UNDERSTANDING DIFFERENCES IN RESIDENTIAL PREFERENCES BETWEEN OWNERSHIP AND RENTING WITH CONSIDERATION OF CHANGES OVER TIME: A CASE STUDY OF LONDON

Md Bashirul Haque, Corresponding Author
Choice Modelling Centre and
Institute for Transport Studies
University of Leeds, UK
36-40 University Road, Leeds LS2 9JT
Email: tsmbh@leeds.ac.uk

Charisma Choudhury
Choice Modelling Centre and
Institute for Transport Studies
University of Leeds, UK
36-40 University Road, Leeds LS2 9JT
Email: C.F.Choudhury@leeds.ac.uk

Stephane Hess
Choice Modelling Centre and
Institute for Transport Studies
University of Leeds, UK
36-40 University Road, Leeds LS2 9JT
Email: s.hess@its.leeds.ac.uk

ABSTRACT

Residential location choices can be long-term (ownership) or medium-term (renting). Most previous studies have however investigated ownership and/or renting decisions in isolation and there is also a distinct research gap in terms of comparing the decisions made at different points in time. This research aims to fill this research gap by investigating the similarities and differences between long-term and medium-term residential choices made by residents of London and also looks at changes in household preferences over time. The Greater London Area, where there is a 56%-40% split in ownership and renting market is considered as the study area. The models are estimated combining the London Household Survey Data, Ward Atlas Data and the London Transport Studies Model outputs. The results indicate that while there are some common factors affecting both long and medium term choices, the sensitivity to commute distance, distance from the central business district (CBD), residential land area, school quality and ethnic preferences are significantly different in the two groups. This leads to the development of a pooled model which is applied to investigate the changes in the sensitivities over time. Comparison of the model parameters across four time periods reveals that there are significant changes in sensitivity towards most of the attributes over time with clear trends in case of sensitivities towards commute distance, distance from the CBD, cost and public transport accessibility. The results provide important insights about land-use and transport policy and planning.

Keywords: Owners, Renters, Long-Term, Medium-Term, Commuters

INTRODUCTION

Due to rapid urbanization and growing transport demand, integrated land use and transport planning with long planning horizon is becoming a key interest of planners and policy makers. Understanding residential location choice is a key element of integrated planning because it has a substantial influence on land use patterns and transport demand. Segregated land use patterns where the residential location is far from other facilities like shopping, healthcare, education and employment results in long commute and non-commute trips and very often leads to increased car dependency (Næss, 2009). Compact development, on the other hand, promotes increased use of transit resulting in healthy urban environments (Brown and Werner, 2008; Ewing and Cervero, 2010; Farber and Li, 2013). Better understanding of the factors driving residential location choices and quantifying the sensitivities is thus an essential pre-requisite for such sustainable land-use planning schemes.

Due to long-term investment and high relocation costs, residential ownership is typically considered to be a long-term decision and has been the subject of extensive research (Bhat and Guo, 2004; Guevara, 2010; Zolfaghari, 2013; Habib and Miller, 2009). However, renting, which is typically a medium term decision due to a higher level of flexibility as a result of lower relocation costs, shorter tenures and duration of agreements etc., also has a substantial share of the residential market especially in large cities. For instance, the average tenure length of owners (living in their current home) in England is 11 years whereas private renters are much more mobile with an average tenure length of around one year, though the tenure length is longer (seven years) in social renting (Randall, 2011).

While both types of residential choices are affected by factors such as household income and other socio-demographics, lifestyle preferences and work stability, the differences in the time scales are likely to lead to differences in the decision process and differing sensitivities towards the influencing factors. For example, owners are likely to consider location attributes (e.g. land use mix, density, green space etc.) to be more important than renters. Factors such as dwelling characteristics, on the other hand, may be important to both owners and renters. Furthermore, the sociodemographic characteristics of these two groups of decision makers are found to be considerably different which also contributes to the heterogeneity in preferences. For instance, high and middle-income households are more likely to be able to afford to own properties while the rest may be more likely to rent (Yates and Mackay, 2006). Due to the high-income range of owners, average car ownership of owners is likely to be higher than renters resulting in different sensitivities to factors such as commute distance and transit accessibility. However, most of the previous residential location choice models have focused on either residential ownership decision (Bhat and Guo, 2004; Guevara, 2010; Zolfaghari, 2013; Habib and Miller, 2009) or renting (Hoshino, 2011). Some studies looked at both but had very limited scope to capture the households' true preference for a wide range of attributes (Liao et al., 2015). The absence of a comprehensive framework for evaluating factors affecting ownership and renting decisions poses a challenge in developing an effective land-use and transport plan.

Further, the sensitivity to different factors affecting residential location choice has been found to change over time due to lifestyle trends, technological advancements and socio-demographic changes. For instance, the average commute distance in the UK has been observed to increase by 12% between the year of 2001 to 2011, reflecting the decreasing trend of commute distance sensitivity over the time (Gower, 2014). Transit oriented mixed land use pattern is also becoming popular rather than living in car dependent suburban areas (Burda, 2014). Ignoring the trends in changes in sensitivity over time can potentially lead to inaccurate model forecasts and inappropriate policy analysis. Although there have been a number of studies on lifestyle (Brown and Werner, 2008; Ewing and Cervero, 2010) and cohort effects (Næss, 2009) on residential location choice, to the best of our knowledge, there has not been any study which systematically investigates how the sensitivities to different factors influencing residential location choice have changed over time.

This research is aimed at addressing these two research gaps by investigating the following research questions:

1. Are there significant differences in long-term (ownership) and medium-term (renting) decisions?
2. Can these decisions be combined in a single model framework?
3. Are there distinct trends in changes of sensitivities towards certain attributes over time?

We answer these questions by developing a revealed preference (RP) based residential location choice model to study the residential ownership and renting decisions of households living in Greater London Area (GLA), where 56 % of households live in owned properties, 40% live in rented properties, while the rest live in shared accommodation (Census, 2001).

The rest of the paper is organized as follows. We start with a literature review followed by data description. The model structure, estimation results are presented after that followed by the trend analyses, the conclusions and directions of future research.

LITERATURE REVIEW

Significant methodological and analytical improvements have been achieved in residential location choice modeling over the last few decades, focussing on choice set generation and sampling of alternatives (Zolfaghari, 2013; Guevara, 2010), the treatment of complex correlation structures (Bhat and Guo, 2004), and endogeneity correction (Guevara, 2010), to name just a few. Due to the availability of high spatial resolution data and computational efficiency, several attempts have been made recently to develop more disaggregate (parcel or dwelling) level residential location choice models. Lee et al. (2009) developed a parcel (a unit piece of land) level residential location choice model using the Puget Sound Region data. They used Multinomial Logit models and adopted a Time Space Prism (locations of an individual over time) approach for measuring accessibility in the residential location choice model. Random sampling techniques were used to generate the choice sets of the alternatives. Lee and Waddell (2010) applied an improved technique to develop a parcel based model using the same data source. They developed a two-tier nested logit model and corrected the log sum of the maximum likelihood estimator to correct for sampling bias. Parcel level models have limitations in capturing the variation of different dwellings within a parcel. For example, the basement floor of a multi-storey building might be less attractive than other floors and a parcel level model could not capture this dissimilarity. Zolfaghari (2013) developed a zone based dwelling level residential location choice model for GLA to overcome the limitation of parcel level model. He used London Household Survey Data (LHSD) and applied a dwelling synthesizing approach for universal choice set generation. Spatial behavioral consideration during the residential location choice was also captured in previous research. Ibraimovic and Hess (2016) developed a Stated Preference (SP) based residential location choice model to explore the household preference structures for a neighborhood's ethnic composition. They found that individuals react negatively to decreases in the share of their co-nationals in the neighborhoods while being indifferent to increases. Differences in commute travel time sensitivity between male and female working members in a household in their residential location choices were investigated by Sermons and Koppelman (2001).

Despite numerous contributions aimed at capturing different methodological and analytical issues in residential location choice modeling, the influence of tenure (ownership and renting) on residential location choice has remained a relatively unexplored area of research. Although several attempts have been made to model the joint choice of tenure and dwelling (Boehm, 1982; Cho, 1997; Skaburskis, 1999; Ho and Hensher, 2014), residential location choice has not been considered in this work. However, the choice of tenure, dwelling and residential location are interdependent on each other (Ho and Hensher, 2014). Studies on joint tenure and dwelling choice have found that income is one of the most important determinants of tenure type choice and a higher level of income increases the probability of owning a house (Boehm, 1982; Cho, 1997; Skaburskis, 1999; Ho and Hensher, 2014). The only attempt at joint estimation of tenure, dwelling and residential location choice was made by Yates and Mackay (2006). Household socio-demographic characteristics and housing cost have been considered as explanatory variables to model the household choice in inner and outer Sydney for both owning and renting a house. Liao, et al. (Liao et al., 2015) developed a latent class model to estimate the preferences for compact, walkable and transit-friendly neighborhoods in residential location choice where tenure types (ownership and renting) were used as an indicator of class membership. A limited set of attributes were tested here in different hypothetical scenarios. To the best of our knowledge, our research is the first attempt to capture the difference in sensitivity between owners and renters and the change in preferences over the time using RP data.

DATA

Study Area

Our study used the London Household Survey Data (LDSD), the Ward Atlas Data (WAD) and an Origin-Destination (OD) matrix from the London Transport Studies (LTS) model as main data sources.

The GLA is divided into 32 boroughs. The total number of electoral wards before 2002 was 773 where 286 were in inner London, 462 were in outer London and the rest were in the city of London. In 2002, the ward boundaries of the GLA were changed significantly and most of the wards were physically affected. The total number of wards was reduced to 649 after reshaping, where 221, 403 and 25 were categorized to be in the inner, outer and city of London, respectively. A map view of inner London, outer London and the City of London (number 6 on the map) is presented in Figure 1.

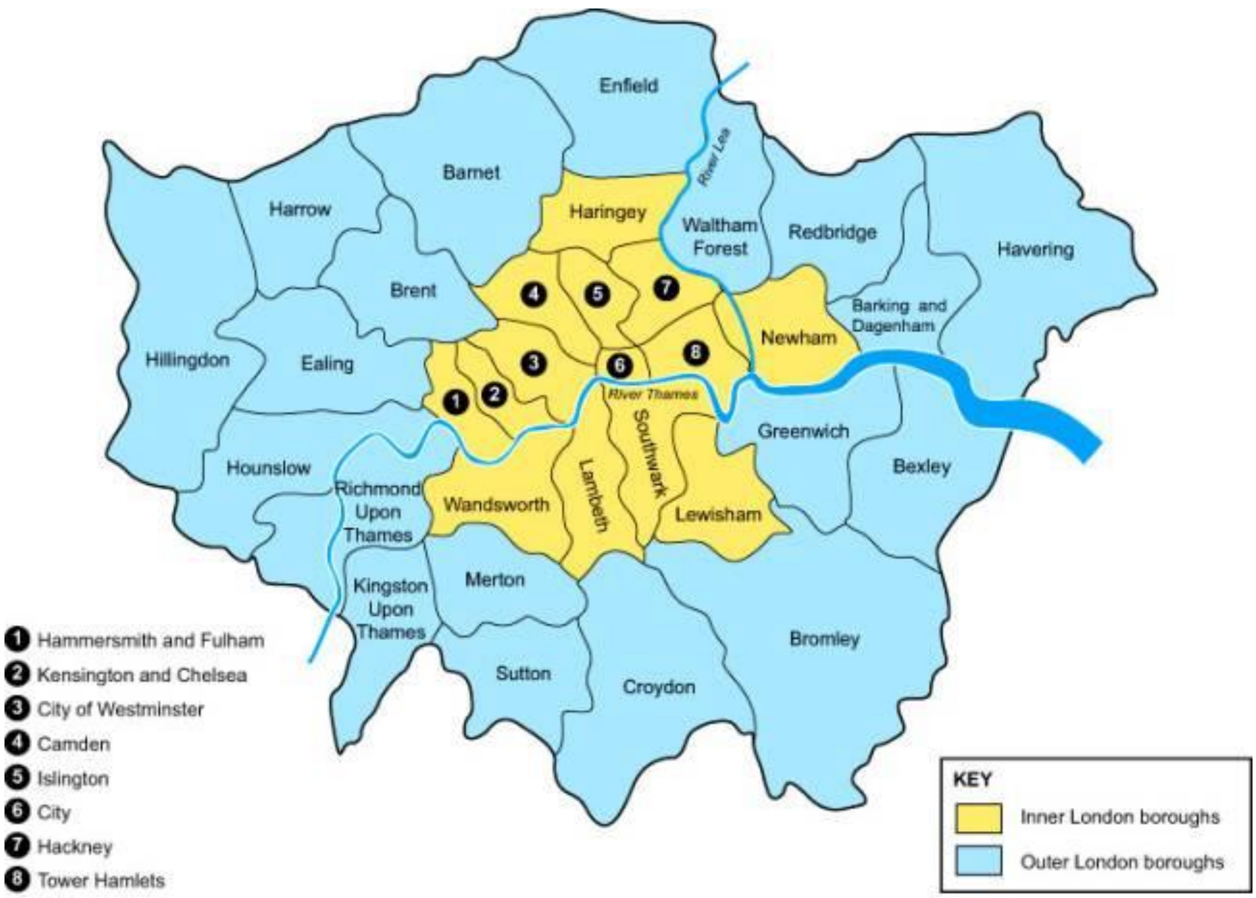


FIGURE 1 Map of Greater London Area. (Source: http://www.geocases.co.uk/**)**

Different ward boundaries were used in different data sets considered for this research which posed a significant challenge in combining the data sources. Ward boundaries before 2002 are termed as old ward boundaries and the updated ones are termed as new ward boundaries in the rest of the paper. The key information about the datasets is presented below.

Data Description

London Household Survey Data

The LHSD collected in 2002 was the main source of disaggregate level household and dwelling information for model estimation. The survey covered 8,158 households and 20,910 individuals from 498 wards (old ward boundaries) in the GLA area. Multistage stratified random sampling was used to collect representative samples from the selected wards. The dataset contains information of 4,491 households living in their owned houses, 3,576 households living in rented houses and 91 households living in shared accommodation. This research focused on households having at least one commute member and used 2,180 owners and 1,293 renters. Detailed information on households' socio-demographic characteristics (household size, income, etc), dwelling information (tenure type, size, price/rent, etc.), employment status, home and work location, car ownership, etc. was also collected in the survey.

Ward Atlas Data (WAD)

WAD was used as a source of zone level (Ward) aggregated demographic, land use and other information. New ward boundaries were used in the dataset. The dataset contains ward level aggregated information of land use pattern, population density, household composition, ethnic proportion, employment and economic activity, household income, crime rates, land use, public transport accessibility, green space, car use, etc.

London Transport Studies (LTS) Model

Information on individual commute distance was missing in the LHSD files which but is clearly of utmost importance as a determinant of household residential location. The origin-destination matrix of GLA from the London Transport Studies (LTS) model was used to extract the commute distances between the reported residential and work locations in the LHSD.

Data Preparation

Extracting relevant information from the multiple large data sources and combining them posed significant challenges. Old ward boundaries (before 2002) were used in the LHSD and new ward boundaries (changed after 2002) were used in the WAD. The zoning system used in the LTS model (called traffic analysis zone, TAZ) was also different from both the old and new ward boundaries.

To be consistent with the different data sets, physical changes of old and new ward boundaries were investigated using the ArcGIS software and all new boundaries were converted to the equivalent old boundaries. TAZ boundaries were also converted to the equivalent old ward boundaries using MapInfo. Information relevant for model estimation was considered from the three data files and data files were merged based on a unique id (old ward boundaries). The information considered for the final data set from the different sources are presented in Figure 2.

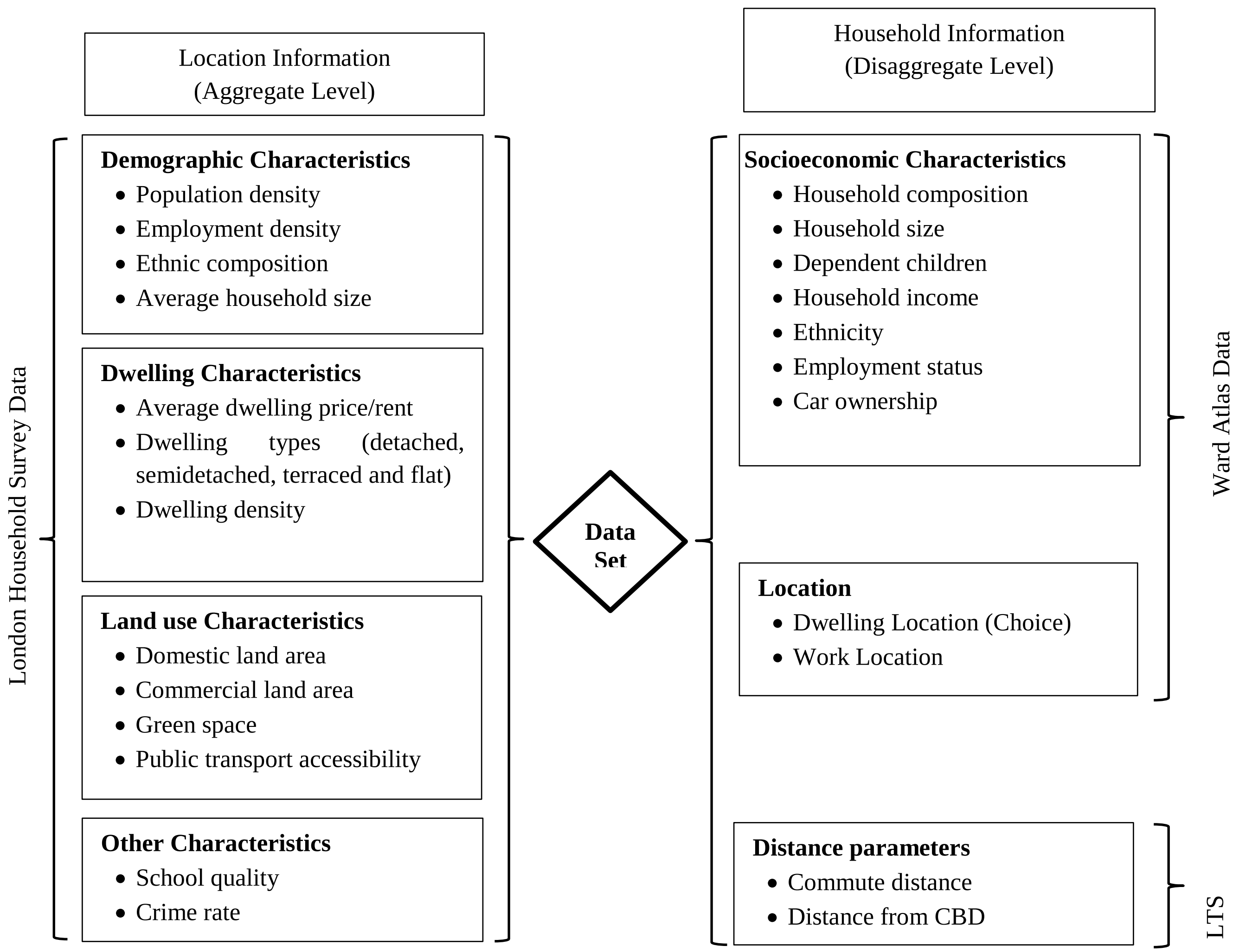


FIGURE 2: List of potential variables considered in the model from different data sources

Data Analysis

A descriptive comparison of socio-demographic, travel behavior and other characteristics of households living in owned and rented properties is presented in Table 1. The market share of residential ownership and renting in inner London (40:60) is quite different from outer London

(67:33). The average tenure length of owners is also higher than renters, where 19.2% of owners have lived in their current houses for more than 12 years, with a comparable figure of only 5.2% for renters. The income structure of owners and renters are also significantly different. The annual average income of 47% of households living in rented property is less than £20,000 while only 15% have an income above £50,000. In comparison, the average annual income of only 13% households living in their owned property is less than £20,000 while 33% have an income of more than £50,000 (Figure 3).

The rate of car ownership for households living in their own properties is two times higher than for households living in rented properties. More than half of the households who live in owned properties are married couples whereas only around one fourth of households who live in rented properties belong to this group. Most of the households living in their own properties are currently employed or retired (87%) while a significant number of households living in rented properties (38%) are unemployed, full-time students or dependents. In terms of ethnicity, 68% of white respondents, 66% of people of Asian origin and 41% of black respondents live in their own houses. There are substantial differences in commute behavior of owners and renters too. Renters are more dependent on public transport, whereas owners are more likely to use cars for commute trips. Importantly, the average commute distance of owners is also higher than that for renters. Our statistical analysis has thus revealed significant differences in location and dwelling attributes, travel behavior and socio-demographic characteristics between households living in owned and rented houses. This serves as the motivation to develop the models in the following section.

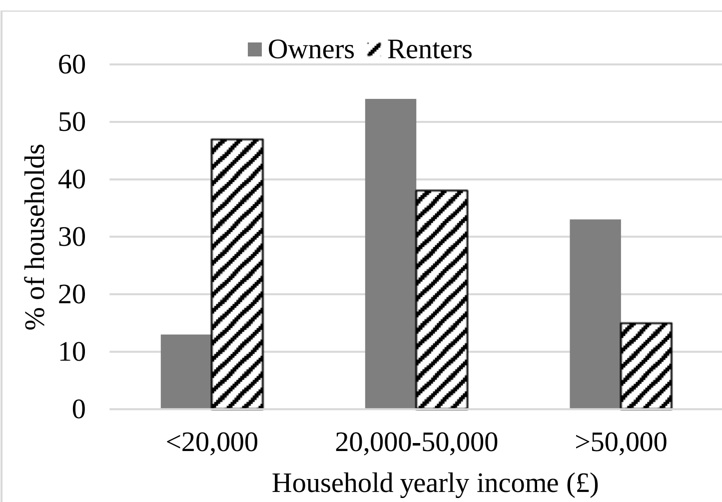


FIGURE 3: Income Differences of Owners and Renters

TABLE 1 Descriptive Statistics of LHSD

Attributes		Tenure Type Owned	Rented
Location and Dwelling Features			
Residential Location (%)	Inner London	40	60
	Outer London	67	33
Household size (bedrooms)	Mean	2.6	2.8
	Standard Deviation	1.39	1.46
Year moved	2001-2002	9.1	15.5
	1997-2000	18.6	10.2
	1990-1996	15.3	7.0
	Before 1990	19.2	5.2
Travel Behaviour			
Car ownership		81	43
Travel Mode (%)	Private vehicle	43	24
	Public Transport (Bus, Train, Tube)	34	54
Commute distance, KM	Mean	10.2	8
	Standard Deviation	8.3	6.86
Socio-demographic Characteristics			
Annual Household Income < £20,000	Proportion	13	47
	Mean	14715	12141
	Standard Deviation	4285	5289
Annual Household Income £20,000-£50,000	Proportion	54	38
	Mean	34656	32032
	Standard Deviation	8394	8618
Annual Household Income >£50,000	Proportion	33	15
	Mean	78878	80241
	Standard Deviation	16213	17084
Household composition (%)	Married couple	55	28
	One Person HH	25	31
Ethnic Composition (%)	White	68	32
	Asian	66	34
	Black	41	59

MODEL DEVELOPMENT

Model Structure

Model parameters were first estimated using a simple multinomial logit model. Mixed logit models were then developed to capture individual taste heterogeneity. Each zone (ward) was considered as a candidate choice and the full choice set is considered. That is, the households evaluated all 498 zones/wards and chose the one they perceive to be the best. This is of course a major assumption but less severe than one making arbitrary assumptions about restricted choice sets.

The utility of household n (n=1…N) for residential location z (z=1…Z) can be expressed as (1)

$$U_{nz} = V_{nz} + \varepsilon_{iz} = \beta x_{nz} + \varepsilon_{nz} \tag{1}$$

where β is the estimated vector of parameters, x_{nz} are observed variables which may include attributes of the alternatives and characteristics of the household. Observed attributes may include the aggregate level dwelling attributes (e.g. average dwelling price, proportion of dwelling types), land use attributes (e.g. land use mix, accessibility) and variables related to the commute (e.g. commute distance) and the location (e.g. north/south/east/west/central London, distance from the CBD, etc.). Household socio-demographic characteristics may include household income, household structure (number and age of children, marital status, etc.), etc. Finally, ε_{iz} is the error term associated with household i and alternative z, capturing any attributes not explicitly specified by the analyst.

Assuming the error term distribution is iid extreme value, the probability of household n choosing alternative z can be express as follows (McFadden, 1978)

$$P_{nz}(\beta) = \frac{\exp(V_{nz})}{\sum_{z=1}^{Z} \exp(V_{nz})} \tag{2}$$

The corresponding log-likelihood function for all the observations is as follows:

$$LL_n(\beta) = \sum_{n=1}^{N}\sum_{z=1}^{Z} y_{nz} \ln(P_{nz}(\beta)) \tag{3}$$

where, y_{nz} = 1 if household n chose zone z and y_{nz} = 0 for all other nonchosen alternatives.

In the Mixed Logit models, random heterogeneity in tastes across households was captured by allowing the coefficients to be distributed randomly across the households with $\beta_n \sim h(\beta|\mu_\beta, \sigma_\beta)$, using for example normal or log-normal distributions. The unconditional likelihood equation would be given by (Train, 2009)

$$LL = \int_{\beta} LL_n(\beta)\, h\left(\beta\middle|\mu_\beta, \sigma_\beta\right) d\beta \tag{4}$$

In contrast with the MNL model, the log-likelihood is now given by a multi-dimensional integral without a closed form solution, and simulation based estimation needs to be used.

The model parameters were selected based on the goodness of fit and the t-stats of the estimated parameters. Separate models were developed first for the owners and renters and the differences in the coefficients of the models were tested using the t-stat difference presented in equation (5)

$$t_{diff} = \frac{\beta_{ik} - \beta_{jk}}{\sqrt{(\frac{\beta_{ik}}{t_{ik}})^2 + (\frac{\beta_{jk}}{t_{jk}})^2}}, \tag{5}$$

where β_{ik} and β_{jk} are the estimates of k th attributes of the model in two different contexts i (owning) and j (renting), t_{ik} and t_{jk} are the respective t ratio of the estimated parameters. The differences in estimated parameters are significant at the 95% level of confidence if the absolute value of t_{diff} exceeds 1.96.

The likelihood ratio (LR) test value was used for comparing competing models where the LR was calculated using equation (6)

$$LR = -2\left[l(\beta_p) - \sum_s l(\beta_s)\right] \tag{6}$$

where $l(\beta_p)$ is the log-likelihood for the pooled model,
$l(\beta_s)$ is the log-likelihood of the model estimated with s_{th} sub market

The LR can then be compared to a critical value from a χ_n^2 distribution with n degrees of freedom, where $n = \sum_s K_s - K$, with K being the number of coefficients in the pooled model, and K_s the number of coefficients in the sth market segment model.

Variable Specification

The variables used in this research are explained below.

Land Use Mix

Land use mix is a widely used index to quantify land use homogeneity. Its scale ranges from 0 to 1 where 0 stands for pure homogeneous land use pattern and 1 stands for uniform mixed land use pattern. Land use mix index can be computed as (Frank et al., 2004)

$$\text{Land use mix} = \sum_j \frac{[P_j \times \ln(P_j)]}{\ln(J)}, \tag{5}$$

where Pj = the proportion of the land area of the jth land-use category, J = total land uses categories considered for the study area. Six land use categories were considered in this research, namely residential use, commercial use, green space, transport facilities and others.

Land use type

The percentage of the area in each zone used by residential and commercial activities is considered in the models. Households who prefer a quiet lifestyle may prefer a residential zone with less commercial activities whereas those who prefer an urban active lifestyle may prefer a residential zone with more commercial activities.

Employment Opportunity

A household having commute sensitive working members is more likely to be inclined towards an area of high employment opportunity. Per person employment opportunity is therefore considered as a candidate variable in this research.

Public transport accessibility

This attribute is likely to have a significant impact on the utility of transit-dependent households (i.e. who do not own cars). Systematic taste heterogeneity of sensitivity towards this variable is tested among households who own cars and those who do not.

Distance from CBD

Household preference for active urban areas or suburban areas is likely to be heterogeneous. It depends on individual household lifestyle, preference and sociodemographic characteristics. Thus, distance of each alternative from the central business district (CBD) was tested in the model. Central London was considered as CBD.

Ethnicity

Households' preferences for residential neighborhoods from the same ethnic community was observed in previous research (Ibraimovic and Hess, 2016). To test ethnic preferences in this research, the proportions of each ethnic group (white ethnicity, black ethnicity and Asian ethnicity) in the zone were interacted with households from the same ethnic background.

Crime Rate

Crime rate is an indicator of the living standard of an area, and an area with higher crime rate is likely to be less attractive to households. This information is available in the data as the total number of crimes per year per thousand of population.

Dwelling Attributes

The average dwelling characteristics considered in the models are the percentage of detached houses, the percentage of flat houses and dwelling density. These variables are estimated separately for people living in inner London and Outer London. Average dwelling price or rent used in the models are interacted with different income groups to capture the potential heterogeneity in price or rent sensitivity.

Commute Attributes

Commute distances between workplace and residential location alternatives (chosen and non-chosen) are considered in the model to estimate household commute distance sensitivity. Commute distance of each respondent was extracted from LTS model.

RESULTS

Preference Heterogeneity between Ownership and Renting

Separate models were developed first for owners and renters. Different model forms were tested with a special focus on investigating the most appropriate structure to capture inter-respondent heterogeneity in the data. Each model was estimated with a choice set of 498 alternative locations for each individual household (i.e. full choice set). The candidate parameters which are significant in at least one of the models (at 90% level of significance) were retained for comparison. Estimation results indicated that mixed logit models with a log normally distributed coefficient for commute distance had the best model fit. It may be noted that taste heterogeneity for all other variables was systematically checked as well but was not found to be statistically significant. The parameters of the final models are presented in Table 2. The results of the t-stat difference test (Equation 5) guided the development of the pooled model which is presented next.

Separate Models

As seen in Table 2, the parameters of the models for owners and renters have the same direction of sensitivity but the magnitude of some of the coefficients was found to be significantly different for the two cases at 95% confidence interval (e.g. commute distance, distance from CBD, school quality, the percentage of residential land-use and preference for ethnic segregation).

Separate constants were estimated for the alternatives in central, north, south east and west London. All else being equal, alternatives in east London were found to be most preferred whereas alternatives in west London were found to be least preferred. The interaction variable between house price and rent with the household income gave a negative sign as expected and revealed that different income groups have different levels of price sensitivities (which is in agreement with the findings of (Habib and Miller, 2009; Zolfaghari, 2013). Households from lower income groups were found to be more price sensitive than higher income groups both for ownership and renting. Preferences for ethnic similarity were found to have a positive and statistically significant effect which suggests that people prefer to live in an area where a higher number of households come from the same ethnic group and is in agreement with findings of previous research (Ibraimovic and Hess, 2016).

Results also showed that both owners and renters dislike higher levels of dwelling density but households living in outer London are significantly more sensitive to dwelling density than households living in inner London. Though households prefer to live in areas of higher residential activities and dislike areas of higher commercial activities, they are also likely to prefer areas with a more balanced mix of land use patterns. Results indicated that households do not prefer an area with a higher percentage of detached houses, both for owning and renting, but this may due to the fact that these houses are not affordable to the respondents. Detached houses in inner London are substantially more expensive and fewer in number than in outer London. Thus, the estimated coefficient for detached houses in inner London showed a higher sensitivity than that in outer London. On the other side, households have positive sensitivity to flats in inner London but a negative one in outer London. Crime rates and household size (absolute difference of zonal average and individual household size) were found to affect the utility negatively.

TABLE 2 Estimation of Long-term and Medium-term Residential Choices

Parameter		Separate Models				t-stat differ-ence	Pooled model					
		Ownership		Renting			Generic		Ownership		Renting	
		Coeff.	t-stat	Coeff.	t-stat		Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Constants	Central London	0.156	1.3	0.086	0.7	0.4	0.147	1.7	-	-	-	-
	South London	0.360	4.1	0.196	1.6	1.1	0.305	4.2	-	-	-	-
	North London	0.432	3.9	0.449	2.9	-0.1	0.449	5.0	-	-	-	-
	East London	0.562	5.7	0.299	2.2	1.6	0.472	6.0	-	-	-	-
	West London	Fixed	-	Fixed	-	-	Fixed	-	-	-	-	-
Housing cost of owners (House Price*0.00001)	None	-0.130	-3.1	-	-	-	-0.153	-3.7	-	-	-	-
	Income < 20,000	-0.382	-4.0	-	-	-	-0.410	-4.3	-	-	-	-
	20,000<Income<50,000	-0.421	-7.9	-	-	-	-0.452	-8.8	-	-	-	-
	Income>50,000	-0.084	-2.7	-	-	-	-0.107	-3.5	-	-	-	-
Housing cost of renters (Monthly Rent * 0.001)	None	-	-	-0.121	-5.1	-	-0.109	-4.8	-	-	-	-
	Income < 20,000	-	-	-0.130	-5.7	-	-0.118	-5.4	-	-	-	-
	20,000<Income<50,000	-	-	-0.095	-4.7	-	-0.084	-4.3	-	-	-	-
	Income>50,000	-	-	-0.024	-1.4	-	-0.015	-1	-	-	-	-
Ethnic Composition	% White people	0.017	8.4	0.008	2.7	2.7	-	-	0.018	8.9	0.007	2.6
	% Asian people	0.042	12.4	0.050	10.4	-1.4	0.044	16.2	-	-	-	-
	% Black people	0.053	7.3	0.042	7.0	1.1	0.047	10.1	-	-	-	-
Dwelling density in outer London (per sqkm)		-0.113	-12.1	-0.096	-7.8	-1.0	-0.107	-14.4	-	-	-	-
Dwelling density in inner London (per sqkm)		-0.017	-3.6	-0.008	-1.9	-1.5	-0.012	-3.8	-	-	-	-
Land use mix(1=balance,0=homogeneous)		1.379	4.3	1.900	4.3	-1.0	1.520	5.9	-	-	-	-
Residential land area in inner London		0.146	10.3	0.098	6.9	2.4	-	-	0.136	12.2	0.103	9.1
Residential land area in outer London		0.210	10.4	0.161	5.6	1.4	0.194	11.9	-	-	-	-
Fraction of commercial land area		-0.045	-4.6	-0.051	-5.2	0.4	-0.047	-6.8	-	-	-	-
% detached house in outer London		-0.035	-7.5	-0.017	-2.0	-1.9	-0.030	-7.4	-	-	-	-
% detached house in inner London		-0.140	-5.8	-0.083	-3.0	-1.5	-0.115	-6.4	-	-	-	-
% flat in outer London		-0.008	-3.2	-0.004	-0.9	-0.9	-0.008	-3.6	-	-	-	-
% flat in inner London		0.032	7.9	0.033	7.0	-0.2	0.032	10.4	-	-	-	-
School quality		0.004	2.8	-0.009	-4.0	4.9	-	-	0.004	2.9	-0.009	-4.2
Crime rate (per 1000 people)		-0.047	-1.0	-0.056	-1.2	0.1	-0.050	-1.5	-	-	-	-
Household Size		-0.391	-4.4	-0.147	-1.4	-1.8	-0.292	-4.2	-	-	-	-
Employment opportunity (per person)		0.100	2.3	0.142	3.3	-0.7	0.123	4.1	-	-	-	-
Distance from CBD in km		0.065	8.4	0.022	2.2	3.3	-	-	0.066	9.4	0.018	2
Public transport accessibility (No Car)		0.343	5.2	0.258	4.5	1.0	0.322	7.7	-	-	-	-
Public transport accessibility (Car owner)		-0.210	-4.6	-0.087	-1.5	-1.7	-0.164	-4.7	-	-	-	-
Commute distance in km	Mean	-0.202	-38.0	-0.251	-28.3	4.7	-	-	-0.187	-38.6	-0.222	-45.4
	Standard deviation	0.094	20.7	0.166	27.2	-9.5	-	-	0.053	9.9	0.067	9.2
Number of Observations		2180		1293		-	3473					
Initial LL		-13538.11		-8030.31			-21569.41					
Final LL		-10437.23		-6214.47			-16672.95					
Adjusted ρ2		0.229		0.226			0.227					
Likelihood ratio test (χ2, DF, p)		-					42.5, 20, 0.001					

On the other hand, households were found to be inclined to areas having greater employment opportunities and those further from the central business district (CBD). Increase in public transport accessibility also increases the utility of carless households but decreases the utility for car owning households.

The school quality (only considered for households with children) was found to have a positive effect for owners (as expected). However, it has a non-intuitive negative sign for renters. This may be due to the fact that in the UK schooling system, residential location at the year before a child starts primary and high schools are critical and it is not uncommon for people to rent a house close to a good school (which are much more expensive) only for the critical time period and then move to less expensive areas. Given that the detailed age of the child was not available in the data, it was not possible to investigate this hypothesis further. The coefficient of commute distance was allowed to vary randomly across households, with estimates revealing significant taste heterogeneity. As expected, increased commute distance adds disutility to residential location alternatives.

The estimated coefficients of five parameters in the separate models were found significantly different at 95% confidence interval, namely: commute distance, distance from CBD, school quality, the percentage of residential land-use (in inner London) and preference for ethnic similarity among the white ethnic respondents. The t-stat difference of the coefficient of public transport accessibility (car owners), household size and percentage of detached house (in outer London) were found to be statistically significant at the 90% confidence interval (t_{diff}>1.65).

Pooled model

In order to investigate the second research question, a fully pooled model was developed first where all coefficients were assumed to be common for owners and renters. However, this resulted in a significant loss of fit compared to the separate models. A likelihood ratio test (χ^2=219.2, degree of freedom (DF)=26, P=0.001) strongly rejected the null hypothesis. It confirmed the existence of preference heterogeneity of owners and renters in their residential location choice and led us to estimate separate parameter coefficients for owners and renters. Though only five out of thirty parameters (commute distance, distance from CBD, school quality, the percentage of residential land-use in inner London and preference for ethnic similarity among the white ethnic respondents) were found to be significantly different between ownership and renting at the 95% confidence interval, a revised pooled model was estimated allowing for separate coefficients for owners and renters of the parameters significantly different between ownership and renting. A likelihood ratio test ((χ^2=42.5, DF=26, P=0.001)) then accepted the hypothesis of revised pooled model. Pooling helps to reduce the estimation time through minimizing the number of parameters estimated without effecting the model goodness of fit.

In the final model, renters were found to be more sensitive to commute distance than owners. This is likely to be due to the fact that renters have lower car ownership and a greater tendency to minimize commute cost due to lower household income. On the other hand, owners are found to be more inclined to be sensitive to the location attributes (fraction of residential land use in inner London, distance from CBD) and ethnic segregation (white people) which supports our prior hypothesis. More importantly, the coefficient of school quality was found to be statistically significant for owners and renters but gives opposite signs. The coefficient of school quality for owners was positive as excepted (Kim et al., 2005) but for renters, it gave a negative sign. As mentioned, this may be due to the higher demand and increased rents of houses in the area of better schools and likely to be driven by the exact age of the children in the household (which was not available in the data).

Change of Preference over Time

In order to investigate the third research question (change of sensitivities to attributes over time), separate models were estimated for households who had moved in a given time period. Data was split into four groups: TP1 (households moved before 1990), TP2 (households moved between 1990 to 1996), TP3 (households moved between 1997-2000) and TP4 (households moved between 2001-2002). It may be noted that given the continuous nature of change in preferences, it was difficult to identify any intuitive breakpoints of the sample subdivision. The samples were therefore subdivided to ensure representative samples in each data set. Slight variations of the sample subdivisions were tested and the one with the best total goodness of fit was selected. ECL models were estimated using the same specification of the final pooled model in the previous section. Parameters which were statistically significant in at least one model were kept for comparison. The model findings are presented in Table 3.

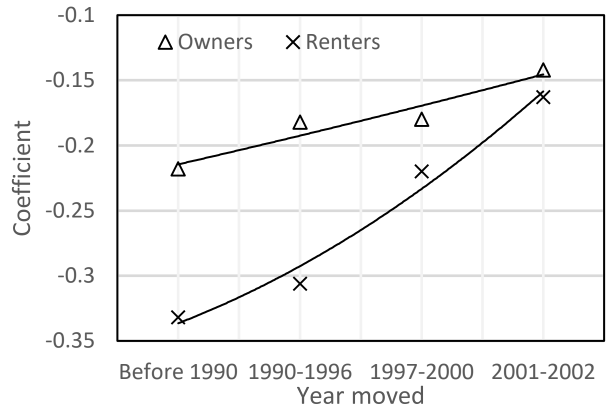
The models showed a trend (either increasing or decreasing) of changing sensitivity over time (from TP1 to TP4) for most of the parameters. The statistical differences of these changes were tested using t-stat difference tests. Results of the t-stat difference tests indicated that the sensitivity to 9 parameters (commute distance, distance from CBD, housing cost of low-income people, school quality, public transport accessibility, crime rate, household size and percentage of flat in outer London) was found to change significantly at the 90% to 95% confidence interval over the years. Sensitivities to two additional attributes (employment opportunity and land use mix) changed at the 80% confidence interval while the rest of them were almost stable over time. The parameters which had significant changes in sensitivities over the years are presented graphically in Figure 4 and further discussed below:

(a) Parameters which had the different sensitivities of ownership and renting

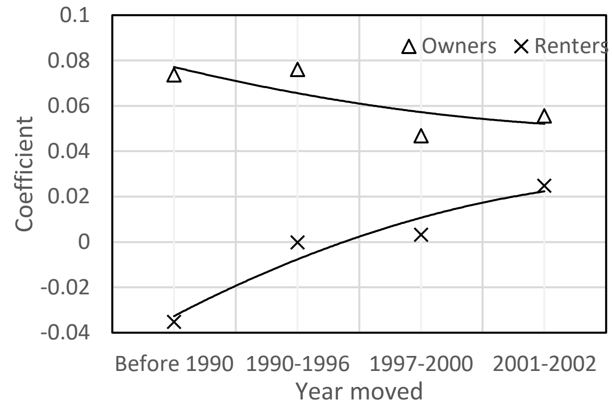
- Sensitivity to commute distance was found to decrease over time but the sensitivity of renters was found to change more sharply than owners. The sharp changes in housing cost in the city centre or commercial zones in the Greater London area, better living facilities in suburbs and strong competition in the job market are likely to have contributed to this.
- Before 1990, owners were more inclined to live far from the CBD whereas renters were more likely to live close to the CBD. However, the trend indicates that there is decreased sensitivity to the distance from CBD in the owner group (which is in line with other studies, such as (Næss, 2009). Interestingly renters showed the opposite trend potentially due to the fact that rents have skyrocketed closer to the CBD in that period.
- Sensitivity to housing cost of low-income households increased sharply for both owning and renting. This is due to the imbalance of household income and housing cost. Housing cost has increased at a higher rate over time compared to the increase of earning of low-income households. Interestingly, for other income groups, the difference in sensitivity over time was not significant.

TABLE 3 Estimation of Models for Households Moved in Different Time Periods (TP)

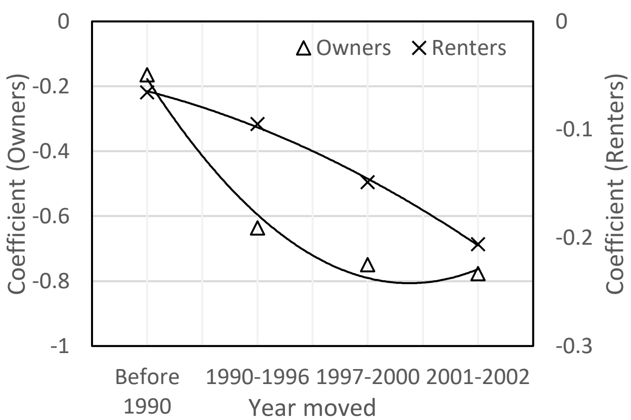
Parameter		Year moved								t-diff (TP-1 and TP-4)
		TP-1 (Before1990)		TP-2 (1990-1996)		TP-3 (1997-2000)		TP-4 (2001-2002)		
		Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	
Constants	Central London	0.333	1.7	0.114	0.6	0.160	1	-0.016	-0.1	1.4
	South London	0.007	0.1	0.214	1.3	0.345	2.5	0.487	3.6	-3.1
	North London	0.698	3.6	0.407	2	0.284	1.6	0.288	1.6	1.5
	East London	0.641	3.6	0.568	3.3	0.335	2.2	0.265	1.7	1.6
Housing cost of owners (House Price*0.00001)	None	-0.140	-1.9	-0.057	-0.8	-0.214	-2.1	-0.122	-1.2	-0.1
	Income < 20,000	-0.164	-1.3	-0.636	-2.3	-0.750	-3	-0.777	-1.8	1.4
	20,000<Income<50,000	-0.442	-4.5	-0.299	-3.3	-0.553	-5.6	-0.564	-3.8	0.7
	Income>50,000	-0.114	-1.6	-0.079	-1.4	-0.156	-2.8	-0.082	-1.1	-0.3
Housing cost of renters (Monthly Rent * 0.001)	None	-0.327	-3.5	-0.094	-1.4	-0.123	-2.4	-0.087	-3	-2.4
	Income < 20,000	-0.065	-1.8	-0.095	-1.9	-0.149	-2.9	-0.206	-4.5	2.4
	20,000<Income<50,000	-0.114	-2.1	-0.064	-1.2	-0.092	-2.3	-0.105	-3.4	-0.2
	Income>50,000	-0.058	-0.7	-0.052	-1.4	-0.018	-0.5	-0.020	-1	-0.5
Ethnic Composition	% White people, owners	0.016	4.2	0.022	5.1	0.020	5.3	0.015	2.8	0.2
	% White people, renters	0.022	2.6	0.010	1.5	0.000	0	0.014	3.1	0.9
	% Asian people	0.046	8.3	0.044	7.5	0.039	7	0.043	7.5	0.3
	% Black people	0.045	4.3	0.045	4.6	0.037	4.3	0.052	5	-0.4
Dwelling density in outer London (per sqkm)		-0.112	-6.9	-0.094	-6.1	-0.114	-7.6	-0.111	-7.4	-0.1
Dwelling density in inner London (per sqkm)		-0.017	-2.3	-0.005	-0.7	-0.006	-1.1	-0.018	-3.1	0.1
Land use mix(1=balance,0=homogeneous)		1.119	2.1	1.092	2	1.729	3.5	2.093	3.8	-1.3
Residential land area in inner London	Owners	0.142	6.1	0.096	3.9	0.158	7.8	0.136	5.3	0.2
	Renters	0.128	4.7	0.064	2.4	0.080	3.6	0.136	6.7	-0.2
Residential land area in outer London		0.217	6.2	0.173	5	0.191	5.7	0.201	6	0.3
Fraction of commercial land area		-0.041	-2.7	-0.048	-3.1	-0.037	-2.8	-0.064	-4.8	1.1
% detached house in outer London		-0.037	-4.4	-0.031	-3.6	-0.030	-3.8	-0.024	-2.6	-1.1
% detached house in inner London		-0.100	-2.7	-0.119	-2.9	-0.102	-3	-0.134	-3.6	0.6
% flat in outer London		-0.014	-3.2	-0.005	-1.1	-0.006	-1.5	-0.004	-0.1	-2.4
% flat in inner London		0.027	4.2	0.032	4.7	0.031	5.3	0.036	6	-1.0
School quality	Owners	0.0001	0.1	0.006	2.2	0.007	2.5	0.004	0.9	-0.9
	Renters	-0.014	-2.2	-0.015	-3.7	-0.007	-1.8	-0.001	-0.4	-1.7
Crime rate (per 1000 people)		-0.018	-0.3	-0.079	-1.1	-0.043	-0.7	-0.160	-2.6	1.7
Household Size		-0.477	-3.2	-0.234	-1.6	-0.274	-2.1	-0.141	-1	-1.7
Employment opportunity (per person)		0.111	1.7	0.027	0.4	0.095	1.6	0.232	4	-1.4
Distance from CBD in km	Owners	0.074	5.2	0.076	5.2	0.047	3.5	0.056	3.4	0.8
	Renters	-0.035	-1.3	-0.002	0	0.003	0.2	0.025	1.7	-2.0
Public transport accessibility (No Car)		0.122	1.2	0.268	2.7	0.429	5.1	0.386	5.2	-2.1
Public transport accessibility (Car owner)		-0.134	-1.8	-0.217	-2.8	-0.284	-4.3	-0.082	-1.1	-0.5
Commute distance of owners in km	Mean	-0.218	-21.1	-0.182	-18.8	-0.180	-19.8	-0.142	-13.5	-5.2
	Standard deviation	0.051	3.3	0.041	3.6	0.056	5.7	0.031	3.6	1.1
Commute distance of renters in km	Mean	-0.332	-11.8	-0.306	-13.8	-0.220	-14.3	-0.163	-16.2	5.7
	Standard deviation	0.111	5.7	0.083	4.0	0.060	4.3	0.034	1.82	2.1
Number of Observations		800		731		943		805		-
Initial LL		-4968		-4540		-5857		-5000		
Final LL		-3580		-3409		-4570		-4100		
Adjusted ρ2		0.279		0.249		0.220		0.180		



a. Commute distance

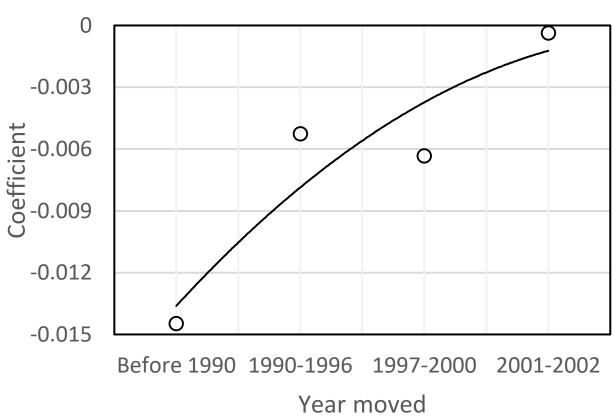


b. Distance from CBD



c. Housing cost (Low-income households)

d. PT accessibility (No car)



e. Percentage of flat in outer London

f. Household size and crime rate

FIGURE 4 Sensitivity changes of parameters over the years

(b) Parameters which had the same values for owners and renters

- The sensitivity of households without cars to public transport accessibility increased with time. Carless households are normally more dependent on public transport and increased commute distance over the years potentially led them to be more inclined to better public transport accessibility.
- Household sensitivity to crime rate increased over time. Due to the increasing trend of crime in UK including the greater London area (Home office, 2012), households preferred the area with reduced crime rate.
- Household became less sensitive to household size and the proportion of flats in outer London over time.

CONCLUSION AND IMPLICATION

In this research, we investigated three research questions using RP data combined with multiple other data sources. We made innovative use of publicly available real world data and were able to estimate residential location models without requiring sampling of alternatives. The key findings are as follows:

1. Estimation results indicate that both owners and renters have similar preferences (same signs of parameters) but the sensitivity to several attributes are significantly different. These include commute distance, distance from CBD, school quality, the percentage of residential land-use (in inner London) and preference for ethnic similarity among the white ethnic respondents, public transport accessibility (car owners), household size and percentage of detached house (in outer London). This confirms our hypothesis regarding significant differences between long-term and medium-term residential location decisions.

2. It is possible to develop a pooled model for owners and renters – but this should acknowledge the differences in sensitivities towards the attributes identified above, as done in our work.

3. There is a significant difference in sensitivity towards attributes over time (1971-2002, in particular). Linking it with the full range of broader changes in the economy, technology, etc. is beyond the scope of the current research but an interesting direction of future research.

A key difference between the current studies and the previous studies is the fact that we have combined detailed data from a range of sources which has enabled us to capture a wider range of attributes compared to previous state-of-the-art models (which had mostly dealt with a smaller subset of variables in isolation due to data limitation). The developed models are therefore expected to lead to better predictions. There are however several limitations of the research which need to be addressed in future work. Firstly, the full choice set has been considered for each respondent which is very large [498 alternatives in this case]. While efficient software has allowed us to avoid sampling of alternatives and the necessary corrections, it could be argued that the use of such a full set is behaviourally non-representative. Future research will generate restricted choice sets for each respondent based on behavioral rules rather than considering a full choice set and test the impact of this on results. Better treatment of the choice set is likely to make the models computationally easier as well as behaviourally more representative, although it will potentially divert explanatory power away from the choice model itself, a point alluded to earlier in the paper. Secondly, residential location choice models are more likely to suffer from endogeneity problems, primarily caused by the omission of attributes correlated with price (Guevara, 2010). In further research, we also have the aim to correct the models for endogeneity for consistent model parameter estimation. However, even in its current form, the models provide important behavioral insights on how people trade-off differently when making location choices in different time scales and the trend of change of these sensitivities.

The model findings can provide useful insights for land use and transport related policy formulation on existing and future land development. The distinct nature of these two-different time scale decision (ownership and renting) observed in this research may lead to differences in policy as well. For instance, according to our model outcomes, renters are more inclined to transit oriented compact development to minimize housing and travel cost whereas owners are less likely to be inclined to compact development. Though policy makers are more interested in compact development nowadays to restrain urban sprawl, they need to consider the differences in preferences of these different decision-making units. The dynamics of household preferences over time in their residential location choice found in this research are more likely to be a case for any other cities but can be driven by different parameters. Therefore, the dynamics of choice over time also need to be recognised for a long term and sustainable policy formulation.

ACKNOWLEDGEMENT

The authors are thankful to Greater London Authority and Transport for London authority for providing data used in this research. This research is funded by the Economic and Social Research Council and Institute for Transport Studies, University of Leeds. The authors acknowledge the financial support by the European Research Council through the consolidator grant 615596-DECISIONS. The authors also acknowledge Andrew Daly’s contribution of using ALOGIT software in model estimation.

REFERENCES

Bhat CR and Guo J. (2004) A mixed spatially correlated logit model: formulation and application to residential choice modeling. *Transportation Research Part B: Methodological* 38: 147-168.

Boehm TP. (1982) A hierarchical model of housing choice. *Urban studies* 19: 17-31.

Brown BB and Werner CM. (2008) Before and after a new light rail stop: resident attitudes, travel behavior, and obesity. *Journal of the American Planning Association* 75: 5-12.

Burda C. (2014) Home Location Preference Survey: Understanding where GTA residents prefer to live and commute. Royal Bank of Canada and the Pembina Institute.

Census. (2001) Office for national statistics, United Kingdom.

Cho C-J. (1997) Joint choice of tenure and dwelling type: a multinomial logit analysis for the city of Chongju. *Urban Studies* 34: 1459-1473.

Ewing R and Cervero R. (2010) Travel and the built environment: a meta-analysis. *Journal of the American Planning Association* 76: 265-294.

Farber S and Li X. (2013) Urban sprawl and social interaction potential: an empirical analysis of large metropolitan regions in the United States. *Journal of Transport Geography* 31: 267-277.

Frank LD, Andresen MA and Schmid TL. (2004) Obesity relationships with community design, physical activity, and time spent in cars. *American journal of preventive medicine* 27: 87-96.

Gower TL. (2014) 2011 Census Analysis-Distance Travelled to Work.

Guevara CA. (2010) Endogeneity and sampling of alternatives in spatial choice models. *Unpublished doctoral dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA.*

Habib M and Miller E. (2009) Reference-dependent residential location choice model within a relocation context. *Transportation Research Record: Journal of the Transportation Research Board*: 92-99.

Ho C and Hensher D. (2014) Housing prices and price endogeneity in tenure and dwelling type choice models. *Case Studies on Transport Policy* 2: 107-115.

Home office U. (2012) A summary of recorded crime data from 1898 to 2001/02.

Hoshino T. (2011) Estimation and analysis of preference heterogeneity in residential choice behaviour. *Urban Studies* 48: 363-382.

Ibraimovic T and Hess S. (2016) Changes in the ethnic composition of neighbourhoods: Analysis of household's response and asymmetric preference structures. *Papers in Regional Science*.

Kim JH, Pagliara F and Preston J. (2005) The intention to move and residential location choice behaviour. *Urban Studies* 42: 1621-1636.

Lee BH and Waddell P. (2010) Residential mobility and location choice: a nested logit model with sampling of alternatives. *Transportation* 37: 587-601.

Lee BH, Waddell P, Wang L, et al. (2009) Operationalizing Time-Space Prism Accessibility in a Building-Level Residential Choice Model. *Proceedings of the 88th Annual Meeting of the TRB.*

Liao FH, Farber S and Ewing R. (2015) Compact development and preference heterogeneity in residential location choice behaviour: A latent class analysis. *Urban Studies* 52: 314-337.

McFadden D. (1978) Modeling the choice of residential location. *Transportation Research Record*.

Næss P. (2009) Residential self-selection and appropriate control variables in land use: Travel studies. *Transport Reviews* 29: 293-324.

Sermons MW and Koppelman FS. (2001) Representing the differences between female and male commute behavior in residential location choice models. *Journal of Transport Geography* 9: 101-110.

Skaburskis A. (1999) Modelling the choice of tenure and building type. *Urban studies* 36: 2199-2215.

Train KE. (2009) *Discrete choice methods with simulation*: Cambridge university press.

Yates J and Mackay DF. (2006) Discrete choice modelling of urban housing markets: a critical review and an application. *Urban studies* 43: 559-581.

Zolfaghari A. (2013) Methodological and empirical challenges in modelling residential location choices.

Randall C. (2011) Housing-Social Trends 41. Office for National Statistics. UK