ** | 1.12 | 0.1359 | (3.15) | *** | 1.13 | 0.1288 | (3.05) | *** | 1.13 |
| LnLotsize | 0.0989 | (10.89) | *** | 1.65 | 0.0971 | (10.19) | *** | 1.90 | 0.1040 | (11.72) | *** | 1.72 |
| Quality | 0.0898 | (8.55) | *** | 1.24 | 0.0844 | (8.16) | *** | 1.26 | 0.0862 | (8.49) | *** | 1.26 |
| InferiorFacing | -0.0755 | (-4.03) | *** | 1.11 | - | - | - | - | - | - | - | - |
| SuperiorRoofQual | 0.0939 | (2.27) | ** | 1.07 | 0.1382 | (3.41) | *** | 1.07 | 0.1937 | (4.81) | *** | 1.10 |
| SuperiorFloorQual | 0.0515 | (7.35) | *** | 1.41 | 0.0449 | (6.49) | *** | 1.43 | 0.0419 | (6.17) | *** | 1.44 |
| FacingStoneBrick51%+ | 0.0488 | (7.36) | *** | 1.31 | 0.0486 | (7.5) | *** | 1.31 | 0.0517 | (8.14) | *** | 1.31 |
| FinishedBasement | 0.0548 | (8.72) | *** | 1.18 | 0.0530 | (8.62) | *** | 1.17 | 0.0506 | (8.4) | *** | 1.17 |
| Storey | 0.0938 | (5.42) | *** | 2.10 | 0.0840 | (4.97) | *** | 2.08 | 0.0703 | (4.23) | *** | 2.10 |
| Stair | 0.0258 | (3.87) | *** | 1.44 | 0.0255 | (3.88) | *** | 1.45 | 0.0281 | (4.36) | *** | 1.45 |
| Fireplace | 0.0509 | (8.01) | *** | 1.38 | 0.0485 | (7.72) | *** | 1.40 | 0.0529 | (8.61) | *** | 1.40 |
| Dishwasher | 0.0303 | (4.68) | *** | 1.15 | 0.0318 | (4.99) | *** | 1.15 | 0.0324 | (5.19) | *** | 1.15 |
| AttGarage | 0.0747 | (9.48) | *** | 1.25 | 0.0771 | (9.91) | *** | 1.27 | 0.0700 | (9.17) | *** | 1.28 |
| DetGarage | 0.0624 | (7.91) | *** | 1.22 | 0.0602 | (7.78) | *** | 1.23 | 0.0568 | (7.48) | *** | 1.23 |
| InGroundPool | 0.0746 | (6.77) | *** | 1.11 | 0.0758 | (7) | *** | 1.12 | 0.0755 | (7.12) | *** | 1.12 |
| WaterSewer | 0.1545 | (6.73) | *** | 1.20 | 0.1364 | (5.89) | *** | 1.27 | 0.1229 | (5.39) | *** | 1.28 |
| CSF2 FamCycle | -0.0155 | (-3.91) | *** | 2.81 | - | - | - | - | - | - | - | - |
| CSF3 Socio-EconomicStatus | 0.0990 | (20.78) | *** | 2.05 | 0.0782 | (15.09) | *** | 2.52 | 0.0766 | (14.82) | *** | 2.62 |
| CSF4 Replacement | -0.0468 | (-4.69) | *** | 1.66 | - | - | - | - | - | - | - | - |
| CTD to MACs | -0.0147 | (-12.82) | *** | 2.28 | -0.0146 | (-12) | *** | 2.70 | -0.0135 | (-10.98) | *** | 2.85 |
| CTD to MACs_Sq | 0.0016 | (9.64) | *** | 1.37 | 0.0012 | (6.48) | *** | 1.71 | 0.0008 | (3.42) | *** | 3.16 |
| Highway150m | -0.0516 | (-2.4) | ** | 1.03 | -0.0438 | (-2.07) | ** | 1.03 | - | - | - | - |
| ResidMatureTrees 100m | | | | | 0.0011 | (3.54) | *** | 3.38 | 0.0007 | (2.57) | ** | 3.34 |
| ResidMatureTrees 500m | | | | | 0.0025 | (4.14) | *** | 4.81 | 0.0023 | (3.92) | *** | 4.84 |
| ResidLowTreeDensity 500m | | | | | -0.0019 | (-2.95) | *** | 2.13 | -0.0021 | (-3.29) | *** | 2.17 |
| Woodlands 500m | | | | | -0.0016 | (-6.11) | *** | 2.21 | -0.0010 | (-3.51) | *** | 2.75 |
| Agriculture/Disp.Trees 100m | | | | | -0.0023 | (-3.23) | *** | 1.31 | -0.0024 | (-3.55) | *** | 1.21 |
| NDVI StdDev 1km (Heterogeneity) | | | | | 0.2514 | (4.35) | *** | 1.92 | 0.2322 | (4.04) | *** | 1.98 |
| Mean NDVI 40m | | | | | 0.0538 | (2.33) | ** | 1.87 | - | - | - | - |
| Agric/BarrenLand500m * CTDtoMACs | | | | | | | | | -0.0008 | (-4.04) | *** | 1.59 |
| ResidMatureTrees100m * CTDtoMACs | | | | | | | | | -0.0004 | (-8.03) | *** | 2.35 |
| Woodlands500m * CTDtoMACs | | | | | | | | | -0.0003 | (-3.21) | *** | 3.07 |
| Agric/DispTrees100m * CTDtoMACs | | | | | | | | | 0.0007 | (3.31) | *** | 1.28 |

*** significant at the 1 per cent level; ** significant at the 5 per cent level; * significant at the 10 per cent level.

| | | Model $2_A$ | Model $2_B$ | Model $2_C$ |
|---|---|---|---|---|
| **Model Specification** | Nb of cases | 2 058 | 2 058 | 2 058 |
| | R-square | 0.865 | 0.870 | 0.877 |
| | Adj. R-Square | 0.864 | 0.869 | 0.875 |
| | SEE | 0.126 | 0.124 | 0.121 |
| | SEE in % | 13.5% | 13.2% | 12.9% |
| | F ratio | 504 | 457 | 450 |
| | Sig. | 0.0000 | 0.0000 | 0.0000 |
| | Df1/Df2 | 26/2 031 | 30/2 027 | 32/2 025 |
| | Ind. Variables | 26 | 30 | 32 |
| | Maximum VIF value | 4.1 | 4.8 | 4.8 |
| **Spatial Autocorrelation** | 1500m Moran's I | 0.215 | 0.190 | 0.136 |
| | Sig. | 0.003 | 0.009 | 0.044 |
| **Heteroskedasticity** | G-Q test Price (Sig.) | 0.000 | 0.000 | 0.000 |
| | G-Q test Appage (Sig.) | 0.000 | 0.000 | 0.000 |
| **Validation Sample** | Nb of cases | 220 | 220 | 220 |
| | R-square | 0.821 | 0.826 | 0.826 |
| | Adj. R-Square | 0.797 | 0.799 | 0.795 |
| | SEE | 0.136 | 0.128 | 0.125 |
| | SEE in % | 14.6% | 13.7% | 13.3% |
| | Max. Abs. Residual | 53 189 $ | 49 066 $ | 46 549 $ |
| | Mean Abs. Residual | 12 225 $ | 11 578 $ | 11 316 $ |
| | Stdd of Residuals | 10 020 $ | 9 527 $ | 9 202 $ |

**Model $2_A$.** A first standard model explaining 86.4% of the price variance integrates 19 property specifics, three Census factors, two accessibility factors, one location attribute and one taxation variable. With a SEE of 13.5% and an *F*-value of 504, this model already performs well. Coefficients' signs and magnitudes are in accordance with theoretical expectations, whereas significance tests are all below the .05 level. Low VIF values (maximum VIF value of 4.07) indicate the absence of severe multicollinearity.

**Model $2_B$.** The second model integrates seven additional environmental variables, of which two are related to the Normalised Difference Vegetation Index (NDVI), that is, the density of green vegetation. Two Census factors from the first model are excluded, i.e. *CSF2 Old/New Suburbs* and *CSF4 Young Adults / Retired*. Adjusted *R*-square rises slightly to .869, with a SEE of 13.2% and an *F*-value of 457.

Major findings can be summarised as follows:

1. Four variables entering significantly into the model relate to the presence of trees: the proportion of (i) residential areas with mature trees, both in the 100 and 500-metres radii, (ii) residential areas with low tree density and (iii) woodlands within 500 metres. The presence of mature trees has a positive impact both on a very local scale, with a premium of roughly 1% for each additional 10% in coverage, and on a larger scale, with a premium of roughly 2.5% for each additional 10%. Conversely, residential land use with low tree density has a negative impact on property values of roughly –1.9% for each additional 10% of coverage. Woodlands, here too, impact negatively, when considering a 500-metre radius around the property.
2. Agricultural land with dispersed trees has an overall negative impact on property value of –2.3% per 10% additional coverage in close surroundings (100 metres).
3. The two significant variables integrating NDVI values, with positive coefficient signs, are the amount of green density within 40 metres of the property (on the property lot and in the immediate neighbouring areas), and the standard deviation of NDVI values within one kilometre, showing that the diversity in land use is valued positively.

**Model $2_C$.** As for the first set, the last step in the modelling process introduces the interactions between environmental variables and CTD to the MACs. This third model integrates four new variables, all variables now being significant at the .01 level. Adjusted *R*-square reaches .878, with a SEE of 12.9%, and an *F*-ratio of 457. All the variables of the previous model except *Highway150* and *MeanNDVI40m* remain significant (Table 7).

Three environmental variables are present both in their original form and in interaction with the distance factor. This means that for these characteristics, an average effect is measured for the whole area of study, and an adjustment must be made considering location within the city, using distance to the MACs as a proxy. This holds for *ResidMatureTrees100m,* *Woodlands500m* and *Agric/DispTrees100m*. *Agric/BarrenLand500m* does not affect property values for the entire city, but becomes significant at some specific locations.

### 1.4.3 Methodological Issues

The validation of each set of models using the previously separate sub-samples proved positive, as indicate the high *R*-square values, the reasonable SEEs and the mean absolute values of residuals given in Table 6 and Table 7. Multicollinearity is well under control, VIF values for all models being below 5.4 and only reaching this high among quadratic terms. Concerning heteroskedasticity, a visual control of the residuals shows a relative homogeneity of their variance. However, the Goldfeld-Quandt test is significant for models of both sets, showing that the addition of land-use locational externalities did not solve the heteroskedasticity problem. Although beyond the scope of this paper, different specifications should be tested in further research. Interesting results concerning the heteroskedasticity problem have been achieved using Generalised Least Squares specification (Fletcher *et al.*, 2000).

Spatial autocorrelation (SA), measured using the Moran (1950) index, was computed using the 15 nearest neighbours in the immediate vicinity (maximum distance of 1 500 metres). The limit of 15 was adopted to ease computation. Due to the inverse squared distance weighting, considering more neighbours does not change Moran's index significantly. In the first set, SA is non-significant for all three models (See Table 6). In the second set, although SA of the residuals remains significant at the 5% level for the three models, the level of significance decreases from model $2_A$ to $2_C$ as locational variables are included (Table 7). These encouraging results must however be considered with caution, as local forms of spatial autocorrelation could still be present (Brunsdon *et* *al.*, 2002; Fotheringham *et al.*, 2002; Getis and Ord, 1992; Ord and Getis, 1995; Páez *et al.*, 2001).

## 1.5 Discussion

Both classification methods and data sources produced interesting and significant results. However, the models using remote sensing data integrated more land-use locational variables, and NDVI measures offered additional information. The lower efficiency of the map based on aerial photographs could be linked to the Modifiable Area Unit Problem (MAUP). The grain (or resolution) of both land-use maps indeed differs: whereas the satellite image is a regular grid of 625 square-metre pixels, the polygons of the aerial photograph map are heterogeneous in size as their construction relies on the urban structure – with a mean area of 8 100 square metres, and a standard deviation of 48 100. The grain variability induces heterogeneity in measurements using small circular buffers. This bias leads, in the first set of models, to a lower efficiency when integrating land-use and vegetation externalities. Therefore, we consider that the satellite image is probably a better source of information for integrating land-use locational externalities for hedonic modelling purposes.

Land-use locational externalities are significant on four different scales, i.e. for distances of 40, 100, 500, and 1 000 metres around the properties. Significant vegetation-related variables are numerous and play a role on all scales. The mean NDVI has a positive impact on a 40-metre scale, indicating a premium for vegetation in the immediate vicinity of the property. Although to our knowledge this is the first time NDVI data is used for residential hedonic modelling, the results confirm previous findings concerning the premium associated with the presence of trees on the property lot (Anderson and Cordell, 1985 & 1988; Payne, 1973; Seila and Anderson, 1982). The positive impact of mature trees is also significant at the 100-metre and 500-metre scales, in accordance with previous findings (Anderson and Cordell, 1988; Thériault *et al.*, 2002). Inversely and logically, a low-tree density bears a negative impact. The negative impact of woodland, on the 1 000-metre scale for set 1 and at 500 metres for set 2, concords with findings by Paterson and Boyle (2002), but is in contradiction with Tyrvainen and

Miettinen (2000), whereas Garrod and Willis (1992) find both positive and negative impacts depending on how the presence of woodland is measured. In our case, most of the woodland areas are located in the outer suburbs, where, hypothetically, a probable excess of visible woodland affects prices negatively. Furthermore, the significant negative impact of woodland and agricultural land could reflect a negative premium due to a lack of urbanisation, meaning an indication of price drop due to highly rural areas and lower levels of proximity service.

On a 1 000-metre scale, the standard deviation of the NDVI values is positively associated with property values. This variable expresses the diversity of land use in terms of vegetation cover, and the findings are consistent with those of a previous study that reports a positive sign for a diversity index (also measured within a 1 000-metre radius), indicating the premium associated with a diversified landscape in terms of land use (Geoghegan *et al.*, 1997).

Other land-use locational externalities proved significant as such, within a 100-metre radius (water [set 1, +] built [set 1, -] and agricultural [set 2, -] surfaces) and within a 500-metre radius (Industrial surfaces [set 1, -]). Coefficient signs are in line with expectations. Agricultural surfaces did not show an eventual premium that could be associated with open space. However, a recent study held in Portland, Oregon, U.S.A., showed that the significant positive impact of open space becomes non-significant when distance is inferior to 100 metres (Bolitzer and Netusil, 2000).

The use of several measurement scales with buffer functions in order to partly integrate the hierarchical structure of the perceptual regions produced interesting results. Most of the significant land-use locational externalities are significant on the 100-metre and 500-metre scales. The first could be associated with the vista space, and the second with the local-displacement space. These results confirm Geoghegan's study (1997) held in the Patuxtent Watershed, where some land-use characteristics measured within a 100-metre radius prove significant indeed; and Tyrvainen's work (2000) shows that the premium associated with proximity to urban forest is significant up to 600 metres, or within walking distance. Beyond the limitations of buffer functions (fixed boundaries and isotropic view), the results obtained are satisfactory and prove the effectiveness of integrating hierarchical perception patterns in hedonic modelling.

Furthermore, the significant integration of interactive relations proved that in some cases, the overall effect of a land-use externality has to be adjusted considering location in the city space. Taking a closer look at set 2 models, we observe that three of the four significant interactive relationships (Model $2_C$) concern environmental attributes that were already present in the previous model (Model $2_B$). Two 3-D diagrams help to illustrate the phenomenon for mature trees within a 100-metre radius and woodland within 500 metres (Figure 1 and Figure 2). It is important to restrict visualisation to observed combinations in order to avoid hazardous extrapolations. Therefore, non-existent combinations have been blanked out in these figures.

Figure 1. Interaction effect between mature trees and car-time distance to MACs

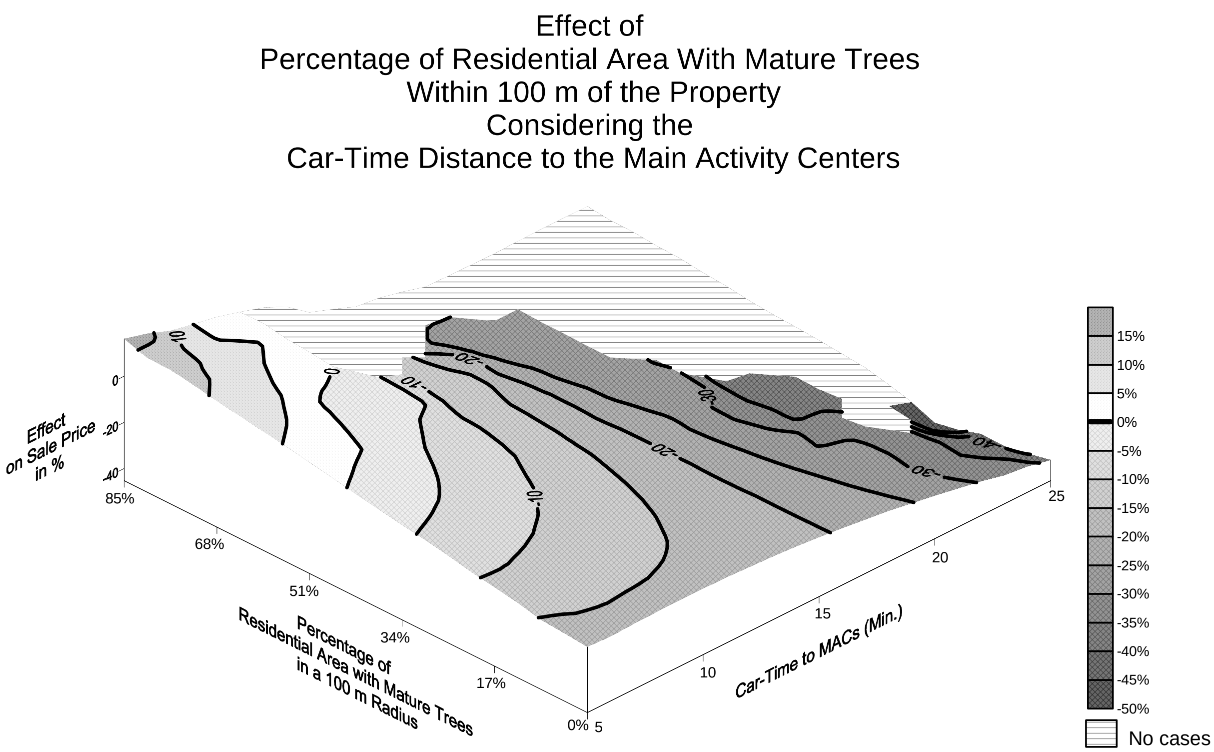
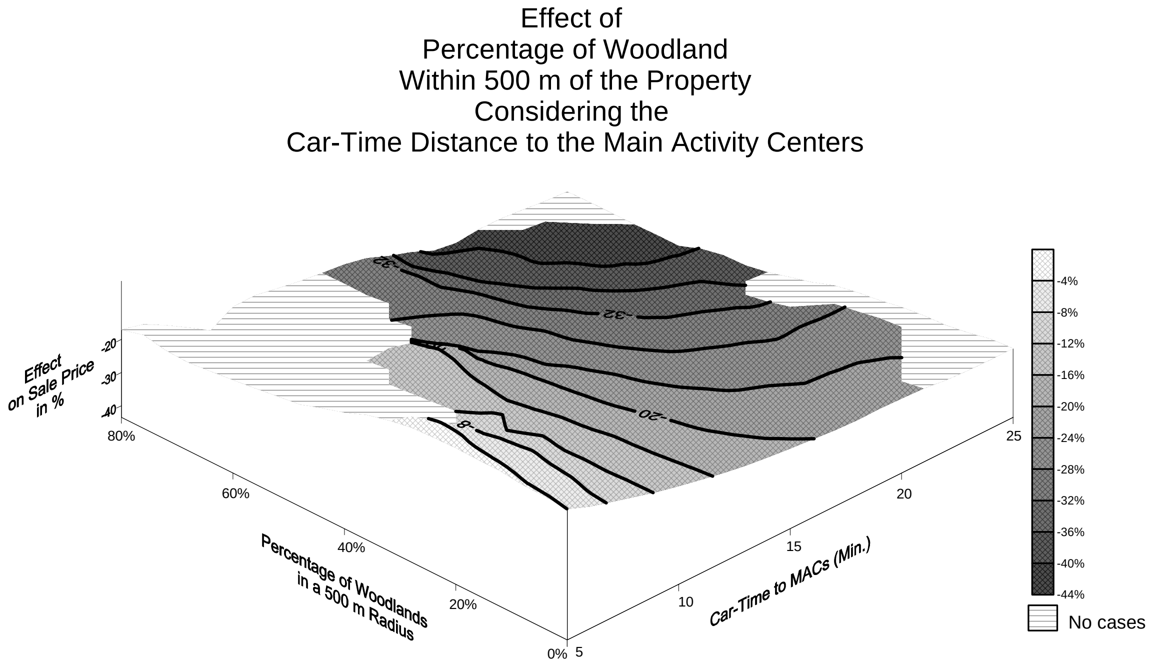


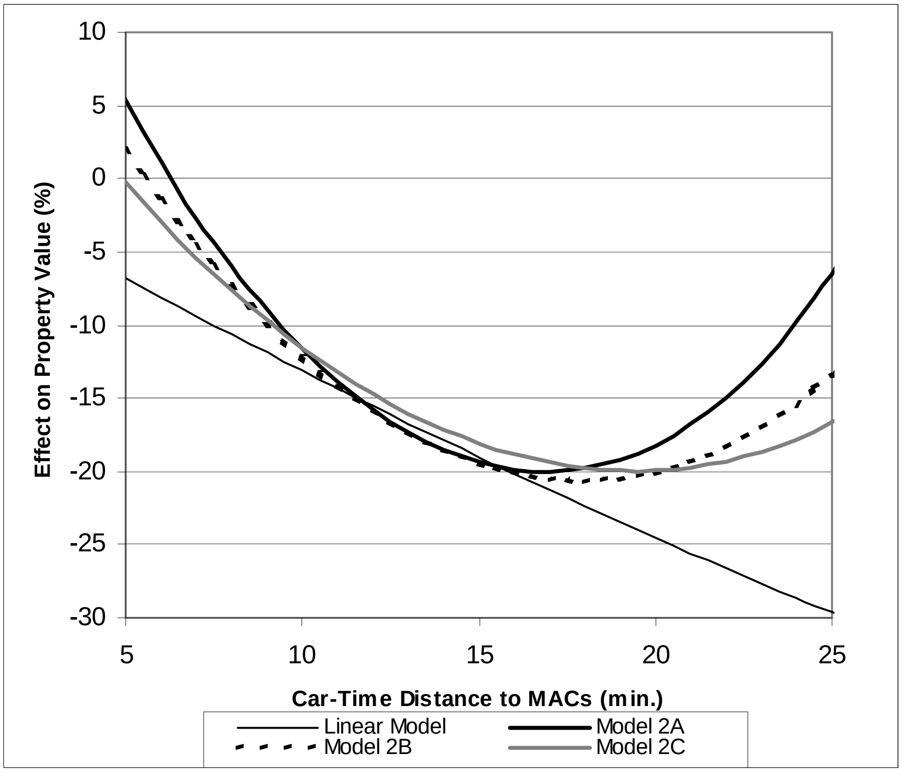
Figure 2. Interaction effect between woodland and car-time distance to MACs




Near the MACs, the positive trend associated with mature trees is important. In fact, being located in a central neighbourhood with less than roughly 50% of residential area with mature trees within 100 metres of the properties, has a negative effect on house values, whereas very green areas can add a premium of up to 15%. As distance to the MACs increases, this trend becomes less important, but the variety of mature tree coverage decreases concomitantly. At more than 17 minutes' driving distance from the MACs, the highest proportion of mature trees drops under 50%, which makes the statement of penalty associated with mature trees in the urban fringes hazardous. However, these findings indicate that the appreciation of trees is not homogeneous and may depend on surrounding characteristics; i.e. in remote areas where woodland is abundant, the presence of trees on the property lot – potentially affecting the extent of the view – may not be as highly valued. Concerning the effect of woodlands, the devaluation trend related to higher proportions is insignificant when located close to the MACs, and increases with distance. As previously stated, these findings are in line with a study held in Central England and the Welsh borders where the view on woodlands had a negative impact. However, Garrod and Willis (1992) also showed that a significant tract of woodland within one kilometre had a positive effect on property values. Our study could not show a positive premium associated with woodland presence. However, it is important to remember that although the method is universal and can be applied in diverse urban situations, the resulting coefficients and the significance of the variables hold true for Quebec City only.

The progressive integration of land-use locational externalities has interesting consequences on the effect of the CTD to the MACs. In model $2_A$, both the linear and squared form of the distance to MACs are significant, the first with a negative and the second with a positive coefficient sign. This shows that the effect of distance is not linear, but forms a U-shaped curve, with a highest negative effect at approximately 17 minutes away from the MACs (Figure 3). For locations at a greater distance, the negative effect becomes less important. In fact, considering this model, the negative effect of distance (-6% of property value) is the same for a property located in the outer suburbs (25 minutes) or near the centre (eight minutes). However, when environmental locational externalities are added, and even more so when the spatial interaction effects are considered, the coefficient value of the squared term of the CTD progressively drops, from 0.0016 (Model $2_A$) to 0.0012 (Model $2_B$) to 0.0008 (Model $2_C$), and the U-shaped curve becomes a rather linear trend. This indicates that the positive effect of land-use locational externalities is partly internalised in the distance coefficients of the first model, and is later explained by the integration of land use, vegetation and spatial interaction attributes. Finally, the positive marginal contribution of vegetation attributes is primarily significant in the central areas, where we find most of the residential areas with mature trees, in older neighbourhoods with high-level income, and in the distant outer suburbs, where the benefits of proximity to open spaces and nature counterbalance the loss of accessibility. In our case, the premium associated with land-use externalities is therefore twofold, concerning both the MACs and the urban fringes.

Figure 3. Effect of car-time distance to MACs on property prices




Lastly, all other things being equal, when integrating land-use locational externalities, the maximum negative effect (20% drop in property value) shifts from around 17 minutes to MACs (Model $2_A$) to 18 minutes (Model $2_B$) to 20 minutes (Model $2_C$). However, it is important to note that the squared term of the CTD to the MACs is still significant in the last model, showing some positive premium for remote locations. This has not yet been explained. We believe that the proximity to natural parks and to specific externalities such as ski resorts and lakes, located north of the Quebec City region, could partly explain this additional premium and should explicitly be modelled in future research. Moreover, sight attributes should be considered, as some areas located more than 20 minutes from the MACs are hilly, providing better landscape views as well as views of the Quebec City skyline.

## 1.6 Conclusion

This paper presents a straightforward method for integrating land-use locational externalities in residential hedonic models. Aerial photographs and a Landsat TM 5 image are categorised, and the obtained land-use data, measured on various scales, are used to estimate the marginal contribution of land-use locational externalities on property values. Applied to two distinct sets of residential sales in Quebec City in 1986-1987 and 1993-1996, the models progressively integrate the following data: (i) property-specifics, Census factors and location attributes (*Models* $1_A$ and $1_B$); (ii) land-use and land-cover data, measured around houses on various scales using buffer functions (*Models 2A* and 2B); (iii) interaction effects between land-use locational externalities and location within the city (CTD to the MACs). Special attention is given to (i) the scale effect, e.g. how the hierarchical structure of the perceptual region does or does not appear significant; (ii) the interaction effect, e.g. how the impact of land-use locational externalities varies through space; and (iii) the consequence of integrating land-use data on another major determinant of price, e.g. the distance to the MACs.

The significant integration of land-use data on various scales (40, 100, 500 and 1 000 metres) shows that a hierarchical structure of perception has to be considered when analysing locational externalities. Furthermore, the significance of interaction effects emphasises the importance of location in the valuation of externalities. Considering the interaction between land-use locational externalities and location not only indicated that the impact of land use varies through space, but it also showed how the effect of other attributes, such as the distance to the MACs, can be inaccurately estimated when locational externalities are omitted.

Further research is needed in order to improve our understanding of the impact of locational externalities on perception and residential choice behaviour. As was shown, impacts are uneven through space; but are perceptions homogeneous among people? It would be interesting to further the analysis by integrating data characterising the buyers' socio-economic profile, in order to investigate whether locational externalities are evenly valued among people *and* through space. Furthermore, a closer look at local spatial statistics could better our understanding of the phenomena that remain unexplained.

# References

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

Anderson L M, Cordell H K, 1985, "Residential Property Values Improved by Landscaping with Trees" S. J. Applied For. **9** 162-166

Anderson L M, Cordell H K, 1988, "Influence of Trees on Residential Property Values in Athens, Georgia (U.S.A.): A Survey based on Actual Sales Prices" Landscape and Urban Planning **15** 153-164

Anderson R J, Crocker T D, 1971, "Air Pollution and Residential Property Values" Urban Studies **8** 171-180

Anselin L, 1990, "Spatial Dependence and Spatial Structural Instability in Applied Regression Analysis" Journal of Regional Science **30** 185-207

Anselin L, Can A, 1986, "Model Comparison and Model Validation Issues in Empirical Work on Urban Density Functions" Geographical Analysis **18** 179-197

Anselin L, Rey S, 1991, "Properties of Tests for Spatial Dependence in Linear Regression Models" Geographical Analysis **23** 112-131

Bender A R, Din A, Hoesli M, Brocher S, 2000, "Environmental Preferences of Homeowners: Further Evidence Using the AHP Method" Journal of Property Investment & Finance **18** 445-455

Benson E D, Hansen J L, Schwartz A L, 2000, "Water Views and Residential Property Values" The Appraisal Journal **68** 260-271

Benson E D, Hansen J L, Schwartz A L, Smersh G T, 1998, "Pricing Residential Amenities: The Value of a View" Journal of Real Estate Finance and Economics **16** 55-73

Beron K, Hanson Y, Murdoch J, Thayer M, 2001, "Hedonic Price Functions and Spatial Dependence: Implications for the Demand for Urban Air Quality", in New Advances in Spatial Econometrics Eds L Anselin, R Florax (Springer-Verlag, New-York)

Besner C, 2002, "A Spatial Autoregressive Specification with a Comparable Sales Weighting Scheme" Journal of Real Estate Research **24** 193-211

Bolitzer B, Netusil N, 2000, "The Impact of Open Spaces on Property Values in Portland, Oregon" Journal of Environmental Management **59** 185-193

Boyle M A, Kiel K A, 2001, "A Survey of House Price Hedonic Studies of the Impact of Environmental Externalities" Journal of Real Estate Literature **9** 117-144

Brunsdon C, Fotheringham A S, Charlton M, 2002, "Geographically Weighted Summary Statistics - A Framework for Localised Exploratory Data Analysis" Computers, Environment and Urban Systems **26** 501-524

Buyhoff G J, Gauthier L J, Wellman J D, 1984, "Predicting Scenic Quality for Urban Forests Using Vegetation Measurements" Forest Science **30** 71-82

Can A, 1990, "The Measurement of Neighbourhood Dynamics in Urban House Prices" Economic Geography **66** 254-272

Can A, 1992, "Specification and Estimation of Hedonic Housing Price Models" Regional Science and Urban Economics **22** 453-477

Can A, Megbolugbe I, 1997, "Spatial Dependence and House Price Index Construction" Journal of Real Estate Finance and Economics **14** 203-222

Casetti E, 1972, "Generating Models by the Expansion Method: Applications to Geographical Research" Geographical Analysis **4** 81-91

Cho C-J, 2001, "Amenities and Urban Residential Structure: An Amenity-Embedded Model of Residential Choice" Papers in Regional Science **80** 483-498

Cliff A D, Ord J K, 1981 Spatial Processes, Models and Applications (Pion Ltd., London)

Cooper-Marcus C, 1982, "The Aesthetic of Family Housing: The Residents' View-Point" Landscape Research **7** 9-13

Davy B, 1997 Essential Injustice: When Legal Institutions Cannot Resolve Environmental and Land Use Disputes (Springer, New York)

Delaney C J, Timmons D, 1992, "High Voltage Power Lines: Do they Affect Residential Property Value?" Journal of Real Estate Research **3** 315-329

Des Rosiers F, 2002, "Power Lines, Visual Encumbrance and House Values: A Microspatial Approach to Impact Measurement" Journal of Real Estate Research **23** 275-302

Des Rosiers F, Bolduc A, Thériault M, 1999, "Environment and Value: Does drinking Water Quality affect House Prices?" Journal of Property Investment & Finance **17** 444-463

Des Rosiers F, Thériault M, 1999 House Prices and Spatial Dependence: Towards an Integrated Procedure to Model Neighbourhood Dynamics Université Laval, Faculté des Sciences de l'Administration, Québec

Des Rosiers F, Thériault M, Kestens Y, Villeneuve P, 2002, "Landscaping and House Values: An Empirical Investigation" Journal of Real Estate Research **23** 139-161

Des Rosiers F, Thériault M, Villeneuve P-Y, 2000, "Sorting Out Access and Neighbourhood Factors in Hedonic Price Modelling" Journal of Property Valuation and Investment **18** 291-315

Din A, Hoesli M, Bender A, 2001, "Environmental Variables and Real Estate Prices" Urban studies **38** 1989-2000

Do A Q, Sirmans C F, 1994, "Residential Property Tax Capitalization: Discount Rate Evidence from California" National Tax Journal **47** 341-348

Dubin R A, 1998, "Predicting House Prices Using Multiple Listings Data" Journal of Real Estate Finance and Economics **17** 35-60

Duda R O, Hart P E, 1973 Pattern Classification and Scene Analysis (Wiley and Sons, New York, NY)

Eichenbaum J, 1989, "Location as a Factor in Determining Property Values" Property Tax Journal **8** 151-169

Fletcher M, Gallimore P, Mangan J, 2000, "Heteroscedasticity in Hedonic House Price Models" Journal of Property Research **17** 93-108

Foldvary F E, 1994 Public Goods and Private Communities: The Market Provision of Social Services (E. Elgar Pub., Brookfield, VT)

Fotheringham A S, Brundson F C, Charlton M, 2002 Geographically Weighted Regression: The Analysis of Spatially Varying Relationships (John Wiley & Sons, Chichester)

Freeman A M I, 1979, "Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues" Scandinavian Journal of Economics **81** 154-173

Gallimore P, Fletcher M, Carter M, 1996, "Modelling the Influence of Location on Value" Journal of Property Valuation and Investment **14** 6-19

Garrod G D, Willis K G, 1992, "Valuing Goods' Characteristics: An Application of the Hedonic Price Method to Environmental Attributes" Journal of Environmental Management **34** 59-76

Geoghegan J, Wainger L A, Bockstael N E, 1997, "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS" Ecological Economics **23** 251-264

Getis A, 1990, "Screening for Spatial Dependence in Regression Analysis" Papers of the Regional Science Association **69** 69-81

Getis A, Griffith D A, 2002, "Comparative Spatial Filtering in Regression Analysis" Geographical Analysis **34** 130-140

Getis A, Ord J K, 1992, "The Analysis of Spatial Association by Use of Distance Statistics" Geographical Analysis **24** 189-206

Goodman A C, Thibodeau T G, 1995, "Age-Related Heteroskedasticity in Hedonic House Price Equations" Journal of Housing Research **6** 25-42

Goodman A C, Thibodeau T G, 1997, "Dwelling-Age-Related Heteroskedasticity in Hedonic House Price Equations: An Extension" Journal of Housing Research **8** 299-317

Graves P, Murdoch J, Waldman D, 1988, "The Robustness of Hedonic Price Estimation: Urban Air Quality" Land Economics **64** 220-233

Griffith D A, 1996, "Spatial Autocorrelation and Eigenfunctions of the Geographic Weights Matrix Accompanying Geo-Referenced Data" The Canadian Geographer **40** 351-367

Hay G J, Marceau D J, Dubé P, Bouchard A, 2001, "A Multiscale Framework for Landscape Analysis: Object-Specific Analysis and Upscaling" Landscape Ecology **16** 471-490

Hoch I, Waddell P, 1993, "Apartment Rents: Another Challenge to the Monocentric Model" Geographical Analysis **25** 20-34

Huang J C, Palmquist R B, 2001, "Environmental Conditions, Reservation Prices, and Time on the Market for Housing" Journal of Real Estate Finance and Economics **22** 206-219

Jaccard J, Turrisi R, Wan C K, 1990 Interaction Effects in Multiple Regression (Sage Publications, Newbury Park, CA)

Kaplan R, 1983, "The Role of Nature in the Urban Context", in Behavior and the Natural Environment Eds I Altman, J F Wohlwill (Plenum Press, New York) pp 127-161

Krantz D P, Waever R D, Alter T R, 1982, "Residential Tax Capitalization: Consistent Estimates Using Micro-Level Data" Land Economics **58** 488-496

Lake I R, Lovett A A, Bateman I J, Day B, 2000, "Using GIS and Large-Scale Digital Data to Implement Hedonic Pricing Studies" International Journal of Geographical Information Science **14** 521-541

Lake I R, Lovett A A, Bateman I J, Langford I H, 1998, "Modelling Environmental Influences on Property Prices in an Urban Environment" Computers, Environment and Urban Systems **22** 121-136

Lancaster K J, 1966, "A New Approach to Consumer Theory" Journal of Political Economics **74** 132-157

Latham R, Yeates M, 1970, "Population Density in Metropolitan Toronto" Geographical Analysis **2** 177-185

Leggett C G, Bockstael N E, 2000, "Evidence of the Effects of Water Quality on Residential Land Prices" Journal of Environmental Economics and Management **39** 121-144

Luttik J, 2000, "The Value of Trees, Water and Open Space as Reflected by House Prices in the Netherlands" Landscape and Urban Planning **48** 161-167

McAvoy G E, 1999 Controlling Technocracy: Citizen Rationality and the Nimby Syndrome (Georgetown University Press, Washington)

McLeod P B, 1984, "The Demand for Local Amenity: An Hedonic Price Analysis" Environment and Planning A **1** 398-400

Mesarovic M D, Macko D, Takahara Y, 1970 Theory of Hierarchical Multilevel Systems (Academic Press, New York)

Michael H J, Boyle K J, Bouchard R, 1996 Water Quality Affects Property Prices: A Case Study of Selected Maine Lakes University of Maine, Orono

Morales D, Boyce B N, Favretti R J, 1976, "The Contribution of Trees to Residential Property Value: Manchester, Connecticut" Valuation **23** 26-43

Moran P A P, 1950, "Notes on Continuous Stochastic Phenomena" Biometrika **37** 17-23

Murdoch J, Thayer M, 1988, "Hedonic Price Estimation of Variable Urban Air Quality" Journal of Environmental Economics and Management **15** 143-146

Musterd S, Van Zelm I, 2001, "Polycentricity, Households and the Identity of Places" Urban Studies **38** 679-696

Nasar J L, 1983, "Adult's Viewer Preferences in Residential Scenes: A Study of the Relationship of Environmental Attributes to Preference" Environment and Behavior **15** 589-614

Neter J, Wasserman W, Kutner M H, 1990 Applied Linear Statistical Models : Regression, Analysis of Variance, and Experimental Designs (Irwin, Burr Ridge)

Newling B E, 1966, "Urban Growth and Spatial Structure: Mathematical Models and Empirical Evidence" The Geographical Review **56** 213-225

Nguyen N, Cripps A, 2001, "Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networs" Journal of Real Estate Research **22** 313-336

Nijkamp P, L., Van Wissen L, Rima A M, 1993, "A Household Life Cycle Model for Residential Relocation Behaviour" Socio-Economic Planning Science **27** 35-53

Ord J K, Getis A, 1995, "Local Spatial Autocorrelation Statistics" Geographical Analysis **27** 286-306

Orford S, 1999 Valuing the Built Environment: GIS and House Price Analysis (Aldershot, Ashgateh)

Orford S, 2000, "Modelling Spatial Structures in Local Housing Market Dynamics: A Multilevel Perspective" Urban Studies **37** 1643-1671

Orford S, 2002, "Valuing Locational Externalities: A GIS and Multilevel Modelling Approach" Environment and Planning B: Planning and Design **29** 105-127

Pace K R, Gilley O W, 1997, "Using the Spatial Configuration of the Data to Improve Estimation" Journal of Real Estate Finance and Economics **14** 333-340

Páez A, Uchida T, Miyamoto K, 2001, "Spatial Association and Heterogeneity Issues in Land Price Models" Urban Studies **38** 1493-1508

Papageorgiou Y Y, Pines D, 1998 An Essay on Urban Economic Theory (Kluwer Academic Publishers, Dordrecht)

Paterson R W, Boyle K J, 2002, "Out of Sight, Out of Mind? Using GIS to Incorporate Visibility in Hedonic Property Value Models" Land Economics **78** 417-425

Payne B R, 1973, "The Twenty-Nine Tree Home Improvement Plan" Natur. Hist. **82** 74-75

Powe N A, Garrod G D, Brunsdon C F, Willis K G, 1997, "Using a Geographical Information System to Estimate an Hedonic Price Model of the Benefits of Woodland Access" Forestry **70** 139-149

Powe N A, Garrod G D, Willis K G, 1995, "Valuation of Urban Amenities Using and Hedonic Price Model" Journal of Property Research **12** 137-147

Reginster I, Edwards G, 2001, "The Concept and Implementation of Perceptual Regions as Hierarchical Spatial Units for Evaluating Environmental Sensitivity" URISA Journal **13** 5-16

Remy J, Voyé L, 1992 La ville: vers une nouvelle définition? (L'Harmattan, Paris)

Rodriguez M, Sirmans C F, 1994, "Quantifying the Value of a View in Single-Family Housing Markets" The Appraisal Journal **62** 600-603

Rosen S, 1974, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" Journal of Political Economy **82** 34-55

Seila A F, Anderson L M, 1982, "Estimating Costs of Tree Preservation on Residential Lots" Journal of Arboriculture **8** 182-185

Shi M, Yu K K, Siu S S H, 2000, "Using Response Surface Analysis in Computer Assisted Mass Appraisal to Examine the Influence of Location on Property Values" Journal of Property Tax Assessment and Administration **5** 3-16

Smith V K, Huang J-C, 1995, "Can Markets Value Air Quality? A Meta Analysis of Hedonic Property Value Models" Journal of Political Economy **103** 209-227

Strange W, 1992, "Overlapping Neighborhoods and Housing Externalities" Journal of Urban Economics **32** 17-39

Tay D P H, Ho D K H, 1992, "Artificial Intelligence and the Mass Appraisal of Residential Apartments" Journal of Property Valuation and Investment **10** 525-539

Thériault M, Des Rosiers F, Vandersmissen M H, 1999 GIS-Based Simulation of Accessibility to Enhance Hedonic Modelling and Property Value Appraisal : An Application to the Quebec City Metropolitan Area Université Laval, Faculté des Sciences de l'Administration, Québec

Thériault M, Kestens Y, Des Rosiers F, 2002 "The Impact of Mature Trees on House Values and on Residential Location Choices in Quebec City", in Meeting of the iEMSs Society, Ed. i Society Lugano, Switzerland) pp 478-483.

Tse R Y C, 2002, "Estimating Neighbourhood Effects in House Prices: Towards a New Hedonic Model Approach" Urban Studies **39** 1165-1180

Tucker C J, 1979, "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation" Remote Sensing and Environment **8** 127-150

Tueller P T, 1989, "Remote Sensing Technology for Rangeland Management Applications" Journal of Range Management **42** 442-452

Tyrvainen L, Miettinen A, 2000, "Property Prices and Urban Forest Amenities" Journal of Environmental Economics and Management **39** 205-223

Von Thünen J H, 1850 Der isolierte Staat in Beziehung auf Landwirschaft und Nationalökonomie (Scientia Verlag, Aalen)

Walmsley D J, Lewis G J, 1993 People and Environment: Behavioral Approaches in Human Geography (John Wiley, New York)

Weinstein N D, 1976, "Human Evaluations of Environmental Noise", in Perceiving Environmental Quality: Research and Application Eds K H Craik, E H Zube (Plenum Press, New York) pp 229-252

Worzala E, Lenk M, Silva A, 1995, "An Exploration of Neural Networks and its Application to Real Estate Valuation" Journal of Real Estate Research **10** 185-201

Wu J, Gao W, Tueller P T, 1997, "Effects of Changing Spatial Scale on the Results of Statistical Analysis with Landscape Data: A Case Study" Geographic Information Sciences **3** 30-41

Yuan M, 2000 "Representation of Dynamic Geographic Phenomena Based on Hierarchical Theory", in 9th International Symposium on Spatial Data Handling, Ed. P Foster pp 2a.19-2a.29.

# Chapter 2. Why Families Move and What They Choose: An Analysis of Single-Family Property Buyers

**Résumé : Ce chapitre analyse en détail les motivations liées au déménagement et les critères de choix d'acheteurs de propriétés unifamiliales. À partir des résultats d'une vaste enquête téléphonique réalisée à Québec auprès d'acheteurs de propriétés unifamiliales, nous avons analysé les critères de choix en fonction du type de ménage, de l'âge, du revenu et du niveau d'éducation. Une attention particulière accordée à la dimension spatiale des critères vise à observer la variabilité géographique du choix résidentiel. Une analyse des correspondances réalisée sur les critères de choix du quartier et de la propriété permet d'identifier les principaux choix des acheteurs, qui peuvent être mis en relation avec les cadres conceptuels de la théorie de la cognition spatiale et de la théorie psychologique de place-identité. De plus, des régressions logistiques estiment la probabilité d'évoquer un critère de déménagement ou un critère de choix considérant le profil du ménage et sa localisation. Les résultats procurent des éclairages pertinents pour l'aménagement du territoire, soulignant les liens entre cycle de vie et choix résidentiels, et explorant la complexité spatiale des choix résidentiels.**

**Abstract: The purpose of this chapter is to better understand the motivation of single-family home buyers with regard to moving as well as to neighbourhood and property choice criteria. Based on a vast telephone survey of single-family property buyers in Quebec City, we analysed the stated criteria using detailed information broken down by household type, age, income, and educational attainment. The spatial considerations included in the survey also highlight the geographical variability of residential behaviour. First, a correspondence analysis of both property and neighbourhood choice criteria identifies the main choice constructs which are related to the psychological place-identity and spatial cognitive conceptual frameworks. Furthermore, a series of logistic regressions measure the likelihood of evoking a move or choice criterion depending on household profile and location. The findings provide additional insights for urban planning and research by underscoring the life cycle determinants and spatial complexity of residential choice.**

## 2.1 Introduction

Since Rossi's (1955) pioneering work on life cycle changes and relocation decisions, "*Why families move,*" much research has been done to try to disentangle the complexity of residential behaviour. The author apprehends the residential mobility process, which encompasses the act of choice. Residential behaviour as such is manifold, as is shown in the vast literature on the subject. Residential mobility studies concentrate on the propensity to move and the reasons underlying the act of moving. Residential choice studies concern preferences, choice or satisfaction, these three aspects forming a temporal continuum: preferences lead to choices, which are the foundations of satisfaction. Preferences, choice or satisfaction are analysed using various methods for stated or revealed data, such as contingent valuation, conjoint analysis, discrete choice modelling, hedonic modelling and others. Whereas methods based on stated data suffer from the critique of relying on hypothetical data, revealed data analysis methods may suffer from sample selection biases. However, both questions – what do people think they would do and what do people actually do ? – fully merit attention. Furthermore, it seems important to explore what people think of their actual residential choices, for example not only by exploring why they moved, but also by analysing what their residential choice criteria were when they actually chose their residence. In order to answer these questions, a vast phone survey was held in Quebec City involving 774 households that bought a single-family house between 1993 and 2001. The information collected describes motivation as to moving, neighbourhood and property choice criteria, as well as type of household, age, income, educational attainment and previous tenure type. We hypothesise that these incentives – motivation as to moving and choice criteria – differ significantly among households, and that this variability can be properly modelled applying logistic regressions to household-level data. This paper therefore analyses the stated criteria of actual choices depending on the socio-demographic profile of buyers and on their attachment to the neighbourhood.

First, a correspondence analysis held on all choice criteria identifies the main constructs that can be related to the psychological place-identity and spatial cognitive conceptual frameworks. Then, a series of logistic regressions measures the likelihood of mentioning a move or housing choice criterion depending on household profile and home location. A spatial analysis also explores the relative geographic variability of the expressed choice criteria. For this purpose, two spatial partitioning methods of the city are used and compared for this purpose, giving further insight into the complexity and multiplicity of residential geography. In Section 0, some geographical and psychological theoretical concepts are discussed and previous studies having analysed revealed residential criteria at the household level are reviewed. Section 0 describes the data bank and the analytical approach, whereas results are presented in Section 0. Section 0 concludes and opens an agenda for further research.

## 2.2 Conceptual Framework and Literature Review

### 2.2.1 Conceptual Framework

The choice process may be viewed as an individual reaction to an identified problem that must be solved or to a need that must be fullfilled. According to the means-end model, people's cognitive structure links values to categories of objects/attributes. The consumption of the object/attribute – the consequence – represents the intermediate level between mean (object/attribute) and end (value). The act of choosing – and this applies to residential choice – is therefore a value-oriented and goal-directed form of behaviour, evolving through time and space (Bettman, 1979; Coolen and Van Montfort, 2001; Rubinstein and Pamelee, 1992).

The specificity of the residential choice process is that beyond the acquisition of material goods, the inhabitant settles at a location, and through this process, acquires its related amenities. It is therefore important to consider the spatial dimension of residential choice by integrating the spatial cognition of the potential buyer or renter. Following Gibson (1950) and Gärling and Golledge (1993), Reginster and Edwards (2001) propose a conceptual framework for spatial perception integrating both the notions of location and activities. Location can be characterised by a set of externalities or environmental amenities. Furthermore, residential location is central to a set of activities taking place through the urban and suburban space. The perceptual region concept relies on the combination of the environmental characteristics of a location and the displacements deriving from it. Furthermore, these activities and their associated moves generate a sense of belonging proportionate to their frequencies. The combination – for various spaces – of environmental characteristics, linked activities, and frequency of use, is the essence of the concept of hierarchical perceptual regions. Reginster and Edwards (2001) identify three levels of hierarchical perceptual regions. The vista space is a "spatial region with perceptually similar characteristics apprehended from a single place, but not determined by vision alone, and which corresponds to a sense of belonging resulting from activities carried out in that region" (residence, work, school, etc.). The local-displacement space surrounds the vista space and its representation is reinforced with the frequency of visits and trips. Finally, the enlarged-displacement space relates to the large region enclosing the different local-displacement spaces. This region is principally perceived as a network, and therefore contains numerous unknown spaces. As the authors say, it is full of holes!

Similarly, a spatial conceptual framework derived from geographical concepts and adapted to the location decision process is proposed by Filion *et al*. (1999). The authors distinguish space, proximity and place. Space refers to the location in terms of potential accessibility to activities that take place in the activity catchment area, for example, in an urban context, in the whole metropolitan region. The choice of location relating to space relies on the need to maximise the possibilities for accessing activity places, while reducing travel-times and costs. Place, on the contrary, relates to the close spatial region encompassing the property (Duncan and Ley, 1993). Place is principally characterised by the physical attributes of the site, environment and buildings, which are also good indicators of the socio-economic context of the neighbourhood. Place and space do not cover the entire range of spatial factors tied to residential location choice. Proximity, an intermediate principle, refers to the need to be close to frequently visited activity places, within reasonable travel times, within a long walk or short drive, for example.

In line with the geographical concepts of site and situation (Dieleman and Mulder, 2002), it is the opinion of the authors that both the Perceptual Region (PR) and the Space-Proximity-Place (SPP) models are appropriate to better apprehend the decision-making process of residential choice. Furthermore, the impact of attributes measured through hedonic modelling fits easily into these conceptual frameworks. Based on the principle that goods are valued on the marginal utility of their attributes, residential hedonic modelling (HM) makes it possible to estimate the marginal monetary value of the property's specifics, neighbourhood attributes and externalities (Rosen, 1974). Although the limit between proximity and space in the case of the SPP model, or between local- and large-displacement space for the PR model is somewhat fuzzy, two types of accessibility – local accessibility and regional accessibility – appear distinctly and significantly in recent hedonic modelling work (Des Rosiers *et al.*, 2000). The authors apply principal component analysis (PCA) to GIS-measured distances and travel times to the nearest service poles, based on car and walking travel-times to a set of the nearest 17 amenities. Two highly significant accessibility factors clearly appear, confirming the two scales of accessibility- and activity-based spaces.

Concomitantly with the geographical attempt to identify spaces of perception, environmental psychologists have studied the question of residential attachment (Altman and Low, 1992; Feldman, 1996; Fried, 1982; Giuliani, 1991; Giuliani and Feldman, 1993; Twigger-Ross and Uzzell, 1996). As cited in Sundstrom *et al.* (1996 p. 493), "research is increasingly focussed on psychological attachment to places, often in the context of home and neighbourhood (Altman and Low, 1992)". The purpose is to better understand how affective bonds between people and residential environments develop, and how those contribute to one's place-identity (Bonaiuto *et al.*, 1999; Low and Altman, 1992; Proshansky *et al.*, 1983). According to Breakwell's model of identity (1986; 1992), the key concepts of identity rely on four principles: distinctiveness, continuity, self-esteem, and self-efficacy – one's perception of the ability to be effective in achieving one's goals. Concerning the desire to preserve continuity of the self-concept, two distinct self-environment relationships are discussed in the literature: the place-referent continuity, whereby specific places that have emotional significance play the role of continuity markers between, on the one hand, past and present and, on the other, present and future, and the place-congruent continuity, referring to the generic features of places assuring continuity from one place to the next. In fact, the affective bonds between self and environment may transcend the relationship with a unique or specific place, and attachment may be developed throughout space(s) for types of places with similar characteristics (Proshansky, 1978; Proshansky *et al.*, 1983; Twigger-Ross and Uzzell, 1996). Feldman (1990) has extended this notion to the idea of settlement-identity.

For the purpose of this paper, we consider that the PR or SPP paradigms are appropriate theoretical frameworks in order to consider the cognitive process of place-identity development. Furthermore, in accordance with Feldman's concept of settlement-identity, the sense of belonging or attachment is reinforced through frequency of use and activities, and partly inherited from previous place attachments. This transfer of a sense of belonging from one place to another – in accordance with the place-congruent continuity principle – explains why people feel "at home," even after having just visited a property for eventual acquisition. Part of the place-identity associated with the newly acquired property is inherited from previous residential locations, in accordance with the notion of settlement-identity (Feldman, 1990). Our contention is that the geographical hierarchical spatial concepts and the psychological dimensions of space-identity should be considered jointly in order to better understand and model residential location choices.

### 2.2.2 Residential Mobility

In line with the thesis of Rossi (1955) – people move to adapt their housing to the life cycle evolution of their household needs –, numerous studies have analysed the moving process in urban areas. Most studies analyse the propensity for moving considering various socio-demographic characteristics, at the neighbourhood or Census tract, but also at the household level. For a review of the main work in this area, see Dieleman (2001) and Quigley (1977). Clark (1983) distinguishes forced moves from adjustment moves (relating to housing, neighbourhood, and accessibility) from induced moves (relating to employment and life cycle changes). The major impact of life cycle on residential mobility has been largely recognised, and numerous studies are based on the life cycle model of the demand for housing proposed by Artle and Varaiya (1978) and Henderson and Ioannides (1983). Dieleman (2001) identifies three regularities in the residential mobility literature: (i) the strong correlation between rate of mobility and life cycle, (ii) the strong correlation between residential mobility and size and tenure of dwelling, and (iii) the interrelationships between the housing career and other aspects of the life course, such as educational and job career, and family history (Dieleman and Mulder, 2002; Mulder and Hooimeijer, 1999; Van Ommeren *et al.*, 1999). The behaviour of specific types of households in moving is studied in detail, whether concerning young adults (Clark and Mulder, 2000), elderly households (Megbolugbe *et al.*, 1999), divorcees (Timmermans *et al.*, 1996) or ethnic groups (Deng *et al.*, 2003; Gabriel and Painter, 2003) (see Dieleman, 2001 for additional references). In his multiple-attribute housing disequilibrium model, Onaka (1983) shows the extent to which specific attributes of the household and property are related to the decision to move. More recently, a major survey held in Scotland among households that acquired a property in 1990 gives an indication of motivation for moving and choosing housing (Forster, 2001). Among the ten proposed reasons for moving, wanting a larger home, wishing to own a house and changing the type of house ranked as the top three.

As pointed out by Rossi (1955), who stresses the difficulty of disentangling the reasons underlying the moving decision, “a general ‘why’ question usually produces a congeries of answers” since respondents often confuse the events or motivation leading to the move and the reasons associated with the property and location choice. This is why the three aspects of residential choice – motivation for moving, property choice and more generally spatial location choice – have to be addressed concomitantly, in order to sort out the various dimensions of residential behaviour.

### 2.2.3 Residential Choice

Residential studies on preference, choice or satisfaction are based on either stated or revealed data. The first use hypothetical or intended statements chosen from a constructed and often controlled range of possibilities, and are mainly used for preference and choice studies. The second are based on surveys or on actual sale or rent price analyses, and apply mostly to choice and satisfaction. The choice process is central to preference and satisfaction, as choices result from preferences while satisfaction relies on past choices.

Among the main stated preference and choice analysis methods are contingent valuation estimates of the willingness to pay (WTP) – mostly applied to the valuation of environmental amenities (see Cummings *et al.*, 1986) –, conjoint analysis methods, which relies on ranking or scaling various goods and attributes (Goodman, 1989), and choice-based methods, whereby respondents choose one combination of attributes from a constructed and controlled set of possibilities (Timmermans *et al.*, 1992; Timmermans and Van Noortwijk, 1995). Choice-based conjoint analysis (based on stated choices), has been derived from discrete choice modelling (based on actual choices), in turn derived from the random utility theory first developed by Thurstone (1927), and further put in the context of the multinomial logit model (MNL) by McFadden (1978). As Earnhart (1998) points out, a few authors only have used this framework for actual residential choice studies (Friedman, 1981; Longley, 1984; Nechyba and Strauss, 1998; Quigley, 1976 & 1985; William, 1979). Pellegrini and Fotheringham (2002) provide an interesting discussion about discrete choice models and their use in a spatial context. Whereas some critics consider that actual choice sets induce sample selection biases, the discrete choice method is extended to stated preferences and hypothetical choices and is termed choice-based conjoint analysis (Hauser and Rao, 2002). However, in order to be able to consider the numerous potential combinations of attributes of complex goods – and more specifically of residential property –, various refinements were developed. Considering the Hierarchical Information Integration (HII) method proposed by Louvière (1984), Louvière and Timmermans adapted a choice-based HII to residential choice (1990). More recently, Oppewal *et al*. (1994) proposed an integrated conjoint choice experiments approach (IHII), later tested in the residential context by van de Vyvere *et al*. (1998), and later adapted to the study of group preferences (Molin *et al.*, 2001). HII relies on the assumption that, when confronted with a complex decision or evaluations involving numerous elements, people group attributes in various constructs, that are valued separately. Combining these construct evaluations leads to overall preference, satisfaction, or choice decisions. Studies using this framework for residential choice analysis generally distinguish between two hierarchical levels: housing and location constructs. Whether using stated or revealed-choice methods, we support the opinion that the hierarchical structure of perceived spaces should be considered in the residential choice process, using the behavioural-based SPP or PR theoretical frameworks.

Another interesting alternative to exploring revealed residential choices is the hedonic framework. Based on the principle that goods are valued on the marginal utility of their attributes, hedonic modelling (HM) makes it possible to estimate the marginal monetary value of the property specifics, neighbourhood attributes and externalities (Rosen, 1974). Most hedonic models estimate one general coefficient for each measured attribute. However, the expansion method (Casetti, 1972 & 1997), using interactive variables, makes it possible to estimate the variation of the marginal value of any attribute according to the context, that is, for example, the spatial location (Kestens *et al.*, At Press; Thériault *et al.*, 2003) or the socio-demographic characteristics of the buyer's households. However, to the best of our knowledge, no hedonic model has so far incorporated the interactions at the household level, due to the relative scarcity of appropriate data bases as well as to conceptual issues regarding the very nature of the hedonic function. This paper is an attempt to analyse the heterogeneity in the importance accorded to various residential choice criteria considering the household profile, the relative location within the metropolitan area, and the attachment to the neighbourhood. In a forthcoming paper, we plan to verify – using the same disaggregated databases – whether the HM framework can statistically reveal some variability in the marginal values of property and neighbourhood characteristics depending on the household profiles.

Although many studies in residential choice have analysed the influence of housing attributes on residential choice using stated or revealed data – or both (Earnhart, 1998) –, very few studies have addressed the question of how homogeneous the household choice criteria are depending on both the socio-demographic profiles and the final location choice of property buyers. Heterogeneity in tastes is difficult to measure within the random utility framework model and has rarely been addressed within the residential choice literature (see Adamowic (2002) and Boxall (1999) for a review of the main methods of heterogeneity measures on random utility models). A study of environmental perception conducted on households in Geneva revealed some determinants of perceived housing environmental quality (Bender *et al.* 1997 & 2000). Quietness and greenness of the area are the main factors, while accessibility to the city centre and the social value of the neighbourhood appeared to be the less decisive factor. In Zurich, a spatial analysis of responses showed that the importance devoted to the distance to the city centre, the distance to school, and the social standing of the neighbourhood varies according to location, whereas the importance of other environmental quality factors was similar in the four postal-code defined areas.

Recently, Molin and Timmermans (2003) measured the links between socio-demographic characteristics of the household – age, education, income, daily activities – and the actual housing and location attributes, within a larger structural equation model aimed at validating the causal relationships between household characteristics, construct attributes, construct valuations and overall preference. Primary findings regarding housing underline the positive link between education and size (number of bedrooms), education and type of tenure, as well as income and housing costs. Concerning location, the few significant attributes are "frequency of transit transport" – negatively linked to the husband's income – and "travel time of wife" – negatively associated with age and positively linked to the wife's educational attainment.

Considering that additional research is needed in order to better understand the heterogeneity in residential choice criteria, this paper uses logistic regression to analyse the motivation to move and expressed choice criteria of 774 single-family property buyers in Quebec City, Canada.

## 2.3 Data Bank and Analytical Approach

### 2.3.1 Data Bank

A computer-assisted phone survey was carried out, between 2001 and 2002, of single-family property owners who bought their homes (1993 to 2001) in Quebec City, mostly over the 1993-1996 period (88%). Some 2521 people answered calls, 1134 (45%) agreed to respond to the survey, and 774 answered all the questions, including income. The 1134 occurrence sample was stratified spatially, in proportion to all single-property transactions which occurred in the 13 former municipalities now amalgamated into Quebec City. This paper analyses the sample of 774 homeowners for whom we have complete answers, which represents 6.3% of all single-family property transactions over the 1993-1996 period. The main topic of the survey concerned motivation for moving, residential choice criteria and sensitivity to the immediate surroundings. We are presenting readers with our analysis of the answers to three questions: (i) what was the motivation for moving, (ii) what were the property choice criteria and (iii) what were the neighbourhood choice criteria? These questions were asked in an open format: no answer list was suggested, and the number of answers was unlimited. Afterwards, the answers were grouped by category, resulting in 21 possible motivations for moving, 19 neighbourhood choice criteria, and 20 property choice criteria. Additional socio-economic data describing the household were collected. They concern the type of household, the occupation and educational attainment of the respondent and (eventually) his or her partner, the income of the household, and the age of the respondent.

Table 8. Socio-demographic profile of property buyers' sample (N=774)

| | Mean (Std dev.) / Proportion | Age (Mean) | Income (Mean, Cad $) |
|---|---|---|---|
| Age | 42 (8) | | |
| University degree | 55% | 36 | 77 482 |
| University education without degree | 4% | 37 | 65 714 |
| College degree | 29% | 36 | 71 452 |
| High-school degree | 11% | 38 | 55 960 |
| Single-parent family | 7% | 38 | 44 035 |
| Single-person household | 6% | 42 | 49 286 |
| Couple without child | 18% | 39 | 71 844 |
| Couple with child | 67% | 34 | 73 480 |
| Dual workers | 70% | 34 | 75 595 |
| Single worker | 30% | 40 | 54 830 |
| Ex-owner | 52% | 39 | 72 040 |
| New owner | 48% | 33 | 66 263 |
| Income in Cad $ (Std. Dev.) | 69 264 (22 705) | | |

Table 8 contains key numbers describing the socio-demographic profiles of the households surveyed. The majority of the respondents have university degree (55%), and nearly one-third (29%) holds a college degree. Households are predominantly couples (85%), among which 79% have children. Single-parent families represent 7% of the sample, as opposed to 6% for single-person households. Some 70% of households are dual-workers while the mean household income is close to \$70 000. This value is underestimated, however, since income questions were asked in \$10 000 brackets, with the highest being \$100 000 or more. Single-parent families have the lowest income, around \$44 000 on average. The mean age at transaction date is 36; however, when considering new and former owners, mean ages would be respectively 33 and 39. Nearly half of the sample is comprised of new owners (48%).

## 2.3.2 Analytical Approach

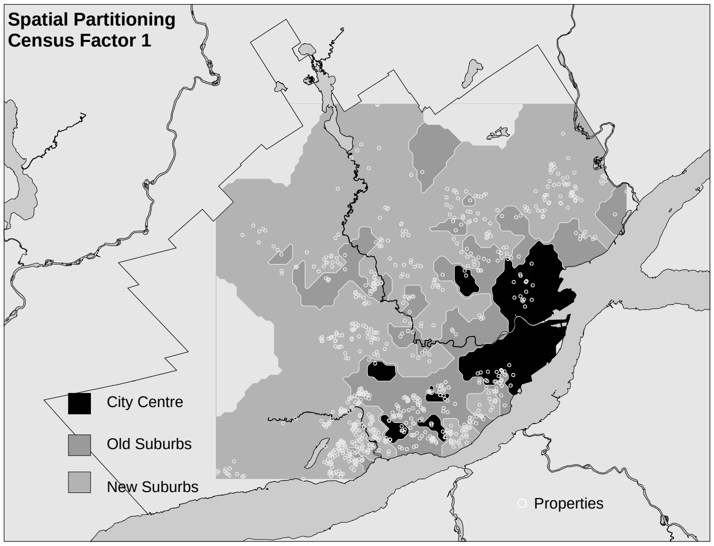
First, a simple frequency analysis for each criterion gives us insight into overall moving and residential choice motivation. Next, a correspondence analysis was performed on both property and neighbourhood choice criteria in order to verify which groups of criteria emerge and to bring out their eventual concordance with the concepts of the place-proximity-space model. Thirdly, several binary logistic regression models were built using a forward stepwise procedure. For each moving or choice criterion, a logistic regression estimates the likelihood of being mentioned depending on (i) the household profile – age, income, dual worker, education, household type –, (ii) the location of the property and (iii) whether people felt attached to the neighbourhood or not at the time of purchase. In order to take into consideration interaction effects, several two-dimensional interactive variables have also been included in the model. In order to ease interpretation, the household profile variables were categorised as shown in Table 9. Concerning the location within the CMA, two types of spatial division were considered. These areas are based on a PCA performed on 1996 Census data on the Census tract level, which resulted in two major socio-demographic factors (See Des Rosiers *et al.*, 2000 for detailed procedure). For each of the two Census factors, three categories were constructed, that is, low (<-0.5), medium (between –0.5 and 0.5), and high factor scores (>0.5), resulting in two spatial divisions of the territory (see Figure 4). The first factor expresses centrality, distinguishing the city centre with a majority of tenants from, on the one hand, the old suburbs and, on the other, more recent developments with low densities in the more remote parts of the city. The second factor is mainly related to family life cycle, with a well-educated Upper Town, a Lower Town characterised by low incomes, and mixed suburbs with young families. Consideration of two distinct spatial dimensions based on orthogonal PCA factors represents an attempt to both determine the geographical division which is the most appropriate for explaining differences in residential strategies and to account for the socio-spatial segmentation of the city.

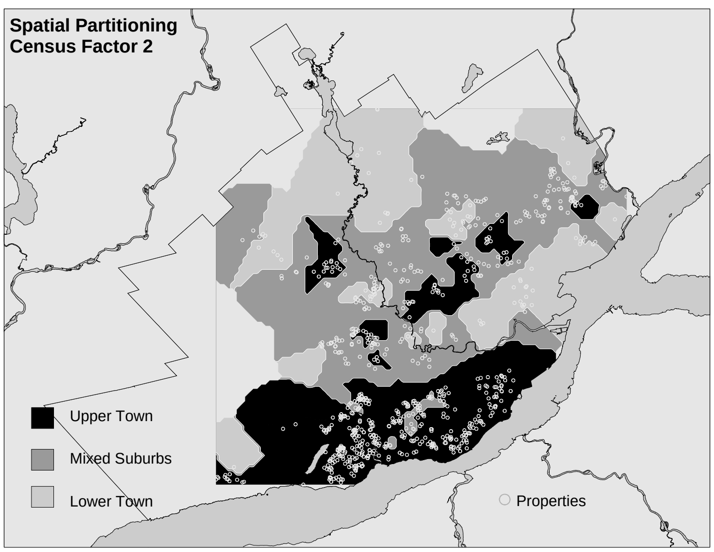
Table 9. Variables used in logistic regression models

| Variable Type | Variable | Definition/Categories | n |
|---|---|---|---|
| Categorical | Household type | Couple with children | 520 |
| | | Couple without child | 141 |
| | | Single-parent family | 57 |
| | | Single-person household | 47 |
| | Age | Less than 30 | 178 |
| | | 30 to 39 | 350 |
| | | 40 to 49 | 197 |
| | | Over 49 | 49 |
| | Income / 10 000 $ | Income in 10 000 $ intervals | |
| | Income | Less than 50 000 $ | 127 |
| | | From 50 to 80 0000 $ | 217 |
| | | Over 80 000 $ | 430 |
| | Education | University degree | 425 |
| | | College degree | 250 |
| | | Secondary and below | 99 |
| | Location Census Factor 1 | City centre | 31 |
| | | Old suburbs | 215 |
| | | New suburbs, fringes | 528 |
| | Location Census Factor 2 | Upper town | 416 |
| | | Lower town | 45 |
| | | Mixed suburbs | 313 |
| Binary | Single | 1= Single household or Single-parent family; 0=Couple | 104 |
| | Child | 1=Child in household; 0=No child | 577 |
| | Attached | 1=Stated to choose the neighbourhood because of attachment | 206 |
| | Dual Worker | 1=Dual worker household; 0=non dual worker household | 538 |
| | First-time owner | 1=First-time owner, 0=Former owner | 372 |

Grey cells: Reference category used in logistic regression

Figure 4. Spatial partitioning using Census factors




## 2.3.3 Logistic Regression: A Few Interpretation Keys

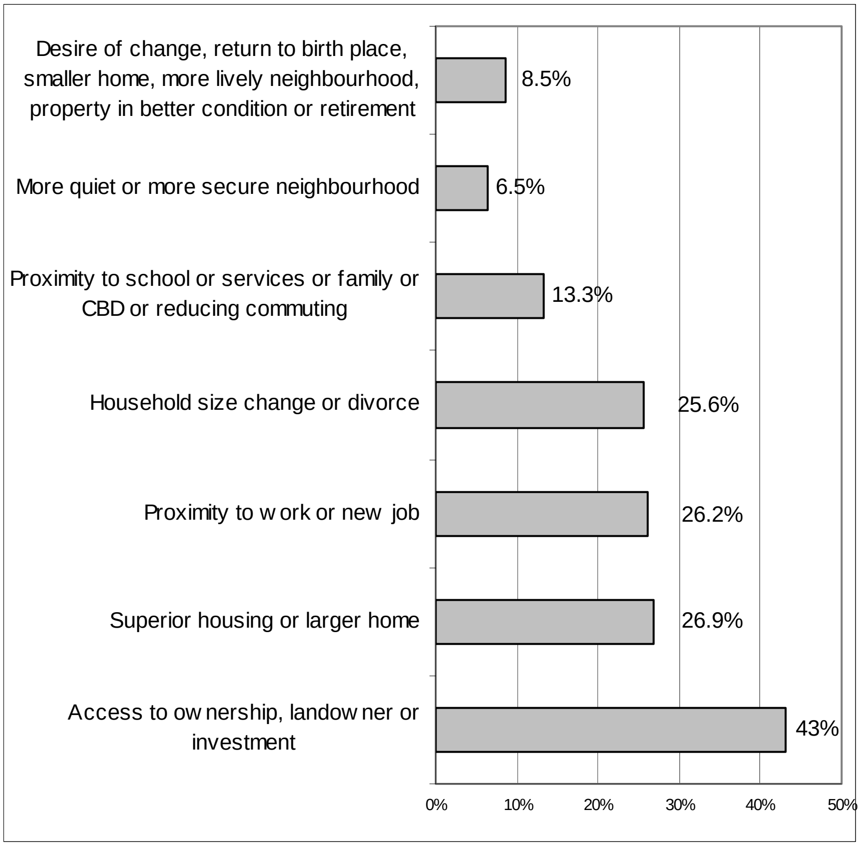
In logistic regression, the global test of parameters is a chi-square test. The probability of the observed results given the parameter estimates is known as the *likelihood*. As it is a small number, inferior to 1, -2 times the log of the likelihood (-*2LL*) is generally used. This measures how well the estimated model fits the data, and is analogous to the sum of squared errors (SSE) in the OLS regression model. The chi-square is the difference between the –*2LL* of the initial log likelihood function, in which only the constant is included, and the –*2LL* of the final model. We also present the Nagelkerke *R*-square, adapted from Cox and Snell's *R*-square (Cox and Snell, 1989), and which is presented by Nagelkerke (1991, p. 691) as the equivalent of the "classical *R*-square" of linear regression. The Wald statistic (W), based on a chi-square distribution, tests the significance of the coefficients associated with each variable. Furthermore, the exponential of the B coefficients expresses the odds ratio (probability of an event / probability of non-event). If Exp(B) is greater than one, the odds are increased, and vice-versa. The model performance presented in Table 13, Table 14 and Table 15 are the chi-square values with their associated probability, ending *-2LL*, and the Nagelkerke *R*-square. The main idea behind this paper being to better understand the *links* between household profile and moving and choice criteria, the focus is on the identification of significant variables. Therefore, even models yielding low *R*-squares are presented if significant variables emerge. The most important is to measure and determine which impact the proposed variables have on the likelihood in order for the choice criteria to be mentioned. Due to the lack of space, Tables 8, 9 and 10 present only the result of a selection of logistic models, giving the above-mentioned model performances as well as the odd-ratio and associated probability for each significant independent variable.

## 2.4 Moving Incentives: Some Results

### 2.4.1 Overview

Figure 5 provides a picture of the main motivation for moving. As can be seen, access to ownership or investment ranks first (43%). The second most frequent goal concerns the will to better one's housing, mainly in terms of house size (26.9%), followed by proximity to work (26.2%). Roughly the same percentage of buyers (25.6%) mention certain aspects of the household's lifecycle as a reason for moving, either a change in the size of the household or a divorce. Proximity to school, services, family, CBD (Central Business District) or the will to reduce commuting time comes fifth (13.3%). Only 6.5% of the buyers indicated a desire for a quieter or more secure neighbourhood as a motivation for moving. This minor concern about security issues is likely to be specific to Quebec City, which has the lowest crime rate among the 25 largest CMAs in Canada (Statistics Canada, 2001a).

Figure 5. Frequency of expressed moving motivations




Concerning the residential choice, four main groups of criteria emerge, both as regards the neighbourhood and the property. Details of criteria in each group and the frequency of respondents who cited at least one of the criteria are given in Table 10. For the choice of the neighbourhood, these main groups are related to accessibility (at least one of the criteria in this group cited by 60% of respondents), the socio-economic and urban context (43%), a psychological dimension, attachment (27%) and aesthetics (25%). Regarding the property, the size factor is the most prevalent (48%), followed by interior features (37%), style (36%) and environmental considerations (15%). Detailed frequencies for all expressed criteria are given in Table 11.

Table 10. Classification of residential choice criteria

| | Group | List of criteria in the group | At least one of the criteria cited by ...% of the respondents |
|---|---|---|---|
| Property choice | Size | Lot size, house size, number of rooms | 48% |
| | Interior | Interior architecture, floor quality, functionality, interior decoration, garage | 37% |
| | Style | Architectural style, condition | 36% |
| | Environment | Trees, landscaping | 15% |
| Neighbourhood choice | Accessibility | proximity to services, job, school, highway, CBD, public transit system | 60% |
| | Socio-economic context | quietness, young nbhd, security, lively | 43% |
| | Attachment | attachment | 27% |
| | Aesthetics | cachet, trees | 25% |

Table 11. Detail of frequencies of expressed neighbourhood and property choice criteria

| Neighbourhood choice criteria | Frequency | Property choice criteria | Frequency |
|---|---|---|---|
| Services | 37% | Property price | 39% |
| Quietness | 35% | Lot size | 29% |
| Attachment | 27% | Interior architecture | 27% |
| Proximity to work | 19% | Architectural style | 26% |
| Proximity to school | 19% | Property size | 17% |
| Cachet | 16% | Number of rooms | 16% |
| Trees | 12% | General condition | 14% |
| Highway accessibility | 11% | Trees | 13% |
| Proximity to CBD | 8% | Floor quality | 6% |
| Young neighbourhood | 7% | Commod/functn | 5% |
| Transit network | 6% | Interior decoration | 4% |
| Security | 5% | Landscaping | 3% |
| Lively neighbourhood | 5% | Garage | 2% |
| Taxes | 4% | No neighbour | 2% |
| Other (park, low traffic, prox. to bridge, suburb, view) | <3% each | Other (finishing, in-ground pool, above-ground pool, orientation, view, surface material) | <2% each |

In order to go further than the main groups of criteria identified in Table 10, a correspondence analysis (CA) was performed on the residential choice criteria, combining both property and neighbourhood. Correspondence Analysis (Benzecri, 1973; Greenacre, 1984 & 1993) is similar to principal component analysis extracting eigenvectors (PCA), but is better adapted for presence/absence data. It is widely used in ecology for searching associations among species and focus on relations between species and environment (Hill, 1974; Legendre and Legendre, 1998). This CA extracted eight factors, of which the eigenvalues represent the correlation coefficient between "species scores" and "sample scores", that is, between weighted values of variables and observations. The cumulative percentage of explained variance is 48.3%. Looking at the factor scores in Table 12, it appears that the first factor relates to the neighbourhood choice criteria, whereas factor four and eight are tied to the property specifics. The first factor opposes cachet and environmental quality of the neighbourhood (trees) on the one hand, to proximity to work and services on the other. Factor 4 is clearly an indication of the quality of the property specifics, and Factor 8 opposes inside and outside property attributes (style and floor quality vs. landscaping). Moreover, Factor 6 reveals the trade off between property size and centrality, and Factor 7 the trade-off between location (highway proximity and cachet) and property specifics (number of rooms and interior decoration). Factor 2 contrasts objective (property quality and young neighbourhood) and subjective (attachment) criteria. Finally, Factor 3 relates to the trade-off between financial ability to pay (price) and aesthetic criteria (trees and landscaping).

In the perspective of the Space-Proximity-Place and Place-Identity models, this correspondence analysis clearly distinguishes the importance of place (factors 3, 4 and 8) and proximity (Factor 5), while certain trade off considerations are given by Factor 6 (place/space trade-off) as well as factors 1 and 7 (place/proximity trade-off). Additionally, Factor 2 relates to place identity, underlining the opposition between psychological attachment and objective choice criteria.

Table 12. Correspondence analysis on property and neighbourhood choice criteria: factor scores

| | COMPONENTS | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| | Eigenvalues | 0.306 | 0.295 | 0.284 | 0.273 | 0.254 | 0.235 | 0.23 | 0.214 |
| | % of expl. variance | 7.06 | 6.82 | 6.57 | 6.31 | 5.86 | 5.42 | 5.31 | 4.95 |
| | Cumulative % | 7.1 | 13.9 | 20.4 | 26.8 | 32.6 | 38.0 | 43.3 | 48.3 |
| **Property** | Price | | | **0.900** | | | | | |
| | Lotsize | | | | | | | | |
| | Design | | | | | | **0.916** | | |
| | Style | | | | | | | | **-0.635** |
| | Size | | | | **0.689** | | **1.386** | | |
| | Nb Rooms | | | | | | | **-0.579** | |
| | Condition | | | | **0.632** | | | | |
| | Trees | | | **-0.868** | | | | | |
| | Floor Qual. | | **-1.025** | | **1.401** | **1.671** | | | **-0.615** |
| | Functionality | | | | **0.662** | | | | |
| | Inter. Deco. | | | | | **1.080** | | **-0.943** | |
| | Landscaping | | | **-1.171** | | | | | **3.173** |
| **Neighbourhood** | Services | **0.457** | | | | | | | |
| | Quiet | | | | | | | | |
| | Attachment | | **1.212** | | | | | | |
| | Work | **1.098** | | | | **1.476** | | | |
| | School | | | | | **-0.545** | | | |
| | Cachet | **-1.086** | | | | | | **1.585** | |
| | Trees | **-1.095** | | | | | | | |
| | Highway | | | | | | | **1.452** | |
| | CBD | | | | | | **-1.068** | | |
| | Young | | **-1.154** | | | **-0.995** | | | |
| | Trans. Network | **1.365** | | | | | | | |
| | **Interpretation** | Proximity / Cachet trade-off | Objective (Prop. Qual. and Neighbd) vs Subjective criteria (Attachment) | Landscaping & Trees / Price trade-off | Property quality and size | Young neighbourhood / Work proximity and prop. quality (life-cycle) | Centrality / Size trade-off | Highway proximity / Property quality | Interior / Exterior Style |
| | **In the context of theoretical models (PPS and Place-Identity)** | Place / Proximity trade-off | Place Identity | Place | Place | Proximity | Place / Space trade off | Place / Proximity trade-off | Place |

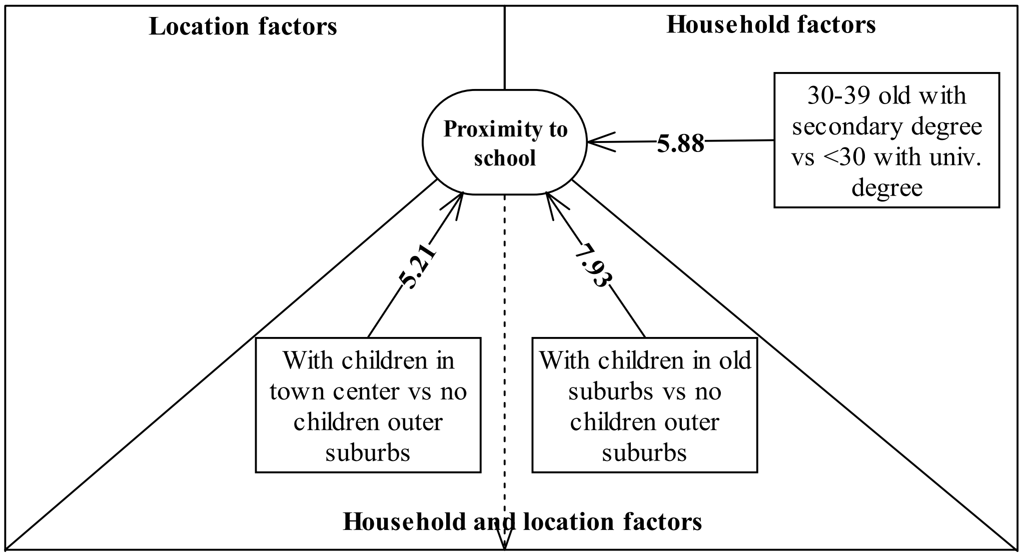
Whether for motivations for moving or for neighbourhood and property choice, we then estimated the likelihood for a criterion to be mentioned depending on the household profile, the psychological attachment, and the final location choice. Table 13, Table 14 and Table 15 present a summary of the logistic regression models built for this purpose for each of the categories identified in Table 10 and for various specific criteria. The complete results for all the logistic regressions could not be shown here due to space limitations. However, the significant relationships will be reported when contributing to a better understanding of the residential strategies, even if all corresponding models are not actually shown in the tables. Although the overall fit of the models (chi-square and associated probability) is an indicator of their significance, our attention is mainly focussed on the *identification* of the significant relationships between variables. Some variables are significant as such, but consideration of interactions adds much to the explanatory power. In order to facilitate the interpretation of the logistic regression results, the dependent variables are written in *italics* and the odd ratios are given in parentheses in the following section. It is important to bear in mind that these odd ratios express the likelihood for a criterion to be mentioned in contrast to the reference categories, which are defined in the first column. For categorical variables, the latter account for respondents under 30 years old, households with a yearly income of more than \$80 000, university-degree holders, couples with children (household type), new suburbs (Census Factor 1 socio-spatial division), and mixed suburbs (Census Factor 2 socio-spatial division). For binary variables, the odds ratio measures the likelihood of being cited over that of being omitted.

### 2.4.2 Moving Incentives

The most frequent moving incentive, the desire to *own property* or to make an investment, is obviously closely related to the previous ownership status. Furthermore, low-income households both between 30-39 (odds ratio of 0.214) and 40-49 (0.085) are less likely to report this motivation than younger households with high incomes (Table 13). Although the size factor – wanting a *larger home* or a *superior housing* – is often brought up, few significant differences emerge considering household profiles. New owners, as well as low-income single-adult households are less likely to mention this motivation for moving (0.676 and 0.330, respectively).

Motivation for moving relating to job issues (*new job* or *moving closer to job*) is mainly associated with age and household size. Childless couples over 50 are 3.1 times less prone to report this reason than young couples with children (0.322), whereas single parents between 30 and 39 advance it 4.5 times less often (0.22). Also, first-time owners are much less concerned with work issues (0.35). However, if we look at each criterion in detail, *new job* is five times less likely to be mentioned by first-time owners (0.21), as opposed to only two times for *proximity to work* (0.49). Furthermore, the *new job* argument is directly related to income, each additional $10 000 increasing the probability of this moving incentive by 13% (1.13). The other proximity issues, grouped in the proximity factor – *proximity to services, schools, CBD* – are mainly associated with household status and presence of children: dual-worker households are twice less concerned with proximity issues (0.50), but people with children who settle in the old suburbs, are more than twice more concerned (2.64). Here again, bringing up proximity issues is nearly three times less likely by first-time owners (0.36). P*roximity to school* is much more of a concern for households with children settling in the city centre, and even more so for those settling in the old suburbs (5.21 and 7.93). Middle-aged families with little education are *more* likely to mention proximity to school as an incentive to move than younger parents with university degrees (5.88), probably because new parents are less preoccupied with school issues as their children are not yet of school-age (Figure 6).

Figure 6. Links between location, household profile and proximity to school as a moving motivation




Numbers in arrows are odd-ratios given by logistic regression

## Table 13. Logistic regression models for moving motivations

| Household Attribute \ Moving Criterion | | Ownership | Size Group | Job Group | New Job | Proximity to Work | Proximity Group | Proximity to School | Proximity Family | Household Size Change | Divorce | Secure Neighbhd |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Age (vs under 30)** | 30-39 | ¤Inc.<50K 0.214*** | | ¤SgleParFam 0.22** | | ¤OldSuburbs 2.21*** | | ¤Secondary 5.88*** | | ¤DualWork 0.499*** | ¤UpperTown 38.05***<br>¤NewOwner 0.028*** | |
| | 40-49 | ¤Inc.<50K 0.085*** | | | | | | | | ¤DualWork 0.517** | ¤UpperTown 5.61**<br>¤NewOwner 0.07**<br>¤CollDegree 10.93*** | |
| | 50 and plus | | | | | | | | ¤Inc.<50K 6.53**<br>¤Inc.50-80K 10.57*** | | | |
| **Income** | Income (/10 000$) | 0.811*** | | | 1.13** | | | | | | | |
| **Income (vs >80K)** | <50K | ¤Age30-39 0.214***<br>¤Age40-49 0.085***<br>¤SgleParFam 3.93**<br>¤CpleNoChild 4.85**<br>¤SglePersHld 8.24*** | ¤SgleAdult 0.330** | | | | | | ¤Age>49 6.53** | | | |
| | 50K-80K | | | | | | | | ¤Age>49 10.57*** | | | ¤Secondary 5.91** |
| **Dual worker** | Dual-worker | 1.96** | | | | | 0.503*** | | | ¤Age30-39 0.499***<br>¤Age40-49 0.517** | | |
| **Education (vs Univ.Degree)** | College degr. | | | | | | | | ¤Attached 4.32** | | ¤Age40-49 10.93*** | |
| | Secondary | | | | | | | ¤Age30-39 5.88*** | | | | |
| **Household Type (vs couple with child)** | Single Par. Fam. | ¤Inc.<50K 3.93** | | ¤Age30-39 0.22** | ¤NewOwner 16.07** | | | | | | 93.6*** | |
| | Couple without child | ¤Inc.<50K 4.85** | | | | | | | | | | |
| | Single-Pers. Household | ¤Inc.<50K 8.24*** | | | | | | | | | 60.01*** | |
| **Child at home** | With child | | | | | | ¤OldSuburbs 2.64*** | ¤Centre 5.21**<br>¤OldSuburbs 7.93*** | | 2.91 *** | | |
| **Couple vs Single Adult** | Single Adult (with or without child) | | ¤Inc.<50K 0.330** | | 0.124** | | | | | ¤NewOwner 2.73*** | | |
| **Loc. CSF1 (vs new suburbs)** | Centre | | | | | | | ¤WithChild 5.21** | | | | |
| | Old suburbs | | | | | ¤Age30-39 2.21** | ¤WithChild 2.64*** | ¤WithChild 7.93*** | | | | |
| **Loc. CSF2 (vs mixed suburbs)** | High income upper town | | | | | | | | | | ¤Age30-39 38.05***<br>¤Age40-49 5.61** | |
| **Previous ownership** | New owner | 20.53*** | 0.676** | 0.243*** | 0.211***<br>¤SgleParFam 16.07** | 0.486*** | 0.357*** | | 0.29** | ¤Single 2.73*** | ¤Age30-39 0.028***<br>¤Age40-49 0.073** | |
| **Attachment** | Attached | | | 0.35*** | 0.374*** | 0.402*** | 0.482*** | | ¤CollDegree 4.32** | | Not proposed | |
| **Model performance** | Chi-square | 36.7 | 17.5 | 120.0 | 76.7 | 36.5 | 51.7 | 36.6 | 33.2 | 36.9 | 114.6 | 24.6 |
| | Sig | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 |
| | Ending -2LL | 675.1 | 883.48 | 770.78 | 499.7 | 627.8 | 555.44 | 236.2 | 153.23 | 785.55 | 139.29 | 146.35 |
| | Nagelkerke $R^2$ | 0.052 | 0.019 | 0.135 | 0.133 | 0.055 | 0.119 | 0.156 | 0.197 | 0.071 | 0.492 | 0.158 |

**Values in cells are odds ratios; interactions identified by ¤NameofVariable; *** sig<0.01; ** sig<0.05.**

The motivation to move closer to family (*proximity to family*) shows a different pattern. New owners are much less concerned (0.29), but older households are far more likely to mention the family criterion, especially when the income is low, meaning less than $50 000 (6.53), but even more so when the income is average (10.57). Furthermore, the family proximity is 4.3 times more frequently cited by college-degree holders who feel attached to the neighbourhood they ended up choosing.

Concerning life-cycle-related motivation for moving, the *household size change* argument is given mainly by families who now have children at home (2.73), and is half less frequently given by dual-worker households aged 30 to 39 (0.55) or aged 40 to 49 (0.517). Logically, single-person households are 6.5 times less likely to refer to this criterion (0.16). However, bringing up *divorce* (or separation) as a moving incentive is more likely for single-parent families and single-person households. We did not use the variable "attached to the neighbourhood" for this model, because other more relevant variables were added while this dimension was omitted. In fact, as we shall see later, the feeling of attachment to the neighbourhood is in itself partly linked to single-parent families, who do probably move to known places after a separation in order to improve integration and minimise stress. The divorce criterion is mainly associated with age, the chosen home location and the educational attainment. People who settle in the Upper Town are 38 times more likely to cite this criterion if they are 30 to 39 years old, and 5.6 times more likely for those in their forties. Also, 40 to 49–year old persons holding college degrees are 10.9 times more likely to be concerned with divorce as reason for moving than younger persons holding university degrees. The divorce rate in Canada is one of the highest among the western countries, following close behind the United States and the U.K (Ambert, 1998). Furthermore, Quebec City's rate is 20% higher than the country's average value (Statistics Canada, 2001b). A Pan-Canadian study on divorce held in the beginning of the 1990s showed that the divorce rate is at its highest five years after marriage and swiftly diminishes afterwards, mainly affecting 25- to 29-year-old couples (Gentleman and Park, 1997). Our findings indicate a higher probability of mentioning divorce as a moving motivation for couples in their thirties, this holding true for those settling in the higher socio-economic status area of the Upper Town.

Interestingly, one interactive variable only is linked to the likelihood of a concern for security, persons with less education and an income of between \$50 000 and \$80 000 being nearly six times more concerned than high-income university-degree holders. Not only is the security argument itself brought up very rarely (2.3%), but no differences on security sensitivity could be linked to age or household type. As Quebec City has a very low crime rate, these findings are not surprising. Similarly, the desire to move to a quieter neighbourhood could not be linked to any household descriptor.

As a concluding remark, it appears that the previous ownership status (first-time owners vs. former owners) is one of the most important determinants of the motivation to move. Former owners are much more motivated by housing quality and location issues, such as proximity to work and services. Age, which can be considered a proxy of life cycle, also plays an important role, both for proximity issues (work and school) and for family-career arguments (household size change or divorce). The educational attainment has nearly no impact on the motivation to move. Furthermore, it is interesting to point out that the psychological dimension of attachment to the neighbourhood has important ramifications with regard to moving incentives. People who do feel attached to the neighbourhood are less likely to cite *proximity to work* (0.40) or *new job* (0.36) as a motivation for moving than people who do not, but more likely to consider *proximity to family* when holding a college rather than a university degree (4.32).

## 2.5 Neighbourhood Choice Criteria

Two location variables are used among the independent variables in the logistic regression models (Table 14). At first, the introduction of location variables in order to explain neighbourhood choice criteria can seem tautological. People who have the same desires concerning location will settle close to each other. Therefore, estimating the likelihood of mentioning a neighbourhood criterion by using the final location may seem awkward. However, we deliberately decided to include final location choice among the explanatory variables in order to underscore the links between neighbourhood criteria and location. In fact, knowing for example that people settling in the old suburbs are twice as likely to evoke the proximity to services as a choice criterion gives us insights into people's perception of the relative quality of these areas within the city. Furthermore, the differences in criteria frequency can be significant with one or the other of the two spatial divisions, the first being based on centrality and the second mainly on life cycle factors. This duality can better our understanding of the multiplicity of the spatial context – or at least of the *perceived* spatial context – within the city. Furthermore, very little further significant information was obtained concerning the impact of household attributes when re-running the logistic regressions without location factors. The dichotomous variable "choosing the neighbourhood because of a feeling of belonging" was used as a predictor, except of course for explaining the attachment criterion itself.

Table 14. Logistic regression models for neighbourhood choice criteria

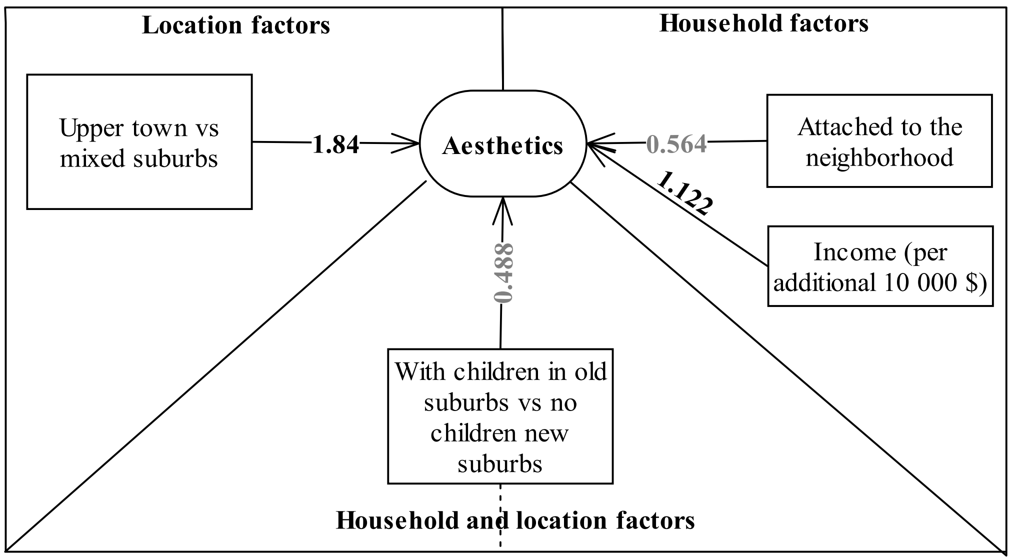
| Household Attribute | Neighb. choice Criteria | Proximity Group | Services | Transit | School | Job | Socio-econ. Group | Cachet | Trees | Aesthetic Group | Attachment |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Age (vs under 30)** | 30-39 | | | | | ¤CpleNoChild 0.299** | ¤Inc.50-80K 1.92*** | | | | |
| | 40-49 | | | ¤Withchild 3.38*** | ¤UpperTown 2.14*** | | | | | | |
| | 50 and plus | ¤Inc.50-80K 11.47*** ¤Income 0.832*** | ¤Inc.50-80K 3.17** | | | | | | | | ¤Income<50K 9.72*** |
| **Income** | Income (/10000$) | ¤Age>49 0.832*** | | | | | | 1.15*** | | 1.12*** | |
| **Income (vs >80K)** | <50K | | | | | | | | | | *Age>49 9.72*** |
| | 50K-80K | ¤Age>49 11.47*** | ¤Age>49 3.17** | | | | ¤Age30-39 1.92*** | | | | |
| **Household Type (vs couple with child)** | Single Par. Fam. | | | | | | | | | | ¤LowerTown 4.87*** |
| | Couple without child | | | | | ¤Age30-39 0.299** | | | | | |
| **Couple vs Single Adult** | Single Adult (with or without child) | | | | | | 0.436*** | | ¤NewOwner 0.235** | | |
| **Child at home** | With child | | ¤OldSuburbs 2.36*** | ¤Age40-49 3.38*** | 4.33*** | 0.397** | | | ¤OldSuburbs 0.46** | ¤OldSuburbs 0.488*** | |
| **Previous ownership** | New owner | | | | | | | | ¤SingleAdult 0.235** | | |
| **Location/CSF1 (vs new suburbs)** | Centre | | | | | | 0.276*** | | | | |
| | Old suburbs | 3.14*** | ¤WithChild 2.36*** | 2.27*** | | 2.43*** | 0.683** | | ¤WithChild 0.46** | ¤Withchild 0.488*** | |
| **Location/CSF2 (vs mixed suburbs)** | Upper town. high income | | | | ¤Age 40-49 2.14*** | | | | 2.54*** | 1.84*** | |
| | Lower town. low education | | 0.418** | | | | | | | | ¤SglParFam 4.87*** |
| **Attachment** | Attached | 0.492*** | 0.682** | | 0.490*** | 0.537*** | | 0.571** | | 0.564*** | Not proposed |
| **Model performance** | Chi-square | 82.2 | 55.7 | 23.1 | 58.4 | 54.1 | 39.9 | 16.1 | 27.5 | 37.6 | 35.4 |
| | Sig | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Ending -2LL | 959.9 | 968.2 | 320.2 | 691.2 | 701.3 | 1018.0 | 661.7 | 552.8 | 825.1 | 861.6 |
| | Nagelkerke $R^2$ | 0.136 | 0.095 | 0.082 | 0.117 | 0.108 | 0.067 | 0.035 | 0.066 | 0.071 | 0.065 |

**Values in cells are odds ratios; interactions identified by ¤NameofVariable; *** sig<0.01; ** sig<0.05.**

The major significant predictors for differences of neighbourhood choice criteria are the type of final geographical location, age, type of household (mainly presence or absence of children), income, and the attachment to the neighbourhood (Table 14). We did not find any significant link to the educational attainment. The proximity-factor – *proximity to services*, *school*, *CBD*, *public transit*– is positively linked to people who settle in the old suburbs, which is an area with proper accessibility to services and the workplace, as previous hedonic modelling work has shown. Interestingly, proximity issues are primarily related to age and income, older people feeling less concerned as their income increases. *Accessibility to services* is rather a concern for parents who chose to live near the city centre in the old suburbs (2.36). Both Lower-Town homeowners and those attached to their neighbourhood are less likely to mention proximity to services as a neighbourhood choice criterion. Proximity to the *public transit* is mainly a preoccupation for parents in their forties, underlining their sensitivity to the accessibility of their teenagers who do not own a driver's license or have access to a car. Those who are the most aware about transit issues mainly choose to buy a house in the old suburbs, an area where the public transit system is in fact the most efficient. The pattern is similar for *school* proximity, mainly a concern for parents (4.33), and 40 to 49-year old respondents who settle in the Upper Town (2.14). The first spatial division, based on centrality, explains more differences in the frequency of *accessibility to services* and *job* criteria. The aesthetic dimension of the *presence of trees*, or the question of *school* quality is merely associated with the second spatial division based on life cycle. The socio-economic group of variables, including *security*, *quiet*, *lively* or *young neighbourhoods*, is significantly related to relative centrality. While *quietness* is four times less frequently mentioned by people who move into the city centre (0.23), the desire of a *young neighbourhood* is more frequent among families (2.66), as is the desire of a *lively neighbourhood*, particularly for 30- to 39-year-old parents (3.97). However, dual-worker couples of this age are less prone to want a *lively neighbourhood* (0.228), but the relationship is inversed for dual-worker households with a medium rather than a high combined income (3.92).

*Cachet* is positively linked to income, each additional $10 000 of income multiplying the likelihood of mentioning this aesthetic criterion by 1.15. The presence of *trees* in the neighbourhood is of particular interest to people who settle in the high-income Upper Town area (2.62), but half less frequently mentioned by parents who choose the old suburbs area (0.54). Single-person first-time-owner households are also less sensitive to the presence of vegetation (0.24). Furthermore, people who feel *attached* to the neighbourhood are nearly twice less influenced by aesthetic criteria (Figure 7).

Figure 7. Links between location. household profile and aesthetic criteria for neighbourhood choice



Numbers in arrows are odd-ratios given by logistic regression

## 2.6 Property Choice Criteria

Although price emerges as the most frequent property-choice criterion respondents give, this information does not give much insights into a better understanding of the very criteria underlying the residential choice strategies. As can be seen, the likelihood of mentioning the price criterion is inversely related to income with a factor of 0.84 for each additional 10 000 $ of income (Table 15). Furthermore, new owners are more likely to mention the price argument (odd ratio of 1.54), as are people who settle in the old suburbs, rather than in the new suburbs (1.40). More interestingly, too, the group of variables relating to size – *property size*, *number of rooms*, and *lot size* taken together – is largely associated with education, university-degree holders being 1.7 times more likely to consider this group of variables than people holding a college degree or high-school diploma. Naturally, couples are twice as much concerned with size as single-person households, as are former owners compared with first-time owners (1.59). Examining each criterion in detail, it appears that the likelihood of giving the criterion *size of the property* is positively linked to income, with a factor of 1.16 for each additional ten thousand dollars of income. However, childless couples with the same income accord less importance to size, as shows the combination of both characteristics, i.e. $1.16^n$*0.884. *Lot size* is more frequently mentioned by people who settle in the Upper Town compared with those who settle in the more remote suburbs (1.55), but is less of a concern for single-person households (0.57). The *number of rooms* is less important for either first-time owners (0.59), or single-person households (0.297), but is more likely to influence families in the old suburbs (2.01), people who settle in Upper Town (2.38), and middle-income households that are attached to the neighbourhood (3.81). The importance of interior attributes (*interior group*), meaning the *interior architecture* and *decoration*, *floor quality*, *functionality*, or the presence of a *garage*, is mainly associated with age and schooling,– people over 49 whether holding a college degree or attached to the neighbourhood feeling more concerned (odds ratios respectively 13.05 and 6.47) than young households with university degrees. Furthermore, households with children are 1.56 times more likely to refer to interior attributes criteria than households without children.

Table 15. Logistic regression models for property choice criteria

| Household Attribute | Prop. Choice criteria | Price | Size Group | Size | Nb Rooms | Lot Size | Interior Group | Style | Trees |
|---|---|---|---|---|---|---|---|---|---|
| **Age (vs under 30)** | 40-49 | | | | | | ¤Income 1.08** | | |
| | 50 and plus | | | | | | ¤CollDegree 13.05** ¤Attachment 6.47** | ¤WithChild 7.32*** | |
| **Income (continuous)** | Income (/10 000$) | 0.84*** | | 1.16*** ¤CpleNoChild 0.884** ¤SglePersHld 1.33** | | | ¤Age40-49 1.08** | 1.21*** | |
| **Income (vs >80K)** | <50K | | | | | | | ¤DualWorker 2.38** | |
| | 50K-80K | | | | ¤Attached 3.81*** | | | | ¤Centre 16.44*** |
| **Dual worker** | Dual-worker | | | | | | | ¤Income<50K 2.38** | |
| **Education (vs Univ.Degree)** | College degr. | | 0.579*** | | | | ¤Age>49 13.05*** | | ¤Attached 0.220** |
| | Secondary | | 0.579** | | | | | | |
| **Household type (vs couple with child)** | Couple without child | | | ¤Income 0.884** | | | | | |
| | Single –Pers. Household | | | ¤UpperTown 0.074** ¤Income 1.33** | | | | | |
| **Couple vs Single Adult** | Single Adult (with or without child) | | 0.485*** | | 0.297*** | 0.57** | | | |
| **Child at home** | With child | | | | ¤OldSuburbs 2.01*** | | 1.56** | ¤Age>49 7.32*** | |
| **Previous ownership** | First-time owner | 1.54*** | 0.628*** | | 0.589** | | | | |
| **Location/CSF1 (vs new suburbs)** | Centre | | | | | | | | ¤Inc.50-80K 16.44*** |
| | Old suburbs | 1.40** | | | ¤WithChild 2.01*** | | | | |
| **Location/CSF2 (vs mixed suburbs)** | Upper town High income | | | ¤SgleHld 0.074** | 2.38*** | 1.55*** | | | 2.69*** |
| **Attachment** | Attached | | | | ¤Inc.50-80K 3.81*** | | ¤Age>49 6.47** | 1.96*** | ¤CollDegree 0.220** |
| **Model performance** | Chi-square | 42.1 | 48.8 | 48.1 | 74.6 | 16.4 | 40.6 | 46.7 | 36.2 |
| | Sig | 0.000 | 0.000 | 48.140 | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 |
| | Ending -2LL | 995.0 | 1022.6 | 665.2 | 606.5 | 913.1 | 983.3 | 835.7 | 574.6 |
| | Nagelkerke $R^2$ | 0.072 | 0.081 | 0.100 | 0.157 | 0.030 | 0.070 | 0.086 | 0.084 |

**Values in cells are odds ratios; interactions identified by ¤NameofVariable; *** sig<0.01; ** sig<0.05.**

The importance of the *style* of the property is positively linked to the household income, each $10 000 increasing the likelihood by 1.21 time. *Style* is also more frequently mentioned by aged parents (2.02), low-income dual-worker households (2.38) and people who feel attached to the neighbourhood (1.96). The frequency of the other exterior attribute, such as the presence of *trees*, is more important for people who settle in the Upper Town, but also for middle-income households that settle in the city centre. College-degree holders who are attached to the neighbourhood put less emphasis on the presence of trees on the property (0.22), as opposed to university-degree holders who, while less attached to the neighbourhood, will consider that factor as important.

## 2.7 Conclusions

This study, conducted in Quebec City, explores both the motivation for moving and property choice criteria of actual single-family property buyers. Using logistic regression, the likelihood of considering a criterion depending on the household profile, the psychological dimension of attachment to the neighbourhood, and the final location choice, is measured. Detailed studies held with household-level data are useful in order to better understand needs and aspirations in terms of housing strategies. Since the end of the baby-boom period, western cities have been through major societal changes which have had a strong impact on land use and residential behaviour. In North-America, the strong growth rate following World War II induced an increasing demand from young families for new low-density single-family housing, thereby generating an important decline of city-centre densities and causing growing urban sprawl. Concomitantly, the shifting from industrial to post-industrial service-oriented economies and the accession of women to the workplace prompted new and increasing mobility needs. Furthermore, massive investments during the '60s and '70s in road networks helped to shape the evolution of the cities' land use. Recent trends in family structures – an increase in single-parent families and reconstituted households, for example – as well as the accession of a large baby-boomer cohort to retirement has also led to specific residential needs and behaviours. In this context, understanding the motivation for moving and choice criteria is relevant for urban research and can enhance future planning policies.

The correspondence analysis performed in this study sorted out the various factors of residential choice. Each of the eight extracted factors could be linked to the theoretical dimensions of the Place-Proximity-Space and the Place-Identity models, underlining certain previously identified aspects of residential choice strategies, such as the distance-size trade-off, the environmental-quality location trade-off (Kestens *et al.*, At Press), but also the importance of Place attributes, such as property quality, trees or landscaping features (Des Rosiers *et al.*, 2002).

The logistic regressions could sort out the numerous links between the different aspects of household profile and the multidimensionality of residential behaviour. The latter encompasses both motivation for moving and choice. This paper corroborates Rossi's pioneering findings (1955); see the links between life cycle and motivation for moving, in accordance with other recent studies (See Clark and Dieleman, 1996 for a review; Dieleman and Mulder, 2002). However, most of these papers focus on explaining the *propensity* to move, whereas this paper analyses the *stated reasons* underlying actual moving decisions. More as yet unconsidered household-related dimensions also have important implications on motivations for moving, such as the type of previous tenure – i.e., tenancy as opposed to ownership. However, as the previous ownership status is intricately associated with age and type of household, all these dimensions of life have to be considered concomitantly. In line with housing economics, the importance of previous ownership status is particularly significant with regard to desire for improved housing and the will to choose a better location in terms of proximity, two features to which former owners appear to be more sensitive. Interestingly, the educational attainment could not, at first, explain any differences in the motivation to move. However, schooling appeared positively significant for explaining the likelihood of mentioning *new job* as a criterion once income had been controlled for. Such a finding confirms recent results obtained in a nation-wide study held in the United States, which showed that highly educated people are more likely to move for employment-related reasons (U.S. Census Bureau, 2001). These findings can be explained by the fact that not only are high-income or highly-educated households more motivated to move for more interesting jobs, but they also have a greater ability to pay for newer or better housing. It would be highly insightful to analyse data on the successive temporal and spatial moving trajectories in the light of job and family histories of individuals. This would enable us to better understand the intricate implications of wealth, education, job and family career on residential mobility. Also, considering the intra- or inter-urban dimension of the move would perhaps highlight differences in residential choice criteria.

The neighbourhood choice criteria are primarily linked to type of household, age and income. Here too, schooling does not seem to have any impact, and does not appear significant provided income is included in the models. Security issues – although not very frequently mentioned overall – are a lesser preoccupation for low-income households. This is a surprising result, as low-income areas are positively associated with higher crime rates in the city, *a priori* leading to the conclusion that low-income households are more exposed and therefore more sensitive to security issues. Findings suggest, however, that households with a higher income are more sensitive to security issues, even if they are in fact less exposed. However, as the sample studied is only comprised of single-family properties, which are scarce in the very central high-density area, any interpretation regarding security issues must be considered as approximate. Proximity to the public transit system is mainly a concern for parents in their forties, suggesting that they are sensitive to the urban accessibility of their teenagers who, for the most part, do not have a driver's license or have only limited access to a car.

Both spatial divisions, based on centrality and on socio-economic status, appeared to be highly linked to the neighbourhood and property choice criteria. The spatial division based on centrality is mainly linked to *accessibility to services* and *jobs*, but *proximity to schools* and the aesthetic dimension of the *presence of trees* are simply associated with the second division based on income and life cycle. Furthermore, and keeping in mind that only few single-family properties of the sample are found in the very central high-density area, it appears the households which are the most sensitive to accessibility do actually locate in the old suburbs and not in the city centre. This perception of accessibility corroborates recent accessibility measures underlining the higher accessibility in the old-suburbs, due to the major highway infrastructures developed in this area during the '60s and '70s (Thériault *et al.*, 1999). As to property choice, *lot size* is more frequently cited by people who settle in the more central Upper Town than by those who settle in the more remote suburbs, in contradiction with the intuitive distance-size trade off assumption. However, it appears that the actual lot sizes of the observed sample are quite similar in both areas (around 600 square metres), in contradiction with the generally accepted assumption that more distant properties benefit from larger lots. In fact, the correlation between distance to city centre and lot size is marginal, with a Pearson correlation of only 0.123, but significant at the 5% level. The explanation may lie with the fact that small lots in remote suburbs allow for even lower house prices, which is what remote-location households are looking for, while higher income household living in the Upper Town are ready to spend more on increasingly scarce land. Finally, this raises the bias issue related to surveys. Although the properties close to the city centre have lot sizes similar to those located in newly built remote suburbs, buyers who settle near the downtown area seem to be more sensitive to the lot size and are more prone to mention this criterion. Similarly, high-income households that are looking for properties in well maintained areas with abundant vegetation do not even think about mentioning that they are looking for neighbourhoods with trees, as this assumption is implicit in their choice behaviour. The bias linked to the various levels of awareness of what people really want, or the gap between what people think they have considered and actual choices they have in fact made represent a few limitations of this type of household-level survey. Further research focusing on the link between stated choice criteria and actual characteristics of property and neighbourhood attributes would give us additional material to further this debate. Similarly, a systematic comparison of findings derived from both stated-choice and hedonic methodologies may help in sorting out such limitations and biases.

Among the major findings of this study, the strong links between, on the one hand, the psychological dimension of attachment and, on the other, the three aspects of residential behaviour considered here: the motivation for moving and property and neighbourhood choice, are worth emphasizing. Attachment to the neighbourhood is the third household-related reason – next to the type of household and income bracket – explaining the moving or choice criteria. The previous ownership status also has a strong influence on the motivation for moving, although neighbourhood and property choice criteria of first-time and former owners are quite similar. Whereas first-time owners seem to focus mainly on access to ownership, experienced owners are much more concerned with size and location (proximity to services and workplace). These two dimensions – the psychological dimension of attachment and previous tenure type – relate to the buyer's past experiences. Therefore, additional research at an individual level is needed to better understand the temporal succession and intertwine of events underlying residential behaviour. Also, it would be of great interest to check whether the differences in the values people assign to property-specific or neighbourhood attributes can be properly measured through revealed choice methods such as hedonic modelling.

# References

Adamowicz V, Boxall P, 2002, "Future Directions of Stated Choice Methods for Environment Valuation", in *Choice Experiments: A New Approach to Environmental Evaluation* Eds M Christie, P Midmore (Edward Elgar, Cheltenham) p. at press

Altman I, Low S M, Eds, 1992 *Place Attachment.* (Plenum Press, New York), Vol. 12. (Human Behavior and Environment), 314 p.

Ambert A-M, 1998 *Divorce: Facts, Figures and Consequences* York University, Vanier Institute of the Family, York

Artle R, Varaiya P, 1978, "Life Cycle Consumption and Homeownership" *Journal of Economic Theory* **18** 38-58

Bender A, Din A, Favarger P, Hoesli M, Laakso J, 1997, "An Analysis of Perceptions Concerning the Environmental Quality of Housing in Geneva" *Urban Studies* **34** 503-513

Bender A R, Din A, Hoesli M, Brocher S, 2000, "Environmental Preferences of Homeowners: Further Evidence Using the AHP Method" *Journal of Property Investment & Finance* **18** 445-455

Benzecri J P, 1973 *L'analyse des données, Tome 2: L'analyse des correspondances* (Dunod, Paris)

Bettman J R, 1979 *An Information Processing Theory of Consumer Choice* (Addison-Wesley, Reading)

Bonaiuto M, Aiello A, Perugini M, Bonnes M, Ercolani A P, 1999, "Multidimensional Perception of Residential Environment Quality and Neighbourhood Attachment in the Urban Environment" *Journal of Environmental Psychology* **19** 331-352

Boxall P C, Adamowicz W L, 1999 *Understanding Heterogeneous Preferences in Random Utility Models: The Use of Latent Class Analysis* University of Alberta, Department of Rural Economy, Edmonton, Alberta

Breakwell G M, 1986 *Coping with Threatened Identity* (Methuen, London)

Breakwell G M, 1992, "Processes of Self-evaluation: Efficacy and Estrangement", in *Social Psychology of Identity and the Self-concept* Ed. G M Breakwell (Surrey University Press, Surrey)

Casetti E, 1972, "Generating Models by the Expansion Method: Applications to Geographical Research" *Geographical Analysis* **4** 81-91

Casetti E, 1997, "The Expansion Method, Mathematical Modelling, and Spatial Econometrics" *International Regional Science Review* **20** 9-32

Clark W A V, Dieleman F M, 1996 *Households and Housing: Choice and Outcomes in the Housing Market* (Center for Urban Research Policy, New Brunswick, NJ)

Clark W A V, Mulder C H, 2000, "Leaving Home and Entering the Housing Market" *Environment and Planning A* **32** 1657-1671

Clark W A V, Onaka J L, 1983, "Life Cycle and Housing Adjustment as Explanations of Residential Mobility" *Urban Studies* **20** 47-57

Coolen H, Van Montfort X, 2001, *Meaning-based Representations of Preferences for Housing Attributes*. *Housing Imaginations: New Concepts, New Theories, New Researchers*. Cardiff,

Cox D R, Snell E J, 1989 *The Analysis of Binary Data* (Chapman and Hall, London)

Cummings R G, Brookshire D S, Schulze W D, 1986 *Valuing Environmental Goods: A State of the Arts Assessment of the Contingent Method* (Roman and Allanheld, Totowa, NJ)

Deng Y, Ross S L, Wachter S M, 2003, "Racial Differences in Homeownership: The Effect of Residential Location" *Regional Science and Urban Economics* **Article at Press**

Des Rosiers F, Thériault M, Kestens Y, Villeneuve P, 2002, "Landscaping and House Values: An Empirical Investigation" *Journal of Real Estate Research* **23** 139-161

Des Rosiers F, Thériault M, Villeneuve P-Y, 2000, "Sorting Out Access and Neighbourhood Factors in Hedonic Price Modelling" *Journal of Property Valuation and Investment* **18** 291-315

Dieleman F M, 2001, "Modelling Residential Mobility: A Review of Recent Trends in Research" *Journal of Housing and the Built Environment* **16** 249-265

Dieleman F M, Mulder C H, 2002, "The Geography of Residential Choice", in *Residential Environments: Choice, Satisfaction, and Behavior* Eds J Aragonés, G Francescato, T Gärling (Bergin and Garvey, Westport, Connecticut) p. 265

Duncan J, Ley D, 1993 *Culture/Place/Representation* (Routledge, New York)

Earnhart D, 1998 *Combining Revealed and Stated Data to Examine Decisions of Housing Location: Discrete-Choice Hedonic and Conjoint Analysis* University of Kansas, Lawrence, KS

Feldman R M, 1990, "Settlement-Identity: Psychological Bonds with Home Places in a Mobile Society" *Environment and Behavior* **22** 183-229

Feldman R M, 1996, "Constancy and Change in Attachments to Types of Settlements" *Environment and Behavior* **28** 419-445

Filion P, Bunting T, Warriner K, 1999, "The Entrenchment of Urban Dispersion: Residential Preferences and Location Patterns in the Dispersed City" *Urban Studies* **36** 1317-1347

Forster E, 2001 *Migration and Housing Choice Survey* Napier University, Edinburgh

Fried M, 1982, "Residential Attachment: Sources of Residential and Community Satisfaction" *Journal of Social Issues* **38** 107-119

Friedman J, 1981, "A Conditional Logit Model of the Role of Local Public Services in Residential Choice" *Urban Studies* **18** 347-358

Gabriel S, Painter G, 2003, "Pathways to Homeownership: An Analysis of the Residential Location and Homeownership Choices of Black Households in Los Angeles" *Journal of Real Estate Finance and Economics* **27** 87-109

Gärling T, Golledge R G, 1993, "Understanding Behavior and Environment: A Joint Challenge to Psychology and Geography", in *Behavior and Environment: Psychological and Geographical Approaches* Eds T Gärling, R G Golledge (Elsevier Science Publishers B. V., Amsterdam) pp 2-13

Gentleman J F, Park E, 1997, "Divorce in the 1990s" *Rapports sur la santé* **9** 57-62

Gibson J J, 1950 *The Perception of the Visual World* (Houghton Mifflin, Boston)

Giuliani M V, 1991, "Towards an Analysis of Mental Representations of Attachment to the Home" *Journal of Architectural and Planning Research* **8** 133-146

Giuliani M V, Feldman R M, 1993, "Place Attachment in a Developmental and Cultural Context" *Journal of Environmental Psychology* **13** 267-274

Goodman A C, 1989, "Identifying Willingness-to-Pay for Heterogeneous Goods with Factorial Survey Methods" *Journal of Environmental Economics and Management* **16** 58-79

Greenacre M J, 1984 *Theory and Applications of Correspondence Analysis* (Academic Press, London)

Greenacre M J, 1993 *Correspondence Analysis in Practice* (Academic Press, London)

Hauser J R, Rao V R, forthcoming, "Conjoint Analysis, Related Modelling, and Applications", in *Advances in Marketing Research: Progress and Prospects* Ed. J Wind

Henderson J V, Ioannides Y M, 1983, "A Model of Housing Tenure Choice" *American Economic Review* **73** 98-113

Hill M O, 1974, "Correspondence Analysis: A Neglected Multivariate Method" *Applied Statistics* **23** 340-354

Kestens Y, Thériault M, Des Rosiers F, At Press, "The Impact of Land Use and Surrounding Vegetation on Single-Family Property Prices" *Environment and Planning B: Planning and Design* **At Press**

Legendre P, Legendre L, 1998 *Numerical Ecology* (Elsevier Science, Amsterdam)

Longley P A, 1984, "Comparing Discrete Choice Models: Some Housing Market Examples", in *Discrete Choice Models in Regional Science* Ed. D E Pitfield (Pion Limited, London) pp 163-180

Louvière J J, 1984, "Hierarchical Information Integration: A New Method for the Design and Analysis of Complex Multiattribute Judgment Problems" *Advances in Consumer Research* **11** 148-155

Louvière J J, Timmermans H, 1990, "Hierarchical Integration Theory Applied to Residential Choice Behavior" *Geographical Analysis* **22** 127-144

Low S, Altman I, 1992, "Place Attachment: A Conceptual Inquiry", in *Place Attachment* Eds I Altman, S Low (Plenum Press, New York) pp 1-12

McFadden D, 1974, "Conditional Logit Analysis of Qualitative Choice Behavior", in *Frontiers in Econometrics* Ed. P Zarembka (Academic Press, New York) pp 105-142

McFadden D, 1978, "Modelling the Choice of Residential Location", in *Spatial Interaction Theory and Planning Models* Eds A Karlquist, L Lundquist, F Snickars, J Weibul (Elsevier North-Holland, New York) pp 75-96

Megbolugbe I, Sa-Aadu J, Shilling J D, 1999, "Elderly Female-Headed Households and the Decision to Trade Down" *Journal of Housing Economics* **8** 285-300

Molin E J E, Oppewal H, Timmermans H J P, 2001, "Analysing Heterogeneity in Conjoint Estimates of Residential Preferences" *Journal of Housing and the Built Environment* **16** 267-284

Molin E J E, Timmermans H J P, 2003, "Testing Hierarchical Information Integration Theory: The Causal Structure of Household Residential Satisfaction" *Environment and Planning A* **35** 43-58

Mulder C H, Hooimeijer P, 1999, "Residential Relocations in the Life Course", in *Population Issues, an Interdisciplinary Focus* Eds L J G Van Wissen, P A Dykstra (Kluwer Academic / Plenum Publishers, New York)

Nagelkerke N J D, 1991, "A Note on a General Definition of the Coefficient of Determination" *Biometrika* **78** 691-692

Nechyba T, Strauss R, 1998, "Community Choice and Local Public Services: A Discrete Choice Approach" *Regional Science and Urban Economics* **28** 51-73

Onaka J L, 1983, "A Multiple-attribute Housing Desequilibrium Model of Residential Mobility" *Environment and Planning A* **15** 751-765

Oppewal H, Louvière J J, Timmermans H, 1994, "Modelling Hierarchical Conjoint Processes with Integrated Choice Experiments" *Journal of Marketing Research* **31** 92-105

Pellegrini P A, Fotheringham A S, 2002, "Modelling Spatial Choice: A Review and Synthesis in a Migration Context" *Progress in Human Geography* **26** 487-510

Proshansky H, 1978, "The City and Self-Identity" *Environment and Behavior* **10** 147-170

Proshansky H, Fabian A K, Kaminoff R, 1983, "Place-Identity: Physical World Socialization of the Self" *Journal of Environmental Psychology* **3** 57-83

Quigley J, 1976, "Housing Demand in the Short Run: An Analysis of Polytomous Choice" *Explorations in Economic Research* **3** 76-102

Quigley J, 1985, "Consumer Choice of Dwelling, Neighborhood, and Public Services" *Regional Science and Urban Economics* **15** 41-63

Quigley J, Weinberg D, 1977, "Intraurban Residential Mobility: A Review and Synthesis" *International Regional Science Review* **1** 41-66

Reginster I, Edwards G, 2001, "The Concept and Implementation of Perceptual Regions as Hierarchical Spatial Units for Evaluating Environmental Sensitivity" *URISA Journal* **13** 5-16

Rosen S, 1974, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" *Journal of Political Economy* **82** 34-55

Rossi P H, 1955 *Why Families Move: A Study in the Social Psychology of Urban Residential Mobility* (Free Press, Glencoe, Illinois)

Rubinstein R, Pamelee P A, 1992, "Attachment to Place and the Representation of the Life Course by the Elderly", in *Place Attachment* Eds I Altman, S Low (Plenum, New York) pp 139-164

Statistics Canada, 2001a, *Crime rates*. Statistics Canada, 2003,

Statistics Canada, 2001b, *Divorce rates*. 2003,

Sundstrom E, Bell P A, Busby P L, Asmus C, 1996, "Environmental Psychology 1989-1994" *Annual Review of Psychology* **47** 485-512

Thériault M, Des Rosiers F, Kestens Y, Villeneuve P, 2003, "Modelling Interactions of Location with Specific Value of Housing Attributes" *Property Management* **21** 25-62

Thériault M, Des Rosiers F, Vandersmissen M H, 1999, *GIS-Based Simulation of Accessibility to Enhance Hedonic Modelling and Property Value Appraisal*. Joint IAAO-URISA 1999 Conference, New Orleans, 151-160

Thurstone L L, 1927, "A Law of Comparative Judgment" *Psychological Review* **34** 273-286

Timmermans H, Borgers A, Van Dijk J, Oppewal H, 1992, "Residential Choice Behaviour of Dual Earner Households: A Decompositional Joint Choice Model" *Environment and Planning A* **24** 517-533

Timmermans H, Van Noortwijk L, 1995, "Context Dependencies in Housing Choice Behavior" *Environment and Planning A* **27** 181-192

Timmermans H, Van Noortwijk L, Oppewal H, Van der Waerden P, 1996, "Modelling Constrained Choice Behaviour in Regulated Housing Markets by Means of Discrete Choice Experiments and Universal Logit Models: An Application to the Residential Choice Behaviour of Divorcees" *Environment and Planning A* **28** 1095-1112

Twigger-Ross C L, Uzzell D L, 1996, "Place and Identity Processes" *Journal of Environmental Psychology* **16** 205-220

U.S. Census Bureau, 2001 *Why People Move: Exploring the March 2000 Current Population Survey* U.S. Department of Commerce, Economic and Statistics Administration,

Van De Vyvere Y, Oppewal H, Timmermans H, 1998, "The Validity of Hierarchical Information Integration Choice Experiments to Model Residential Preference and Choice" *Geographical Analysis* **30** 254-272

Van Ommeren J, Rietveld P, Nijkamp P, 1999, "Job Moving, Residential Moving and Commuting: A Search Perspective" *Journal of Urban Economics* **46** 230-253

William R J, 1979, "A Logit Model of Demand for Neighborhood", in *A Logit Model of Demand for Neighborhood* Ed. D Segal (Academic Press, New York) pp 17-42

# Chapter 3. Heterogeneity in Hedonic Modelling of House Values: What Can Be Explained by Household Profiles?

**Résumé: Ce chapitre aborde la question de l'hétérogénéité des valeurs marginales en introduisant des données désagrégées à l'échelle des ménages acheteurs de propriétés dans le contexte de l'analyse par modélisation hédonique. Le recours à deux méthodes d'analyse spatiale permet de mesurer la variation des valeurs implicites en fonction du type de ménage, de l'âge, du niveau d'éducation et le profil antérieur en terme de propriété. Les méthodes de l'expansion spatiale et des Geographically Weighted Regressions sont appliquées aux mêmes échantillons. Les deux approches donnent des résultats concluants, et montrent que la valeur implicite de plusieurs variables de propriété et de localisation varie en fonction du profil du ménage acheteur. Ainsi, il a notamment pu être démontré que le revenu de l'acheteur avait un impact sur le prix que celui-ci paye pour une propriété, et que les ménages possédant un diplôme universitaire paient un supplément pour résider dans des quartiers au niveau d'éducation élevé.**

**Abstract : This chapter introduces household-level data into hedonic models in order to measure the heterogeneity of implicit prices regarding household type, age, educational attainment, income, and the previous tenure status. Two methods are used for this purpose: a first series of models uses expansion terms, whereas a second series applies Geographically Weighted Regressions. Both methods yield conclusive results, showing that the marginal value given to certain property specifics and location attributes do vary regarding the characteristics of the buyer's household. Particularly, major findings concern the significant effect of income on the location rent as well as the premium paid by highly-educated households in order to fulfill social homogeneity.**

## 3.1 Introduction

The analysis of house values using hedonic modelling makes it possible to estimate the marginal monetary contribution of property attributes and neighbourhood externalities (Rosen, 1974). In most hedonic models, one unique coefficient is derived for each observed attribute. It is entirely possible that this coefficient may vary according to some systematic pattern. Various methods have been designed to handle such variation (Anselin, 1988; Brunsdon *et al.*, 1996; Casetti, 1972; Fotheringham *et al.*, 2002; Griffith, 1988). Explicitly integrating heterogeneity – which may be spatial – should improve the calibration of the models while enhancing the understanding of the residential market structure.

This paper presents an empirical case study analysing the spatial and social structure of residential property markets by combining single-family property sales and household-level socio-economic data. Through the use of two context-sensitive hedonic methods – the Casetti expansion method (Casetti, 1972 & 1997) and Geographically Weighted Regression (GWR) (Brunsdon *et al.*, 1996; Fotheringham *et al.*, 2002) – and trough the incorporation of the socio-economic profile of actual property buyers, we have attempted to validate the following hypothesis: the variability of the implicit prices of certain property and location attributes is partly linked to preference. In a preceding paper, Kestens *et al*. (Submitted) showed that the residential choice criteria – both as regards property and neighbourhood – vary significantly with the household profile, that is, with the type of household, age, income, educational attainment, the type of previous tenure (first-time owner vs. former owner), and even with the sense of belonging to the neighbourhood.

In order to investigate these questions, this paper analyses the variation of the impact of property-specifics and neighbourhood attributes considering household socio-economic profiles using hedonic modelling. Thereby, we hope to contribute to Starret's (1981) debate on homogeneity of preferences and capitalisation. As pointed out by Tyrvainen (1997), according to Starret, the capitalisation of an attribute is complete "if: (i) there is enough variation within the variable" – e.g. in order to measure the effect of proximity to power lines, it is important to account for cases where people live at such distance to prevent an effect on house prices – and "if (ii) the residents' preferences are homogeneous. If the preferences are heterogeneous, capitalisation is only partial" (Tyrvainen, 1997, p. 220). Whereas the first condition can easily be controlled, the second has been the object of little research. Thus, we hypothesise that the capitalisation is partial in that the value given to an attribute differs with household preferences. While such an assumption may seem to challenge the traditional interpretation of an hedonic function and to question the identification problem addressed by Rosen (1974), it is supported by empirical evidence about the existence of sub-markets and the heterogeneity of hedonic prices over space (Goodman and Thibodeau, 2003). We therefore feel that in order for hedonic modelling to adequately measure the capitalisation of an attribute, residents' preferences for this attribute have to be homogeneous, or they have to vary in a systematic way. In other words, we argue that part of the non-stationarity of the value of property and location attributes is linked to differences among the buyer's household profiles. Furthermore, we argue that when data is available at the household level, appropriate drift-sensitive regression techniques can be used to validate this hypothesis. Of course, the object of this research is mainly to analyse the processes underlying the market dynamics. The methods presented here should therefore not be considered a valuation tool but merely a way to better understand urban dynamics. Furthermore, it must be kept in mind that this paper's results are specific to the area of study and to the socio-economic conditions of the market for the observed period. In the property market of Quebec City, most of the 1993-2001 period is characterised by high vacancy rates and the abundance of sellers. Advantages in the negotiation process are therefore largely given to the buyers, which can be the cause for some of our findings.

Two sets of hedonic models are built using some 761 single-property values sold in Quebec City between 1993 and 2001. The first set uses Casetti-type interactive terms, while the second relies on Geographically Weighted Regression (GWR). Special attention is given to Local Spatial Autocorrelation (LSA) (Anselin, 1995), as it is expected that the introduction of disaggregated household-level data reduces the number of local spatial autocorrelation "hot spots". Section 0 discusses the hedonic modelling technique, the spatial dimensions of property markets, and presents Casetti's expansion method and GWR. Section 0 presents the data bank and the modelling procedure, whereas the results are given in Section 0. Finally, a summary of the main findings and further research possibilities are presented in Section 0.

# 3.2 Literature

## 3.2.1 Hedonic Modelling

The hedonic framework relies on Lancaster's consumer theory, stating that utility is derived from the properties or characteristics of a good (Lancaster, 1966). Since this theory has been extended to the residential market by Rosen (1974), residential hedonic analysis has become widely used as an assessment tool and for property market and urban analysis. The regression of house values on a variety of property specific and neighbourhood descriptors evaluates their marginal contribution, also called implicit or hedonic prices. In their basic form, hedonic regressions assume each parameter to be fixed in space, which means that each identified attribute has the same intrinsic contribution throughout the submarket under study:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon \tag{1}$$

where **y** is a vector of selling prices, **X** a matrix of explanatory variables, **β** a vector of regression coefficients, and ε the error term.

However, property markets are very much tied as well as inherent in the spatial structure of the urban landscape. In fact, although capital is mobile, supply may be quite inelastic (Goodman and Thibodeau, 1998), and a property, once constructed, becomes immovable, or spatially "rooted". As a result, the value of a property is largely defined by its location attributes, that is, by its relative location compared with urban infrastructure and services. Furthermore, as pointed out by Goodman and Thibodeau (1998), inelasticities in both supply and demand contribute to market segmentation. As a previous paper has shown, the choice criteria concerning both location and property choice vary depending on the household profile (Kestens *et al.*, Submitted). This market segmentation may lead to heterogeneous implicit prices, which should be explicitly considered in the residential hedonic price function. In fact, the implicit prices of the hedonic function reflect both supply- and demand-driven forces. In an equilibrium situation, it is assumed that these forces cannot be distinguished within a hedonic function. However, we believe that when the market conditions are not in equilibrium, but instead those of a seller market (much supply for low demand), it becomes possible for the buyers to influence the price they pay for an amenity. If the conditions were reversed, that is, if it were a buyer market (much demand for low supply), the sellers would have more power to impact upon the selling price, and the seller's characteristics could then be significantly linked to the drift of the implicit prices. Therefore, we assume that the introduction of household-level variables within the hedonic function using appropriate methods like the Casetti expansions may make it possible to estimate the drift in the coefficients associated with certain characteristics of the buyers.

### 3.2.2 Spatial Dimensions of Property Markets

Can (1992) distinguishes two types of spatial effects: neighbourhood effects and adjacency effects. The former refers to internalised values of geographical features (exogenous effects), while the latter refers to spatial spill-over effects; that is, the impact of the characteristics of close surrounding properties (endogenous effects). Exogenous effects can be manifold, ranging from city-wide structural factors (e.g. location rent) to local externalities (e.g. view on a high-voltage tower). These geographical features induce trends into housing expenditures that have to be explicitly incorporated into the hedonic function, if they are not removed before modelling.

Classical hedonic modelling would estimate 'fixed' coefficients, however, above-mentioned market segmentation may lead to spatial heterogeneity, that is, to possible 'drifts' in the estimated coefficients.

Independently from this contextual variation of the impact of housing attributes, similarity of prices between close properties may also be partly linked to spatial spill-over (endogenous effect). Spatial spill-over occurs when characteristics of surrounding or adjacent properties are internalised in the property value, leading to spatial dependence or association. This spatial dependence cannot be modeled adequately using additional descriptive geographical variables, and necessitates the introduction of spatial autoregressive (SAR) terms into the hedonic function:

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\rho}\mathbf{Wy} + \varepsilon \tag{2}$$

$$\mathbf{y} = \mathbf{X\beta} + \boldsymbol{\alpha}\mathbf{W(y - X\beta)} + \varepsilon \tag{3}$$

where ***X*** is the matrix of explanatory variables, ε the error term, ***Wy*** a spatially lagged dependent variable, with ***W*** as the weight matrix, **ρ** and **α** the spatial autoregressive parameters, that is, **ρ** the degree to which the values at individual locations depend on their neighbouring values, and **α** the degree to which the values at individual locations depend on their neighbours' residuals (Fotheringham, 2002 p. 23).

SAR terms may take several forms. Most often, however, they are weighted lagged values of the dependent variable (equation 2) or of the error term (equation 3) (Anselin, 1988; Griffith, 1988; Kelejian, 1995). Ordinary Least Squares (OLS) is not appropriate for SAR procedures that necessitate Generalised Least Squares (GLS) or Maximum Likelihood (ML) estimations. However, OLS regression presents several advantages: it "has a well-developed theory, and has available a battery of diagnostic statistics that make interpretations easy and straightforward" (Getis and Griffith, 2002, p. 131). Spatially dependent variables can also be transformed prior to modelling in their spatial and non spatial components, using spatial filtering techniques (Cliff and Ord, 1981; Getis, 1995; Getis and Griffith, 2002; Griffith, 1996). Of course, combinations of these methods can be used. For example, a model integrating geographical features accounting for the spatial drift may also include an autoregressive term controlling for spatial dependence. However, "a two step procedure is considered to be more suitable" (Can, 1990). That means that SAR terms should only be included if spatial dependence is still present after spatial heterogeneity has been fully considered.

### 3.2.3 Methods and Previous Results

In this paper, we use two methods accounting for the spatial heterogeneity of the parameters, namely, the spatial expansion method developed by Casetti (Casetti, 1972 & 1997) and Geographical Weighted Regression (Brunsdon *et al.*, 1996; Fotheringham *et al.*, 2002). Furthermore, we observe how the introduction of detailed household-profile data helps to explain spatial heterogeneity while diminishing spatial dependence.

The spatial expansion method developed by Casetti has first been used to analyse the spatial drift inherent to various geographical phenomena like migration (Casetti, 1986), labor markets (Pandit and Casetti, 1989) or price analyses before being applied to property market and price analysis (Aten and Heston, Forthcoming; Can, 1990 & 1992; Casetti, 1997). The parameter drift refers to the variation of the parameter value depending on the context. In fact, this method "extends" fixed parameters by introducing interactive variables that combine a previously defined (fixed) characteristic with a (spatially) dependent variable relating to the (spatial) context:

$$\mathbf{y} = (\mathbf{C}^{\mathbf{t}}(\mathbf{E}+\mathbf{I})\mathbf{X})\boldsymbol{\beta} + \varepsilon \tag{4}$$

with **C**, a matrix of contextual variables which can be manifold (including a vector of 1 values in the first column), **E** a matrix of expansion indicating which explanatory variables are expanded by the contextual variables, **I**, the identity matrix, and **X**, a matrix of explanatory variables, each one being activated in **E**.

In most models' specifications, the estimation of varying parameters is limited to structural factors and the "contextual" variables mainly relate to neighbourhood characteristics (e.g. neighbourhood quality in Can [1990]). However, the expansion method can be applied more generally, by observing the heterogeneity of any parameter $(X)$ depending on the "context". This "context" may refer to neighbourhood attributes (quality, distance to the city centre, etc.), but also, as is suggested in this paper, to the specific characteristics of the buyers. The significant expansion parameters therefore measure the variation of the implicit prices people assign to attributes. Also, a parameter can be non significant overall, but may become significant once contextualised. This is only a special case of equation (4), that is, when $\beta_0$ is null and $\beta_1$ is not.

Can (1990) measures the drift of several property specific parameters in relation to the neighbourhood quality for a sample of 577 single-family houses of the Columbus metropolitan area. The two final models consider both the spatial heterogeneity of property specifics (using spatial expansion to neighbourhood quality) and the spatial dependence (using a spatially lagged dependent variable). The parameters that vary significantly through space are the following: the type of exterior, the lot size, the presence of a two-car garage and the presence of a utility room. Recently, a model built with single-family properties transacted during the 1990-1991 period in Quebec City includes several expansion variables (Thériault *et al.*, 2003). Various property attributes are spatially expanded using indicators of relative centrality, family cycle and socio-economic status (derived from census data) as well as using measures of accessibility to regional and local services (computed within a GIS). In addition to age, lot size and connection to the sewer system, three property specifics present spatial drifts: inferior ceiling quality, kitchen cabinets made of hard wood, and the number of washrooms. It seems important to us to verify whether further drifts in the implicit prices can be related to the household- rather than to the census-tract- social profile. This question follows a previous paper that showed that the odds-ratio of mentioning a property or neighbourhood choice criteria – i.e., a proxy of their preference for certain types of attributes – is significantly linked to the household profile (Des Rosiers *et al.*, 2002). To the best of our knowledge, no research has yet integrated household profile data into hedonic modelling.

Concomitantly with the expansion method, we ran several GWRs, which gave additional indications on the spatial non-stationarity of the parameters. GWR is an adaptation of moving regressions. Moving regression functions are calibrated for every point of a regular grid, using all data within a certain region around this point. The resulting parameters are site-specific and can therefore vary through space. However, this method is discontinuous, as no weighting schemes are applied to the data used for calibration.

GWRs calibrate local models for every sampling point. However, a weighting scheme (spatial kernel) is applied in order to give greater influence to close data points. Furthermore, the spatial kernel may be fixed (identical for all locations) or adaptive; that is, its bandwidth may vary with the density of the data:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \qquad (5)$$

where $(u_i, v_i)$ denotes the coordinates of the *i*th point in space and $\beta_k(u_i, v_i)$ is a realisation of the continuous function $\beta_k(u, v)$ at point *i* (Fotheringham, 2002; p. 52).

Various methods can be used to derive the bandwidth that provides a trade-off between goodness-of fit and degrees of freedom: the generalised cross-validation criterion (GCV)

(Craven and Wahba, 1979; Loader, 1999), the Schwartz Information Criterion (Schwartz, 1978) or the Akaike Information Criterion (AIC) (Akaike, 1973; Hurvich *et al.*, 1998). For further details on the spatial weighting function calibration, see Fotheringham (Fotheringham *et al.*, 2002, p. 59-62). Furthermore, the stationarity of each estimated parameter can be tested using either a Monte Carlo approach (Hope, 1968) or the Leung test (See Fotheringham *et al.*, 2002, pp. 92-94; Leung *et al.*, 2000).

In a GWR application on residential value analysis, Brunsdon *et al.* (1999) showed that the relationship between house price and size varies significantly through space in the town of Deal in south-eastern England.

## 3.3 Modelling Procedure

All models were built with 761 single-family properties transacted between 1993 and 2001 in Quebec City, Canada (mainly between 1993 and 1996). Property-specific variables were extracted from the valuation role (See variable description in Table 16). The characteristics of the vegetation around each property were extracted from remote-sensing data. A Landsat TM-5 image shot in 1999 was categorised using the semi-automated ISODATA (Iterative Self-Organising Data Analysis) technique, widely used and implemented in some GIS packages (Duda and Hart, 1973). Furthermore, the Normalised Difference Vegetation Index (NDVI), a sensitive indicator of the green biomass (Tucker, 1979; Tueller, 1989; Wu *et al.*, 1997), was derived. For more details about the extraction of vegetation data from remote sensing images and its integration into hedonic models, see Kestens *et al.* (At Press). NDVI is a measure of vegetation density, whereas its standard deviation indicates land-use heterogeneity. An additional variable identifies properties with more than 29 trees (according to the number of trees mentioned by the owners during a phone survey, as described below). Previous work by Payne identified this number as the limit upon which the premium accorded to trees was reversed (Payne, 1973). Centrality – the mean car-time distance to the Main Activity Centres (MACs) – was computed within a GIS (Thériault *et al.*, 1999). Furthermore, a major phone survey carried out from 2000 to 2003 provided detailed information about each buyer household. The survey concerned the household's moving motivations and property choice criteria, and provided additional data on the household profiles and on specific attributes of the property, like the number of trees on the lot. A detailed description of the survey and the relations between the motivation to move, choice criteria and household profile are given in Kestens *et al.* (Submitted).

# Table 16. Variables description

| | Variable name | Description | Type* | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|---|---|
| Property Specifics | SPRICE | Sale price of the property (Cad $) | N | 53000 | 290000 | 114459 | 40249 |
| | LNSPRICE | Natural logarithm of the sale price (Cad $) | N | 10.9 | 12.6 | 11.6 | 0.31 |
| | Local Tax Rate | Local tax rate ($/100$ of assessed value) | N | 1.20 | 2.76 | 2.06 | 0.39 |
| | Living Area M2 | Living area (sq.m.) | N | 65.7 | 265.0 | 122.0 | 35.3 |
| | Age30-39* Living Area | Living area (sq.m.) * Age 30-39 | | -54.9 | 59.3 | -1.8 | 16.4 |
| | Living Area Bung. | Living area of a bungalow (sq.m.) | | 0 | 206.6 | 49.7 | 52.2 |
| | Living Area Bung. * Income 80K up | Living area of a bungalow (sq.m.) * Yearly income 80,000 Cad $ up | | -37.7 | 113.4 | -3.6 | 24.2 |
| | LnLotsiz | Natural logarithm of the lot size (sq. m.) | N | 5.3 | 7.5 | 6.3 | 0.3 |
| | App. Age | Apparent age (years) | N | 0 | 54 | 16.6 | 12 |
| | App. Age * Sgle Houshld | Apparent age (years) * Single houshld | | -15.6 | 22 | 0.1 | 2.7 |
| | App. Age Cd Sqd | Apparent age centreed squared | N | 0.185 | 1401.0 | 143.8 | 160.1 |
| | App. Age Cd Sqd * Houshld Income | Apparent age centreed squared * Household Income | N | -3253.4 | 2662.2 | 30.4 | 501.1 |
| | Quality | House quality index | N | -1 | 2 | 0.00 | 0.2 |
| | Finished Basement | Finished basement | B | 0 | 1 | 0.56 | 0.50 |
| | Superior Floor Qual. | Superior floor quality | B | 0 | 1 | 0.49 | 0.50 |
| | Facing 51%+ Brick | More than 50% of facing made of stone or brick | B | 0 | 1 | 0.38 | 0.49 |
| | Built-in Oven | Built-in Oven | | 0 | 1 | 0.14 | 0.34 |
| | Built-in Oven * Age 40-49 | Built-in Oven * Age 40-49 | B | -0.33 | 0.54 | 0.00 | 0.17 |
| | Fireplace | Numbers of fireplaces | N | 0 | 6 | 0.34 | 0.53 |
| | Fireplace * First-time Owner | Superior floor quality | N | -2.72 | 0.86 | -0.03 | 0.26 |
| | In-ground Pool | Presence of an in-ground pool | B | 0 | 1 | 0.06 | 0.23 |
| | In-ground Pool * Houshld Income | Presence of an in-ground pool * Household income | N | -4.65 | 2.90 | 0.07 | 0.58 |
| | In-ground pool * Sgle Houshld | Presence of an in-ground pool * Single household | N | -0.13 | 0.82 | -0.01 | 0.05 |
| | Detached Garage | Presence of a detached garage | B | 0 | 1 | 0.12 | 0.33 |
| | Det. Garage * Univ. Degree Holders | Presence of a detached garage * University degree holder | N | -0.48 | 0.40 | 0.00 | 0.16 |
| | Det. Garage * Couple without Child | Presence of a detached garage * Couple without children | B | -0.16 | 0.72 | 0.00 | 0.13 |
| | Det. Garage * Age 40-49 | Presence of a detached garage * Age 40-49 | B | -0.33 | 0.55 | -0.01 | 0.15 |
| | Att Garage | Presence of an attached garage | B | 0 | 1 | 0.08 | 0.28 |

| | Variable name | Description | Type* | Min. | Max. | Mean | Std. Dev. |
|---|---|---|---|---|---|---|---|
| Accessibility | CBD Car Time | Car time distance to CBD (min) | B | 4.685 | 21.0 | 12.6 | 3.51 |
| | CBD Car Time Cd Sqd | Car time distance to CBD centreed squared (min) | B | 1E-06 | 70.9 | 12.3 | 14.7 |
| | CBD Car Time Cd Sqd * Houshld Income | Car time distance to CBD centreed squared (min) * Houshld Income | N | -217.2 | 154.8 | 3.37 | 36.3 |
| | Highw. Exit | Car time distance to nearest highway exit (min.) | N | 0.04 | 7.53 | 2.74 | 1.46 |
| **Household-level data** | **Household Income** | Household income in 10,000 $ ranges up to 100,000 and up | N | 1 | 10 | 6.93 | 2.27 |
| | **First-Time Owners** | The owner of this property is owner for the first time | N | 0 | 1 | 0.48 | 0.50 |
| | **Age under 30** | Aged under 30 at date of transaction | N | 0 | 1 | 0.04 | 0.20 |
| Vegetation | Mature Trees 100 m | Percentage of area in a 100 m radius covered by residential use with mature trees | N | 0 | 76 | 18.5 | 17.20 |
| | Mature Trees 100 m * Age 30-39 | Percentage of area in a 100 m radius covered by residential use with mature trees * Age 30-39 | | -20.3 | 35.4 | -1.19 | 8.11 |
| | Mature Trees 500 m | Percentage of area in a 500 m radius covered by residential use with mature trees | N | 0.5 | 45.0 | 14.1 | 9.68 |
| | Low Tree Dens. 500 m | Percentage of area in a 500 m radius covered by residential use with low tree density | N | 2.6 | 34.3 | 16.8 | 5.97 |
| | Nb Trees 29up | Number of trees on the property is 29 and up | N | 0 | 1 | 0.03 | 0.18 |
| | Nb Trees 29up * Age 40up | Number of trees on the property is 29 and up * Age 40 and up | N | -0.31 | 0.66 | 0.003 | 0.09 |
| | NDVI Std. Dev. 1 km | NDVI standard deviation in a 1 km radius (landuse heterogeneity measure) | N | 0.18 | 0.60 | 0.32 | 0.06 |
| | Agricult. Land 100 m * Univ. Degr. Holders | Percentage of agricultural land with dispersed trees in a 100 m radius * University degree holders | N | -21.5 | 6.68 | -0.13 | 1.63 |
| | Agricult. Land 100 m * Age 30-39 | Percentage of agricultural land with dispersed trees in a 100 m radius * Age 30-39 | N | -9.5 | 24.2 | 0.1209 | 1.7 |
| | Agricult. Land 100 m * Age 40up | Percentage of agricultural land with dispersed trees in a 100 m radius * Age 40 and up | N | -12.5 | 16.9 | -0.1 | 1.4 |
| | Agricult. Land 100 m * Houshld with Children | Percentage of agricultural land with dispersed trees in a 100 m radius * Household with children | N | -11.0 | 10.0 | 0.0 | 1.3 |
| | Agricult. Land 100 m * First-time Owner | Percentage of agricultural land with dispersed trees in a 100 m radius * First-time owner | N | -11.9 | 20.4 | 0.07 | 1.6 |
| | NDVI 40 m * Age 30-39 | Normalised Deviation Vegetation Index within a 40 m radius (density of vegetation) * Age30-39 | N | -0.36 | 0.32 | 0.00 | 0.07 |
| | Woodlands 500 m * Single Household | Percentage of woodlands in a 500 m radius * Single households | N | -10.6 | 26.1 | -0.3 | 3.5 |
| **Census data** | % of Univ. Degree Holders | Percentage of university degree holders in census tract | N | 4.7 | 60.1 | 33.0 | 14.7 |
| | % of Univ. Degree Holders * Houshld Univ. Holders | Percentage of university degree holders in census tract * Household data: university holders | N | -14.8 | 15.6 | 2.5 | 6.9 |
| | % aged 65 up | Percentage of people aged 65 and up in census tract | N | 1.3 | 58.9 | 8.18 | 7.33 |
| | Developped Area (Ha) | Developed area in census tract | N | 10.4 | 407 | 69.6 | 32.1 |
| | % Dwellings 1946-60 | Percentage of dwellings built between 146 and 1960 in census tract | N | 0 | 62.7 | 12.3 | 15.4 |
| | % Unemployed Aged 15-24 | Percentage of unemployed aged between 15 and 24 in census tract | N | 0 | 55.2 | 18.1 | 11.4 |
| | Nb Persons per Room | Average number of persons per room in census tract | N | 0.3 | 0.6 | 0.42 | 0.06 |

### 3.3.1 Expansion Models

In this paper, a first group of models, referred to as global models, is built using the expansion method. All models are in the semi-log functional form (the dependent variable is the logarithm of the selling price) using OLS specification. The first four models (M) omit census variables, whereas the last three (N) include them (see Table 17). A time-drift variable was introduced but did not prove significant. Concerning the M models, a basic model (M1) contains property specifics, vegetation attributes derived from remote-sensing data, and centrality measures, whereas homebuyers' socio-economic variables are added in a second step (M2). Expansion terms (all attributes being "expanded" with regard to the socio-economic profile of the buyers' households) are then added on to both model M1 (resulting in model M3a) and model M2 (resulting in model M3b). The N series is distinctive in that it contains additional socio-economic Census variables, with N1 as the basic model (including property specifics, vegetation, centrality and Census data), N2 including household profile variables, while expansion terms are introduced in N3. In order to avoid multicollinearity, all expansion terms are built with the previously centered original variables, thereby reflecting the departure from the overall average market values (Jaccard *et al.*, 1990, p. 31).

Table 17. Results of regression models

| | | Without Census data | | | | With Census data | | |
|---|---|---|---|---|---|---|---|---|
| **Global Models** | | **Model M1** | **Model M2** | **Model M3a** | **Model M3b** | **Model N1** | **Model N2** | **Model N3** |
| **Model Specifications** | Nb of cases | 761 | 761 | 761 | 761 | 761 | 761 | 761 |
| | R-square | 0.853 | 0.867 | 0.876 | 0.881 | 0.870 | 0.882 | 0.894 |
| | Adj. R-Square | 0.848 | 0.863 | 0.870 | 0.876 | 0.866 | 0.878 | 0.889 |
| | SEE | 0.121 | 0.115 | 0.112 | 0.110 | 0.114 | 0.109 | 0.104 |
| | SEE in % | 12.9% | 12.2% | 11.9% | 11.6% | 12.1% | 11.5% | 10.9% |
| | F ratio | 504 | 200 | 142 | 159 | 197 | 203 | 161 |
| | Sig. | 0.0000 | 0.0000 | 0.0000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | Df1/Df2 | 23/737 | 24/736 | 36/723 | 34/727 | 25/735 | 27/733 | 38/722 |
| | Interactive Variables / Total Variables | 0/23 | 0/24 | 13/36 | 12/34 | 0/25 | 0/27 | 11/38 |
| | Maximum VIF value | 4.1 | 3.0 | 3.6 | 3.1 | 5.1 | 5.1 | 3.9 |
| **Spatial Auto-correlation of residuals** | Moran's I (1500 m lag) | 0.172 | 0.130 | 0.084 | 0.034 | 0.176 | 0.159 | 0.102 |
| | Sig. | 0.096 | 0.162 | 0.262 | 0.397 | 0.092 | 0.114 | 0.218 |
| | Most sig. Moran's I SA range (300m lags) | 600 m | 600 m | 600 m | 600 m | 600m | 600 m | 600 m |
| | Nb of significant LSA zG*i statistics (600 m lag, sig. 0.05) | 90 | 61 | 41 | 24 | 46 | 35 | 26 |
| | Nb of significant LSA zGi statistics (600 m lag, sig. 0.05) | 67 | 41 | 34 | 28 | 42 | 26 | 17 |
| **Variables in model** | Property specifics | X | X | X | X | X | X | X |
| | Vegetation data | X | X | X | X | X | X | X |
| | Centrality | X | X | X | X | X | X | X |
| | Census data | | | | | X | X | X |
| | Household variables | | X | | X | | X | X |
| | Interactions (household var. * others) | | | X | X | | | X |

| **GWR Models** | | **Model GWR_M1** | **Model GWR_M2** | | | **Model GWR_N1** | **Model GWR_N2** | |
|---|---|---|---|---|---|---|---|---|
| **GWR Models** | Nb of cases | 761 | 761 | | | 761 | 761 | |
| | R-square | 0.902 | 0.902 | | | 0.885 | 0.892 | |
| | SEE | 0.1061 | 0.1043 | | | 0.1098 | 0.1059 | |
| | Kernel bandwidth | 320.5 | 412.5 | | | 661.49 | 706.5 | |
| | F statistic of GWR Improvement (sig.) | 3.36 (0.002) | 3.15 (0.004) | | | 2.78 (0.008) | 2.51 (0.013) | |
| **Spatial Auto-correlation of residuals** | Moran's I (1500 m lag) | 0.045 | 0.049 | | | 0.063 | 0.082 | |
| | Sig. | 0.364 | 0.352 | | | 0.316 | 0.265 | |
| | Nb of significant LSA zG*i statistics (600 m lag, sig. 0.05) | 21 | 26 | | | 26 | 26 | |
| | Nb of significant LSA zGi statistics (600 m lag, sig. 0.05) | 22 | 21 | | | 22 | 20 | |

### 3.3.2 GWR Models

Concomitantly, using the same dependent and explanatory variables as in M1, M2, N1 and N2, four Geographically Weighted Regressions are built (GWR_M1, GWR_M2, GWR_N1, and GWR_N2). The limitation of the GWR software we used, constraining the spreadsheet to a maximum of 35 variables, made it impossible to derive further GWR versions of models M3a, M3b or N3. However, the interest of GWR relies in the possibility of deriving local statistics and a significance test for the stationarity of individual parameters. For a description of further local descriptive statistics that can be obtained using the geographically weighting framework, see Brunsdon *et al.* (2002). An *F*-statistic also indicates the significance of improvement between the global and the GWR models. Furthermore, as M3a, M3b and N3 are the "expanded" versions of M1, M2 and N2, they can easily be compared with their GWR counterparts, GWR_M1, GWR_M2 and GWR_N2.

All the GWRs were computed with adaptive bi-square spatial kernels, using all data and the Akaike Information Criterion minimisation for calibration of the spatial weighting function (Fotheringham *et al.*, 2002, p. 61). The significance test for the heterogeneity of the parameters was made using the Monte Carlo approach (Hope, 1968).

For each model, global and local spatial autocorrelation of the residuals are measured, using Moran's I for the former (Moran, 1950) and Getis and Ord's zG*I (Getis and Ord, 1992; Ord and Getis, 1995) for the latter.

The significant variables are described in Table 16, whereas Table 17 contains the specifications and performance of all models. The estimated parameters, their significance and the Variance Inflation Factor (VIF) values – indicating eventual multicollinearity – are detailed in Table 18 (M series) and Table 19 (N series).

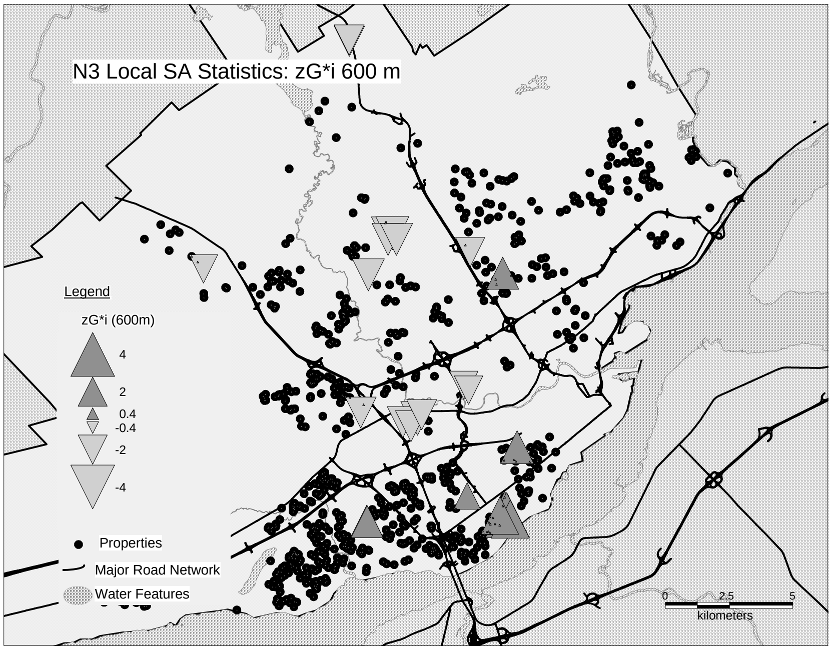
# 3.4 Results

## 3.4.1 Performance of the Global Models

Each of the global models explains at least 84% of the house price variation. Model N3, with an adjusted *R*-square of 0.889, a SEE of 10.9%, and an *F*-value of 161 achieves the best performance. Collinearity is well under control in all models, with only one VIF value slightly exceeding 5 (*Car time to MACs*, model N1).

No model presents significant global autocorrelation at the 95% level (Moran's I ranges from 0.034 [M3b] to 0.172 [M1]). Local autocorrelation is present, but decreases when household-level data is included, and further more when expansion terms are introduced. The number of "hot spots", that we defined as the significant zG*i statistics given a 600 m lag (which is the most significant autocorrelation range according to the correlogram), drops from 90 (M1) to 61 (M2) to 41 (M3a) to 24 (M3b). Results are similar for the N series that includes Census variables: the number of hot spots is already low for N1 (46), and still decreases for N2 (35) and N3 (26). The remaining local spatial autocorrelation in M3b and N3, as defined before, concerns less than 5% of the sample (respectively 24 and 26 cases out of 761, or 3.15% of all cases), and is as such not significant at the 95% confidence level (see Figure 8).

Figure 8. Local spatial autocorrelation: Significant zG*i statistics for N3




The basic models M1 and N1 include classic descriptors as well as several significant variables relating to vegetation, confirming the impact of environmental factors and surrounding land use on house values (Geoghegan *et al.*, 1997; Kestens *et al.*, At Press).The percentage of trees has a global positive impact, however, when the socio-economic condition of the neighbourhood is considered (Census data in Model N1), the impact of vegetation within a 500 m range becomes non-significant. This stresses the links between the socio-economic status of the neighbourhood and the land use, mainly with regard to vegetation. Although mature trees in the close surroundings (100 m around the property) represent a premium, the presence of trees becomes detrimental when exceeding a threshold. In fact, the coefficient for the binary variable identifying properties with more than 29 trees is significantly negative (-5.90%, M1), in accordance with previous findings by Payne (1973).

Accessibility to the Main Activity Centres (MAC) is highly significant (*t*-value of – 11.02), but the negative effect on property values is not strictly linear, as proved by the presence of the squared form of the parameter (previously centered to avoid collinearity), with a positive sign (*t*-value 4.41). Hence, the location rent follows a quadratic function and takes the form of a U-shaped curve, with positive premiums both in the city centre and in the outer suburbs, *ceteris paribus*. A previous study showed that land-use and vegetation attributes significantly explain part of these premiums, reducing the value and significance of the squared distance term (Kestens *et al.*, At Press). Therefore, if vegetation descriptors were absent, this parameter would be even higher and more significant.

Table 18. Coefficients of M models

| Dependent Variable: LnSprice | Model M1 | | | | Model M2 | | | | Model M3a | | | | Model M3b | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | (t-value) | Sig | VIF | B | (t-value) | Sig | VIF | B | (t-value) | Sig | VIF | B | (t-value) | Sig | VIF |
| (Constant) | 10.5479 | (90.57) | *** | | 10.6027 | (97.15) | *** | | 10.8351 | (107.78) | *** | | 10.7353 | (108.71) | *** | |
| Local Tax Rate | -0.1333 | (-9.47) | *** | 1.5 | -0.1145 | (-8.47) | *** | 1.5 | -0.1349 | (-11.15) | *** | 1.3 | -0.1211 | (-10.15) | *** | 1.3 |
| Living Area M2 | 0.0042 | (20.6) | *** | 2.6 | 0.0039 | (19.94) | *** | 2.7 | 0.0042 | (22.35) | *** | 2.7 | 0.0039 | (21.16) | *** | 2.7 |
| Age30-39* Living Area | -- | -- | | -- | -- | -- | | -- | 0.0007 | (2.98) | *** | 1.1 | -- | -- | | -- |
| Living Area Bung. | 0.0007 | (6.15) | *** | 2.1 | 0.0007 | (6.16) | *** | 2.1 | 0.0008 | (7.52) | *** | 2.1 | 0.0008 | (7.30) | *** | 2.1 |
| Living Area Bung. * Income 80K up | -- | -- | | -- | -- | -- | | -- | -0.0006 | (-3.56) | *** | 1.1 | -0.0005 | (-3.12) | *** | 1.1 |
| LnLotsiz | 0.1705 | (9.17) | *** | 1.9 | 0.1511 | (8.54) | *** | 1.9 | 0.1319 | (7.59) | *** | 1.9 | 0.1346 | (7.93) | *** | 1.9 |
| App. Age | -0.0132 | (-20.01) | *** | 3.2 | -0.0116 | (-19.14) | *** | 3.0 | -0.0123 | (-19.99) | *** | 3.2 | -0.0118 | (-20.24) | *** | 3.0 |
| App. Age * Sgle Houshld | -- | -- | | -- | -- | -- | | -- | 0.0035 | (2.24) | ** | 1.0 | 0.0044 | (2.95) | *** | 1.0 |
| App. Age Cd Sqd | 0.0001 | (2.36) | ** | 1.7 | 0.0001 | (1.81) | * | 1.7 | 0.0001 | (2.00) | ** | 1.8 | 0.0000 | (2.26) | ** | 1.8 |
| App. Age Cd Sqd * Houshld Income | -- | -- | | -- | -- | -- | | -- | 0.0001 | (5.39) | *** | 1.3 | -- | -- | | -- |
| Quality | 0.0933 | (3.67) | *** | 1.1 | 0.0883 | (3.65) | *** | 1.1 | 0.0780 | (3.27) | *** | 1.1 | 0.0624 | (2.68) | *** | 1.1 |
| Finished Basement | 0.0509 | (5.16) | *** | 1.2 | 0.0485 | (5.15) | *** | 1.2 | 0.0577 | (6.24) | *** | 1.2 | 0.0526 | (5.82) | *** | 1.2 |
| Superior Floor Qual. | 0.0566 | (5.7) | *** | 1.2 | 0.0508 | (5.35) | *** | 1.2 | 0.0605 | (6.54) | *** | 1.2 | 0.0529 | (5.85) | *** | 1.2 |
| Facing 51%+ Brick | 0.0268 | (2.76) | *** | 1.1 | 0.0193 | (2.08) | ** | 1.1 | 0.0272 | (-3.00) | *** | 1.1 | 0.0216 | (2.44) | ** | 1.1 |
| Built-in Oven | 0.0414 | (3.04) | *** | 1.1 | 0.0437 | (3.39) | *** | 1.1 | 0.0465 | (3.66) | *** | 1.1 | 0.0534 | (4.33) | *** | 1.1 |
| Built-in Oven * Age 40-49 | -- | -- | | -- | -- | -- | | -- | 0.0581 | (2.3) | ** | 1.0 | 0.0657 | (2.68) | *** | 1.0 |
| Fireplace | 0.0259 | (2.67) | *** | 1.3 | 0.0237 | (2.57) | *** | 1.3 | 0.0368 | (3.98) | *** | 1.4 | 0.0341 | (3.79) | *** | 1.4 |
| Fireplace * First-time Owner | -- | -- | | -- | -- | -- | | -- | 0.0465 | (2.81) | *** | 1.1 | 0.0448 | (2.80) | *** | 1.1 |
| In-ground Pool | 0.1068 | (5.33) | *** | 1.1 | 0.0940 | (4.93) | *** | 1.1 | 0.0750 | (3.51) | *** | 1.4 | 0.0645 | (3.14) | *** | 1.4 |
| In-ground Pool * Houshld Income | -- | -- | | -- | -- | -- | | -- | 0.0358 | (3.93) | *** | 1.7 | 0.0205 | (2.61) | *** | 1.3 |
| In-ground pool * Sgle Houshld | -- | -- | | -- | -- | -- | | -- | 0.2985 | (3.20) | *** | 1.5 | -- | -- | | -- |
| Detached Garage | 0.0510 | (3.42) | *** | 1.2 | 0.0440 | (3.1) | *** | 1.2 | 0.0508 | (3.63) | *** | 1.2 | 0.0427 | (3.12) | *** | 1.2 |
| Det. Garage * Univ. Degree Holders | -- | -- | | -- | -- | -- | | -- | 0.0714 | (2.72) | *** | 1.0 | 0.0864 | (3.37) | *** | 1.0 |
| Det. Garage * Couple without Child | -- | -- | | -- | -- | -- | | -- | -0.0810 | (-2.37) | ** | 1.0 | -0.0812 | (-2.42) | ** | 1.0 |
| Det. Garage * Age 40-49 | -- | -- | | -- | -- | -- | | -- | -0.0559 | (-1.99) | ** | 1.0 | -0.0800 | (-2.92) | *** | 1.0 |
| AttGarage | 0.0770 | (4.34) | *** | 1.2 | 0.0687 | (4.08) | *** | 1.2 | 0.0682 | (4.09) | *** | 1.2 | 0.0592 | (3.64) | *** | 1.2 |
| CBD Car Time | -0.0250 | (-11.02) | *** | 3.2 | -0.0271 | (-13.1) | *** | 2.9 | -0.0223 | (-12.49) | *** | 2.3 | -0.0233 | (-13.44) | *** | 2.3 |
| CBD Car Time Cd Sqd | 0.0017 | (4.41) | *** | 1.6 | 0.0018 | (5.06) | *** | 1.6 | 0.0011 | (3.24) | *** | 1.6 | 0.0015 | (4.51) | *** | 1.6 |
| CBD Car Time Cd Sqd * Houshld Income | -- | -- | | -- | -- | -- | | -- | 0.0004 | (3.12) | *** | 1.2 | 0.0005 | (4.23) | *** | 1.1 |
| Mature Trees 100 m | 0.0010 | (2.42) | ** | 2.9 | -- | -- | | -- | 0.0007 | (1.79) | * | 3.0 | -- | -- | | -- |
| Mature Trees 100 m * Age 30-39 | -- | -- | | -- | -- | -- | | -- | | | | | -- | -- | | -- |
| Mature Trees 500 m | 0.0041 | (4.44) | *** | 4.0 | 0.0036 | (5.41) | *** | 2.4 | 0.0037 | (4.59) | *** | 3.6 | 0.0039 | (6.22) | *** | 2.3 |
| Low Tree Dens. 500 m | -0.0034 | (-3.32) | *** | 2.0 | -0.0026 | (-2.9) | *** | 1.6 | -0.0033 | (-3.88) | *** | 1.6 | -0.0037 | (-4.49) | *** | 1.5 |
| Nb Trees 29up | -0.0609 | (-2.25) | ** | 1.1 | -0.0629 | (-2.45) | ** | 1.1 | -0.0689 | (-2.71) | *** | 1.1 | -0.0617 | (-2.47) | ** | 1.1 |
| Nb Trees 29up * Age 40up | -- | -- | | -- | -- | -- | | -- | -- | -- | -- | -- | -0.1653 | (-3.37) | *** | 1.0 |
| NDVI Std. Dev. 1 km | 0.2893 | (2.87) | *** | 2.1 | 0.2217 | (2.46) | ** | 1.8 | -- | -- | -- | -- | -- | -- | | -- |
| **Household Income** | -- | -- | | -- | 0.0166 | (7.91) | *** | 1.3 | -- | -- | -- | -- | 0.0172 | (8.47) | *** | 1.3 |
| **First-Time Owners** | -- | -- | | -- | -0.0427 | (-4.72) | *** | 1.1 | -- | -- | -- | -- | -0.0402 | (-4.69) | *** | 1.1 |
| **Age under 30** | -- | -- | | -- | -0.0390 | (-1.75) | * | 1.0 | -- | -- | -- | -- | -- | -- | | -- |
| Agricult. Land 100 m * Univ. Degr. Holders | -- | -- | | -- | | -- | | -- | -0.0093 | (-3.54) | *** | 1.1 | -- | -- | | -- |
| Agricult. Land 100 m * Age 30-39 | -- | -- | | -- | | -- | | -- | -0.0091 | (-3.50) | *** | 1.1 | -- | -- | | -- |
| Agricult. Land 100 m * Age 40up | -- | -- | | -- | | -- | | -- | -- | -- | | -- | 0.0098 | (3.21) | *** | 1.1 |
| Agricult. Land 100 m * Houshld with Children | -- | -- | | -- | | -- | | -- | -- | -- | | -- | 0.0099 | (3.09) | *** | 1.0 |

**Grey boxes: Interactions with buyers' household characteristics; Bold: buyers' household variables**
*****: sig. <0.01; **: sig. <0.05; *: sig. <0.1**

Table 19. Coefficients of N models

| Dependent Variable: LnSprice | Model N1 | | | | Model N2 | | | | Model N3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | B | (t-value) | Sig | VIF | B | (t-value) | Sig | VIF | B | (t-value) | Sig | VIF |
| (Constant) | 10.1201 | (82.85) | *** | | 10.1710 | (86.95) | *** | | 10.2443 | (91.49) | *** | |
| Local Tax Rate | -0.1113 | (-7.28) | *** | 2.0 | -0.0963 | (-6.54) | *** | 2.0 | -0.0911 | (-6.58) | *** | 2.0 |
| Living Area M2 | 0.0039 | (20.18) | *** | 2.7 | 0.0036 | (19.39) | *** | 2.8 | 0.0036 | (20.41) | *** | 2.8 |
| Living Area of Bungalow | 0.0007 | (6.23) | *** | 2.1 | 0.0006 | (6.17) | *** | 2.1 | 0.0007 | (7) | *** | 2.1 |
| Ln Lot Size | 0.1800 | (10.28) | *** | 1.9 | 0.1618 | (9.59) | *** | 1.9 | 0.1550 | (9.61) | *** | 1.9 |
| Apparent Age | -0.0132 | (-20.88) | *** | 3.3 | -0.0122 | (-19.91) | *** | 3.4 | -0.0113 | (-20.01) | *** | 3.2 |
| App. Age Centered Squared | 0.0001 | (2.72) | *** | 1.7 | 0.0001 | (2.39) | ** | 1.7 | -- | -- | | -- |
| Quality | 0.0821 | (3.44) | *** | 1.1 | 0.0767 | (3.36) | *** | 1.1 | 0.0690 | (3.15) | *** | 1.1 |
| Finished Basement | 0.0528 | (5.69) | *** | 1.2 | 0.0489 | (5.5) | *** | 1.2 | 0.0483 | (5.75) | *** | 1.2 |
| Superior Floor Quality | 0.0528 | (5.64) | *** | 1.2 | 0.0487 | (5.44) | *** | 1.2 | 0.0550 | (6.59) | *** | 1.2 |
| Built-in Oven | 0.0371 | (2.89) | *** | 1.1 | 0.0415 | (3.38) | *** | 1.1 | 0.0426 | (3.6) | *** | 1.1 |
| Built-in Oven * First-Time Owner | -- | -- | | -- | -- | -- | | -- | -0.0701 | (-3.02) | *** | 1.1 |
| Fireplace | 0.0272 | (2.99) | *** | 1.3 | 0.0262 | (3.01) | *** | 1.3 | 0.0306 | (3.61) | *** | 1.4 |
| Fireplace * First-Time Owner | -- | -- | | -- | -- | -- | | -- | 0.0379 | (2.49) | ** | 1.1 |
| In-ground Pool | 0.1013 | (5.38) | *** | 1.1 | 0.0907 | (5.04) | *** | 1.1 | 0.0746 | (4.3) | *** | 1.1 |
| Detached Garage | 0.0596 | (4.24) | *** | 1.2 | 0.0546 | (4.07) | *** | 1.2 | 0.0513 | (3.93) | *** | 1.2 |
| Det. Garage * Univ. Degree Holders | -- | -- | | -- | -- | -- | | -- | 0.0799 | (3.25) | *** | 1.1 |
| Det. Garage * Couple without Child | -- | -- | | -- | -- | -- | | -- | -0.0790 | (-2.5) | ** | 1.0 |
| Det. Garage * Age 40-49 | -- | -- | | -- | -- | -- | | -- | -0.0636 | (-2.44) | ** | 1.1 |
| Attached Garage | 0.0807 | (4.84) | *** | 1.2 | 0.0770 | (4.82) | *** | 1.2 | 0.0763 | (4.97) | *** | 1.2 |
| CBD Car Time | -0.0202 | (-7.55) | *** | 5.0 | -0.0215 | (-8.4) | *** | 5.1 | -0.0180 | (-8.46) | *** | 3.8 |
| CBD Car Time Centered Squared | 0.0013 | (3.45) | *** | 1.9 | 0.0014 | (3.88) | *** | 1.9 | 0.0013 | (3.94) | *** | 1.8 |
| CBD Car Time C. Sq. * Househld Income | -- | -- | | -- | -- | -- | | -- | 0.0004 | (3.76) | *** | 1.1 |
| HIGHWEXT | 0.0090 | (2.13) | ** | 2.2 | 0.0085 | (2.1) | ** | 2.2 | -- | -- | | -- |
| Mature Trees 100 m | 0.0014 | (4.05) | *** | 2.1 | 0.0010 | (3.18) | *** | 2.1 | 0.0006 | (2.02) | ** | 2.2 |
| Low Tree Density 500 m | -- | -- | | -- | -- | -- | | -- | -0.0018 | (-2.58) | *** | 1.3 |
| Nb Trees 29 up | -0.0514 | (-2.02) | ** | 1.1 | -0.0475 | (-1.96) | ** | 1.1 | -0.0553 | (-2.34) | ** | 1.2 |
| Nb Trees 29 up * Age 40 up | -- | -- | | -- | -- | -- | | -- | -0.1462 | (-3.17) | *** | 1.0 |
| % of Univ. Degree Holders | 0.0050 | (9.68) | *** | 3.4 | 0.0047 | (9.36) | *** | 3.5 | 0.0044 | (9.17) | *** | 3.5 |
| % of Univ. Degree Holders * Univ. Degree Holders | -- | -- | | -- | -- | -- | | -- | 0.0018 | (3.14) | *** | 1.1 |
| % aged 65 up | 0.0043 | (4.57) | *** | 2.9 | 0.0043 | (4.7) | *** | 2.9 | 0.0038 | (4.4) | *** | 2.9 |
| Developed Area (Ha) | -0.0004 | (-2.82) | *** | 1.4 | -0.0003 | (-2.15) | ** | 1.4 | -0.0003 | (-2.26) | ** | 1.4 |
| % Dwellings 1946-60 | 0.0016 | (3.31) | *** | 3.3 | 0.0013 | (2.82) | *** | 3.3 | 0.0011 | (2.63) | *** | 3.4 |
| % Unemployed Aged 15-24 | -0.0013 | (-3.31) | *** | 1.1 | -0.0012 | (-3.38) | *** | 1.1 | -0.0011 | (-3.02) | *** | 1.2 |
| Nb Persons per Room | 0.4368 | (3.72) | *** | 3.0 | 0.4574 | (4.07) | *** | 3.0 | 0.3692 | (3.37) | *** | 3.1 |
| **Household Income** | -- | -- | | -- | 0.0145 | (7.23) | *** | 1.3 | 0.0155 | (7.99) | *** | 1.3 |
| **First-Time Owners** | -- | -- | | -- | -0.0396 | (-4.68) | *** | 1.1 | -0.0417 | (-5.1) | *** | 1.1 |
| NDVI 40 m * Age 30-39 | -- | -- | | -- | -- | -- | | -- | -0.2110 | (-3.45) | *** | 1.1 |
| Woodlands 500 m * Single Household | -- | -- | | -- | -- | -- | | -- | -0.0027 | (-2.49) | ** | 1.0 |
| Agricult. Land 100 m * Univ. Degr. Holders | -- | -- | | -- | -- | -- | | -- | -0.0083 | (-3.42) | *** | 1.1 |
| Agricult. Land 100 m * First-time Owner | -- | -- | | -- | -- | -- | | -- | -0.0054 | (-2.14) | ** | 1.1 |

**Dark Grey boxes: Interactions with buyers' household variables; Light grey boxes: Census variables; Bold: Buyers' household variables**

*****: sig. <0.01; **: sig. <0.05; *: sig. <0.1**

### 3.4.2 Introduction of Socio-economic Variables Describing the Household

Three variables describing the household are significant : the household income and the previous tenure status (Models M2, M3b, N2 and N3) as well as the age of the respondent at transaction date (*under 30*) (Model M2 only).

*Ceteris paribus,*

- For each additional $10,000 of income, buyers pay an average premium of 1.61% (1.46 to 1.73%, depending on the model);
- First-time owners pay between 4.04 to 4.36% less than former owners;
- Young households, under 30 years of age, pay 3.98% less than older buyers for the same property (only model M2, and sig 0.1).

Whether Census variables – describing the socio-economic profile of the neighbourhood at the Census-tract level – are included or not in the model, the two household-level variables *Household Income* and *First-time Owners* stay significant, with similar and high t-values (t-values ranging from 7.23 to 8.47 and from 4.68 to 5.1 respectively, depending on the model). Furthermore, no significant collinearity is detected between the two levels of socio-economic measures (Census data and household data), the maximum VIF value among these variables being 3.5 (*Percentage of university degree holders in the Census tract,* model N2).

Concerning the dichotomous age variable (*Under 30*), it is present in one model only (M2), with a low significance test (*t*-value –1.75, sig. 0.1). Although it does not present any collinearity with *Household Income* or *Previous Tenure Status* as could have been expected, this variable drops out when Census data (N2) or further expansion terms are included (M3a, M3b, N3).

### 3.4.3 Adding Expansion Terms: Controlling for heterogeneity

In a last step, we introduced expansion terms allowing for the basic parameters (property specifics, accessibility, vegetation [M3a and M3b] and Census data [N3]) to vary with regard to the household profile. Several expansion terms are significant, showing that the value given to certain property specifics or location attributes is not homogeneous among buyers. Table 20 presents the list of the parameters that are heterogeneous considering the household characteristics of the buyers.

While a majority of expansion terms (15) is significant when both Census data and raw household profile variables are omitted (Model M3a), only a few drop out when these are included (12 interactive variables in both models M3a and N3). Also, some parameters are only significant when their non-stationarity is considered, as *NDVI 40 m, Woodlands 500 m* and *Agricultural Land 100 m*. These variables are not significant as such but need to be expanded to enter the model. This shows that for some attributes, estimating a unique coefficient for the whole area of study is not possible, and that the spatial variability must be considered in order to properly measure their impact.

Table 20. Synthetic table of significant expansion terms

| The value given to the… | …varies regarding the buyer's… | Age | Income | Household Type | Educational Attainment | Previous tenure status |
|---|---|---|---|---|---|---|
| Property specifics | Living area | X | | | | |
| | Living area of a bungalow | | X | | | |
| | Apparent age | | X | X | | |
| | Built-in oven | X | | | | X |
| | Fireplace | | | | | X |
| | In-ground pool | | X | X | | |
| | Detached garage | X | | X | X | |
| Centrality | Car-time to MACs | | X | | | |
| Vegetation | Mature trees 100 m | X | | | | |
| | Nb trees 29up | X | | | | |
| | Agricultural land 100 m | X | | X | X | X |
| | Woodlands 500 m | | | X | | |
| | NDVI 40 m (greenness) | X | | | | |
| Neighbourhood profile (Census) | % of Univ. degree holders | | | | X | |

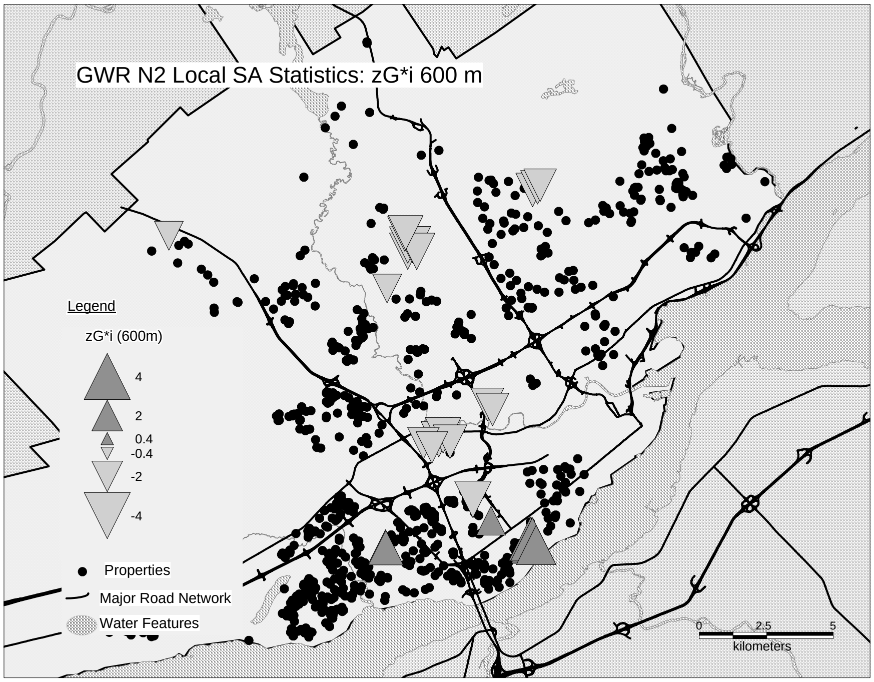
Nb: The significant buyer's household variables may vary depending on the interaction considered. For example, Age may refer to several categories (age 30-39, age 40-49, age 40 up, etc.). See Table 18 and Table 19 for complete details.

### 3.4.3 GWR Models

The variables of the four models M1, M2, N1 and N2 were introduced in four GWRs, resulting in GWR_M1, GWR_M2, GWR_N1 and GWR_N2. These models performed well, with *R*-squares ranging from 0.885 to 0.902 (see Table 17). The *F*-statistics of improvement between global and GWR models, however low (values ranging from 2.51 to 3.36), are significant.

As expected, no global autocorrelation is left in the models. Some local "hot spots" are still significant here too, but represent less than 5% of the sample (21 to 26 significant zG*i statistics for a spatial lag of 600 m.). Figure 9 shows a map of significant zG*i statistics for GWR_N2, which are presenting a similar pattern than for model N3 (Figure 8).

Figure 9. Local spatial autocorrelation: Significant zG*i statistics for GWR_N2




GWR gives the possibility of deriving local regression statistics, for example the local significance of a parameter. As GWR calculates distinct regressions for each point of the sample, the variability of the significance can be mapped. Furthermore, the non-stationarity can be tested using a Monte Carlo approach. That is, the question is to know whether the observed variation is sufficient to say that the parameter is not globally fixed. *P*-values testing for non-stationarity are given in Table 21. For the parameters with non-significant *p*-values, it is assumed that a unique coefficient holds true. The parameters that are considered non-stationary are therefore the following: *Local Tax Rate*, *Apparent Age* (Figure 10), *Car Time to MACs* (Figure 11), *NDVI Stdd. 1km* (GWR_M1 and GWR_M2), and *% Univ. Degree Holders* (GWR_N1). Also, local *R*-squares give further indication about the fit of the model depending on location. However, the value of the local *R*-square is also influenced by the stationarity of the process that is modeled. Therefore, this statistic should be interpreted with care (See Figure 12 as an example of a map of local *R*-squares for GWR_M2).

Figure 10. GWR_M1: Spatial variation of apparent age parameter

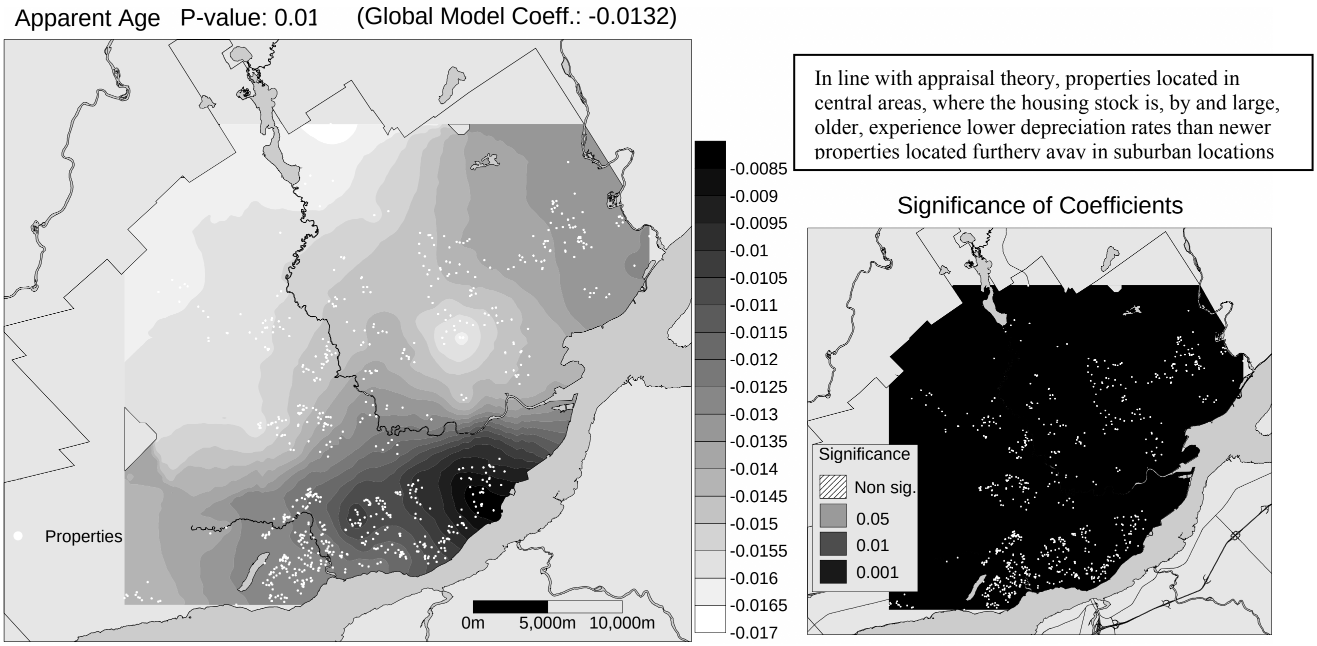
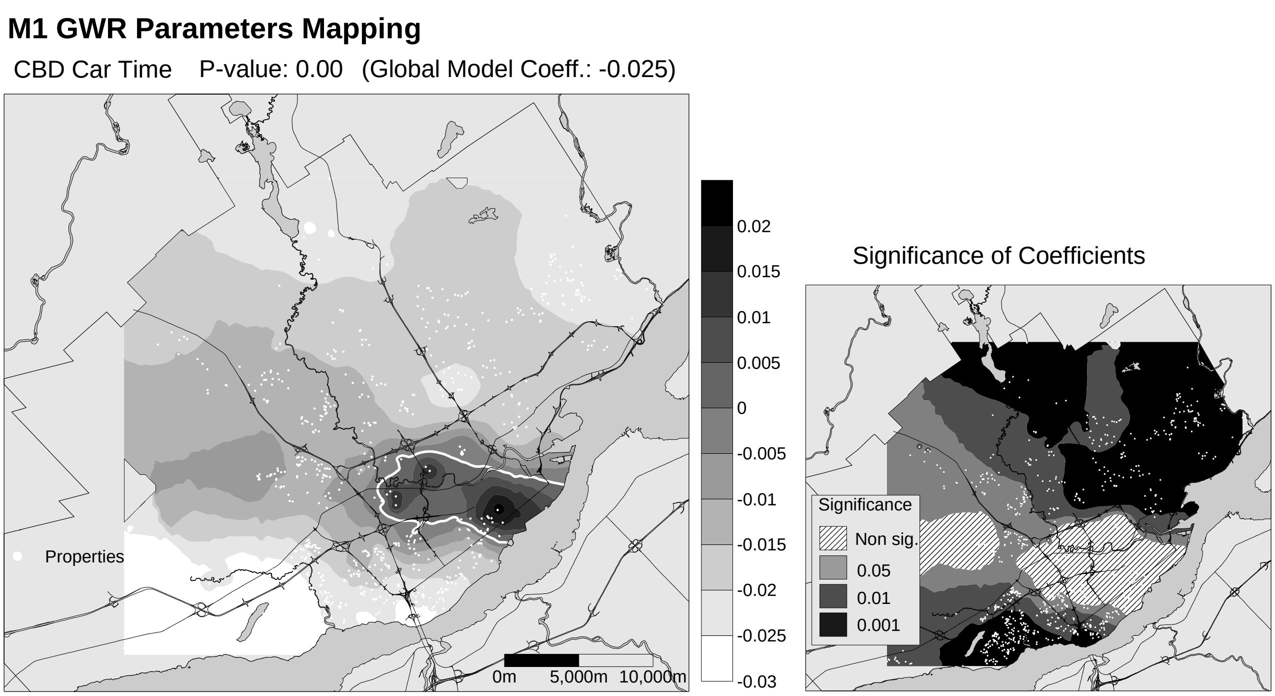
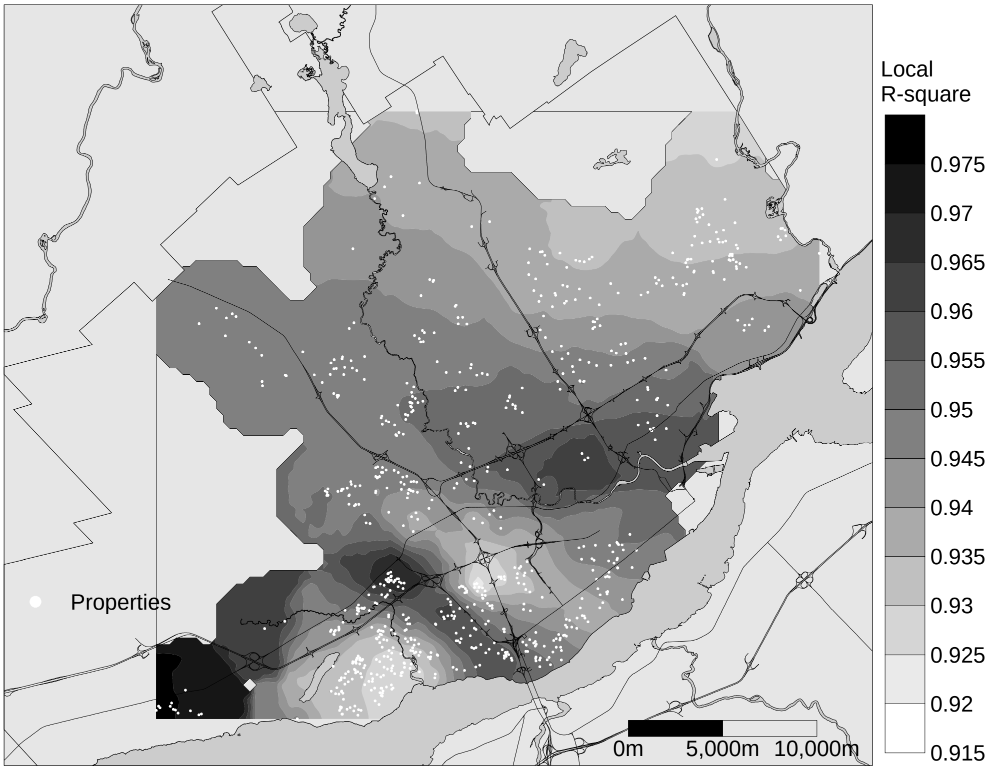



Figure 11. GWR_M1: Spatial variation of car-time to MACs coefficients




**NB: The non-significance of the coefficients in certain areas is partly due to the scarce presence of single-family properties, as for example in the more central positive-sign area.**

Figure 12. Local *R*-squares for GWR_M2




### 3.4.4 A Comparison of Global and GWR Models

Although the GWR models must be compared with their global counterpart (that is, with the global models built with the same variables, M1, M2, N1 and N2), it is also of interest to compare the GWRs with the expanded versions of the global specifications. For example, let us compare the two "drift"-sensitive versions of N2, that is N3 and GWR_N2. In both cases, the percentage of explanation of the variance is similar (0.894 for the global version, vs. 0.892 for the GWR), as is the global autocorrelation of the residuals (Moran's I values respectively 0.102 and 0.0802). Concerning the local autocorrelation, the number of significant zG*i statistics (26) is identical, although these hot spots do not strictly match spatially (See Figure 8 andFigure 9). In the end, these models are similar in terms of explanation power and for their ability to handle spatial autocorrelation.

Let us compare more precisely how these models handle heterogeneity. For N3, the coefficients that vary spatially are identified by the significant expansion terms. These expansions refer to the following variables: *Built-in Oven, Fireplace, Detached Garage, Car Time to MAC, Nb of Trees 29 up, % of Univ. Degree Holders, NDVI 40 m* (greenness)*, Woodlands 500 m* and *Agricultural Land 100 m*. The statistical significance of expansion terms indicates that for these variables, a single coefficient is not a valid alternative. In fact, we know that the impact of these variables varies according to age, income, educational attainment and type of household. However, no local measure of significance is available.

For the GWRs, the heterogeneity of the parameters is given by the *p*-values measured through a Monte Carlo procedure (Table 21). According to these *p*-values, five parameters vary significantly at the 95% confidence level for GWR_M1 and GWR_M2 (*Local Tax Rate, Apparent Age, Car Time to MACs* [linear and squared form] and *NDVI Stdd. 1 km* [heterogeneity of land use]), one for GWR_N1 (*% of Univ. Degree Holders*), and none for GWR_N2. It is interesting to note that each of these variables identified as non-stationary is also strongly spatially structured, as indicate the corresponding high Moran's I statistics (Table 21, fourth column). Also, the findings suggest that for the variables with non-significant *p*-values, a unique coefficient is adapted, that is, the implicit price is homogeneous among the observations. This is *a priori* in contradiction with the findings of the global models using expansion terms. One could argue that the heterogeneity identified in the expansion models refers to the household heterogeneity, and not specifically to *spatial* heterogeneity, as it would have been had the attributes been expanded according to their coordinates (through the use of trend surface analysis for example).

In fact, some of the variables describing the household profile are not spatially structured, as indicate the Moran's I values shown in Table 22. For the attributes that have been expanded with these "non-spatial" household characteristics, it is to be expected that they are not identified in the GWR framework as spatially heterogeneous (although other dimensions than household profile and preferences could be the cause of heterogeneity). However, both the income (*Household Income* and *Income 80K up*) and the educational attainment of the households (*University degree holders*) do present a spatial structure, with significant Moran's I values at the 95% confidence level. The attributes that are significantly expanded in the global models with these two characteristics should also be identified in the GWR models as heterogeneous, that is, with significant *p*-values. This concerns the following: *Living Area of a Bungalow, In-ground Pool, Detached Garage, Car Time to MAC* and *% of Univ. Degree Holders.* Whereas the two latter values are identified in the GWR as heterogeneous, the three former ones are not.

Table 21. Non-stationarity of parameters in GWR Models (*p*-values) and Moran's I statistic

| Parameter | *p*-value GWR-M1 | GWR-M2 | GWR-N1 | GWR-N2 | Moran's I (1500 m) |
|---|---|---|---|---|---|
| (Constant) | 0.02 | 0.06 | 0.39 | 0.32 | |
| Local Tax Rate | **0.00** | **0.00** | 0.08 | 0.10 | **0.66** |
| Living Area m2 | 0.10 | 0.17 | 0.58 | 0.53 | **0.47** |
| Living Area Bung. | 0.27 | 0.39 | 0.56 | 0.79 | **0.53** |
| Ln Lot Size | 0.07 | 0.10 | 0.16 | 0.11 | **0.94** |
| App. Age | **0.01** | **0.02** | 0.27 | 0.40 | **1.10** |
| App. Age Cd Sqd | 0.50 | 0.49 | 0.60 | 0.77 | **0.98** |
| Quality | 0.24 | 0.32 | 0.49 | 0.38 | 0.04 |
| Finished Basement | 0.26 | 0.17 | 0.35 | 0.34 | 0.08 |
| Superior Floor Qual. | 0.32 | 0.30 | 0.25 | 0.38 | **0.31** |
| Facing 51%+ Brick | 0.50 | 0.24 | -- | -- | **0.44** |
| Built-in Oven | 0.15 | 0.04 | 0.38 | 0.37 | 0.08 |
| Fireplace | 0.64 | 0.76 | 0.36 | 0.20 | **0.29** |
| In-ground Pool | 0.88 | 0.90 | 0.45 | 0.62 | 0.07 |
| Det. Garage | 0.38 | 0.48 | 0.32 | 0.59 | **0.29** |
| Att. Garage | 0.90 | 0.67 | 0.71 | 0.40 | 0.02 |
| CBD Car Time | **0.00** | **0.02** | 0.17 | 0.32 | **0.78** |
| CBD Car Time Cd. Sqd. | **0.00** | **0.00** | 0.13 | 0.21 | **0.96** |
| Highway Exit | -- | -- | 0.09 | 0.22 | **0.81** |
| Mature Trees 100 m | 0.75 | -- | 0.37 | 0.28 | **0.88** |
| Mature Trees 500 m | 0.49 | 0.32 | -- | -- | **1.09** |
| Low Tree Dens. 500 m | 0.15 | 0.08 | -- | -- | **0.60** |
| Nb Trees 29 up | 0.58 | 0.58 | 0.41 | 0.28 | 0.02 |
| NDVI Std. Dev. 1 km | **0.00** | **0.00** | -- | -- | **0.95** |
| **Household Income** | -- | 0.45 | -- | 0.41 | **0.27** |
| **First-Time Owners** | -- | 0.73 | -- | 0.97 | -0.11 |
| **Age under 30** | -- | 0.41 | -- | -- | -0.05 |
| % of Univ. Degree Holders | -- | -- | **0.01** | 0.13 | **0.80** |
| % Aged 65 up | -- | -- | 0.89 | 0.46 | **0.75** |
| Developped Area (Ha) | -- | -- | 0.11 | 0.17 | **0.81** |
| % Dwellings 1946-60 | -- | -- | 0.87 | 0.90 | **0.90** |
| % Unemployed Aged 15-24 | -- | -- | 0.22 | 0.28 | **1.30** |
| Nb Persons per Room | -- | -- | 0.66 | 0.60 | **0.59** |

**Grey boxes: Census data variables; Bold: Buyers' household variables**
**Bold: Significant at the 95% confidence level.**

Concomitantly, two variables are considered heterogeneous within the GWRs, but are not significantly expanded in the global models (*Local Tax Rate* and *NDVI Stdd. 1 km* [land-use heterogeneity]). We can assume that the heterogeneity associated with these two attributes is not related to variations in the household profiles.

Table 22. Spatial structure of the household characteristics that explain the heterogeneity of parameters (Expansion models)

| | Moran's I (1500 m) | Prob. |
|---|---|---|
| Household Income | **0.271** | **0.02** |
| Income 80K up | **0.231** | **0.04** |
| Age 30-39 | -0.109 | 0.21 |
| Age 40-49 | -0.158 | 0.12 |
| Age 40 up | -0.022 | 0.44 |
| Household with Children | -0.219 | 0.05 |
| Couple without Child | -0.053 | 0.35 |
| Single Household | 0.145 | 0.14 |
| Univ. Degree Holders | **0.360** | **0.00** |
| First-time Owner | -0.105 | 0.22 |

**Bold: significant at the 95% level.**

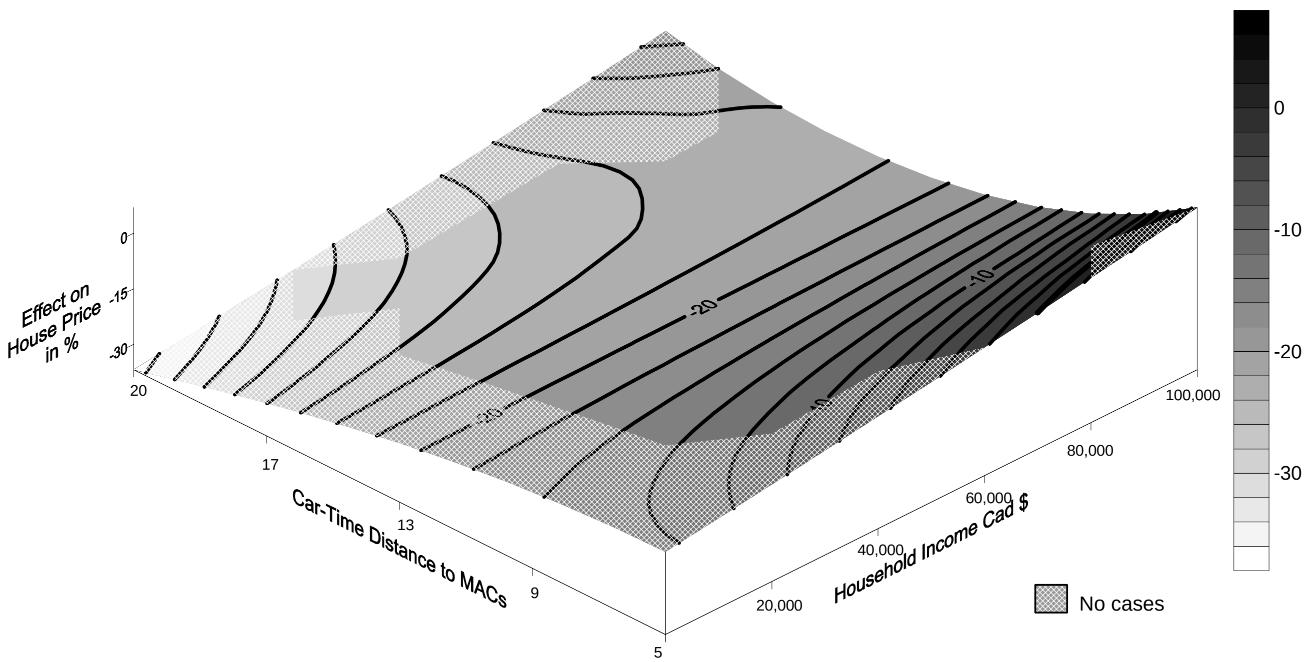
Both methods yield highly interesting results. Whereas spatial expansion makes it possible to consider both the spatial and the non-spatial heterogeneity of parameters, GWR provides interesting information through local regression statistics. However, although GWR is an interesting tool to identify and spatially describe non-stationary processes, it does not identify the cause of the parameter drift. Spatial expansion on the contrary, although less precise locally, makes it possible to integrate the cause of non-stationarity, and thereby helps disentangle the complex interactions influencing property values.

### 3.4.5 Some Provocative Findings About…

#### …Accessibility and Income

It is worthwhile to underline the significant drift of accessibility (*Car Time to MACs*, under its squared form) regarding household income. The *Car Time to MACs* is negatively linked to property values: each additional minute away from the city centre lowers the property value of 1.82%. However, this relation is not strictly linear but rather follows a U-shaped curve form, as shown by the significant integration of the squared form of the variable, with a positive sign. Furthermore, this squared term significantly interacts with the household income, with a positive sign too. This shows that the higher the income, the higher the squared term. Therefore, the devaluation associated with distance is more important for low-income than for high-income households, as shown on the three-dimensional surface of Figure 13. This tends to corroborate the distance-cost trade-off theory, stating that high-income households can afford additional transportation costs and are ready to pay more for properties located in the outer-city limits. Also, the increasing practice of telework, which particularly concerns managers and professionals, may have an effect on the propensity of the most highly educated people to locate in more remote areas of the urban scene. In fact, those who can spend some working days at home may be willing to pay more for non-central locations, thus benefiting more from premium environments than typical commuters.

Figure 13. Effect of car time distance to MACs considering household income




**…Social Homogeneity**

As foreseen, the percentage of university degree holders in the census tract has a global positive effect on the property value, each additional 10% adding a premium of 4.41%. This variable is among the most significant ones, with a *t*-value of 9.17. Additionally, the expansion with the household-level binary variable “Holding a university degree” proved significant, with an additional 1.81% premium. This shows that all things being equal, highly educated buyers who select single-family housing, are ready to pay more in their quest for social homogeneity, thereby influencing market values upward.

## 3.5 Summary and Conclusions

This paper aims at understanding how the marginal value given to property and location attributes may vary among buyers. A telephone survey was conducted in order to obtain detailed information about 761 households that acquired single-family properties in

Quebec, Canada, during the 1993-2001 period. Household-level variables were introduced into hedonic functions to measure the effect of the homebuyer's socio-economic context on implicit prices. Both the expansion method (Casetti, 1972 & 1997) and Geographically Weighted Regressions (GWR) (Fotheringham *et al.*, 2002; Fotheringham *et al.*, 1998) are used to assess the eventual heterogeneity of the impact of property specifics and location attributes.

A major finding is that certain characteristics of the buyer's household have a direct impact on property prices: this concerns the household income, the previous tenure status, and age. These findings must be put into the perspective of a specific location (Quebec City) and specific market conditions, that is, mainly a seller market with high supply and rather low demand for housing. Under these particular conditions, and using appropriate space-sensitive interaction methods, we could show that for each additional $10,000 of income, a buyer pays a premium of 1.61% on average (+1.46% to 1.73%), all other things being equal. Also, the marginal effect of the household income is the fifth most significant parameter after the size (*living area*), the age of the property (*apparent age*), the social status of the neighbourhood (*percentage of university degree holders in the Census tract*), and accessibility (*Car Time to MACs*) (N3). Several hypotheses can explain the parameter significance and its positive sign. First, it is possible that the lack of descriptors defining the luxury attributes of the higher segment of the property market may result in a premium appearing as associated to the buyer's income. However, as their ability to pay is increased, high-income buyers may also be less willing to engage in lengthy price negotiations, and may accept higher selling prices. Concomitantly, households with more restricted financial means may take more time to find the "best" deal as their budget is inflexible. While taking more time, they may visit more houses and thereby increase their chances to find sellers who on the contrary, have time constraints, and may want to sell rapidly. It would be interesting to obtain information about the seller's profile, which could also impact on the property value. These findings should be compared with information on the time elapsed between the decision to look for a piece of property and the actual act of buying one. It is probable that potential buyers who are well off may be more prone to materialise their housing needs as budget constraints do not represent a serious impediment. Furthermore, the argument that the property's price (as well as the desire to make an investment) was a criterion for buying the property is significantly more frequent on the part of low-income households (See Kestens *et al.*, Submitted).

First-time owners, that is, households that were previously tenants, "save" an average of 4.2% (3.88 to 4.18%, depending on the models) compared with former-owner households, all other things being equal. Again, first-time buyers may obtain a better price by waiting longer to close a deal, and former owners can afford a more substantial down-payment due to the sale proceeds from the previous home.

The age variable did enter in as such in one of the models (M3a), however with a low *t*-value. Furthermore, this criterion was dropped when additional expansion terms or Census data were included. Some collinearity may still be at stake here, and any direct interpretation about the direct link between age and price is therefore risky.

The integration of numerous expansion terms shows how the marginal value of certain property specifics and location attributes varies with the household profile. These findings partly complete Starret's statement (Starret, 1981). He hypothesised that capitalisation of an attribute is only complete if the residents' preferences are homogeneous. In fact, the significant drift of parameters according to the household characteristics shows that the capitalisation of an attribute *does vary* according to the household profile. Certain characteristics of the household profile are also significantly linked to the odds-ratio of mentioning certain property or neighbourhood choice criteria (See Kestens *et al.*, Submitted), that is, to the household preference, as far as the choice criteria can be interpreted as a proxy for preference. Certain choice criteria are difficult to translate into measurable determinants of value. In fact, among those choice criteria for which the odds-ratio of being mentioned is linked to the household profile, few find their equivalent as expansion terms. For example, among the neighbourhood choice criteria, the odds-ratio of mentioning "Proximity to services" is significantly linked to age, household type, or income. Educational attainment has no impact on the propensity to mention this criterion. However, this paper suggests that the drift of the value assigned to accessibility to the main activity centres (MACs) is linked to educational attainment, and not age, household type, or income. Similarly, this paper shows that the value given to vegetation in the close surroundings of the property varies significantly with age (*Nb Trees 29 up* expanded by *Age 40 and over* and *NDVI 40 m* expanded by *Age 30-39*). However, the odds-ratio to mention the presence of trees as a choice criterion is not linked to age but to the previous tenure status and the household type (for trees in the neighbourhood) and educational attainment and income (for trees on the property).

Although this paper has stressed that the marginal value of certain attributes varies with the household profile, the links between the coefficient's drift and preference (or choice criteria) need further exploration. Straightforward relations between stated choice criteria and heterogeneity of implicit prices could not be established.

More specifically, two significant expansion terms are worth underscoring. The first shows that the marginal value of accessibility varies with the household's income. Whereas the location rent is linearly negative for low-income households, it has more of the form of a U-shaped curve for high-income households, who tend to add a premium to remote locations, *ceteris paribus.* The recent and growing development of telework may be part of the explanation. In the U.S., home-based telework has grown nearly 40% since 2001, concerning some 23.5 million employees in 2003 (Pratt, 2003). In Canada, the 2001 Census reported some 8% of teleworkers. Furthermore, a recent study showed that, out of a sample of salaried teleworkers working at home and using information technology like the internet, 60.6% hold a university degree (Tremblay, 2003), this number being far over the national average (22.6% [Statistics Canada, 2001]). This paper's findings are coherent with the hypothesis that highly-educated teleworkers are prepared to pay a premium for remote locations, as compared with daily commuters. Additional research is needed, however. The insertion within the hedonic framework of Origin-Destination survey data, which procures detailed information on work, shopping and leisure trips, could further our understanding of this phenomena. In fact, the concomitant development of Information Technology and the trend toward more balanced relations between work and family redefines our notions and limits of space and location.

The interaction, too, between the effect of the percentage of university degree holders in the Census tract and the educational attainment of the buyer provides insight into social homogeneity processes. With a positive sign, this parameter indicates that highly-educated households do pay a premium to fulfill their seek for social homogeneity. This partially confirms Goodman and Thibodeau's (2003) hypothesis, that "Higher income households may be willing to pay more for housing (per unit of housing services) to maintain neighbourhood homogeneity" (p. 123). This paper showed it to be true regarding educational attainment, and not directly the household income, although these two dimensions are correlated.

Methodologically, the two methods that were used proved efficient. Expansion terms make it possible to analyse and to fully explain the cause of the parameter heterogeneity, whether its structure be spatial or not. Geographically Weighted Regressions provide additional insight by measuring local regression statistics. Some inconsistencies about non-stationary parameters were detected and need further investigation. However, we feel that both methods are complementary rather than substitutes for each other, and that the use of additional methods like Seemingly Unrelated Regressions (SUR) (Knight *et al.*, 1995; Zellner, 1962) may further our understanding of the complexity of property markets and urban dynamics.

# References


Akaike H, 1973 "Information Theory and an extension of the maximum likelihood principle", in *2nd Symposium on Information Theory,* Eds B Petrov, F Csaki *(*Akadémiai Kiado, Budapest) pp 267-281.

Anselin L, 1988 *Spatial Econometrics: Methods and Models* (Kluwer Academic Press, Dordrecht)

Anselin L, 1995, "Local Indicators of Spatial Association - LISA" *Geographical Analysis* **27** 93-115

Aten B, Heston A, Forthcoming, "Regional Output Differences in International Perspective", in *Spatial Inequality and Development* Eds R Kanbur, A J Venables (World Institute for Development Economics Research (WIDER), London)

Brunsdon C, Fotheringham A S, Charlton M, 1999, "Some notes on parametric significance tests for geographically weighted regression" *Journal of Regional Science* **39** 497-524

Brunsdon C, Fotheringham A S, Charlton M, 2002, "Geographically Weighted Summary Statistics - A Framework for Localised Exploratory Data Analysis" *Computers, Environment and Urban Systems* **26** 501-524

Brunsdon C F, Fotheringham A S, Charlton M E, 1996, "Geographically Weighted Regression: A Method for Exploring Spatial Non-Stationarity" *Geographical Analysis* **28** 281-298

Can A, 1990, "The Measurement of Neighbourhood Dynamics in Urban House Prices" *Economic Geography* **66** 254-272

Can A, 1992, "Specification and Estimation of Hedonic Housing Price Models" *Regional Science and Urban Economics* **22** 453-477

Casetti E, 1972, "Generating Models by the Expansion Method: Applications to Geographical Research" *Geographical Analysis* **4** 81-91

Casetti E, 1986, "The dual expansion method: an application for evaluating the effects of population growth on development" *IEEE Transactions on Systems, Man, and Cybernetics* **16** 29-39

Casetti E, 1997, "The Expansion Method, Mathematical Modelling, and Spatial Econometrics" *International Regional Science Review* **20** 9-32

Cliff A D, Ord J K, 1981 *Spatial Processes, Models and Applications* (Pion Ltd., London)

Craven P, Wahba G, 1979, "Smoothing Noisy Data with Spline Functions" *Numerische Mathematik* **31** 377-403

Des Rosiers F, Thériault M, Kestens Y, Villeneuve P, 2002 *Landscaping Attributes and House Prices: Looking at Interactions Between Property Features and Homeowners' Profiles* Faculté des Sciences de l'Administration, Université Laval, Québec

Duda R O, Hart P E, 1973 *Pattern Classification and Scene Analysis* (Wiley and Sons, New York, NY)

Fotheringham A S, Brundson F C, Charlton M, 2002 *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships* (John Wiley & Sons, Chichester)

Fotheringham A S, Charlton M E, Brundson C F, 1998, "Geographical weighted regression: A natural evolution of the expansion method for spatial data analysis" *Environment and Planning A* **30** 1905-1927

Geoghegan J, Wainger L A, Bockstael N E, 1997, "Spatial Landscape Indices in a Hedonic Framework: An Ecological Economics Analysis Using GIS" *Ecological Economics* **23** 251-264

Getis A, 1995, "Spatial Filtering in a Regression Framework: Experiments on Regional Inequality, Government Expenditures, and Urban Crime", in *New Directions in Spatial Econometrics* Eds L Anselin, R J G M Florax (Springer Verlag, Berlin) pp 172-188

Getis A, Griffith D A, 2002, "Comparative Spatial Filtering in Regression Analysis" *Geographical Analysis* **34** 130-140

Getis A, Ord J K, 1992, "The Analysis of Spatial Association by Use of Distance Statistics" *Geographical Analysis* **24** 189-206

Goodman A C, Thibodeau T G, 1998, "Housing market segmentation" *Journal of Housing Economics* **7** 121-143

Goodman A C, Thibodeau T G, 2003, "Housing Market Segmentation and Hedonic Prediction Accuracy" *Journal of Housing Economics* **12** 181-201

Griffith D A, 1988 *Advanced Spatial Statistics: Special Topics in the Exploration of Quantitative Spatial Data Series* (Kluwer, Dordrecht)

Griffith D A, 1996, "Spatial Autocorrelation and Eigenfunctions of the Geographic Weights Matrix Accompanying Geo-Referenced Data" *The Canadian Geographer* **40** 351-367

Hope A C A, 1968, "A Simplified Monte Carlo significance test procedure" *Journal of the Royal Statistical Society Series B* **30** 582-598

Hurvich C M, Simonoff J S, Tsai C-L, 1998, "Smoothing parameter selection in nonparametric regression using an improved Akaike Information Criterion" *Journal of the Royal Statistical Society Series B* **60** 271-293

Jaccard J, Turrisi R, Wan C K, 1990 *Interaction Effects in Multiple Regression* (Sage Publications, Newbury Park, CA)

Kelejian H, 1995, "Spatial Autocorrelation: A suggested alternative to the autoregressive model", in *New directions in spatial econometrics* Eds L Anselin, R J G M Florax (Springer Verlag, Berlin)

Kestens Y, Thériault M, Des Rosiers F, At Press, "The Impact of Land Use and Surrounding Vegetation on Single-Family Property Prices" *Environment and Planning B: Planning and Design* **At Press**

Kestens Y, Thériault M, Des Rosiers F, Submitted, "Why families move and what they choose: An analysis of single-family property buyers" *Urban Studies* **Submitted**

Knight J R, Dombrow J, Sirmans C F, 1995, "A Varying Parameters Approach to Constructing House Price Indexes" *Real Estate Economics* **23** 187-205

Lancaster K J, 1966, "A New Approach to Consumer Theory" *Journal of Political Economics* **74** 132-157

Leung Y, Mei C-L, Zhang W-X, 2000, "Statistical tests for spatial nonstationarity based on the geographically weighted regression model" *Environment and Planning A* **32** 9-32

Loader C, 1999 *Local Regression and Likelihood* (Springer, New York)

Moran P A P, 1950, "Notes on Continuous Stochastic Phenomena" *Biometrika* **37** 17-23

Ord J K, Getis A, 1995, "Local Spatial Autocorrelation Statistics" *Geographical Analysis* **27** 286-306

Pandit K, Casetti E, 1989, "The shifting patterns of sectoral labor allocation during development: Developed versus developing countries" *Annals of the Association of American Geographers* **79** 329-344

Payne B R, 1973, "The Twenty-Nine Tree Home Improvement Plan" *Natur. Hist.* **82** 74-75

Pratt J, 2003 "Teleworking Trends: Frome Sneakernet to Internet", in *ITAC's Annual Conference,* Baltimore)

Rosen S, 1974, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition" *Journal of Political Economy* **82** 34-55

Schwartz G, 1978, "Estimating the dimension of a model" *The Annals of Statistics* **6** 461-464

Starret D A, 1981, "Land Value Capitalization in Local Public Finance" *Journal of Political Economy* **89** 306-328

Statistics Canada, 2001, *Educational Attainment, 2001 Census*. Statistics Canada, 2003,

Thériault M, Des Rosiers F, Kestens Y, Villeneuve P, 2003, "Modelling Interactions of Location with Specific Value of Housing Attributes" *Property Management* **21** 25-62

Thériault M, Des Rosiers F, Vandersmissen M H, 1999, *GIS-Based Simulation of Accessibility to Enhance Hedonic Modelling and Property Value Appraisal*. Joint IAAO-URISA 1999 Conference, New Orleans, 151-160

Tremblay D-G, 2003 *Telework: A New Mode of Gendered-Segmentation? Results from a Survey in Canada* Chaire Bell en Technologies et Organisation du Travail, Télé-Université, Québec

Tucker C J, 1979, "Red and Photographic Infrared Linear Combinations for Monitoring Vegetation" *Remote Sensing and Environment* **8** 127-150

Tueller P T, 1989, "Remote Sensing Technology for Rangeland Management Applications" *Journal of Range Management* **42** 442-452

Tyrvainen L, 1997, "The Amenity Value of the Urban Forest: An Application of the Hedonic Pricing Method" *Landscape and Urban Planning* **37** 211-222

Wu J, Gao W, Tueller P T, 1997, "Effects of Changing Spatial Scale on the Results of Statistical Analysis with Landscape Data: A Case Study" *Geographic Information Sciences* **3** 30-41

Zellner A, 1962, "An Efficient Model of Estimating Seemingly Unrelated Regressions and Testing for Aggregation Bias" *Journal of the American Statistical Association* **57** 348-368

# General Conclusion

This thesis combines various statistical and geographical tools within defined theoretical frameworks in order to better understand the complex links between residential choice, residential markets and urban structure. Through the analysis of residential markets, it is possible to determine the impact of land use and vegetation on property values, thereby approximating people's preference regarding their home environment. In order to further understand the impact of externalities on property values and residential choice, precise information at the household level concerning over 800 actual buyers is gathered through an extensive phone survey. First, motivations for moving and residential and neighbourhood choice criteria are analysed using correspondence analysis and logistic regressions. The links between the household profile and the choice criteria regarding previous tenure status, age, income, educational attainment, household type and actual location are observed. Then, the household-level data is introduced within the hedonic modelling framework, in order to measure the possible heterogeneity of implicit prices regarding the household profile. It appears that the marginal impact of certain property specifics and externalities varies with the buyer's socio-demographic characteristics. Furthermore, income and previous tenure status have a direct impact on the price at which a property is sold. Although the profile of the buyer's household could be related to choice criteria and heterogeneity of implicit prices, the links between the latter two have not been established.

At the very beginning of this research, the main objective was to link the vegetation and the visual quality observed from and around the property to the market values and the residential choice criteria, using GIS and 3D spatial analysis tools. However, the literature showed that although theoretical frameworks about the visual quality of landscapes have been established, appropriate modelling tools which could accurately measure the multiple dimensions of a visible landscape were still lacking. Furthermore, precise 3D databases, such as for example those resulting from LIDAR (LIght Distance And Ranging) surveys, were at that time not available for the area of study at reasonable costs.

The landscape and its aesthetic valuation have been the object of much research, as demonstrated by the abundant literature both in psychology and geography, ranging from spatial cognition – that is, how people integrate visual information – to landscape valuation – that is, establishing an objective quantification of inherently subjective qualitative criteria, aesthetics. Here again, the core of the subject consists of observing the interaction between human beings and their environment. Wong (cited in Han, 1999) defines landscape valuation as “a general conceptual and methodological framework for describing and predicting how attributes of environments are related to a wide range of cognitive, affective, and behavioural responses.” Although the task to integrate the visual landscape quality in a hedonic framework appeared at that moment illusory, it was possible to integrate one of the central components of the quality of landscape, namely vegetation. In fact, vegetation plays a particular role in the urban landscape. Its living, changing and moving capacities clearly contrast with the unchanging and mineral aspects of humanly built objects.

One objective of this thesis being to develop efficient modelling methods, the measure of vegetation was done within a GIS. Two approaches were tested. The first uses colour aerial photographs, whereas the second is based on a Landsat-TM5 image. The aerial photographs were computerised and combined into a continual mosaic, which was then manually categorised into 20 land-use and -cover classes. The remote sensing image was semi-automatically categorised using the ISODATA technique, resulting in some 9 land-cover types, 6 of which related to vegetation. Furthermore, the NDVI (Normalised Differentiated Vegetation Index) could be computed in order to estimate the density of vegetation, as well as the homogeneity of the land use, using the standard deviation of the NDVI. In order to integrate the theoretical cognitive framework of hierarchical space, several measures of the property’s surrounding land use were computed within the GIS for various distance lags. This information was then integrated into two series of hedonic models, revealing the positive impact of mature trees and land-use heterogeneity, the negative impact of forests, agricultural land and low tree density, for various spatial scales. Furthermore, it appeared that (i) the effect of vegetation varies regarding the relative centrality to the main activity centre; and (ii) as locational externalities relating to land use and vegetation are integrated in the models, the effect of certain attributes, among others the distance to the city centre, may change. This first part of the thesis showed that land use and vegetation have a significant impact on property values, and that this impact may vary spatially. From this point, and seeing that the effect of a locational externality varies through space, it seemed important to verify whether the perception and the residential choice criteria of the property buyers were homogeneous, and if not, whether these preferences and choice criteria could somehow be linked to the households’ characteristics.

An important phone survey was therefore realised, asking over 800 buyers who bought their property in the 1993-2001 period to mention their motivation for moving, their property and neighbourhood choice criteria, and procuring additional socio-demographic information on their household. The residential choice criteria were introduced in a correspondence analysis, which summarised the dataset within eight main factors. The interpretation of these factors could be related to the Place-Proximity-Space and to the Place-Identity theoretical frameworks. Also, logistic regressions showed that the odd-ratios of mentioning a criterion vary with certain characteristics of the household profile and the buyer's actual location. This research confirmed and extended the pioneering findings of Rossi (1955) who first underscored the relation between life-cycle factors, motivations for moving and residential choice. Significant differences in choice criteria could be related to the location of the buyer, and to the buyer's feeling of attachment to the neighbourhood. Having said that, and bearing in mind that one of the objectives of this research aims at better understanding the interaction between people's choices (actions) and the actual spatial structure of the urban setting, it seemed important to verify whether the residential market would reflect the heterogeneity observed among people's choice criteria. These choice criteria can be considered as proxies to people's preference, in the same way that the implicit prices are measured through hedonic analysis. Therefore, the characteristics of the buyers' households were introduced in hedonic models. Also, using two spatial-sensitive methods, the heterogeneity of the implicit prices was estimated regarding the age, the income, the educational attainment, the household type and the previous tenure status of the buyer. The Casetti expansion method made it possible to measure the drift of any property-specific or locational attribute regarding these socio-demographic characteristics. Geographically Weighted Regressions provided additional information on the spatial heterogeneity of the parameters. Major results indicate that the income and the previous tenure status of the buyer have a direct impact on the property value. Also, the marginal value of several property specifics and externalities varies with the characteristics of the buyer. These findings, although challenging the traditional interpretation of the hedonic function, corroborate various statements on the existence of sub-markets (Goodman and Thibodeau, 2003), which may lead to spatial heterogeneity in implicit prices.

In short, the main findings of this thesis can be summarised as follows:

- Land use and vegetation have a significant impact on property values, and this impact varies through space, mainly regarding relative centrality. The use of remote-sensing data, integrated within a GIS, proved efficient. Also, the measure of vegetation at various spatial scales could partly integrate the hierarchy of spatial perception.
- Whereas the impact of environmental attributes varies through space, the motivations for moving and the property and neighbourhood choice criteria vary regarding the socio-economic and life-cycle factors of the buyer's household. Correspondence analysis could identify the main residential choice factors, and relate them to the Place-Proximity-Space and Place-Identity conceptual frameworks.
- Finally, the buyers' household characteristics are partly linked to the heterogeneity of implicit prices. In order to measure this social non-stationarity, both Casetti's expansion method and Geographically Weighted Regressions proved efficient and complementary.

However, and this is an apparent limit of this research, the findings concerning the choice criteria/household relation and those concerning the heterogeneity of the implicit prices with regard to the household characteristics are not fully concordant. In fact, among the choice criteria for which the odds-ratio of being mentioned is linked to the household profile, few find their equivalent as expansion terms. This apparently contradictory result suggests that additional research is needed to better understand the links between what people think and say they prefer or value, and the actual sense of their actions. In our case, the respondents mentioned the criteria for choosing their actual property, and the real value they paid for and a complete description of the property's specifics and externalities were then used in the hedonic context. Biases which could have been resulting from the observation of stated choices were therefore avoided. However, the links could not be clearly established, and further analysis is required. It would be highly instructive to integrate additional information on the location of the working places of property buyers. In fact, centrality was here integrated as the distance to the Main Activity Centres, which is roughly the Laval University-Historic Downtown axis. However, previous findings by Vandersmissen *et al.* (2003) could demonstrate that extra-centre commuters – that is, people traveling from a suburban home to a suburban workplace – make longer trips than workers who commute to the city centre. Therefore, and keeping in mind that getting closer to the working place is a frequently cited motivation for moving, additional understanding would be achieved if considering both residential and workplace location. However, the "dispersal of job opportunities has created a much more complicated behavioral response to the linkage between work and residence" (Clark *et al.*, 2003). Furthermore, the spatial equation becomes even more intricate concerning dual-earner households (Timmermans *et al.*, 1992; Green, 1997).

This thesis has explored the determinants of residential markets by observing on one hand the moving motivations and residential choice criteria of actual property buyers, and on the other hand by measuring the impact of environmental and socio-demographic attributes on property values through the use of spatial-sensitive modelling techniques. Special attention was paid to properly integrate the spatial dimension, at each step of the research. The combination of GIS and econometric statistical modelling techniques proved very effective. However, as it clearly appeared during the first phase of the research, the availability and accuracy of spatial data is a strong requirement in order to be able to properly model any geographical phenomena. There is no doubt that with the ongoing development of spatial databases, the consideration of space within econometric models will grow. Furthermore, as it appeared with the successful integration of the Landsat-TM5 derived data, certain low-cost methods are efficient and could easily be used by planning agencies for valuation or policy purposes. Concerning the integration of the three dimensions of space, further advances are still needed. As stated earlier, the impact of the aesthetic quality of the visible landscape could not be properly measured because of the lack of data. The quality of the urban landscape is however central to our quality of life and would deserve undivided attention.

As it appeared in this research, simple straightforward links are not the rule. Most frequently, interaction effects make it difficult to get a clear and simple picture of the observed phenomena. The complex interactions occurring through space, time and people must therefore not be neglected, as streamlined representations of real-world phenomena lead to biased interpretations and misunderstanding. According to Lancaster (1966), “Goods in combination may possess characteristics different from those pertaining to the goods separately”… and this applies to all attributes of life. In order to better understand the spatial challenges we face in our cities, the aspects that should be considered simultaneously are of multiple nature, from housing to social structure, from transportation to pollution, from health issues to education. Various space-sensitive tools have been developed lastly in order to better integrate the spatial component that is common to these geographical phenomena. Geographically Weighted Regressions, multi-level analyses and other techniques offer promising avenues for future spatial-sensitive and multi-disciplinary research.

Finally, although econometric modelling is a powerful tool to better understand the world we live in, it seems important to mention that issues beyond economic interests have to be considered, although they may be in conflict with people's desire and hedonism. Indeed, it might be "falsely assumed that value can be reduced to price" (Shiva, 1996).

# References

Clark W A V, Huang Y, Withers S, 2003, "Does Commuting Distance Matter? Commuting Tolerance and Residential Change" *Regional Science and Urban Economics* **33** 199-221

Goodman A C, Thibodeau T G, 2003, "Housing Market Segmentation and Hedonic Prediction Accuracy" *Journal of Housing Economics* **12** 181-201

Green A E, 1997, "A Question of Compromise? Case Study Evidence on the Location and Mobility Strategies of Dual Career Households" *Regional Studies* **31** 641-657

Han K-T, 1999, "A Proposed Landscape Assessment Framework: A Connection of Theories and Practical Techniques" *Journal of Architectural and Planning Research* **16** 313 - 327

Lancaster K J, 1966, "A New Approach to Consumer Theory" *Journal of Political Economics* **74** 132-157

Rossi P H, 1955 *Why Families Move: A Study in the Social Psychology of Urban Residential Mobility* (Free Press, Glencoe, Illinois)

Shiva V, 1996, "Viewpoint: Values Beyond Price" *Our Planet, UNEP* **8**

Timmermans H, Borgers A, Van Dijk J, Oppewal H, 1992, "Residential Choice Behaviour of Dual Earner Households: A Decompositional Joint Choice Model" *Environment and Planning A* **24** 517-533

Vandersmissen M-H, Villeneuve P, Thériault M, 2003, "Analyzing Changes in Urban Form and Commuting Time" *The Professional Geographer* **55** 446-463

# List of Acronyms

| | |
|---|---|
| CBD | Central Business District |
| CSF | Census Factor |
| CTD | Car-Time Distance |
| GWR | Geographically Weighted Regression |
| IR | Infrared |
| MAC | Main Activity Center |
| NDVI | Normalized Differentiation Vegetation Index |
| OLS | Ordinary Least Squares |
| PCA | Principal Component Analysis |
| PR Model | Perceptual Region Model |
| SPP Model | Space Proximity Place Model |
| VIF | Variance Inflation Factor |
| VIS | Visible |