UNDERSTANDING DIFFERENCES IN RESIDENTIAL LOCATION PREFERENCES BETWEEN OWNERSHIP AND RENTING: A CASE STUDY OF LONDON

Md Bashirul Haque
Choice Modelling Centre and
Institute for Transport Studies
University of Leeds, UK
Email: tsmbh@leeds.ac.uk
and
Department of Civil and Environmental Engineering
Shahjalal University of Science and Technology, Bangladesh
Email: bashir-cee@sust.edu

Charisma Choudhury
Choice Modelling Centre and
Institute for Transport Studies
University of Leeds, UK
Email: C.F.Choudhury@leeds.ac.uk

Stephane Hess, Corresponding Author
Choice Modelling Centre and
Institute for Transport Studies
University of Leeds, UK
Email: s.hess@its.leeds.ac.uk

ABSTRACT
This study aims to investigate the similarities and differences between residential location choices of owners and renters. The models are estimated for commuter households living in owned or privately rented dwellings in the Greater London Area. The London Household Survey Data, Ward Atlas Data and the London Transport Studies Model outputs are used in this regard. Model parameters are estimated using an error components specification of the mixed logit model to capture differences in error variance. Willingness to Pay (WTP) and elasticity analyses are also performed for interpretation of the estimated model parameters and quantifying the differences between the two groups. The results indicate that while some common factors are affecting the choices of owners and renters, there are significant differences in response to several factors such as public transport accessibility (for 'car-owning' households) and percentage of detached houses in the area. Accounting for preference heterogeneity between the sub-groups is expected to lead to better planning and investment decisions.

Keywords: residential location choice; London; owners; renters; discrete choice model; commute distance; willingness-to-pay

1 INTRODUCTION

Due to rapid urbanization and growing transport demand, integrated land use and transport planning has become a key interest of planners and policymakers. There is evidence that people who live in low-density residential areas tend to commute longer distances for work and other trips and are typically more car-dependent (Alexander and Tomalty, 2002; Næss, 2009). However, people often prefer low-density areas as they offer more green and open spaces, larger homes, greater ease of parking, and other benefits (Masnavi, 2000). On the other hand, people living in mixed and compact developments have better access to facilities and are typically less car-dependent (Masnavi, 2000; Ewing and Cervero, 2010; Farber and Li, 2013). Although compact developments have been criticized in the literature for the lower quality of the local environment, high congestion and limited recreational opportunities (which often leads to long weekend trips) (Holden and Norland, 2005; Boyko and Cooper, 2011), there is evidence that specific urban forms can mitigate some of these problems and ensure sustainability (Arundel and Ronald, 2017). These contradictions between the planning paradigms and user preferences make it crucial to better understand residential location choices at a disaggregate level.

Discrete choice analyses (DCA) enable us to disentangle the effects of different factors affecting household residential location choices and to determine the elasticity of demand in response to changes and willingness-to-pay (WTP)[1] for different attributes. The results of DCA can hence be used to predict the implications of alternative policy scenarios. The factors driving the choice of residential location are however unlikely to be the same for the two key segments of owners and renters, with possible differences in which attributes matter, and also differences in how much each attribute matters. The potential differences can be explained by the distinct nature of ownership and renting decisions. Residential ownership is a long-term decision that involves huge investment and high relocation costs while private renting is typically a medium to short-term decision due to the higher level of flexibility associated with the lower relocation costs, shorter lengths of agreements, and other factors. For instance, the average tenure length in England is 11 years for owners but only around 1 year for private renters and 7 years for social renters (Randall, 2011). The socio-demographic characteristics of owners and private renters are also typically different. For example, high and middle-income households are more likely to be able to afford to buy properties while others may be more likely to rent (Yates and Mackay, 2006). With these clear distinctions, it is important to analyse the ownership and renting choices in detail and identify the similarities and differences in sensitivities to different factors in the two markets. Although there have been some studies focusing on residential ownership (e.g. Guevara, 2010; Zolfaghari, 2013) or renting decisions (e.g. Hoshino, 2011) in isolation or jointly (Ho, Hensher, & Ellison, 2017), to the best of our knowledge, none of these studies have quantified the differences in

[1] WTP analysis is a widely used tool in marketing and environmental economics, giving a measure of how much an individual is willing to pay to acquire desirable attributes and/or to avoid undesirable attributes of a product or service. It has also been investigated in the context of residential location choice – although only in a relatively limited way (e.g. Jara Diaz et al. 1999; Small et al. 2012, etc.).

sensitivities towards different factors (or the systematic heterogeneity in elasticity and WTP values) between owners and renters. Our study hence aims to address this gap by investigating the similarities and differences between residential location choices of owners and renters. We address this study objective by developing Revealed Preference (RP) based residential location choice models for people living in the Greater London Area (GLA)[2]. We combine several different data sources and make use of detailed econometric models to analyse the residential location decisions in those datasets.

The remainder of the paper is organized as follows. We start with a literature review followed by data description. The model structure and estimation results are presented next, followed by the concluding remarks.

2 LITERATURE REVIEW

In recent years, significant methodological and analytical improvements have been achieved in residential location choice modelling. These include methodological work on choice set generation and sampling of alternatives (e.g. Guevara, 2010, Zolfaghari, 2013), the treatment of complex correlation structures (e.g. 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 to develop parcel (e.g. Lee and Waddell, 2010) or dwelling level (e.g. Zolfaghari, 2013) disaggregate residential location choice models. Furthermore, a large number of studies have focussed on factors influencing these choices (see e.g. Sermons and Koppelman, 2001, Ibraimovic and Hess, 2016 for details). Most of the previous studies have however looked at residential ownership decisions (e.g. Bhat and Guo, 2004; Habib and Miller, 2009; Guevara, 2010; Zolfaghari, 2013) or renting (e.g. Hoshino, 2011) in isolation. Those rare studies that have focused on joint choices of tenure and dwelling revealed 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; Yates and Mackay 2006; Ho and Hensher, 2014). Liao et al. (2015) developed a combined model of ownership and renting decisions but this relied on Stated Preference (SP) data which is prone to hypothetical bias and behavioural incongruence. Ho, Hensher, & Ellison (2017) developed a comprehensive residential location choice model where logsums from several models (tenure and dwelling type choice, work and non-work location choice, and others.) were used to capture the interdependencies of location choice with other choices. Although this approach captured how the tenure choice (e.g. choice of residential ownership, renting) influences the choice of location, it

[2] It may be noted that in the Greate London Area, 23% of the housing market is made up of 'social renters' who have a constrained choice set and are not able to exercise their residential choices in the same way as private renters. In order to capture the preferences of social renters, it is critical to know the choice set of each decision maker (i.e. available alternatives while making the decision) and any associated constraints (arising from the allocation policy). However, the London Household Survey dataset did not include these pieces of information. We therefore excluded social renters from our analysis and only focused on private renting (referred as 'renting' in the rest of the paper) and owning.

did not disentangle how these two groups of decision makers (owners and renters) evaluate the attributes of residential location choice and if there are significant differences in their sensitivities towards different attributes. The present research aims to fill this research gap. The specific focus of this study is on investigating the similarities and differences in the residential location choice preferences of owners and renters using large scale RP data. Combining several RP datasets enabled us to capture a wider range of attributes compared to previous state-of-the-art models, and, therefore, was expected to lead to better predictions. Moreover, the owner and renter specific elasticity and WTP values can be used directly for evaluating alternate policies.

Continued gentrification and social change in London over time and their impact on migration, tenure and location choice have been investigated in multiple studies (e.g. Atkinson, 2000; Butler et al., 2008; Hamnett, 2010, Paccoud and Mace 2017). Butler et al. (2008) claimed that the proportion of residents in the higher social classes (who control most of the wealth and power of the society) is increasing in Inner London while the proportion of intermediate social classes[3] is increasing in Outer London. Paccoud and Mace (2017) report that there has also been a marked tenure shift to the private renting sector in Outer London. Reductions in welfare benefits, on the other hand, have resulted in a large shift of low-income renters from Inner London to lower-price Outer London areas (Hamnett, 2010; Fenton, 2011). Hamnett and Butler (2010) report an increasing trend of ownership in Outer London among certain ethnic minority groups as well as large concentrations of ethnic minorities in social and privately rented housing.

In the US, many cities have experienced a noticeable rebound in population in urban areas or near the CBD since 2000 (Baum-Snow and Hartley, 2019). This trend has been interpreted in the literature as a result of the stronger new urban preferences of young adults or millennials (Couture and Handbury, 2017; Lee, 2018; Baum-Snow and Hartley, 2019; Lee, Lee, and Shubho, 2019). The preference of Millennials for living near the CBD has been found to be mostly driven by their generational characteristics, such as their racial diversity and marital attributes, amongst others. They have also been observed to prefer increased walkability, transportation facilities, and the proximity of neighbourhood amenities compared with earlier generations (Lee, 2018). Therefore, the Millennial generation in the US is likely to have a higher propensity to prefer compact, transit-oriented, and mixed land-use developments in the future.

In the context of Australia, an increasing number of low-income households have been found to move to higher density housing in outer areas and to rental housing due to the rise in housing cost (Yates, 2001). Urban consolidation has opened the scope for low-income households to find affordable housing and high-income people to fulfil their aim of owning properties and living in

[3] The intermediate social class includes the ‘Established middle class’ and the ‘Technical middle class’ of the Great British Class Survey (Savage 2013). The first group, which has a good representation of graduates, are described as ‘comfortably off, secure, and established’. The other group includes people who show ‘high economic capital, very high status of social contacts, but relatively few contacts reported, and moderate cultural capital’.

the inner-city (Yates, 2001). From the above discussion, it can be concluded that the residential preferences of (different groups of) people are changing over time and are likely to reshape the future urban form. However, due to data limitations, it has not been possible to capture the longitudinal changes in preferences in this study.

3 DATA

3.1 Study area

The Greater London Area (GLA) is considered as the study area. The GLA is divided into 32 boroughs and the City of London. The total number of electoral wards before 2002 was 773, where 286 were in Inner London, 462 were in Outer London and the remainder were in the city of London. In 2002, the ward boundaries of the GLA were changed significantly and the majority 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 Inner, Outer and the city of London, respectively. A map view of Inner, Outer and the City of London is presented in Figure 1.

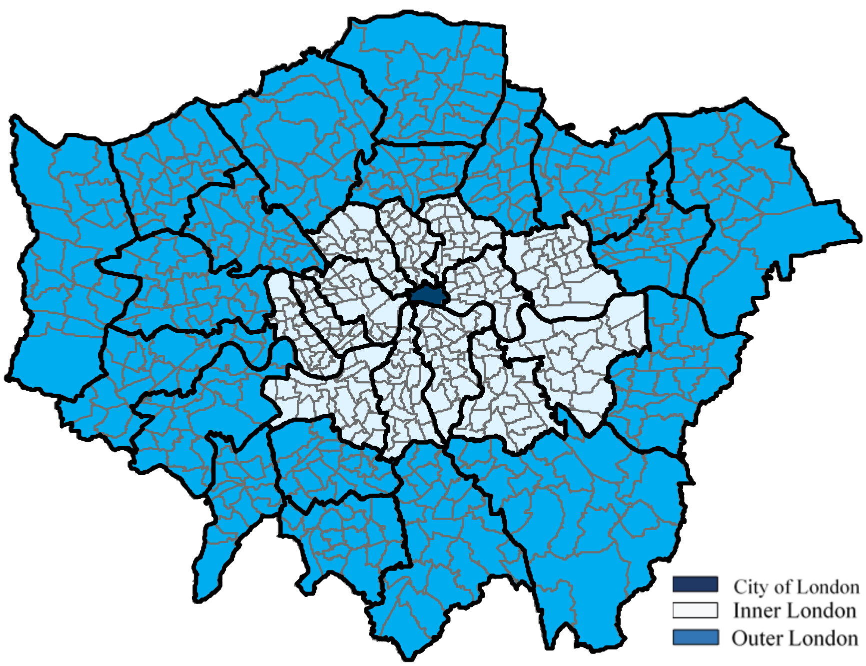

FIGURE 1 Map of the Greater London Area. (Source: https://data.london.gov.uk/**)**

3.2 Data description

To estimate the residential location choice model, both household-level data (e.g. residential location, demographic characteristics, dwelling characteristics, travel behaviour) and location characteristics (e.g. land-use, transport accessibility, employment opportunity) are essential. However, in the context of London, no single dataset contains all the required information. Therefore, we used several datasets as summarised in Figure 2 and detailed below. Figure 2 shows how the inputs for the model come from three different sources. The dependent variable of the model is the chosen residential location, which comes from the London Household Survey data.

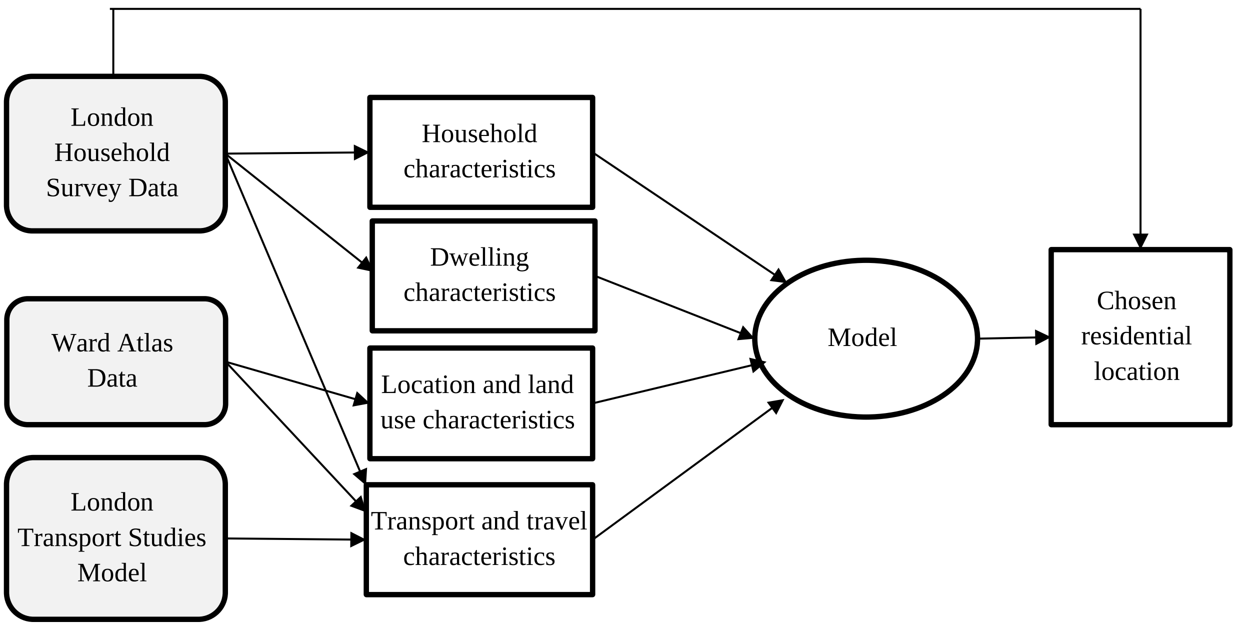

FIGURE 2: Sources of data used for deriving the factors affecting the residential location choices.

London Household Survey Data (LHSD)

The LHSD serves as the main source of disaggregate level household and dwelling information used for model estimation. This dataset was collected in 2002 and contains detailed information (e.g. socio-demographic characteristics, dwelling information, employment status, home and work location, car ownership, etc.) of 8,158 households and 20,910 individuals from 498 wards in the GLA. Multistage stratified random sampling was used in this dataset to ensure representative samples from the selected wards. The dataset contains information on 4,491 households living in privately owned dwellings, 2,489 households living in dwellings rented from councils or housing authorities, 1,087 households living in privately rented dwellings, and 91 households living in shared accommodation. Since only a low number of households have very long tenure length, we retained all households who moved between 1971 and 2002. This study focuses on households living in owned and privately rented dwellings and having at least one member commuting to work which left us with observations from 2,180 owners and 520 private renters.

Ward Atlas Data (WAD)

WAD includes ward level aggregated information of land use pattern, population density, household composition, ethnic composition, employment and economic activity, household income, crime rates, land use, public transport accessibility, green space, car use, and other related variables. Data for the year 2002 was used in this study as the source of location attributes used in the model.

London Transport Studies Model (LTSM)

Information about the distance of the location alternatives (ward) from workplaces and the CBD[4] is missing in the LHSD files, but these are clearly of utmost importance as a determinant of household residential location. The origin-destination (OD) matrix of the GLA from the London Transport Studies Model (LTSM) was used to extract these distances. Since distances did not change for the areas between 2002-2011, distance data from 2011 can be used consistently for the year 2002.

3.3 Data preparation

Combining the datasets was challenging due to the use of different boundaries in different datasets. Old ward boundaries (pre-2002) were used in the LHSD, new ward boundaries (post-2002) were used in the WAD, and traffic analysis zones (TAZ) were used in the LTSM. With the help of GIS map matching, WAD and LTSM data were converted under equivalent old ward boundaries. The layer function in ArcGIS was used to investigate the physical changes of old and new ward boundaries and to convert the WAD to equivalent old boundaries. In some cases where old ward boundaries were found to be similar or wards formed part of a new ward area, attributes of the WAD file in new ward boundaries were kept the same for old ward boundaries. In other cases where the old ward area was found to be shared across multiple new ward areas, the weighted averages of attributes across shared new wards were used for the old ward[5]. We therefore, assume that the attributes are constant within each new ward. This is not expected to introduce substantial inaccuracies since the ward-level data is already based on an assumption of homogeneity. It may also be noted that a similar approach has been used in previous studies (e.g. Habib and Miller, 2009). TAZ boundaries are also converted to equivalent old ward boundaries with the help of ArcGIS. The centre to centre distances between the converted ward boundaries was then used to extract the distance between household work and home locations. We acknowledge that considering the centre to centre distance instead of the actual disaggregate level distance between work and home locations has contributed to some inaccuracies in the values of this variable.

3.4 Data representativeness

A multistage stratified random sampling technique is used in the London Household Survey. Therefore, sampling weights are provided with the dataset to ensure that the weighted data is representative of the London population according to the 2001 census with regards to demographic variables (such as gender distribution, household type, and ethnic composition), economic

[4] The city of London is considered as the central business district (CBD) and the ward Cordwainer is considered as the centre of the CBD. Although the CBD of London is changing over time and extending, the City of London is the oldest and biggest part of the London CBD (Greater London Authority, 2008)

[5] If old ward area p was comprised of 10% of new ward area x, 20% of new ward area y and 70% of new ward area z – for attributes like crime rate, the crime_rate_p was calculated as 0.1*crime_rate_x + 0.2*crime_rate_y +0.7*crime_rate_z

variables (such as household income and employment status) and housing tenure variable (such as owners, private renter, social renters) (Greater London Authority, 2003).

In this study, we have considered a subset of the data that consists of the commuter households (households having at least one working member) who lived in either owned or privately rented houses. The data for the non-commuter households and households living in socially rented houses have been excluded. The characteristics of these two groups (commuter households and non-commuter households) are most likely to be different in the population. For example, the annual income of the households that do not have working member(s) is most likely to be low compare to the households having working member(s). Similarly, the non-commuter householders are most likely to be retired elderly people or unemployed people, therefore the car ownership level of this group can be lower. Since the weighted full sample is representative of the London population, the distribution of the characteristics of the commuter and non-commuter households in the weighted full sample are likely to be different, potentially leading to differences in the characteristics of the households in the subsample and the full sample (see Appendix A for the detailed comparison). As mentioned before, the sampling weights provided with the dataset have been calculated ensuring the representativeness of the different sociodemographic classes (e.g. commuters-noncommuters, owners-renters) in the dataset to the London population. Therefore, the weighted subsample used for analysis in this study is representative of the corresponding group in the population.

3.5 Data analysis

Descriptive analysis of the data reveals significant differences in location and dwelling attributes, travel behaviour and socio-demographic characteristics between owners and renters which are summarised in Table 1. The split of owners in Inner and Outer London is quite different from renters (29.3:70.7 and 56.3:43.7 respectively). The average tenure length of owners is more than three times higher than that of renters. The percentages of owners and renters belonging to the high-income group (more than £60,000 per year) are 28.7% and 27.3% respectively. This is in agreement with previous studies in London which also report substantial portions of high-income people preferring to rent in Inner London due to excessive dwelling prices (Paccoud and Mace, 2017). For the lower-income group (less than £30000 per year), the corresponding shares are 51.0% and 39.4% for renting and ownership respectively. The average household size (number of members in the household) is found to be higher for owners than renters. The rate of car ownership for households living in their own properties is more than 50% higher than for households living in rented properties with a difference of around 20% more for properties owned in Outer London. Around 50% of households who live in owned properties are married couples whereas only 20.5% of households who live in rented properties belong to this group.

Table 1 Descriptive statistics of LHSD

Attributes	Tenure Group	
	Owners	**Renters**
Socio-demographic characteristics		
Annual household income		
Less than £30,000	39.4%	51.0%
Between £30,000 to £60,000	31.9%	21.7%
More than £60,000	28.7%	27.3%
Average household size (members in the household)	2.9	2.7
Household composition		
Married couple with and without kids	51.5%	20.5%
Cohabiting couple with and without kids	14.4%	17.4%
Single member household	24.7%	29.4%
Household having more than one member	9.4%	32.7%
Ethnic composition		
White people	79.5%	74.9%
Asian people	12.9%	14.2%
Black people	7.6%	10.9%
Employment status		
Households have at least one working member	98.9%	90.8%
Households do not have any working member	1.1%	9.2%
Location and dwelling features		
Residential location		
Inner London	29.3%	56.3%
Outer London	70.7%	43.7%
Average dwelling size (number of bedrooms)		
Inner London	2.5	2.4
Outer London	2.9	2.6
Average tenure length (in years)		
Inner London	8.8	2.0
Outer London	10.6	2.8
Travel behaviour		
Car ownership		
Inner London	76.0%	40.8%
Outer London	89.5%	66.4%
Travel mode		
Private car		
Inner London	32.1%	11.7%
Outer London	51.2%	34.7%
Public transport (bus, train, tube)		
Inner London	21.8%	31.9%
Outer London	17.6%	28.3%
Others (motorcycle, pedal cycle, walk, etc.)		
Inner London	46.1%	56.4%
Outer London	31.2%	36.9%
Average commute distance (in kilometres)		
Inner London	7.5	7.2
Outer London	11.1	7.9

There are also substantial differences in the commuting behaviour of owners and renters. Owners are more dependent on private cars (32% in inner London) than renters (12% in inner London), with the percentages again varying largely between Inner and Outer London. Households living in Outer London are found to be more car-dependent whereas households living in Inner London are found to be more transit-oriented. The average commute distances of owners both in Inner and Outer London are higher on average than those for renters. These differences serve as motivation for our work and provide useful insights for the model specifications that are presented in the following section.

4 MODEL DEVELOPMENT

4.1 Model structure

Model parameters were estimated using the discrete choice analysis (DCA) technique. DCA is a widely used technique to analyse consumer choices in which the available options are discrete in nature and mutually exclusive. We started with the most basic version of a discrete choice model: a Multinomial logit (MNL) model. To capture random taste heterogeneity across households as well as differences in error variance between owners and renters, mixed multinomial logit models (MMNL) were estimated later.

A key decision in any study of residential location is the level of disaggregation. We focussed on zone level models, where average dwelling, land use and transport characteristics of each zone are used as independent variables to investigate the households' preferences for their residential zones. Individual zones are considered as location alternatives. This technique clearly has limitations in capturing households' sensitivities to variation in the dwelling attributes within a zone. Dwelling level models, where an individual dwelling is considered as a choice alternative, can capture such variation at the level of the dwelling, but the application of this approach is limited in the literature due to a lack of dwelling supply data for many metropolitan cities, including London. Therefore, a wider application of the zone level approach is observed in the literature (Bhat and Guo, 2004; Zondag and Pieters, 2005; Walker and Li, 2007; Chen et al., 2008; Pinjari et al., 2011; Sener et al., 2011).

A further decision to be made by the analyst concerns the definition of the set of possible options/alternatives. While many studies consider reduced choice sets (e.g. Farooq and Miller, 2012; Rashidi et al., 2012; Zolfaghari, 2013), this requires an analyst to make decisions on what alternatives are considered by each household, potentially biasing the result. In our study, the full choice set is considered for everyone, motivated in part by the observed long-distance mobility of households within the Greater London Area (from one end to another end). That is, households are assumed to evaluate all 498 zones/wards and choose the one they perceive to be the best. This is, of course, a major assumption, but was perceived as a better option than making any arbitrary assumptions about restricted choice sets (which have dominated the residential location choice literature) given the unique mobility patterns in London. This is also in agreement with the

Framework Housing Market Area (HMA) definitions set for London where much of Greater London is identified as a single Framework HMA (Jones et al. 2010).

Our modelling work is based on the principle of utility maximisation, assuming that decision makers choose the alternative that provides them with the greatest utility. The modelling work aims to explain the way individual households choose between mutually exclusive alternatives by estimating the importance they place on the characteristics of these alternatives, where this potentially varies across households. Of course, the actual process of preference formation is not observed by the analyst, and there thus remains a role for an error term in the models, capturing the various influences on decision making not explained by the analyst.

In our analysis, we incorporate a number of potential key effects, as follows:

- heterogeneity in preferences linked to observed characteristics, such as income;
- differences in preferences between owners and renters;
- random (i.e. unexplained) variations in preferences between individual decision makers; and
- differences in the amount of error variance (i.e. unexplained influences on behaviour) for owners and renters.

The utility for zone j for household n is given by:

$$U_{nj} = \sum_{k=1}^{K^f} \left(\beta_k^f + \Delta_{kr} r_n\right) x_{njk}^f + \sum_{l=1}^{L^h} \left(\beta_{lno}^h o_n + \beta_{lnr}^h r_n\right) x_{njl}^h + \xi_{rn} r_n + \xi_{on} o_n + \varepsilon_{nj}, \tag{1}$$

The components of this specification are detailed below.

The first part of the utility specification relates to parameters that do not follow a random distribution across individual households. The model uses K^f such parameters, where these are associated with individual attributes, e.g. x_{njk}^f. In this first part of the utility function, we also incorporate shifts in sensitivity between owners and renters; that is, the marginal utility is β_k^f for owners, and $\beta_{kr}^f = \beta_k^f + \Delta_{kr}$ for renters, where r_n is a dummy for renters (*1* if observation *n* corresponds to a renter, *0* otherwise), and o_n is a dummy for owners (*1* if observation *n* corresponds to an owner, *0* otherwise). Statistical significance of Δ_{kr} thus denotes if the sensitivity for renters is significantly different from that for owners for the attribute x_{njk}^f . The subscript *n* on the attributes relates to the fact that attributes are not just zone-specific but also household-specific given the incorporation of deterministic heterogeneity. For example, for cost, multiple parameters are estimated in the model, with different cost sensitivity for different income groups, and only one of these is used for any given household, with the associated cost attribute set to zero for any income levels that do not apply for that household. The standard errors of the renter specific parameters ($\beta_k^f + \Delta_{kr}$) are obtained using the Delta method that produces exact estimates with full maximum likelihood properties (Daly, Hess and de Jong, 2012). The standard errors for β_{kr}^f are calculated using the formula below (Daly, Hess and de Jong, 2012)

$$\sigma_{\beta_{kr}^f} = \sqrt{var(\beta_k^f) + var(\Delta_{kr}) + 2cov(\beta_k^f, \Delta_{kr})} \quad (2)$$

The second part of the utility specification relates to parameters that follow a random distribution across individual households, i.e. incorporating unobserved heterogeneity in preferences. β_{lno}^h and β_{lnr}^h represent parameters that follow a random distribution across the households (i.e. incorporating unobserved heterogeneity (h) in preferences), with separate groups for owners and renters ; x_{njl}^h are the corresponding attributes. In this case, we explicitly estimate owner and renter specific coefficients (as opposed to shifts), where this is more convenient in the estimation software. We allow for differences between owners in renters in both the mean sensitivities and the level of heterogeneity.

The third component, $\xi_{rn} r_n + \xi_{on} o_n$, allows for differences between owners and renters in the amount of noise in the utility. In the discrete choice technique, the variance of the unobserved factors for one group can be different from that for the other groups – this can reflect a number of different effects, either more noise in the attributes for one group or a greater role for unobserved attributes. If a model specification does not control for this, then the parameters for the two groups cannot be compared other than in the form of relative sensitivities. (e.g. Carrasco & de Dios Ortúzar, 2002; Train, 2003; Hensher, Rose, & Greene, 2015; Hess and Train, 2017). We rely on an error components approach (e.g. Brownstone et al., 2000, Hensher, Rose, & Greene, 2015, Hess and Train, 2017) instead of the nested logit "trick" (e.g. Hensher and Bradley, 1993; Ben-Akiva et al., 1994; Bradley and Daly, 1997; Ho & Mulley, 2013) given that we also incorporate other random heterogeneity through mixing. ξ_{rn} and ξ_{on} are Normally distributed disturbances, with a mean fixed to *0* and an estimated standard deviation. They are shared across all zones and vary randomly across individuals within the group (owners or renters). A larger standard deviation for an error component then indicates more noise.

The final component, ε_{nj}, is the type I extreme value error term, distributed randomly across individuals and across zones.

Some normalisation is required for this model, as follows:

- At least one of the attributes needs to be treated as having generic sensitivity between owners and renters in order to be able to also estimate the difference in the error variance (otherwise the estimation would equate to two separate models which would prevent the estimation of the additional error term). After comparison of group specific models, we fixed the sensitivity to crime to be generic between the two groups as the coefficient was most similar for this attribute.

- Only one of the two error components for differences in noise, i.e. ξ_{on} or ξ_{rn}, can be estimated, with the other fixed to zero. After comparing specifications estimating either ξ_{rn} or ξ_{on}, we found that the noise for renters was higher than for owners, and thus fixed $\xi_{on} = 0$, estimating only ξ_{rn}.

Given the type I extreme value distribution for ε_{nj}, the probabilities in our model are of the Logit form, with the probability of household *n* choosing zone *i* given by:

$$P_{ni}(\beta^f, \Delta_r, \beta_n^h, \xi_{rn}, x_n) = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}} \tag{3}$$

Where $V_{nj} = \sum_{k=1}^{K^f}\left(\beta_k^f + \Delta_k r_n\right)x_{njk}^f + \sum_{l=1}^{L^h}\left(\beta_{lno}^h o_n + \beta_{lnr}^h r_n\right)x_{njl}^h + \xi_{rn} r_n$, i.e. dropping the extreme value error term ε_{nj} and the normalised ξ_{on} term. This probability is conditional on the attributes x_n, estimates for the fixed parameters $\beta^f = < \beta_1^f, \dots, \beta_K^f >$ and shift parameters $\Delta_r = < \Delta_{1r}, \dots, \Delta_{Kr} >$, and specific realisations of the heterogeneous parameters $\beta_n^h = < \beta_{1no}^h, \dots, \beta_{Lno}^h, \beta_{1nr}^h, \dots, \beta_{Lnr}^h >$ and the error term ξ_{rn}. Given the random distribution of these parameters, the unconditional probability is given by:

$$P_{ni}(\beta^f, \Delta_r, \Omega^h, \sigma_r, x_n) = \int_{\beta_n^h}\int_{\xi_{rn}} \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}} h(\beta_n^h \mid \Omega^h)\phi(\xi_{rn}) d\beta_n^h d\xi_{rn} \tag{4}$$

where this is now conditional on the estimated parameters only, i.e. the vector β^f of fixed coefficients and shift parameters Δ_r, the vector of parameters Ω^h for randomly distributed coefficients, and the standard deviation of the error component for renters, i.e. σ_r. In Equation (4), the conditional logit probabilities are then integrated over the distribution of the random terms, with density functions $h(\beta_n^h \mid \Omega^h)$ and $\phi(\xi_{rn})$.

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

$$LL(\beta) = \sum_{n=1}^{N}\sum_{i=1}^{I} y_{ni} \log P_{ni}(\beta^f, \Delta_r, \Omega^h, \sigma_r, x_n) \tag{5}$$

where, $y_{ni} = 1$ if household *n* chose zone *i* and *0* otherwise. Maximisation of this *LL* function yields the maximum likelihood estimates for model parameters. This log-likelihood incorporates the integral in Equation (4), which does not have a closed form solution, and the model is thus estimated using numeric simulation.

Models are estimated with different specifications (e.g. generic coefficients for both owners and renters; owner and renter specific coefficients). The likelihood ratio (LR) test value is used for comparing competing models where the LR was calculated using Equation (6)

$$LR = -2[LL_r - LL_u] \tag{6}$$

Where, LL_r is the log-likelihood for the restricted model and LL_u is the log-likelihood of the unrestricted model. The LR can be compared to a critical value from a χ^2_K distribution with K degrees of freedom, where K=$K_u - K_r$, where K_u and K_r are the numbers of the estimated parameters in the unrestricted and the restricted models, respectively.

4.2 Variable specification

A set of location attributes including land use, dwelling and transport attributes are considered as explanatory variables for this study (see Table 2 for details). The household characteristics are interacted with the location attributes to capture the systematic taste variation (preference heterogeneity) across different groups of households.

A list of potential attributes for residential location choice modelling is identified based on a literature survey (e.g. Bhat and Guo, 2004; Habib and Miller, 2009; Guevara, 2010; Zolfaghari, 2013. Hoshino, 2011). All the potential attributes available in the dataset are tested in the models and the attributes for the final model are selected based on the goodness-of-fit of the model and the t-stats of the estimated parameters. There is a risk of independent variables being strongly correlated to each other which may have serious consequences for the estimated parameters. We checked the correlation between the attributes and found a weak correlation in most of the cases (the correlation matrix is attached in Appendix B). For example, the correlation between the commute distance and distance from the CBD is found to be 0.19 which indicates a weak correlation (Rumsey 2016). The estimation results are thus unlikely to be affected seriously due to the correlation between the independent variables.

TABLE 2 Description of the variables

Variables	Sources	Description*
Dwelling characteristics		
Dwelling cost	WAD & LHSD	This variable represents the median price of the properties in each ward for owning and the median monthly rent of the properties in each ward for renting (in GBP).
Dwelling type	WAD & LHSD	This variable gives the proportion of detached, semidetached, terraced houses and flats in each ward. In estimation, the proportion of the semi-detached and terraced houses are considered as joint base, with parameters estimated for the proportion of detached houses and flats are considered.
Location and land use characteristics		
Land use type	WAD & LHSD	This variable gives the percentage of residential and commercial areas in the alternative wards.
Land use mix	WAD & LHSD	This is an index of the homogeneity/ heterogeneity of land-use in the wards. Its scale ranges from *0* to *1*, where these limits stand for pure homogeneous and uniform mixed land use patterns, respectively. It is computed as $\sum_j \frac{[P_j \times ln\,(P_j)]}{ln\,(J)}$ (Frank et al., 2004), where P_j = the proportion of the land area used for the j^{th} land-use category. A positive coefficient of this variable will indicate a preference for mixed land use patterns.
Ethnic/racial composition	WAD & LHSD	This variable gives the percentage of households in each ward belonging to the different ethnic/racial groups such as white ethnicity, black ethnicity, Asian ethnicity and mixed. Mixed refers to households having members from more than one ethnic/racial classes, and this is considered as the base in estimation.
Dwelling density	WAD & LHSD	This variable gives the total number of dwellings per square kilometre in each ward.
School quality	WAD & LHSD	This variable is reported by the Greater London Authority based on GCSE (General Certificate of Secondary Education) average point score for all wards. This variable is considered for the households having at least one school going child.
Crime rate	WAD & LHSD	This variable gives the total number of crimes reported in each ward per year per thousand members of population.
Average household size	WAD & LHSD	The absolute difference between individual household size and ward level average household size is used for this variable.
Employment opportunity	WAD & LHSD	This is the ratio at the ward level between the total number of job opportunities and the size of the population.
Distance from CBD	LTSM & LHSD	This is the distance between the centre of the City of London and the alternative wards in kms.
Transport and travel characteristics		
Public transport accessibility	WAD & LHSD	This is also a ward level variable that is measured by the Greater London Authority based on walk access time, service availability and network density. The range of this variable is 0 to 8 where 8 represents the highest level of accessibility.
Commute distance	LSTM & LHSD	This is the distance in kms of the individual work location from the centre of the alternative wards. For multiple working member households, the maximum value (distance between work location and potential location alternatives) among the workers in the household is used as in other literatures (e.g. Lee and Waddell, 2010).
Zonal constants		Since the total number of alternatives is very large (498), we use constants for broader areas. Five constants are used, dividing the zones into central, north, south, east and west London where the constant for west London is normalized.

WAD-Ward Atlas Data, LHSD-London Household Survey Data, LTSM-London Transport Studies Model
* All the variables except commute distance present the ward level information for the year 2002.

5 RESULTS

5.1 Estimated model parameters

Mixed multinomial logit (MMNL) models are estimated in this study for investigating the residential location choice behaviour of owners and renters using the professional software ALOGIT[6].

A systematic model specification process was used. After incorporating deterministic heterogeneity (e.g. income effects), we tested for random heterogeneity. With a large choice set and sample size, this is a computationally burdensome process, and was thus carried out prior to the incorporation of differences between owners and renters. Our results indicated significant random heterogeneity only for commute distance, where we used a negative Lognormal distribution.

The scale heterogeneity between the two groups is also captured by means of relative variance where the error variance for the owners is normalized. The estimated standard deviation of the error components of renter specific utilities is found to be very small and not significantly different from zero, indicating no significant scale differences between the owner and renter specific utilities.

Our core focus of the analysis then turned to establishing the differences in behaviour between owners and renters. The estimated parameters of the final models are presented in Table 3.

Two pooled models are developed first where generic coefficients are estimated for all variables assuming equal sensitivity for owners and renters, with the only difference between them being the amount of noise in the utility. In the second model, generic and shift parameters are estimated for all variables assuming different sensitivities of owners and renters (called second pooled model in the remainder of this paper).

The null hypothesis is that 'the model that assumes different sensitivity for owners and renters to all variables is not statistically different from the model that assumes equal sensitivity of all variables for both groups'. The estimated model with generic coefficients for owners and renters for all variables results in a significant loss of fit compared to the model with specific coefficients for owners and renters. A likelihood ratio test (χ^2=160.2, degree of freedom (*DF*) =30, P=0.001) strongly rejects the null hypothesis. It confirms the existence of preference heterogeneity between owners and renters in their residential location choice, even after accounting for differences in the amount of utility variance (where this was not significant in any case). However, among all the parameters, shifts for only five parameters (commute distance, public transport accessibility of the

[6]ALOGIT was found to have significantly shorter run times than the other comparable programs like R which prompted the choice.

households who own car(s), percentage of detached houses in Inner London, percentage of detached houses in Outer London and percentage of flats in Outer London) are found to be statistically significant above the 90% confidence interval.

A third pooled model is then estimated, retaining only those shifts that are statistically significant. The null hypothesis is that 'the model that assumes owner and renter specific sensitivities for a specific subset of variables is not statistically different from the model that assumes different sensitivity of owner and renters to all variables'. A likelihood ratio test (χ^2=11.5, DF=21, P=0.001) then no longer rejects the third pooled model. The final pooled model helps to reduce the estimation time by minimizing the number of parameters estimated without significantly affecting the goodness-of-fit or the accuracy of the estimates of the models. In the following sections, we have discussed the similarities and differences in the owner and renter specific parameters.

Similarities between owner and renter specific parameters

As seen in Table 3, the parameters for owners and renters have the same direction of sensitivity but the magnitudes of some of the coefficients are found to be significantly different. The influences of dwelling attributes on residential location choices are in general found to be significant for both groups. For example, the housing cost sensitivities of both owners and renters are found to be negative as expected and different income groups exhibit different levels of cost sensitivities (which is in agreement with the findings of Habib and Miller, 2009; Zolfaghari, 2013). Households from lower-income groups are observed to be more cost-sensitive than higher-income groups both for ownership and renting. All else being equal, zones having more detached houses are also found to be less preferable options – both for owning and renting. The disutility is found to be higher in Inner London than in Outer London.

The second group of attributes included land-use and location characteristics. These are also found to have considerable influence on residential location decisions. Households are found to have higher utilities for areas with higher concentration of residential properties and less commercial activities for both owning and renting. Although households are found to have a lower preference for areas with higher levels of dwelling density, results indicate that they prefer mixed land use patterns, with a high accessibility to job, shopping, transport and other facilities. This agrees with the findings of other studies – for instance, Arundel and Ronald (2017) have advocated mixed land-use for ensuring the sustainability of a community while absolute density is mentioned not to be effective. Preferences for ethnic/racial similarity are found to have a positive and statistically significant effect for both groups which suggests that people prefer to live in an area where a higher number of households come from the same ethnic/racial group - this is supported by findings of previous studies (e.g. Ibraimovic and Hess, 2016). School quality (only considered for households with children) is found to have a positive effect for both owners and renters which is similar to findings in the literature (Zhou, B. B. & Kockelman, K. M. 2008). Crime rates and household size (absolute difference between each household size and the zonal average) are found to affect the utility of owners and renters negatively. This indicates the clustering of households with similar household size (Zolfaghari, 2013 observed similar finding). Although households are found to be inclined to choose areas having greater employment opportunities, they are found to be less interested to live in and around the central business district (CBD).

The third group of attributes consisted of transport and travel attributes. An increase in public transport accessibility is found to increase the utility of 'car-less' households but decreases the utility for 'car-owning' households. As expected, increased commute distance is found to result is greater disutility. With the use of a negative Lognormal distribution, the estimated parameters are the mean and standard deviation of the logarithm of $-\beta$. The standard deviation reveals significant taste heterogeneity across households. Both the mean and standard deviation of $\log(-\beta)$ are significantly different between owners and renters.

Constants are estimated to capture the utility of all factors that are not explained by the included explanatory variables. Since the total number of alternatives is very large (498), constants are estimated at the aggregate level. In particular, separate constants are estimated for the alternatives in central, north, south, east and west London. The constants were found to be the highest for North London for both owners and renters. It may be noted that a higher value for a constant does not indicate that this is a preferred zone. Rather, the estimated constants capture the effects of factors that are not included in the model (i.e. are unobserved). Therefore, the results indicate that the share of unobserved factors affecting the choice of North London is higher than that of the other four parts.

Based on the data analyses presented earlier indicating substantial differences between Inner and Outer London, statistical tests are conducted to test if the sensitivities to the variables corresponding to Inner and Outer London are statistically different from each other. The results of these tests indicate that the sensitivity towards four of the variables (% detached houses, % flats, % residential area in the ward and dwelling density) are significantly different between Inner and Outer London (Table 3). Parameters which are not significantly different between Inner and Outer London are estimated as generic coefficients for the whole of London.

TABLE 3 Estimation results

Parameters	Pooled model with generic coefficients for owners and renters for all variables		Pooled model with generic coefficients and shifts (for renters) for all variables*				Renter-specific (Computed)		Pooled model with generic coefficients and statistically significant shifts (for renters)				Renter-specific (Computed)	
			Generic (Owner-specific also)		Shift (for renters)				Generic (Owner-specific also)		Shift (for renters)			
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Constants														
Central London	0.274	2.7	0.269	2.1	-0.082	-0.4	*0.187*	*1.1*	0.250	2.4			*0.250*	*2.4*
South London	0.359	4.5	0.319	3.5	-0.047	-0.2	*0.271*	*1.3*	0.304	3.8			*0.304*	*3.8*
North London	0.577	5.9	0.489	4.4	0.103	0.4	*0.592*	*2.6*	0.505	5.1			*0.505*	*5.1*
East London	0.555	6.3	0.481	4.7	-0.150	-0.7	*0.331*	*1.7*	0.447	5.0			*0.447*	*5.0*
Dwelling characteristics														
Dwelling cost (price*0.0001, monthly rent*0.01)														
Household income less than £30,000	-0.259	-9.0	-0.557	-8.9	0.353	5.1	*-0.204*	*-7.1*	-0.567	-9.3	0.367	5.6	*-0.199*	*-7.3*
Household income between £30,000 to £60,000	-0.249	-7.9	-0.444	-6.8	0.283	3.9	*-0.161*	*-5.3*	-0.453	-7.0	0.295	4.2	*-0.158*	*-5.4*
Household income more than £60,000	-0.087	-5.5	-0.200	-6.9	0.127	3.7	*-0.073*	*-4.0*	-0.207	-7.3	0.139	4.3	*-0.068*	*-4.1*
Missing values	-0.023	-2.2	-0.028	-1.2	-0.017	-0.6	*-0.045*	*-3.1*	-0.041	-3.6			*-0.041*	*-3.6*
Dwelling type														
Detached dwelling in inner London	-0.107	-7.3	-0.139	-7.4	0.113	3.8	*-0.026*	*-1.1*	-0.131	-7.2	0.085	3.0	*-0.045*	*-2.0*
Detached dwelling in outer London	-0.027	-6.9	-0.029	-6.7	0.032	3.0	*0.003*	*0.3*	-0.028	-6.6	0.029	3.0	*0.001*	*0.1*
Flat in inner London	0.036	11.4	0.034	9.3	0.004	0.6	*0.038*	*6.2*	0.035	11.0			*0.035*	*11.0*
Flat in outer London	-0.010	-4.2	-0.012	-4.3	0.017	2.8	*0.005*	*1.0*	-0.012	-4.9	0.019	4.3	*0.007*	*1.6*
Location and land use characteristics														
Land use type														
Residential land area in inner London	0.167	13.8	0.170	11.6	-0.006	-0.2	*0.164*	*6.8*	0.166	13.9			*0.166*	*13.9*
Residential land area in outer London	0.239	14.7	0.248	13.5	-0.036	-0.9	*0.212*	*6.0*	0.242	14.9			*0.242*	*14.9*
Commercial land area in inner and outer London	-0.059	-7.1	-0.061	-5.5	0.002	0.1	*-0.059*	*-5.1*	-0.060	-7.2			*-0.060*	*-7.2*
Land use mix	1.630	5.5	1.450	4.3	0.770	1.0	*2.220*	*3.2*	1.589	5.3			*1.589*	*5.3*
Ethnic composition													*0.000*	*0.0*
Ratio of white households × White dummy	0.016	9.4	0.017	8.8	0.002	0.5	*0.019*	*4.7*	0.017	10.3			*0.017*	*10.3*
Ratio of asian households × Asian dummy	0.041	13.8	0.038	11.6	0.005	0.8	*0.044*	*7.4*	0.040	13.6			*0.040*	*13.6*
Ratio of black households × Black dummy	0.058	11.0	0.057	8.9	-0.012	-1.1	*0.045*	*5.3*	0.053	10.3			*0.053*	*10.3*

TABLE 3 Estimation results (cont.)

Parameters	Pooled model with generic coefficients for owners and renters for all variables		Pooled model with generic coefficients and shifts for renters for all variables*				Renter-specific (Computed)		Pooled model with generic coefficients and statistically significant shifts (for renters)				Renter-specific (Computed)	
			Generic (Owner-specific also)		Shift (for renters)				Generic (Owner-specific also)		Shift (for renters)			
	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat	Coeff.	t-stat
Dwelling density														
Inner London	-0.027	-6.5	-0.026	-5.1	0.001	0.1	*-0.025*	*-2.6*	-0.026	-6.3			*-0.026*	*-6.3*
Outer London	-0.125	-17.5	-0.127	-15.2	0.008	0.5	*-0.119*	*-8.7*	-0.126	-17.4			*-0.126*	*-17.4*
School quality	0.007	4.7	0.008	4.7	0.001	0.3	*0.009*	*2.2*	0.008	5.3			*0.008*	*5.3*
Crime rate	-0.117	-3.3	-0.124	-3.3			*-0.124*	*-3.3*	-0.120	-3.3			*-0.120*	*-3.3*
Household size	-0.357	-4.7	-0.404	-5.0	0.183	0.9	*-0.220*	*-1.2*	-0.358	-4.8			*-0.358*	*-4.8*
Employment opportunity	0.176	5.2	0.169	4.6	0.029	0.8	*0.198*	*4.7*	0.185	5.5			*0.185*	*5.5*
Distance from CBD	0.056	8.0	0.050	6.3	-0.017	-1.0	*0.033*	*2.2*	0.046	6.5			*0.046*	*6.5*
Transport and travel characteristics														
Public transport accessibility														
Households own car	-0.156	-3.9	-0.179	-3.8	0.211	2.2	*0.032*	*0.4*	-0.192	-4.5	0.256	3.9	*0.064*	*1.0*
Households do not own car	0.460	9.4	0.359	5.5	-0.013	-0.1	*0.346*	*3.3*	0.354	7.0			*0.354*	*7.0*
Commute distance **														
Mean of *log(-β)*	-1.681	-56.1	-1.711	-63.1	*0.165*	*3.2*	-1.546	-34.5	-1.716	-63.8	*0.183*	*3.6*	-1.533	-35.1
Standard deviation of *log(-β)*	0.185	61.7	0.137	78.9	*0.172*	*5.9*	0.309	10.6	0.134	74.6	*0.175*	*6.0*	0.309	10.6
Standard deviation of the error for renters (inverse function of scale effect)	0.071	0.7			0.057	0.3					0.052	0.4		
Measures of model fit														
Number of observations	2700		2700						2700					
Initial LL	-16768.620		-16768.620						-16768.620					
Final LL	-12944.248		-12864.170						-12869.897					
Adjusted ρ2	0.226		0.229						0.230					

* The parameter crime rate is considered as shared between owners and renters to allow us to capture scale heterogeneity.
** For this random parameter, owner and renter specific coefficients are estimated and the shift parameter is calculated.

Differences between owner and renter specific parameters

This section summarizes the differences in the sensitivities of owners and renters in their residential location choice, based on the estimated parameters. It may be noted that a direct comparison is possible because a) the scale difference between the owner and renter specific parameters is explicitly captured and b) the variables are defined in the same manner for both groups (except cost and commute distance). However, for additional interpretation, we have calculated the elasticity and WTP (Willingness to Pay) values which are discussed in the next section.

The shift parameters for four variables: percentage of detached houses in Inner London, percentage of detached houses in Outer London, percentage of flats in Outer London and public transport accessibility of the households who own cars are found statistically significant at the 95% level of confidence. Owners are found to have a greater dislike than renters for areas having a high percentage of detached houses both in Inner and Outer London which is aligned with previous findings in the literature (e.g. Paccoud and Mace (2017). Owners are found to have a reduced utility for areas with a higher percentage of flats in Outer London whereas the preferences of renters are the opposite but statistically less significant. Dwelling owners who own car(s) are found to have a reduced utility for areas with a high level of public transport accessibility. The preference of renters who own car(s) for public transport accessibility is positive but insignificant. This may be because homeowners (who are likely to be from higher income groups and/or have better parking arrangements) may have multiple cars for active travellers in the household resulting in no/reduced demand for public transport use. On the other hand, many renters may have a single car that is used by one member in the household while the other members need to use public transport. These scenarios may result in differences in the sensitivity of owners and renters (who own a car) to public transport accessibility in their choice of residential zone. The shift parameters for housing cost and commute distance are also found to be statistically significant at the 95% confidence interval but the interpretation cannot be made directly based on the estimated parameters since the variables are different for owners and renters. Elasticity analysis in the next section is used for further interpretation of the sensitivity of owners and renters to the residential location choice attributes including housing cost and commute distance.

5.2 Elasticity analysis

The findings of the models are further analysed by looking at elasticities. Elasticity analysis is the more appropriate tool for interpreting the relative impact of model parameters (Washington et al., 2010). It quantifies the percentage change in the choice probability of one alternative due to changes in the value of an attribute of the same alternative (called direct elasticity) or another alternative (called cross elasticity). The well-known formula for calculating direct elasticity of the MNL model parameters (see Train, 2009 for details) is as follows:

$$E_{ix_{ni}} = \frac{\partial V_{ni}}{\partial x_{ni}} x_{ni}(1 - P_{ni}) \qquad (7)$$

where x_{ni} is the attribute of alternative *i* of household *n*, $\frac{\partial V_{ni}}{\partial x_{ni}}$ indicates the changes in the utility of alternative *i* of household *n* due to changes in the attributes of the corresponding alternative, $P_{ni}(\beta_f, \beta_n^h, \xi_r, x_n)$ is the probability of choosing alternative zone *i* by household *n*. The elasticity for the MMNL model is given by the integration of the MNL elasticity over the distribution of distributed parameters. Therefore, the direct elasticity for the MMNL model (see e.g. Hess et al., 2009) is given by

$$E_{ix_{ni}} = \frac{\int_{\beta_n^h}\int_{\xi_{rn}} \frac{\partial V_{ni}}{\partial x_{ni}} x_{ni}(1 - P_{ni}(\beta_f, \beta_n^h, \xi_{rn}, x_n))P_{ni}(\beta_f, \beta_n^h, \xi_{rn}, x_n)h(\beta_n^h \mid \Omega^h)\phi(\xi_{rn})d\beta_n^h d\xi_{rn}}{\int_{\beta_n^h}\int_{\xi_{rn}} P_{ni}(\beta_f, \beta_n^h, \xi_{rn}, x_n)\, h(\beta_n^h \mid \Omega^h)\phi(\xi_{rn})d\beta_n^h d\xi_{rn}} \quad (8)$$

In this study, direct elasticities are calculated for the MMNL models of residential ownership and renting decisions. The elasticities are calculated for all households for changes in attributes of the chosen alternatives (where households currently live)[7]. Then average elasticities across all households are computed. The results are presented in Table 4.

The direct elasticities calculated in this study reflect the change in the likelihood of choosing a residential zone due to changes in the attributes of the zone where the households currently live. As observed in Table 4, households' residential location choices are found to be more elastic (greater or equal to one) to the parameters associated with housing cost for low income owners and low and middle income renters, some dwelling types (flats in inner London), residential land area in the zone, land use mix, ethnic composition (for white and Asian people), dwelling density, school quality, public transport accessibility for households who do not own car, and commute distance. The household residential choices are found to be less elastic (less than one) for the remaining parameters.

The interpretation of the differences in the sensitivities of owners and renters based on the estimated coefficients remains the same in the elasticity analysis. As in the estimated parameters, the elasticities for the share of detached houses in Inner and Outer London, the share of flats in Inner London and public transport accessibility (for those who own a car) are considerably higher for owners than renters and in some cases the signs are opposite (e.g. for detached dwelling in outer London and flats in outer London). The choice of renters is found to be more elastic to housing cost than that of owners and the opposite applies in the case of commute distance. The elasticities of few other parameters such as commercial land use, land use mix, household size, employment opportunity and distance from CBD are found to vary more than 40% between owners

[7] Computing the elasticity for the chosen alternative only is appropriate when the number of alternatives in the choice set is very high, making the calculation of cross-elasticities too burdensome (Sener et al., 2011).

and renters. Therefore, the elasticity analysis also reflects some significant differences in the sensitivities of owners and renters in their residential location choice attributes.

Table 4. Direct elasticities of the owner specific and renter specific parameters

Parameters	Owner-specific	Renter-specific
Dwelling characteristics		
Dwelling cost (price*0.0001, monthly rent*0.01)		
Household income less than £30,000	-1.076	-1.248
Household income between £30,000 to £60,000	-0.886	-1.100
Household income more than £60,000	-0.543	-0.627
Dwelling type		
Detached dwelling in inner London	-0.392	-0.065
Detached dwelling in outer London	-0.225	0.023
Flat in inner London	2.280	2.743
Flat in outer London	-0.419	0.202
Location and land use characteristics		
Land use type		
Residential land area in inner London	2.613	2.589
Residential land area in outer London	2.463	2.264
Commercial land area in inner and outer London	-0.342	-0.476
Land use mix	1.151	1.841
Ethnic composition		
Ratio of White households × White dummy	1.273	1.360
Ratio of Asian households × Asian dummy	1.055	1.162
Ratio of Black households × Black dummy	0.937	0.774
Dwelling density		
Inner London	-1.219	-1.282
Outer London	-2.567	-2.686
School quality	2.298	2.575
Crime rate	-0.152	-0.208
Household size	-0.458	-0.232
Employment opportunity	0.082	0.176
Distance from CBD	0.839	0.430
Transport and travel characteristics		
Public transport accessibility		
Households own car	-0.576	0.120
Households do not own car	1.442	1.578
Commute distance		
Mean	-1.776	-1.666
Standard deviation	0.195	0.422

5.3 Willingness-to-pay (WTP) values

While the analyses in Table 3 and Table 4 indicate the relative influence of the residential location choice variables, willingness to pay (WTP) analysis can help to translate them into monetary values. As mentioned in the first section, WTP values can be used directly in cost-benefit analysis to evaluate alternate policies, making them a very useful tool for quantifying the monetary value associated with improvement or deterioration in the level of an attribute. For example, WTP for decreasing dwelling density will indicate how much extra rent (or price in the case of ownership) a household is ready to pay for each unit of decrease in dwelling density. With *k* referring to a given non-cost attribute, WTP can be estimated using the following expression:

$$WTP_k = -\frac{\frac{dV_{ni}}{dk}}{\frac{dV_{ni}}{dcost}} = -\frac{\beta_k}{\beta_{cost}}, \tag{9}$$

where, in a linear in attributes model, β_k is the sensitivity to attribute *k* and β_{cost} is the cost coefficient (monthly rent or dwelling price).

In this case, WTP values are calculated for the parameters that influence the residential location decision of owners and renters. The results are presented in Table 5 and explained here. As observed in the table, there is a distinct impact of income. The higher-income households are willing to pay significantly more compared to the lower-income group. The willingness to pay for owners is found to be negative for areas with an increased share of detached houses in Inner London, detached houses in Outer London and flats in Outer London and positive for an increased share of flats in inner London. However, willingness to pay for renters are negative for increases in the share of detached houses in Inner London but positive for increases in detached houses in Outer London and flats in Inner and Outer London. Households are more interested in residential areas in Outer London than in Inner London, therefore, their willingness to pay per unit increase of residential area in Outer London is 1.5 times higher than in Inner London. Similarly, households are more sensitive to dwelling density in Outer London than Inner London. For instance, the willingness to pay for increases in dwelling density is negative for both inner and outer London but the rate is five times higher for outer London than in inner London. The WTP for per kilometre saving in commute distance is much higher compared to the per kilometre increase in distance between residential location and the CBD for all income groups (~4 times for owners and 7 times for renters). Both owners and renters are willing to pay more for an increase in the share of households from the same ethnic group in their neighbourhood, more balanced land use, better school quality, higher employment opportunity, better public transport accessibility (car-less households only). Both groups are willing to pay less for increases in the commercial area, crime rate, and household size in their residential zone.

TABLE 5 Willingness to Pay (WTP) for owners and renters

Parameters	Unit	GLA average	Owners WTP (price in GBP)			Renters WTP (monthly rent in GBP)		
			Low income	Middle income	High income	Low income	Middle income	High income
Dwelling type								
Detached dwelling in inner London	Percentage	2%	-24973	-31351	-69600	-13	-16	-36
Detached dwelling in outer London	Percentage	9%	-5224	-6559	-14560	1	2	4
Flat in inner London	Percentage	74%	6015	7552	16765	18	23	52
Flat in outer London	Percentage	36%	-2070	-2599	-5770	3	3	7
Location and land use characteristics								
Land use type								
Residential land area in inner London	Percentage	14%	30553	38356	85150	81	102	225
Residential land area in outer London	Percentage	11%	44528	55901	124100	104	132	290
Commercial land area in inner and outer London	Percentage	7%	-10874	-13651	-30305	-29	-37	-81
Land use mix	Index	0.81	260136	326577	725000	1088	1377	3041
Ethnic composition								
Ratio of White households × White dummy	Percentage	72%	3025	3797	8430	9	12	26
Ratio of Asian households × Asian dummy	Percentage	12%	6880	8637	19175	21	27	60
Ratio of Black households× Black dummy	Percentage	10.50%	10187	12788	28390	22	28	62
Dwelling density								
Inner London	Dwelling per sq. km.	4956	-47	-59	-132	-0.1	-0.2	-0.3
Outer London	Dwelling per sq. km.	2138	-228	-287	-637	-0.6	-0.7	-1.6
School quality	Score	293	1364	1713	3802	4	6	12
Crime rate	Crime per thousand people	135	-222	-279	-620	-1	-1	-2
Household size	Number	0.4	-72390	-90878	-201750	-108	-137	-302
Employment opportunity	Employment per person	0.6	30319	38063	84500	97	123	272
Distance from CBD	Kilometer	15.1*	9055	11367	25235	16	21	46
Transport and travel characteristics								
Public transport accessibility		3.63						
Households own car	Index		-32060	-40248	-89350	16	20	44
Households do not own car	Index		64424	80878	179550	169	214	474
Commute distance		21.4**						
Mean	Kilometer		-32723	-41081	-91200	-110	-139	-306
Standard deviation	Kilometer		4521	5676	12600	35	44	97

*This is the average distance of the CBD from the alternative locations.
**This is the mean of the average distance of the individual work location from the location alternatives.

6 CONCLUSIONS AND APPLICATIONS

In this study, we have investigated differences between owners and renters in residential location choices using RP data combined with multiple other data sources. We make use of publicly available real-world data and are able to estimate residential location models without requiring sampling of alternatives. The paper contributes to the state-of-the-art by addressing the research gap identified in the introductory section. The results indicate that owners and renters have similar preferences (same signs of parameters) but the sensitivities to several attributes are different. A few parameters are found to be significantly different between owners and renters, such as the percentage of detached houses in Inner London, the percentage of detached houses in Outer London, the percentage of flats in Outer London, and public transport accessibility for households who own cars. When investigating residential location choice behaviour, the potential differences in the sensitivities or preferences of owners and renters towards the attributes should thus be acknowledged, as done in our work.

It may be noted that some of the findings from the study are ‘London-specific’ – the shift in preferences towards renting vs. ownership in inner and outer London for example. However, acknowledging the differences in the elasticities and WTP among renters and owners for different land-use and dwelling attributes serves as an important proof-of-concept that incorporating the heterogeneity and the full range of attributes can add value to the detailed cost-benefit analyses. It will be interesting to use these results in Housing Market Analyses which currently use simplifying assumptions and neglect important dimensions of the housing market– for instance, the fact that households have preferences for different dwelling types and neighbourhoods and areas (Jones et al. 2010). Furthermore, the model outcomes can be used for policy formulation and better predicting the impact of alternative policy scenarios due to explicit consideration of the sub-markets. For instance, a comparison of the WTP for public transport accessibility of renters and owners can be conducted with the corresponding WTP values for land-use mix, dwelling density, etc. Accounting for preference heterogeneity between the sub-groups is also expected to lead to better investment decisions.

It may be noted that the use of revealed preference data in this study helped to capture the true preference of households with accurate parameter estimation avoiding the potential bias associated with hypothetical response in stated preference data. Combining data from a range of sources 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 limitations) and is therefore expected to lead to better predictions. There are however several limitations of this study that need to be addressed in future work. First, the full choice set is considered for each household which is very large (498 alternatives in this case) and may not be behaviourally representative for all households. It may be noted that, in reality, the opportunities and constraints do affect the detailed choices but such information is not available in the dataset. However, in this case, the location choices refer to the choice of wards as opposed to dwellings and such availability effects are likely to be reduced due to the coarser granularity. Future research could generate restricted choice sets for each respondent based on behavioural rules rather than considering a full choice set and test the impact of this on results. Second, the geographically closer location alternatives are likely to be more correlated in their unobserved factors than alternatives that are far from each other. Although, this spatial correlation structure among the residential location choice alternatives has been investigated in the literature (e.g. Bhat and Guo, 2004; Sener et al., 2011), our study is limited in this context. Third, the work location is considered as exogeneous in this study. However, the decisions of residential location and work location can be simultaneous or may have two-way interactions. Ignoring this decision interdependency can under/overestimate the correlations among the decisions neglecting the inherent trade-offs. The decision of residential location can also be interdependent with other decisions as well, such as tenure choice, car ownership, amongst others, which was not tested in this study. Finally, the data sets used in this study is from 2002. Although the absolute sensitivities are likely to have changed over time due to the market dynamics and continuous gentrification, we expect that the direction of sensitivity of the estimated parameters still holds in the current context. This is validated by the results of recent literature in the context of London and other similar cities. However, the framework proposed here can be used to investigate the housing market using recent data. Even in its current form, the models provide important behavioural insights on how people trade-off differently when making location choices in different time scales.

ACKNOWLEDGEMENT

The authors acknowledge Economic and Social Research Council and Institute for Transport Studies, University of Leeds for funding this study. Professor Stephane Hess acknowledges the support of the European Research Council through the consolidator grant 615596-DECISIONS. The authors are very thankful to Andrew Daly for his useful thoughts on the models.

REFERENCES

Alexander D and Tomalty R. (2002) Smart growth and sustainable development: Challenges, solutions and policy directions. *Local Environment* 7: 397-409.

Arundel, R. and R. Ronald. 2017. The role of urban form in sustainability of community: The case of Amsterdam. *Environment and Planning B: Urban Analytics and City Science*, 44(1), pp.33-53.

Atkinson R. (2000) Measuring gentrification and displacement in Greater London. *Urban studies* 37: 149-165.

Baum-Snow, N. and Hartley, D., 2019. Accounting for central neighborhood change, 1980-2010. *Journal of Urban Economics*, p.103228.

Ben-Akiva, M., Bradley, M., Morikawa, T., Benjamin, J., Novak, T., Oppewal, H., & Rao, V. (1994). Combining Revealed and Stated Preferences Data. *Marketing Letters, 5*(4), 335-349.

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

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

Boyko CT and Cooper R. (2011) Clarifying and re-conceptualising density. *Progress in Planning* 76: 1-61.

Bradley, M.A., Daly, A.J., 1997. Estimation of logit choice models using mixed stated preference and revealed preference information. In: Stopher, P.R., Lee-Gosselin, M. (Eds.), Understanding Travel Behaviour in an Era of Change. Pergamon, Oxford, pp. 209–232.

Brownstone, D., Bunch, D.S. and Train, K., 2000. Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B: Methodological*, *34*(5), pp.315-338.

Butler T, Hamnett C and Ramsden M. (2008) Inward and upward: marking out social class change in London, 1981—2001. *Urban studies* 45: 67-88.

Carrasco, J. A., & de Dios Ortúzar, J. (2002). Review and assessment of the nested logit model. *Transport Reviews*, *22*(2), 197-218.

Chen, J., Chen, C. and Timmermans, H. (2008). Accessibility trade-offs in household residential location decisions. *Transportation Research Record: Journal of the Transportation Research Board*, 71-79.

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.

Couture, V., & Handbury, J. (2017). Urban Revival in America, 2000 to 2010. National Bureau of Economic Research. Working Paper Series, No. 24084.

Daly, A., Hess, S. and de Jong, G., 2012. Calculating errors for measures derived from choice modelling estimates. *Transportation Research Part B: Methodological*, *46*(2), pp.333-341.

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

Daly, A., Hess, S. and Train, K., (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation*, *39*(1), pp.19-31.

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.

Farooq, B. and Miller, E. J. (2012). Towards integrated land use and transportation: A dynamic disequilibrium based microsimulation framework for built space markets. *Transportation Research Part A: Policy and Practice*, 46, 1030-1053.

Fenton A. (2011) Housing benefit reform and the spatial segregation of low-income households in London.

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.

Greater London Authority. (2008) London's central business district: Its global importance. London, England: GLA Economics.

Greater London Authority (2003) The 2002 London Household Survey, Data Management and Analysis Group (DMAG).

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.

Hamnett C. (2010) Moving the poor out of central London? The implications of the coalition government 2010 cuts to Housing Benefits. *Environment and Planning A* 42: 2809-2819.

Hamnett C and Butler T. (2010) The Changing Ethnic Structure of Housing Tenures in London, 1991—2001. *Urban studies* 47: 55-74.

Hensher, D.A., Bradley, M., 1993. Using stated response data to enrich revealed preference discrete choice models. Marketing Letters 4 (2), 139–152.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2008). Combining RP and SP data: biases in using the nested logit 'trick' – contrasts with flexible mixed logit incorporating panel and scale effects. *Journal of Transport Geography, 16*(2), 126-133.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). *Applied Choice Analysis* (2 ed.). Cambridge: Cambridge University Press.

Hess, S., Ben-Akiva, M., Gopinath, D., 2009. Taste heterogeneity, correlation and elasticities in latent class choice models. Paper presented at the International Choice Modelling Conference 2009, Harrogate.

Hess, S. and Train, K., 2017. Correlation and scale in mixed logit models. *Journal of Choice Modelling, 23*, pp.1-8.

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.

Ho, C. Q., Hensher, D. A., & Ellison, R. (2017). Endogenous treatment of residential location choices in transport and land use models: Introducing the MetroScan framework. *Journal of Transport Geography, 64*, 120-131.

Ho, C. Q., & Mulley, C. (2013). Multiple purposes at single destination: A key to a better understanding of the relationship between tour complexity and mode choice. *Transportation Research Part A: Policy and Practice, 49*(0), 206-219.

Holden E and Norland IT. (2005) Three challenges for the compact city as a sustainable urban form: household consumption of energy and transport in eight residential areas in the greater Oslo region. *Urban studies* 42: 2145-2166.

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*.

Jones C, Coombes M and Wong C. (2010) Geography of housing market areas: Final report. *London, CLG*.

Lee, H. (2018). Are millennials coming to town? Residential location choice of young adults. Urban Affairs Review, 56(2), pp.565-604.

Lee, Y., Lee, B., & Shubho, M. T. H. (2019). Urban Revival by Millennials? Intra-Urban Net Migration Patterns of Young Adults, 1980-2010. Journal of Regional Science. 59(3), 538-566.

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

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.

Masnavi M-R. (2000) The new millennium and the new urban paradigm: the compact city in practice. *Achieving sustainable urban form*: 64-73.

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

Paccoud A and Mace A. (2017) Tenure change in London's suburbs: Spreading gentrification or suburban upscaling? *Urban studies*: 0042098017712832.

Pinjari AR, Pendyala RM, Bhat CR, et al. (2011) Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. *Transportation* 38: 933.

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

Rashidi, T. H., Auld, J. and Mohammadian, A. K. (2012). A behavioral housing search model: Two-stage hazard-based and multinomial logit approach to choice-set formation and location selection. *Transportation Research Part A: Policy and Practice*, 46, 1097-1107.

Rumsey D J, 2016, Statistics for Dummies, 2nd Edition, ISBN: 978-1-119-29352-1

Savage, M., Devine, F., Cunningham, N., Taylor, M., Li, Y., Hjellbrekke, J., Le Roux, B., Friedman, S. and Miles, A., 2013. A new model of social class? Findings from the BBC's Great British Class Survey experiment. Sociology, 47(2), pp.219-250.

Sener IN, Pendyala RM and Bhat CR. (2011) Accommodating spatial correlation across choice alternatives in discrete choice models: an application to modeling residential location choice behavior. *Journal of Transport Geography* 19: 294-303.

Sermons MW and Koppelman FS. (2001) Representing the differences between female and male commute behaviour 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, K. E. 2009 *Discrete choice methods and simulation*. Cambridge university press.

Walker, J. L and Li J. (2007). Latent lifestyle preferences and household location decisions. *Journal of Geographical Systems*, 9, 77-101.

Washington SP, Karlaftis MG and Mannering F. (2010) *Statistical and econometric methods for transportation data analysis*: Chapman and Hall/CRC.

Yates, J. (2001). The rhetoric and reality of housing choice: The role of urban consolidation. *Urban Policy and Research*, *19*(4), 491-527. doi:10.1080/08111140108727895

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

Zhou, B. B. & Kockelman, K. M. 2008. Microsimulation of Residential Land Development and Household Location Choices: Bidding for Land in Austin, Texas. *Transportation Research Record: Journal of the Transportation Research Board,* 2077**,** 106-112.

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

Zondag, B. and Pieters, M. (2005). Influence of accessibility on residential location choice. *Transportation Research Record: Journal of the Transportation Research Board*, 63-70.

Appendix A: Distribution of the characteristics of the households in the subsample and the full sample

Attributes	Subsample	Full sample
Socio-demographic characteristics		
Annual household income		
Less than £30,000	41.6%	64.0%
Between £30,000 to £60,000	29.9%	18.2%
More than £60,000	28.5%	17.8%
Average household size (members in the household)	2.8	2.6
Household composition		
Married couple with and without kids	45.5%	35.0%
Cohabiting couple with and without kids	15.0%	9.9%
Single member household	25.6%	43.2%
Household having more than one member	13.9%	11.8%
Ethnic composition		
White people	78.6%	76.6%
Asian people	13.1%	10.7%
Black people	8.2%	12.7%
Location and dwelling features		
Residential location		
Inner London	34.5%	40.6%
Outer London	65.5%	59.4%
Average dwelling size (number of bedrooms)		
Inner London	2.5	2.5
Outer London	2.8	2.6
Average tenure length (in years)		
Inner London	7.5	10.1
Outer London	9.0	12.6
Travel behaviour		
Car ownership		
Inner London	69.2%	49.0%
Outer London	85.0%	70.0%
Sample size	2700	8159

Appendix B: Correlation matrix of the independent variables

Variables	Commute distance	Flat house in inner London	Flat house in outer London	Detached house in inner London	Detached house in outer London	Ratio of black people	Ratio of asian people	Ratio of white people	Commer cial land area	Domestic land area in outer London	Domestic land area in inner London	Land use mix	School quality	Crime rate	Dwelling density in inner London	Dwelling density in outer London	Public transport acc_car owner	Public transport acc_non car owner	Distance from CBD	Housing cost for high income people	Housing cost for medium income people	Housing cost for low income people	House hold size	Employ ment density
Commute distance	1																							
Flat house in inner London	0.018	1																						
Flat house in outer London	-0.048	-0.178	1																					
Detached house in inner London	-0.016	-0.273	0.06	1																				
Detached house in outer London	-0.051	0.067	0.106	-0.045	1																			
Ratio of black people	0.02	0.015	0.012	0.057	0.052	1																		
Ratio of asian people	-0.001	-0.037	-0.093	0.089	0.027	-0.035	1																	
Ratio of white people	0.027	0.021	0.043	-0.096	-0.057	0.113	0.111	1																
Commerial land area	-0.008	0.08	0.05	-0.048	0.013	0.078	0.162	-0.13	1															
Domestic land area in outer London	0.033	-0.029	-0.352	-0.011	0.253	0.016	0.049	-0.05	-0.084	1														
Domestic land area in inner London	0.021	-0.367	-0.038	0.061	-0.025	0.046	0.002	-0.016	-0.135	-0.06	1													
Land use mix	0.001	-0.013	0.067	-0.037	-0.138	-0.014	-0.072	0.013	0.472	0.013	-0.138	1												
School quality	0.008	-0.007	0.04	0.048	0.139	-0.065	0.064	0.006	-0.029	0.194	0	-0.048	1											
Crime rate	0.019	0.042	0.094	-0.022	0.027	0.037	0.057	-0.066	0.017	-0.032	0.129	0.089	-0.087	1										
Dwelling density in inner London	-0.015	0.377	0.011	-0.181	-0.007	-0.019	0.005	-0.008	0.166	0.037	0.684	-0.034	-0.038	-0.121	1									
Dwelling density in outer London	-0.041	-0.016	0.462	0.007	-0.318	0.012	0	-0.003	0.057	0.653	0.043	-0.013	-0.156	0.005	-0.063	1								
Public transport acc_car owner	-0.019	0.165	0.21	0.027	-0.07	-0.047	0.049	-0.043	0.2	0.147	0.102	0.038	0.079	0.184	0.067	-0.019	1							
Public transport acc_non car owner	-0.002	-0.026	-0.056	0.011	0.029	-0.018	-0.037	0.034	0.06	-0.043	-0.036	-0.062	0.005	0.01	0.036	0.033	0.293	1						
Distance from CBD	0.19	-0.189	-0.213	0.138	0.166	-0.125	-0.051	0.241	0.06	-0.005	-0.06	-0.091	0.018	-0.027	0.044	-0.067	-0.125	0.013	1					
Housing cost for high income people	-0.029	-0.008	-0.005	-0.014	0.069	-0.069	-0.05	0.202	0.012	0.019	0.037	-0.048	0.074	-0.015	0.03	-0.008	0.061	-0.041	-0.141	1				
Housing cost for medium income people	-0.04	0	0	-0.022	0.089	-0.086	-0.06	0.188	0.013	0.032	0.036	-0.048	0.092	-0.028	0.022	-0.02	0.035	0.01	-0.166	-0.145	1			
Housing cost for low income people	-0.038	0.005	0.007	-0.028	0.096	-0.109	-0.065	0.187	0.009	0.045	0.033	-0.049	0.081	-0.04	0.019	-0.033	0.022	0.046	-0.188	-0.145	-0.151	1		
Household size	-0.011	0.011	0.016	-0.01	-0.009	-0.008	-0.198	0	-0.013	-0.005	-0.004	0.008	0.03	0.006	-0.005	0.001	0.008	-0.047	0.008	0.006	-0.008	-0.02	1	
Employment density	-0.033	-0.02	-0.055	0.006	-0.058	-0.048	-0.094	0.089	0.262	0.047	-0.069	-0.21	0.073	0.808	-0.053	-0.072	-0.073	0.027	0.042	0.065	0.054	0.058	-0.003	1