Steele, F. A., Washbrook, E., Charlton, C., & Browne, W. J. (2016). A Longitudinal Mixed Logit Model for Estimation of Push and Pull Effects in Residential Location Choice. Journal of the American Statistical Association, 111(515), 1061-1074. DOI: 10.1080/01621459.2016.1180984

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://amstat.tandfonline.com/loi/uasa20

A Longitudinal Mixed Logit Model for Estimation of Push and Pull Effects in Residential Location Choice

Fiona Steele, Elizabeth Washbrook, Christopher Charlton & William J. Browne

To cite this article: Fiona Steele, Elizabeth Washbrook, Christopher Charlton & William J. Browne (2016) A Longitudinal Mixed Logit Model for Estimation of Push and Pull Effects in Residential Location Choice, Journal of the American Statistical Association, 111:515, 1061-1074, DOI: 10.1080/01621459.2016.1180984

To link to this article: http://dx.doi.org/10.1080/01621459.2016.1180984

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2016, VOL. 111, NO. 515, 1061–1074, Applications and Case Studies
http://dx.doi.org/10.1080/01621459.2016.1180984

Taylor & Francis
Taylor & Francis Group

A Longitudinal Mixed Logit Model for Estimation of Push and Pull Effects in Residential Location Choice

Fiona Steele[a], Elizabeth Washbrook[b], Christopher Charlton[b], and William J. Browne[b]

[a]Department of Statistics, London School of Economics and Political Science, London, UK; [b]Centre for Multilevel Modelling, University of Bristol, Bristol, UK

ABSTRACT

We develop a random effects discrete choice model for the analysis of households' choice of neighborhood over time. The model is parameterized in a way that exploits longitudinal data to separate the influence of neighborhood characteristics on the decision to move out of the current area ("push" effects) and on the choice of one destination over another ("pull" effects). Random effects are included to allow for unobserved heterogeneity between households in their propensity to move, and in the importance placed on area characteristics. The model also includes area-level random effects. The combination of a large choice set, large sample size, and repeated observations mean that existing estimation approaches are often infeasible. We, therefore, propose an efficient MCMC algorithm for the analysis of large-scale datasets. The model is applied in an analysis of residential choice in England using data from the British Household Panel Survey linked to neighborhood-level census data. We consider how effects of area deprivation and distance from the current area depend on household characteristics and life course transitions in the previous year. We find substantial differences between households in the effects of deprivation on out-mobility and selection of destination, with evidence of severely constrained choices among less-advantaged households. Supplementary materials for this article are available online.

ARTICLE HISTORY
Received April 2014
Revised March 2016

KEYWORDS
Conditional logit model;
Discrete choice model;
Neighborhood choice;
Random effects panel model;
Residential mobility

1. Introduction

The sorting of households into neighborhoods has fundamental implications for many social outcomes. Patterns of residential mobility shape the spatial distribution of populations and the extent to which certain groups, such as ethnic minorities, immigrants and the economically disadvantaged, are geographically concentrated. Any study of neighborhood effects—the causal effect of place on people—must tackle the issue of nonrandom selection into places of residence. Indeed, it can be argued that the selection process is integral to understanding the key issues of interest (Bergström and van Ham 2010).

The availability of geographical identifiers in many datasets means it is relatively easy to document where different types of household are located. It is far more challenging to understand *why* households have chosen the neighborhoods in which they live. The difficulty arises because neighborhoods are multidimensional packages of different attributes, in which types of dwellings, physical geography, and social composition are chosen simultaneously and tradeoffs between different area characteristics when making a decision are inevitable. Of particular interest in residential location choice is the distinction between factors that "push" households to leave a neighborhood and factors that attract or "pull" households toward a particular neighborhood (Lee, Oropesa, and Kanan 1994). However, most previous research has focused exclusively on push factors (e.g., Crowder and South 2008; van Ham and Clark 2009), while studies that have investigated pull factors have restricted analysis to the sample of movers and characterised origin and destination areas along a single dimension of neighborhood quality (e.g., Clark 1992; Clark and Ledwith 2007; Rabe and Taylor 2010). A promising approach adopted in recent research is a conditional logit model, a type of discrete choice model where the categorical response is the neighborhood of residence and neighborhood attributes are included as predictors (Ioannides and Zabel 2008; Hedman, van Ham, and Manley 2011; Bruch and Mare 2012). These models allow location choice to depend on multiple neighborhood characteristics, but their use to date has been confined to cross-sectional data or, when applied to longitudinal data, has not fully exploited information on repeated residential choices.

Longitudinal data of the type collected in household panel studies and population registers provide rich information on changes in households' place of residence and in their demographic and socioeconomic characteristics. A particular advantage of longitudinal data, thus far ignored, is the ability

CONTACT Fiona Steele f.a.steele@lse.ac.uk Department of Statistics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom.

Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JASA.

to disentangle push and pull effects of area characteristics; these effects are confounded when a cross-sectional approach is taken. In this article, we develop a longitudinal discrete choice model parameterized in a way that permits separation and joint estimation of push and pull effects. While mixed logit models have been developed for the analysis of longitudinal discrete choice data (Jain, Vilcassim, and Chintagunta 1994; Bhat and Guo 2004), to our knowledge, the distinction between push and pull effects of alternative-specific attributes has not been considered previously. Our model also allows the importance of area characteristics to depend on observed time-varying household characteristics, and household-level random effects are included to account for unobserved heterogeneity in the effects of neighborhood characteristics which may lead to dependency in household choices over time. The random effects are assumed to follow a multivariate normal distribution which allows for correlation between push and pull effects of a given neighborhood characteristic and between the effects of different characteristics. We show that nonzero random effect correlations have an important role to play in relaxing the "independence of irrelevant alternatives" assumption. Spatial correlation can be incorporated by including random effects at a higher geographical level than neighborhood, for example districts, allowing for correlation in the choice propensities for neighborhoods within the same district.

A potential barrier to the application of random effects discrete choice models to neighborhood choice is the computational challenges presented by typically large analysis datasets. National panel studies usually have several thousands of respondents which, when combined with multiple waves of measurement and a large choice set, renders existing likelihood-based and Bayesian estimation approaches infeasible. Even when cross-sectional models have been used, previous studies have reduced the size of the problem by focusing on choices within a small geographical area (Bruch and Mare 2012) or by taking a small subsample of the full choice set (Hedman, van Ham, and Manley 2011). We propose an efficient, flexible Bayesian procedure for estimation of random effects (mixed) discrete choice models to longitudinal data with large choice sets.

The mixed logit model is illustrated in an analysis of residential choice in England between 1998 and 2008 using data from the British Household Panel Survey linked with neighborhood-level census data. In particular, we consider how the push and pull effects of area deprivation and the pull effect of distance from the current neighborhood depend on observed household characteristics, while allowing for unobserved household heterogeneity. More generally, the methods we propose have applications in other settings where push and pull effects may be of interest. For example, in studies of brand loyalty in market research, a high price could push customers away from a particular brand, while a low price of a rival product could cause customers to switch brands.

2. A Longitudinal Mixed Logit Model for Residential Location Choice

2.1. Preliminaries

The following models are described in terms of household rather than individual choices, while recognising that co-resident nonrelated adults may be independent decision makers with regard to residential mobility and neighborhood choice. Our working definition of a household is given in Section 3.

We begin by setting out the classic discrete choice model before considering extensions that allow separation of push and pull effects of area characteristics, and differential push and pull effects by observed and unobserved household characteristics. Suppose that household i ($i = 1, \ldots, n$) chooses its area of residence at year t ($t = 2, \ldots, T$) from a set $\mathcal{C}_{it}$ containing R_{it} potential areas. Some households will move between years $t-1$ and t, which will typically lead to a change of area. The choice set is permitted to vary across households and time because it is both behaviorally unrealistic and computationally infeasible for households to choose from a common fixed set of areas (Lee and Waddell 2010). For example, the choice set might be restricted to the set of areas within a specified distance of a household's location at $t-1$.

Let y_{it} be the categorical response indicating the observed area of residence for household i at year t. A general discrete choice model for the response probability is

$$\Pr(y_{it} = r|\boldsymbol{\eta}_{it}) = p_{rit} = \frac{\exp(\eta_{rit})}{\sum_{k\in\mathcal{C}_{it}} \exp(\eta_{kit})}, \quad r = 1, \ldots, R_{it}, \tag{1}$$

where η_{rit} is the linear predictor which will usually be a function of area characteristics and their interactions with household characteristics, both defined at year $t-1$, and $\boldsymbol{\eta}_{it}$ is the vector of linear predictors for household-year observation it.

The model in Equation (1) can also be expressed in terms of the log-ratio of the choice probabilities for a pair of areas r and s

$$\log\left(\frac{p_{sit}}{p_{rit}}\right) = \eta_{sit} - \eta_{rit}. \tag{2}$$

2.2. A Multinomial Logit Model with Push and Pull Effects of Area Characteristics

We begin with a model that includes only area characteristics, but allow their effects on the choice between areas r and s at t in Equation (2) to depend on whether a household is resident in one of these areas at $t-1$. We show how this distinction, possible only with longitudinal data, allows estimation of push and pull effects of area characteristics.

Let $w_{ri(t-1)} = I(y_{i(t-1)} = r)$ where $I(.)$ is the indicator function, and denote by $\mathbf{z}_{r(t-1)}$ a p-vector of characteristics of area r defined at wave $t-1$. More generally, area characteristics can also be household specific, for example the distance between a potential area and the current place of residence. The linear predictor for a simple model including only area characteristics is

$$\eta_{rit} = w_{ri(t-1)}\,(\alpha + \boldsymbol{\beta}^T\mathbf{z}_{r(t-1)}) + (1 - w_{ri(t-1)})\,\boldsymbol{\gamma}^T\mathbf{z}_{r(t-1)}, \tag{3}$$

where α is a scalar and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are parameter vectors. The type of model defined by Equations (1) and (3), with covariates relating to response alternatives and fixed coefficients across alternatives, was originally referred to as a conditional logit model (McFadden 1974). However, it was subsequently shown to be equivalent to the multinomial logit model, which traditionally has subject-specific covariates and alternative-specific coefficients (Maddala 1983). In common with most of the discrete choice literature,

we therefore refer to the model as a multinomial logit model hereafter. These models are widely used for analyzing categorical responses where interest lies in the effects of attributes of the response alternatives on individual choice, including applications to brand preference (e.g., Jain, Vilcassim, and Chintagunta 1994) and transportation demand (e.g., Ben-Akiva and Lerman 1985).

In the case of residential choice, Bruch and Mare (2012) first proposed the inclusion of the lagged choice indicator $w_{ri(t-1)}$ and its interaction with area characteristics $\mathbf{z}_{r(t-1)}$ to allow for the possibility that households "evaluate their current location differently from other potential destinations." Equation (3) represents a reparameterization of their model in which $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ may be interpreted as push and pull effects of $\mathbf{z}_{r(t-1)}$. The interpretation of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and our definition of push and pull effects, can be seen more clearly by considering how a household's choice between two areas r and s at t depends on their residence at $t-1$ according to the values of $w_{ri(t-1)}$ and $w_{si(t-1)}$.

Case 1. Resident in r at $t-1$ ($w_{ri(t-1)}=1,\ \ w_{si(t-1)}=0$). From Equations (2) and (3), the log-odds of moving to area s between $t-1$ and t versus remaining in area r are

$$\log\left(\frac{p_{sit}}{p_{rit}}\right)=\boldsymbol{\gamma}^T\mathbf{z}_{s(t-1)}-\left(\alpha+\boldsymbol{\beta}^T\mathbf{z}_{r(t-1)}\right). \tag{4}$$

We can also derive the log-odds of moving out of area r to *any* area as

$$\log\left(\frac{1-p_{rit}}{p_{rit}}\right)=\log\left(\frac{\sum_{s\neq r}p_{sit}}{p_{rit}}\right)=\log\left(\sum_{s\neq r}\exp(\boldsymbol{\gamma}^T\mathbf{z}_{s(t-1)})\right)-(\alpha+\boldsymbol{\beta}^T\mathbf{z}_{r(t-1)}). \tag{5}$$

From Equations (4) and (5) we see that area characteristics influence residential mobility and location choice in two ways. First, there are effects of the characteristics of the current area r on moving out of r, represented by $\boldsymbol{\beta}$. We refer to $\boldsymbol{\beta}$ as the "push" effect of $\mathbf{z}_{r(t-1)}$. Second, the characteristics of a potential area s may attract (or repel) households from choosing s as a new destination. We refer to $\boldsymbol{\gamma}$ as the "pull" effect of $\mathbf{z}_{s(t-1)}$. The push effect of $\mathbf{z}$ is conditional on the level of $\mathbf{z}$ in other areas in the choice set, while the pull effect of $\mathbf{z}$ is conditional on the level of $\mathbf{z}$ in the current area. The model may be extended to allow for interactions between elements of $\mathbf{z}_{r(t-1)}$ and $\mathbf{z}_{s(t-1)}$. For example, a household's sensitivity to area deprivation when choosing a new area may depend on the level of deprivation experienced in the current area.

In Equation (4), each area characteristic z has both a push and pull effect. For some types of z, for example the distance or travel time between a potential destination s and the origin area, only a pull effect is defined. When the origin area is r, this is equivalent to setting $z_{r(t-1)}=0$.

α is the baseline log-odds of staying in r rather than moving to a new area, referred to as the inertia parameter. The estimate of α is expected to be large and positive because most households do not change area between $t-1$ and t.

Case 2. Resident in neither r nor s at $t-1$ ($w_{ri(t-1)}=0$, $w_{si(t-1)}=0$). For a household which is not resident in area r or s at $t-1$, the log-odds of choosing area s over area r are

$$\log\left(\frac{p_{sit}}{p_{rit}}\right)=\boldsymbol{\gamma}^T(\mathbf{z}_{s(t-1)}-\mathbf{z}_{r(t-1)}), \tag{6}$$

where, as in Case 1, $\boldsymbol{\gamma}$ may be interpreted as a pull effect, but now of one new area over another.

2.3. Allowing for Household Heterogeneity in Inertia and the Effects of Area Characteristics

The linear predictor in Equation (3) can be extended to allow inertia and the push and pull effects of $\mathbf{z}_{rt-1}$ to depend on a q-vector of household characteristics $\mathbf{x}_{i(t-1)}$:

$$\begin{aligned}\eta_{rit}=\ & w_{ri(t-1)}\left\{\alpha_0+\boldsymbol{\alpha}_1^T\mathbf{x}_{i(t-1)}+\boldsymbol{\beta}_0^T\mathbf{z}_{r(t-1)}\right.\\ & \left.+\boldsymbol{\beta}_1^T(\mathbf{x}_{i(t-1)}*\mathbf{z}_{r(t-1)})\right\}\\ & +(1-w_{ri(t-1)})\left\{\boldsymbol{\gamma}_0^T\mathbf{z}_{r(t-1)}+\boldsymbol{\gamma}_1^T(\mathbf{x}_{i(t-1)}*\mathbf{z}_{r(t-1)})\right\},\end{aligned} \tag{7}$$

where $\mathbf{x}_{i(t-1)}*\mathbf{z}_{r(t-1)}=[x_{1i(t-1)}\mathbf{z}_{r(t-1)}^T,\ldots,x_{qi(t-1)}\mathbf{z}_{r(t-1)}^T]^T$ is the qp-vector formed by taking the element-wise product of $\mathbf{x}$ and $\mathbf{z}$.

$\boldsymbol{\alpha}_1$ is the effect of $\mathbf{x}$ on the log-odds of a move out of area r between years $t-1$ and t, and $\boldsymbol{\beta}_0$ and $\boldsymbol{\gamma}_0$ are the push and pull effects of $\mathbf{z}$ when $\mathbf{x}=\mathbf{0}$. Writing $\boldsymbol{\beta}_1=[\boldsymbol{\beta}_{11}^T,\ldots,\boldsymbol{\beta}_{1q}^T]^T$, $\boldsymbol{\beta}_{1k}$ is the change in the push effect of $\mathbf{z}$ for a 1-unit change in household characteristic x_k ($k=1,\ldots,q$), and $\boldsymbol{\gamma}_{1k}$ is the corresponding change in the pull effect. For simplicity of notation, the same set of household characteristics is assumed to influence mobility and to moderate the effects of the $\mathbf{z}$, but this restriction may be relaxed.

It is straightforward to extend Equation (7) to include choice-specific random intercepts and coefficients of $\mathbf{x}$. We consider the addition of choice random effects, and random effects defined at a higher level than neighborhood, in Section 2.6.

2.4. The Independence of Irrelevant Alternatives Assumption

It follows from Equations (1) and (2) that the log ratio of the choice probabilities for areas s and r depends on measured characteristics of these areas, but not of other potential areas. This property of the discrete choice model is known as the "independence of irrelevant alternatives" (IIA), and it implies that the choice between s and r is unaffected by the addition or exclusion of other alternatives (Ben-Akiva and Lerman 1985).

The source of the IIA assumption can be seen more clearly if the model is expressed in terms of continuous latent choice propensities (or utilities) y_{rit}^* which underlie the observed choice y_{it} such that $y_{it}=r$ if $y_{rit}^*>y_{sit}^*$ for $r\neq s$. The discrete choice model given by Equation (1) can be written

$$y_{rit}^*=\eta_{rit}+\epsilon_{rit},$$

where ϵ_{rit} are iid residuals, assumed to follow a Type I extreme value distribution with variance $\pi^2/6$. The IIA property stems from the independence of ϵ_{rit}, and will be invalid if the latent propensities to choose areas r and s are correlated. In the context of residential choice, nonzero residual correlation may arise

because of similarity between areas on unmeasured factors used by households in deciding where to live.

Various approaches have been proposed to relax the IIA assumption, including generalized extreme value (GEV) models and the multinomial probit model. The nested logit model is the most widely used type of GEV model (e.g., Hensher, Rose, and Greene 2005), but it requires the researcher to a priori partition subsets (nests) of similar alternatives within which IIA might reasonably hold. While the nested logit model is of limited use in modeling residential choice because of the difficulty in identifying areas that might be similar on unmeasured characteristics, it has been used to model residential location choice jointly with related decisions such as mode of transport and time of travel for various types of activity (Ben-Akiva and Bowman 1998) and residential mobility (Lee and Waddell 2010). Both applications embed a multinomial logit model for residential choice within a nested logit model for the joint decision processes. The multinomial probit model allows explicitly for residual correlation by assuming that ϵ_{rit} follow a multivariate normal distribution, but this is infeasible in situations where the choice set is large. A more flexible way of accommodating similarity between alternatives is to use a mixed model which allows for unobserved heterogeneity between households in the effects of area characteristics.

2.5. Unobserved Household Heterogeneity

The most popular models for unobserved between-subject heterogeneity are the normal-mixed model and the latent class model (see Hensher and Greene (2003) and Keane and Wasi (2013) for reviews of mixed logit models). The normal-mixed model includes normally distributed random coefficients for the effects of $\mathbf{z}_{r(t-1)}$. Log-normal distributions may be assumed for coefficients that are expected to have the same sign for all subjects. The latent class model avoids parametric distributional assumptions and assumes that subjects come from a finite set of subpopulations (Greene and Hensher 2003; Domanski and von Haefen 2010). We focus on normal-mixed models for several reasons. First, a large number of latent classes may be required to capture complex patterns of heterogeneity. In the application to residential choice, for example, there may be between-household differences in push and pull effects of multiple area characteristics, leading to a multidimensional finite mixture distribution. Second, the estimation and interpretation of latent class models is complicated by having a separate set of parameters for each class. Third, the direction and magnitude of the correlations between multivariate normal random effects are of direct substantive interest. For instance, correlated random effects can provide insights into the nature of the association between the household-specific push and pull effects of an area attribute z, or between the push or pull effects of two different area attributes.

While mixed logit models have been applied to cross-sectional data, identification of unobserved household heterogeneity will generally be much improved by the availability of longitudinal data. (See Revelt and Train (1998) for discussion and applications of mixed logit models with normal and log-normal random effects for repeated choice data.) In a normal-mixed model the coefficients of the main effects of $\mathbf{z}$, $\boldsymbol{\beta}_0$ and $\boldsymbol{\gamma}_0$ in Equation (7), are replaced by

$$\boldsymbol{\beta}_{0i} = \boldsymbol{\beta}_0 + \mathbf{u}_{\beta i},$$
$$\boldsymbol{\gamma}_{0i} = \boldsymbol{\gamma}_0 + \mathbf{u}_{\gamma i},$$

where $\mathbf{u}_{\beta i}$ and $\mathbf{u}_{\gamma i}$ are vectors of household-specific random effects which capture variation between households in the importance placed on $\mathbf{z}$ in location decisions. We also allow for unobserved heterogeneity in households' attachment to their current areas by replacing the inertia intercept parameter α_0 by the random coefficient

$$\alpha_{0i} = \alpha_0 + u_{\alpha i}.$$

The random effects $\mathbf{u}_i = [u_{\alpha i}, \mathbf{u}_{\beta i}^T, \mathbf{u}_{\gamma i}^T]^T$ are assumed to follow a multivariate normal distribution with mean $\mathbf{0}$ and variance $\boldsymbol{\Omega}_u$. Thus the linear predictor can be partitioned as $\eta_{rit} = \mu_{rit} + \delta_{rit}$ where μ_{rit} is the systematic (or fixed) component given by Equation (7) and δ_{rit} is a random component which varies across households:

$$\delta_{rit} = w_{ri(t-1)} \left(u_{\alpha i} + \mathbf{u}_{\beta i}^T \mathbf{z}_{r(t-1)}\right) + (1 - w_{ri(t-1)}) \mathbf{u}_{\gamma i}^T \mathbf{z}_{r(t-1)}. \tag{8}$$

From (8) δ_{rit} also varies over time, but only through the observed predictors $w_{ri(t-1)}$ and $\mathbf{z}_{r(t-1)}$.

We now show how the inclusion of household random effects relaxes the IIA assumption by considering how they affect the ratio of (conditional and marginal) choice probabilities and the correlation between latent choice propensities.

Ratio of Choice Probabilities. Equation (1) with η_{rit} now defined as the sum of (7) and (8) gives the probability of choosing area r conditional on the household random effects $\mathbf{u}_i$. The addition of random coefficients $\mathbf{u}_{\beta i}$ and $\mathbf{u}_{\gamma i}$ allows the ratio of the subject-specific choice probabilities for areas r and s to vary across households according to differences in the (unobserved) importance placed on observed area characteristics $\mathbf{z}$. The IIA property is still assumed to hold at the household level because the ratio of subject-specific probabilities for areas r and s does not depend on characteristics of any other area. However, this is not the case for the ratio of unconditional or *marginal* choice probabilities. Letting $\eta_{rit}(\mathbf{u}_i)$ denote the linear predictor for the random effects model, the marginal (population-averaged) response probability is given by

$$\Pr(y_{it} = r|\boldsymbol{\eta}_{it}) = \int \frac{\exp\{\eta_{rit}(\mathbf{u}_i)\}}{\sum_{k \in \mathcal{C}_{it}} \exp\{\eta_{kit}(\mathbf{u}_i)\}} \phi(\mathbf{u}_i)\, \mathrm{d}\mathbf{u}_i, \tag{9}$$

where $\phi(.)$ is a multivariate normal pdf.

The log-ratio of the marginal response probabilities for areas r and s is no longer simply the difference in the linear predictors, as in Equation (2), because the summation in the denominator of (9) does not cancel. Thus the ratio of the marginal probabilities for r and s will depend on characteristics of other areas, and the IIA assumption is relaxed at the population level (Train 2003).

Correlation Between Latent Choice Propensities. The inclusion of household-specific effects induces a correlation between the latent choice propensities for any pair of areas r and s because $\mathbf{u}_i$ is common across the response alternatives faced by household i at time t. Consider a simplified form of Equations (7) and

(8) with one area characteristic $z_{r(t-1)}$, leading to three random effects $(u_{\alpha i}, u_{\beta i}, u_{\gamma i})$ with covariance matrix

$$\mathbf{\Omega}_u = \begin{pmatrix} \sigma_\alpha^2 & & \\ \sigma_{\alpha\beta} & \sigma_\beta^2 & \\ \sigma_{\alpha\gamma} & \sigma_{\beta\gamma} & \sigma_\gamma^2 \end{pmatrix}.$$

Random effect covariances have received little attention in previous applications of mixed logit models and, indeed, random effects are commonly assumed to be independent. A notable exception is Revelt and Train (1998) who contended that correlation between random effects would generally be expected. In the present application, the random effect covariances are of particular interest because they provide information about the associations between households' latent mobility preferences and the importance they place on z in residential decisions. Suppose, for example, that high values of z in an area are associated with an increased probability of moving out of that area ($\beta_0 < 0$) and a reduced probability of moving in ($\gamma_0 < 0$). Then $\sigma_{\beta\gamma} > 0$ implies a positive association between the push and pull effects of z, adjusting for the moderating effects of x: households who attach a higher-than-average importance to z in the decision to move out of an area ($u_{\beta i} < 0$) also tend to have a higher-than-average sensitivity to z when choosing a new destination ($u_{\gamma i} < 0$). The random effect covariances also play a crucial role in relaxing the IIA assumption, as shown below.

The covariance between the propensities to choose areas r and s for household i at t depends on a household's residence at $t-1$ as follows.

Case 1. Resident in r at t − 1 ($w_{ri(t-1)} = 1,\ w_{si(t-1)} = 0$). For a household considering a move from area r to s between $t-1$ and t, the covariance between the latent propensities of remaining in r and moving to s can be derived from Equation (8) as

$$\begin{aligned} \operatorname{cov}(y_{rit}^*, y_{sit}^*) &= \operatorname{cov}(u_{\alpha i} + u_{\beta i} z_{r(t-1)},\ u_{\gamma i} z_{s(t-1)}) \\ &= z_{s(t-1)}\sigma_{\alpha\gamma} + z_{r(t-1)} z_{s(t-1)}\sigma_{\beta\gamma} \end{aligned} \tag{10}$$

when $\operatorname{cov}(\epsilon_{rit}, \epsilon_{sit}) = 0$ and $\operatorname{cov}(\delta_{rit}, \epsilon_{sit}) = 0$ for $r \neq s$.

Thus the covariance depends on two components: (i) the value of z in the potential area, weighted by the covariance between the household-specific mobility propensity ($u_{\alpha i}$) and importance of z as a pull factor ($u_{\gamma i}$), and (ii) the similarity between areas r and s on z, weighted by the covariance between household-specific importance of z as a push and pull factor. If z is mean centred, the second component of the covariance will be highest for two areas with extreme above-average or below-average values on z and 0 for two average areas.

Case 2. Resident in neither r nor s at t − 1 ($w_{ri(t-1)} = 0,\ w_{si(t-1)} = 0$). The covariance between the latent propensities of choosing between two potential areas when currently resident in neither is

$$\operatorname{cov}(y_{rit}^*, y_{sit}^*) = \operatorname{cov}(u_{\gamma i} z_{r(t-1)}, u_{\gamma i} z_{s(t-1)}) = z_{r(t-1)} z_{s(t-1)} \sigma_\gamma^2. \tag{11}$$

Thus the covariance between the latent attractiveness of two potential areas depends on their similarity with respect to z and on the between-household variance in the pull effect of z.

2.6. Unobserved Area Heterogeneity

The model may be further extended to allow for the effects of unmeasured area characteristics on location choice by including choice-specific random effects υ_r in the linear predictor. However, it can be seen from the expressions for the covariance between the latent propensities of choosing neighborhoods r and s given by Equations (10) and (11) that the inclusion of choice-specific random effects does not help to relax the IIA assumption unless $\operatorname{cov}(\upsilon_r, \upsilon_s) \neq 0$. A natural extension would be to impose a spatial autocorrelation structure on the neighborhood effects, for example to allow a nonzero covariance between neighborhoods that share a boundary. Such spatial correlation would arise if neighborhoods in close proximity share unmeasured attributes that influence a household's location choice. Bhat and Guo (2004) proposed a mixed spatially correlated logit model for residential choice at a cross-section that includes a dissimilarity parameter measuring the correlation between adjacent areas.

One issue when considering the addition of choice-specific effects in applications where the choice set is large is that the number of potential choices can exceed the number of decisions actually made, leading to an identification problem because many potential neighborhoods will not be chosen by the survey respondents over the observation period. An alternative approach is to specify random effects for a broader area classification which we refer to as "districts." Denote by $\upsilon_{r(d)} \sim N(0, \sigma_\upsilon^2)$ the random effect for the district d containing neighborhood r. When neighborhoods in the choice set are nested within districts, we have a hierarchical structure and the inclusion of higher-level random effects allows for a form of spatial correlation. Specifically, for neighborhoods r and s in the same district, Equations (10) and (11) will include σ_υ^2 as an additional term. The model assumes that neighborhoods are exchangeable within districts, but this may be reasonable after adjusting for neighborhood characteristics such as distance from the current location.

The area-level random effects capture differences between districts in the mobility propensity of their residents and in their attractiveness as places to live. We might expect the influence of these unmeasured area characteristics to differ for residents and nonresidents, perhaps due to differences in knowledge about area attributes that are not captured by covariates. A possible specification of the area effects, that mirrors the structure of Equation (7), is

$$w_{ri(t-1)}\lambda\upsilon_{r(d)} + (1 - w_{ri(t-1)})\upsilon_{r(d)}, \tag{12}$$

where parameter λ allows the effect of unmeasured district characteristics on household location choice, and thus the between-district variance, to depend on whether the household is currently resident in district d.

2.7. Sampling Alternatives in Large Choice Sets

Estimation of a multinomial logit model with alternative-specific attributes requires the data to be structured so that there is a record for each of the R_{it} response alternatives for household i at year t. This can lead to a prohibitively large

analysis file when the choice set is large, especially when decisions are observed over a long period for a large sample of households. A useful consequence of the IIA property, however, is that consistent parameter estimates can be obtained from a random subset of the full choice set, selected without replacement and including the record corresponding to the chosen alternative (McFadden 1978). For each household-year observation it, denote by q_{rit} the probability that the record for alternative r is selected, where $q_{rit} = 1$ if $y_{it} = r$. McFadden (1978) described a situation where the choice set is fixed and $q_{rit} = q$. More generally, we may wish to include information about the likelihood that household i chooses alternative r at t, referred to as importance sampling (Ben-Akiva and Bowman 1998; Bhat, Govindarajan, and Pulugurta 1998; Brownstone, Bunch, and Train 2000). In residential location choice, for example, the choice set for many households is restricted to areas within commuting distance of the current place of work. This leads to substantial variation in R_{it} across households and time, where R_{it} will typically be considerably larger in metropolitan areas than in rural areas. In such cases q_{rit} will be inversely proportional to R_{it}, and unequal selection probabilities are accommodated in the model by including $-\log(q_{rit})$ as an offset term (e.g., Ben-Akiva and Lerman 1985; Bruch and Mare 2012).

Unfortunately, when the IIA assumption is relaxed, for example by including household-specific random coefficients, McFadden's theoretical result no longer holds and random sampling of choice sets may yield inconsistent estimates (Nerella and Bhat 2004; Keane and Wasi 2012). This has led authors to explore empirically the impact of the size of the sampling fraction q on parameter estimates and standard errors from mixed multinomial logit models. Nerella and Bhat (2004) conducted a simulation study with 750 individuals, a choice set of size 200 and q varying between 0.025 and 0.75. They found a substantial impact of q on the bias and efficiency of the estimated parameters, and suggested that q should be set at 0.25 as a minimum. However, their study was based on a cross-sectional design for which mixed logit models may be weakly identified. Other research, using data on repeated choices as in our application to neighborhood choice, suggests that reliable estimates may be obtained using much smaller sampling fractions. From a potential choice set of 689 alternatives, Brownstone, Bunch, and Train (2000) used a random subset of 28 (4%) and reported that increasing the sampling fraction had little effect on parameter estimates. In the most comprehensive study to date Keane and Wasi (2012) considered the impact of using random subsets of the choice set for three alternative mixed MNL models for panel data, including the normal mixed model, through Monte Carlo simulation and sensitivity analysis of real data. Based on simulations with 200 individuals, 20 choice occasions and 60 alternatives, biases were small when random subsets of 10 or 20 alternatives were used.

2.8. MCMC Estimation

The most commonly used approaches for fitting mixed logit models are maximum simulated likelihood and Markov chain Monte Carlo (MCMC) estimation. Train (2001) compares these approaches and favors MCMC for both theoretical and computational speed reasons. He gives an MCMC algorithm for such models which generalizes work by Allenby (1997), and builds on ideas of Albert and Chib (1993). In this article, we modify Train's algorithm to accommodate parameterizations designed to improve the efficiency of MCMC estimation, which is especially important when the sample size and choice set are large. We consider a combination of hierarchical centering and orthogonal parameterization, adapting algorithms used for estimation of multilevel binary response models (as in Browne et al. 2009). This section provides an overview of estimation of the model with household-level random effects, given by the sum of Equations (7) and (8). Further details of the MCMC algorithm for the extension to area effects are provided in the online supplementary materials.

The fixed part of the model, Equation (7), includes parameters for mobility (α_0 and $\boldsymbol{\alpha}_1$), and push effects ($\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_1$) and pull effects ($\boldsymbol{\gamma}_0$ and $\boldsymbol{\gamma}_1$) of choice-specific attributes $\mathbf{z}$. From an algorithmic point of view, it is convenient to distinguish between coefficients that have an associated household-specific random effect (α_0, $\boldsymbol{\beta}_0$, $\boldsymbol{\gamma}_0$), as described in Section 2.5, and those with a fixed effect only ($\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$, $\boldsymbol{\gamma}_1$). Train (2003) focused on a general model where all coefficients are random at the subject level, but considers the above specification as a special case that may be useful in certain situations, such as when the full random effects covariance matrix cannot be identified. In the application that follows, the variances of the random effects for $\boldsymbol{\beta}_0$ and $\boldsymbol{\gamma}_0$, the main push and pull effects of area attributes $\mathbf{z}_{r(t-1)}$, are of particular interest as measures of between-household heterogeneity in the effects of $\mathbf{z}_{r(t-1)}$ that is unexplained by household covariates $\mathbf{x}_{i(t-1)}$.

Let $\boldsymbol{\theta} = [\alpha_0,\ \boldsymbol{\beta}_0^T,\ \boldsymbol{\gamma}_0^T]^T$ and $\boldsymbol{\theta}_i = \boldsymbol{\theta} + \mathbf{u}_i$ with associated data vector $\mathbf{A}_{ri(t-1)}$, and let $\boldsymbol{\varphi} = [\boldsymbol{\alpha}_1^T,\ \boldsymbol{\beta}_1^T,\ \boldsymbol{\gamma}_1^T]^T$ with data vector $\mathbf{B}_{ri(t-1)}$. The linear predictor for the mixed logit model can be reexpressed as

$$\eta_{rit} = \boldsymbol{\theta}_i^T \mathbf{A}_{ri(t-1)} + \boldsymbol{\varphi}^T \mathbf{B}_{ri(t-1)} - \log(q_{rit}),$$

where $\boldsymbol{\theta}_i \sim \text{MVN}(\boldsymbol{\theta}, \boldsymbol{\Omega}_u)$ and $\log(q_{rit})$ is an offset (see Section 2.7).

It is common to parameterize the model in terms of $\boldsymbol{\theta}_i$ rather than $\boldsymbol{\theta}$ and $\mathbf{u}_i$. This parameterization, known as hierarchical centering (Gelfand, Sahu, and Carlin 1995), can improve mixing when the random effect variances in $\boldsymbol{\Omega}_u$ are not too small, as it allows a Gibbs sampling step for $\boldsymbol{\theta}$ rather than a Metropolis step. The other speed up we consider is an orthogonal reparameterization similar to that of Browne et al. (2009). This involves replacing $\mathbf{B}_{ri(t-1)}$ by an orthogonal vector that spans the same space. This is achieved using a standard orthogonalising algorithm and we can then run MCMC using the transformed predictors. The chains for the parameters in the original parameterization can be retrieved by a simple matrix transform based on the inverse of the transformation of the predictors (see Browne et al. 2009, for details).

The algorithm has been implemented in the Stat-JR package (Charlton et al. 2013) which allows MCMC chains to be run in parallel with both hierarchical centering and orthogonal parameterization of the fixed predictors. The code has also been optimized so that estimation times are significantly

faster than implementing the same model in alternative packages such as WinBUGS (Lunn et al. 2000). Optimization for the mixed logit model involves the storage of intermediate quantities and constants within the likelihood to reduce the number of computationally-expensive calculations, for example exponentiations. The likelihood also contains, for each observation, a linear predictor with many parameters and several steps involve calculation of terms that are equal to the linear predictor minus one element. Storage of the linear predictor for each observation and a technique of subtracting the relevant element, updating it and then adding it back to the linear predictor lead to a substantial reduction in computing time.

2.9. Monte Carlo Simulation Study

A simulation study was conducted to explore the impact of sample size (number of households n) and panel length T on the bias and accuracy of estimates of the regression coefficients and random effect parameters. The linear predictor of the data generating model has the same form as Equations (7) and (8), with one area-level variable z and one household characteristic x and household random effects for inertia and the push and pull effects of z. The simulation conditions included sample sizes of 1000 and 1500 households and panel lengths of 5, 10, and 15. We also investigated the extent to which the correct random effects covariance structure was recovered by fitting models to simulated data with correlated and independent household random effects.

As expected, bias decreases as the number of waves per household increases, with the greatest impact on the inertia variance (σ_α^2), the largest of the random effect variances. However, even in the case of 1000 households and 5 waves, the biases are not large (and are in line with posterior mean estimates from a simulation study by Browne and Draper (2000) that assessed MCMC estimation of random effects binary logistic models with a uniform prior). Increasing the number of households by 50% also leads to bias reduction. Finally, when a model with correlated random effects is fitted to data generated from a model with independent random effects, the correct correlation structure is recovered. There are substantial improvements in accuracy with increases in either the number of households or the number of waves, with the largest impact observed for the random effect parameters and in particular the inertia variance. For the regression coefficients, the mean standard error across replications and the empirical standard error (standard deviation of the parameter estimates) are very similar for all scenarios considered, and there is minimal loss in accuracy when the random effects correlation structure is incorrectly specified. Details of the design of the simulation study and results are given in online supplementary materials.

3. Between-Household Heterogeneity in the Effects of Area Deprivation and Distance on Residential Choice in England

3.1. Data

Neighborhood Definition and Variables. We define a neighborhood as a Lower Super Output Area (LSOA), census areas which contain between 400 and 1200 households and an average of 1500 individuals. There were 32,482 LSOAs in England by the 2001 Census definition. We characterize neighborhoods along two dimensions: deprivation and distance from a person's current residence. Deprivation has been widely used as a measure of neighborhood quality in earlier analyses of residential choice. Previous studies have considered the push effect of deprivation and, among movers, compared the level of deprivation in the origin and destination neighborhood (e.g., Rabe and Taylor 2010). The distance between the current neighborhood and a potential destination has not been previously studied, but is likely to be an important factor in location choice. We anticipate that many households would prefer to remain close to their current residence when considering a local move, especially families with school-age children. However, for new homeowners or for households seeking a larger home, a longer-distance move may be necessary to secure affordable housing.

The measure of deprivation used is the LSOA-level English Index of Multiple Deprivation (IMD), a weighted combination of seven domain indices which capture different aspects of deprivation within a neighborhood, relating to income; employment; health and disability; education, skills, and training; barriers to housing and geographical access to services; crime; and living environment (see Dibben et al. 2007, for details). Three versions of the IMD have been constructed for 2004, 2007, and 2010. The score we allocate to a particular LSOA depends on the household survey year, as described below, and is standardised using the English LSOA-level average and standard deviation from 2007.

The second neighborhood measure is the distance in kilometres between a household's LSOA of current residence and each alternative LSOA. These are the straight-line distances between the population-weighted centroids of each area, calculated from the Ordnance Survey centroid grid references provided by the Office for National Statistics (ONS). An important difference between our distance and deprivation measures is that distance is household specific and, as a consequence, only its pull effect can be defined. If $z_{si(t-1)}$ denotes the distance between area s and the current location of household i at $t-1$, then $z_{ri(t-1)} = 0$ and β vanishes from the expression for the log-ratio of the probabilities of moving to area s rather than remaining in r (Equation 4).

To define the choice set of neighborhoods relevant to a household we use the household's current Travel-to-Work-Area (TTWA). TTWA is a labor market area definition, derived from 2001 Census information on home and work addresses, and used by the ONS to reflect areas where the bulk of the resident population also work within the same area. TTWA boundaries are nonoverlapping and contiguous, and cover the whole of the United Kingdom. TTWAs do cross national boundaries, and of the 243 that cover the entire United Kingdom, 166 contain at least one LSOA in England. In our study a household's choice set of neighborhoods includes all English LSOAs within the current TTWA. The mean number of LSOAs per TTWA is 196, but TTWAs are substantially larger in metropolitan areas such as London with 5467 LSOAs.

Due to the large number of LSOAs, it was not possible to identify choice-specific random effects to allow for between-area heterogeneity. Random effects were therefore specified at a

higher "district" level, as described in Section 2.6, where "districts" are defined as Medium Super Output Areas (MSOA). MSOAs are census areas with between 2000 and 6000 households. There were 6781 MSOAs in England at the 2001 Census, and LSOAs are nested within MSOAs.

Household Panel Data. Data on the characteristics and residential locations of households are taken from the British Household Panel Survey (BHPS) (ISER 2009). The BHPS is a nationally representative sample of about 5500 private households recruited in 1991, containing approximately 10,000 adults who are interviewed annually. The core questionnaire elicits information on topics such as household composition, housing tenure, employment and income at each annual interview. Our analysis uses information from waves 8–18, covering the period 1998–2008. Earlier waves are excluded because of lack of comparable area-level IMD data in this period.

One challenge in the analysis of household panel data is how to define a household longitudinally when its composition may change over time. The usual approach to this problem is to follow individuals, rather than attempt to track households, with analyses based on person-year observations. However, this is inappropriate for couple households where decisions are likely to be made jointly while partners are co-resident (Steele, Clarke, and Washbrook 2013). For this reason, couples contribute only one person-year record to the analysis file while they are together. Using this approach, couples are regarded as a single decision-making unit, and any other individual is treated as an independent decision-maker. Thus an individual living with unrelated adults is treated the same as an individual living on their own.

We model a household's residential location choices for up to 10 years but, due to a combination of late entry to the study and dropout, only 17.7% of households are present at all waves and the mean number of waves is 4.9. (Late entrants include children of original sample members (OSMs) who join the panel at age 18 and new partners of OSMs.) The following analysis is based on records from any household observed for at least two adjacent years $t-1$ and t. Households contribute records to the analysis file up to the time of dropout and, for cases with intermittent nonresponse, after their return to the study. This approach assumes that incomplete data are missing at random, conditional on observed residential location choices together with household and area covariates. In reality, a "missing not at random" (MNAR) mechanism is more plausible because moving house between $t-1$ and t is a cause of dropout at t. Washbrook et al. (2014) investigated the effect of nonignorable dropout in an analysis of residential mobility using two types of selection model applied to BHPS. Although they found strong evidence of nonignorability, modeling the dropout process had little effect on the coefficient estimates of the residential mobility model. Our model of residential choice could also be extended to incorporate a dropout model, but we do not pursue this here because it seems reasonable to expect that MNAR is more of a concern for mobility than for choice of area.

In the residential mobility literature, it is usual to distinguish local or short-distance moves from longer-distance moves because the two types of moves have very different determinants. By restricting the choice set to LSOAs within the TTWA of residence at $t-1$, we focus on local moves within a given labor or housing market which tend to be triggered by family events such as the arrival or departure of a child (Clark and Huang 2003). Because of our focus on movement within labor market areas, the sample is further restricted to individuals of working age, defined as between 18 and 59 years old at t.

Household-level characteristics considered in the analysis are for the most part defined at $t-1$. We define seven categories of household type, distinguishing single males and females; single parents (of either sex) of a child under 16; and couples with a resident youngest child aged 0–4, 5–10, 11–15, and 16 or more. Housing tenure is categorized as owned (outright or with mortgage), private rented, social (council or Housing Association) rented, and living with family. This last group consists of people who are a relative (other than a spouse) of the household reference person (HRP). The HRP is the person legally or financially responsible for the accommodation. Ninety-five percent of individuals in the "living with family" group are children living in the parental home, with siblings of the HRP making up the largest group among the remainder. We measure the gross household income over the previous 12 months as the combined incomes of the two members of a couple, or the individual income of a single person. The log of income in pounds is included in the model. We also include some indicators of life course transitions that occur between $t-1$ and t, that is, contemporaneously with the choice of location at t. These include the birth of a child, a move into or out of home ownership, a move out of the family home into social or private rental, and all other tenure transitions (e.g., between private and social rented accommodation). Indicators of partnership formation and dissolution between $t-1$ and t were considered in preliminary analysis, but were not retained as the additional effects of these variables on location choice were found to be insignificant.

We also tested for period effects on mobility and in the push and pull effects of area characteristics by including year dummies among the predictors of inertia and their interactions with deprivation and distance. As there was little evidence of period effects over our observation period of 1998–2008, we adopted a simpler model specification which assumes time-invariant inertia and push and pull effects.

The analysis sample consists of 30,912 person-wave observations from 6249 individuals (treating couples as a single "individual" as described above). The overall annual mobility rate is 10.8% and 34.8% of households move at least once during the 10-year observation period.

Sampling Choice Sets. Expanding the data to obtain one record for each LSOA in a household's choice set results in a person-wave-LSOA dataset of over 29 million observations. LSOAs were randomly sampled from this expanded dataset with probability inversely proportional to the size of the TTWA, while always retaining the records for the LSOAs of residence at $t-1$ and t. Thus, for household i resident in TTWA j at $t-1$ the probability that LSOA r is selected from their choice set is

$$q_{rit|j} = \begin{cases} \frac{c}{\sqrt{R_{it|j}}} & \text{if } y_{i(t-1)} \neq r \text{ and } y_{it} \neq r \\ 1 & \text{if } y_{i(t-1)} = r \text{ or } y_{it} = r. \end{cases}$$

$R_{it|j}$ is the number of areas in the choice set of household i at year t given residence in TTWA j, and the constant c is chosen so that the number of records in the person-wave-LSOA file is approximately equal to a target of m_{tar} according to

$$c = \frac{m_{tar}}{\sum_j m_j \sqrt{R_j}},$$

where R_j is the total number of LSOAs in TTWA j and m_j is the total number of person-wave-LSOA records in TTWA j. The following results are based on an analysis file with $m_{tar} =$ 800,000. To assess sensitivity of estimates to random sampling of the choice set, the analysis was repeated for two different random subsets of person-wave-LSOA records: the first with the same value of m_{tar} and the second with $m_{tar} = 1{,}600{,}000$. The parameter estimates and credible intervals were found to be very similar for these different random samples.

3.2. Results

As described above, we focus on the effects of two neighborhood (LSOA) characteristics on location choice: area deprivation (IMD), and distance from a household's current residence. We allow for both observed and unobserved heterogeneity in the push and pull effects of IMD and the pull effect of distance through their interactions with the household characteristics $\mathbf{x}_{i(t-1)}$ described in the previous section and the inclusion of household-specific random effects. In addition, mobility (inertia) is modeled as a function of $\mathbf{x}_{i(t-1)}$, again allowing for unobserved heterogeneity. We also consider models with random effects at the MSOA level.

Assessment of Convergence and Model Fit. The results presented below are based on five parallel chains of 100,000 MCMC iterations, each using a different starting value and with a burn-in sample of 10,000. Uniform priors were assumed for all parameters. Convergence was assessed using a range of graphical diagnostics and the potential scale reduction factor (PSRF) (Gelman et al. 2004). Visual inspection of trace plots of each parameter for the multiple chains suggested adequate mixing. Following the iterative graphical approach of Brooks and Gelman (1998), the within-chain variance, a weighted average of the within and between-chain variance, and the PSRF were examined for subsets of the chains. All had stabilized by 100,000 iterations, and the final PSRF estimates were close to 1 for all parameters. Furthermore, increasing the chain length led to little change in the running means of the posterior estimates. Hierarchical centering and orthogonal parameterization were considered, separately and in combination, in an attempt to improve mixing. For the fixed parameters, orthogonal parameterization led to substantial reductions in the effective sample size (ESS) (Kass et al. 1998). As in this case hierarchical centering had almost no impact on the ESS of the random effect variances and covariances, we present results from using only orthogonal parameterization.

Three models were fitted and compared using the Bayesian Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002). A model with household and area covariates, as in Equation (7), has a DIC of 52,525. The DIC decreases to 44,854 with the addition of household random effects, and there is a further decrease to 41,151 when area (MSOA) random effects are included. The specification of the area effects in the selected model is a special case of Equation (12) with $\lambda = 1$. In the more general model with λ unconstrained, the chains for the between-area variance and λ mixed very poorly, suggesting that the data did not support the extra complexity. We, therefore, assume that unmeasured area characteristics have the same influence on moves out of an area and on that area being chosen as a new destination.

The fit of the selected model was assessed using graphical posterior predictive checks (Gelman and Hill 2007). Replicates of the multinomial response y^{rep} were simulated using every 1000th draw from the parameter chains for the fitted model, including predicted values of the random effects. Simulation was carried out sequentially with the lagged choice indicator $w_{ri(t-1)}$ and covariates relating to the current area (distance and deprivation at origin) updated at each wave to reflect the dynamic structure of the model. The replicates were generated from predicted response probabilities computed for each area in the full choice set, not only the random subset selected for analysis. Test statistics $T(y^{\text{rep}})$ were chosen to assess the model's ability to capture the two main components of the model: household mobility and residential choice among movers. Figure 1 shows the distribution across replicates of the proportion of households who never moved over the observation period and the proportion of moves where there was an increase or decrease in deprivation of more than 0.5 standard deviation units between the destination and origin area. Comparison of the posterior predictions of each test statistic with the same statistic computed for the observed data $T(y)$ shows that the model provides a good fit to these aspects of the data.

Effects of Household Characteristics and Differential Push Effects of Area Deprivation on Mobility. The results for the full mixed logit model are shown in Tables 1–3 and Figure 2. The estimates shown in the "main effects" columns of Table 1 are the effects of household characteristics $\mathbf{x}_{i(t-1)}$ on the odds of a move to a new area between years $t-1$ and t; these are the exponents of the α parameters in Equation (7) after reversing their signs. The first row gives $\exp(-\alpha_0)$ the odds of a move for households taking the reference value for each variable in $\mathbf{x}_{i(t-1)}$. The remaining estimates are the odds ratios (OR) for each household characteristic k, $\exp(-\alpha_{1k})$, where an OR greater than 1 implies that a higher value on that covariate (or being in a particular covariate category) is associated with an increased odds of a move. These estimates are for households in an area with an average level of deprivation. As the model includes household and area random effects, all estimates are median ORs, that is the covariate effects for households at the mean of the multivariate random effects distribution and for neighborhoods (LSOA) in an area (MSOA) at the mean of the area-level random effects distribution. Lower household income and the presence of a new baby or young (preschool) child are associated with a higher annual probability of moving to a new neighborhood. There is a strong relationship between housing tenure and neighborhood change (estimates not shown): private renters are the most mobile and, not surprisingly, a change in tenure usually coincides with a move.

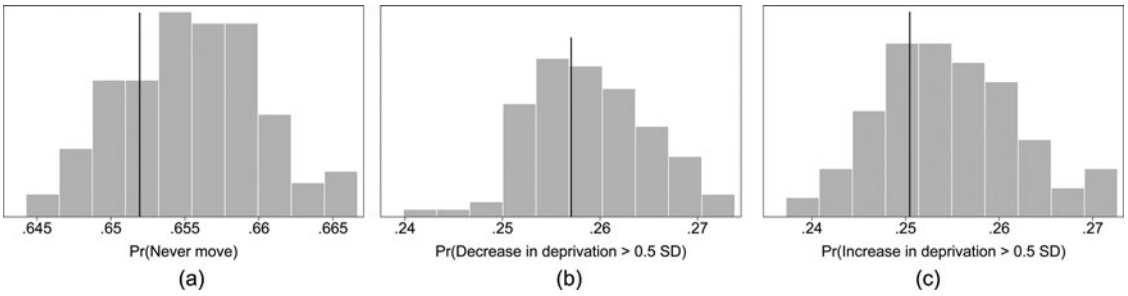

Figure 1. Posterior predictive checks for the probability (a) that a household never moves and, among movers, that deprivation in destination area is more than 0.5 standard deviations (b) lower and (c) higher than in origin area. The vertical lines indicate the observed values of the test statistics $T(y)$. The histograms show $T(y^{rep})$ for 100 replicates under the selected model.

Differential Push and Pull Effects of Deprivation and Distance on Mobility and Area Choice. Also shown in Table 1 are the interaction effects between $\mathbf{x}_{i(t-1)}$ and area deprivation (IMD) on the odds of a move. These estimates are the exponents of the β parameters in Equation (7), after reversing their sign. The first row gives $\exp(-\beta_0)$, which is the multiplicative effect of a 1 standard deviation increase in IMD in the current area on the odds of moving out for the reference group. Two types of IMD effects are presented: the interaction effects $\exp(-\beta_{1k})$, which represent differential push effects of IMD compared to the reference group, and, for categorical household covariates, the total or overall push effects for each households in each category, $\exp(-(\beta_0 + \beta_{1k}))$. Similarly, the estimates shown in Table 2 are the pull effects of IMD and distance from the current area on neighborhood choice for the reference group, $\exp(\boldsymbol{\gamma}_0)$ in Equation (7), and the differential and total pull effects by household characteristics, $\exp(\boldsymbol{\gamma}_{1k})$ and $\exp(\boldsymbol{\gamma}_0 + \boldsymbol{\gamma}_{1k})$. The estimates for IMD are the effects of a 1 standard deviation increase in IMD in area s on the odds of choosing s rather than another area r at year t, holding constant IMD in r, where r could be the area of residence at $t-1$. The estimates for distance are the effects of an increase of 1 km in the distance between area s and the area at $t-1$ on the odds of choosing s rather than r, holding constant the distance between r and the current area.

For the reference group of households we find that, as expected, a higher level of deprivation in a neighborhood is associated with an increase in the probability of out-mobility (Table 1) and a decrease in the probability of being chosen as a destination by movers (Table 2). Households are also less likely to move to neighborhoods that are far from their current place of residence (Table 2). However, the interaction effects in Tables 1 and 2 show that the effects of IMD, and to a lesser extent of distance, vary by household characteristics. In Table 1 an interaction OR > 1 for characteristic k indicates that the positive push effect of IMD for that subgroup of households is stronger than for the reference group, implying a greater aversion to remaining in a deprived area (and a total OR exceeding the OR for the reference group). An interaction OR < 1 may imply a weaker positive effect of deprivation or even a negative effect for a particular group. As the pull effects of IMD and distance are negative for the reference group (Table 2), an interaction OR < 1 implies a stronger negative effect of increasing deprivation or distance in area s on the odds of choosing s over another area.

Although there is little evidence that the push effect of deprivation depends on household income, higher income is associated with a stronger aversion to deprivation when choosing a new area, most likely because higher-income households

Table 1. Effects of selected household characteristics and differential push effects of deprivation (IMD) on odds of moving out of current area of residence.

			Effects of IMD		
	Main effects		Interaction		Total
Household characteristic†	Mean	95% CI	Mean	95% CI	Mean
Reference group‡	**0.0002**	(0.0002, 0.0003)	**1.66**	(1.22, 2.19)	–
log(income)	**0.90**	(0.85, 0.95)	1.05	(0.99, 1.11)	–
Household type $t-1$ (ref=couple no child)					
Single female	0.82	(0.53, 1.21)	0.85	(0.57, 1.22)	1.40
Single male	**0.64**	(0.44, 0.91)	0.77	(0.54, 1.06)	1.27
Single parent	0.99	(0.61, 1.50)	**0.44**	(0.28, 0.65)	0.72
Couple, y 0-4	**1.60**	(1.11, 1.58)	**0.71**	(0.50, 0.97)	1.16
Couple, y 5-10	1.14	(0.72, 1.71)	0.74	(0.47, 1.11)	1.21
Couple, y 11-15	0.90	(0.52, 1.47)	0.76	(0.44, 1.20)	1.24
Couple, y 16+	**0.59**	(0.33, 0.96)	0.96	(0.58, 1.51)	1.58
Birth $(t-1, t]$	**3.28**	(2.16, 4.75)	**1.49**	(1.04, 2.06)	**2.45**

†The model also includes housing tenure and change in tenure. The push effects of IMD by tenure are shown in Figure 2(a). ‡Couple with no children at $t-1$ or birth in $(t-1, t]$, owner-occupiers at $t-1$ with no change in tenure in $(t-1, t]$ and mean log(household income).

Note: Estimates are posterior means and 95% credible intervals (2.5 and 97.5 percentiles) of odds ratios (OR). ORs with 95% CI excluding 1 are highlighted in bold.

Table 2. Differential pull effects of deprivation (IMD) in area s and its distance from the current area on odds of choosing s over another area r.

	IMD			Distance		
	Interaction		Total	Interaction		Total
Household characteristic†	Mean	95% CI	Mean	Mean	95% CI	Mean
Reference group‡	**0.29**	(0.24, 0.36)	–	**0.55**	(0.52, 0.58)	–
log(income)	**0.96**	(0.93, 0.99)	**0.28**	**1.01**	(1.00, 1.02)	–
Household type $t-1$ (ref = couple no child)						
Single female	0.89	(0.72, 1.09)	**0.26**	1.04	(0.99, 1.09)	**0.57**
Single male	1.04	(0.86, 1.26)	**0.31**	**1.07**	(1.03, 1.12)	**0.59**
Single parent	1.20	(0.93, 1.52)	**0.35**	0.97	(0.92, 1.03)	**0.54**
Couple, y 0-4	1.15	(0.92, 1.42)	**0.34**	0.98	(0.94, 1.02)	**0.54**
Couple, y 5-10	1.08	(0.79, 1.42)	**0.31**	**0.94**	(0.89, 0.99)	**0.52**
Couple, y 11-15	0.98	(0.65, 1.40)	**0.29**	**0.93**	(0.87, 0.98)	**0.51**
Couple, y 16+	**1.62**	(1.12, 2.25)	**0.47**	**0.94**	(0.88, 1.00)	**0.52**
Birth $(t-1, t]$	**1.47**	(1.14, 1.85)	**0.43**	**0.94**	(0.90, 0.98)	**0.52**
Deprivation in area at $t-1$	**1.12**	(1.04, 1.19)	**0.33**	–	–	–

†The model also includes housing tenure and change in tenure. The pull effects of IMD and distance by tenure are shown in Figure 2(b)-(c). ‡Couple with no children at $t-1$ or birth in $(t-1, t]$, owner-occupiers at $t-1$ with no change in tenure in $(t-1, t]$ and mean log(household income).

Note: Estimates are posterior means and 95% credible intervals (2.5 and 97.5 percentiles) of odds ratios (OR). ORs with 95% CI excluding 1 are highlighted in bold.

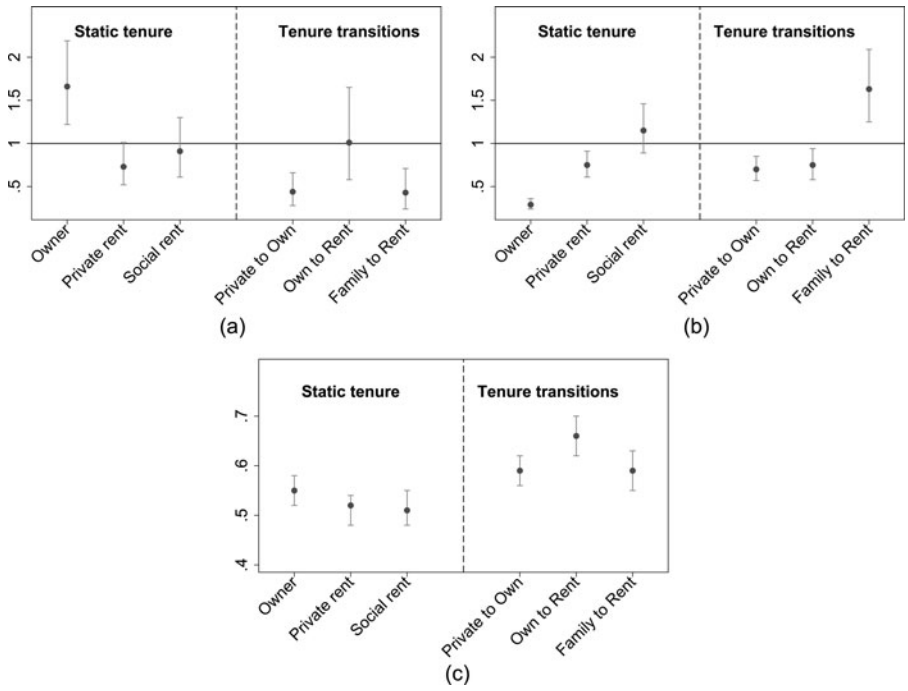

Figure 2. Effects of IMD and distance from current area on mobility and area choice by housing tenure at $t-1$ and t: (a) shows the effect of increasing IMD in the current area on the odds of a move out of that area, while (b) and (c) show the effects of increasing IMD and distance from the current area for area s on the odds of choosing s rather than another area. Estimates are the posterior means and 95% credible intervals of odds ratios for the effects of a 1 SD increase in IMD and a 1 km increase in distance.

are better able to act on preferences toward living in better-off areas. This result is consistent with research in the United States, Sweden and Britain that finds that household income constrains movers' access to more advantaged neighborhoods (Ioannides and Zabel 2008; Hedman, van Ham, and Manley 2011; Clark, van Ham, and Coulter 2013). The differential pull effect of deprivation by the level of area deprivation at origin is in line with the income effect: deprivation is a less important factor when choosing a new area for households currently living in a more deprived area. The effect of distance also depends on income, with a weaker effect for higher-income households.

Previous research on Britain suggests that area deprivation exerts a push effect on mobility among couples but not singles (Rabe and Taylor 2010) while singles are more likely than couples to move to less advantaged areas (Clark, van Ham, and Coulter 2013). However, we find little evidence of a difference in the push or pull effects of deprivation for singles and couples without children: differences between singles and couples only emerge when they have dependent children. Deprivation in the current area of residence has a weaker effect on the decision to move out for single parents and couples with a preschool child than for other household types, and there is also a suggestion that deprivation has a weaker deterrent effect on choosing a new area among couples with older children (aged 16+) than for households with a younger child or no children. Distance is a more important factor in movers' choice of destination for couples with school-age children than for other household types, which may reflect stronger local ties and a reluctance to move far from current schools among these families. A birth between years $t-1$ and t strengthens the push effect of deprivation during the same period, but also weakens aversion to deprivation when choosing a new area to live. These apparently contradictory findings may be due to a desire to move out of a deprived area among some new parents while others compromise on neighborhood quality in the search for an affordable family home.

Another source of heterogeneity in the importance of deprivation and distance in residential location decisions is housing tenure and changes in tenure. Figure 2 shows estimates of the push effect of IMD on mobility and the pull effects of IMD and distance from the current area for households whose tenure remains the same at $t-1$ and t ("static") and for households who change tenure ("transitions"). Starting with households whose tenure did not change, we find that higher deprivation in the current area is associated with increased odds of a move for homeowners, but the effect of deprivation switches direction and becomes nonsignificant for private and

Table 3. Estimated standard deviations (diagonal) and correlations (off-diagonal) of household-level random effects, and SD of area-level random effect.

	Inertia	Push: IMD	Pull: IMD	Pull: Distance
Inertia	3.03 (2.83, 3.23)			
Push: IMD	−0.09 (−0.22, 0.05)	1.56 (1.29, 1.81)		
Pull: IMD	−0.10 (−0.25, 0.05)	**0.23** (0.03, 0.43)	0.68 (0.57, 0.80)	
Pull: Distance	**0.58** (0.54, 0.63)	−0.01 (−0.13, 0.09)	0.08 (−0.03, 0.21)	0.44 (0.42, 0.47)
Random effect means†	8.38	−0.49	−1.23	−0.59
95% plausible range‡	(2.44, 14.32)	(−3.55, 2.57)	(−2.56, 0.10)	(−1.45, 0.27)
Between-area SD	1.38 (1.29, 1.47)			

†Estimates of inertia, push and pull effects of IMD, and pull effect of IMD on the log-odds scale for the reference group, that is α_0, β_0, $\gamma_{0,\texttt{IMD}}$ and $\gamma_{0,\texttt{Distance}}$.
‡Calculated as the random effect mean ± 1.96 times the random effect SD.
Note: Estimates are posterior means and 95% credible intervals (2.5 and 97.5 percentiles). Correlations with 95% CI excluding 0 are highlighted in bold.

social renters (Figure 2(a)). Furthermore, private renters are less averse to deprivation than homeowners when choosing a new area (Figure 2(b)), while the effect of deprivation on the area choices of social renters is not significant. Distance is an important factor for all three tenure groups, but less so for renters than for homeowners (Figure 2(c)). In Britain, private renters tend to be more mobile than owner-occupiers (e.g., Rabe and Taylor 2010; Steele, Clarke, and Washbrook 2013) and may, therefore, be less concerned about neighborhood factors. On the other hand, rented accommodation, especially social housing, tends to be located in the most deprived areas (Clark, van Ham, and Coulter 2013) which limits the chance of a move to an affluent area without a change in tenure. Low mobility within the social housing sector, particularly in high-demand areas in the South, is well documented (e.g., Kearns and Parkes 2003) and previous research has found that social renters have limited opportunities to "move up" to less-advantaged areas or to maintain residence in better-off areas (Clark, van Ham, and Coulter 2013).

Turning to households who changed tenure, higher deprivation in the area of origin is associated with a lower probability of leaving that area for private renters who made the transition into home-ownership (Figure 2(a)), although new homeowners tend to choose less-deprived areas when choosing a new destination (Figure 2(b)). Individuals who left the family home for private rented housing have a tendency either to remain in or move to a deprived area, most likely because of a greater availability of lower-cost accommodation in such areas. Finally, proximity to the current residence is an important factor for all households who change tenure, but less so for new homeowners and home-leavers.

Unobserved Heterogeneity in Push and Pull Effects. Table 3 shows estimates of the variances and correlations between the four household-level random effects representing time-invariant propensities to stay in the same area (inertia), and sensitivities to deprivation as a push or pull factor in residential location choice and distance from the current area as a pull factor. The results presented are from the full model that allows for differential effects of deprivation and distance by the observed household characteristics of Tables 1 and 2 and Figure 2. There is considerable variation between households in the propensity to stay in the same area and in the effects of deprivation and distance on location choice. Based on the normality assumption, 95% of households are expected to have a baseline log-odds of remaining in the same area between 2.44 and 14.32, while the corresponding ranges for the effects of IMD and distance span zero. The proportion of households for whom higher deprivation is associated with higher odds of a move is estimated as $\Pr(-\beta_{0i} > 0) = \Pr(\beta^*_{0i} < 0.49/1.56) = 62\%$, where $\beta^*_{0i} \sim N(0, 1)$, which suggests that a substantial number of households are unable to move out of deprived areas (assuming an underlying preference to live in more prosperous neighborhoods). By similar calculations, the proportion of households with an aversion to deprivation when choosing a new area and with a preference toward areas that are close to their current residence are 96.5% and 91%, respectively.

After accounting for differential effects of deprivation and distance by observed household characteristics, there remains a significant correlation between inertia and the pull effect of distance. This strong, positive correlation suggests that more mobile households have a stronger-than-average preference to remain close to their current neighborhood when moving house. Put another way, households that move less frequently are prepared to move greater distances when they do so. There is also a moderate, positive correlation between the push and pull effects of deprivation, which implies that households with a stronger-than-average aversion to remaining in a deprived neighborhood have a tendency to avoid more deprived areas when choosing a new place to live.

For neighborhoods (LSOAs) r and s in different MSOAs, the log-odds of choosing s versus r depends on the difference in their unmeasured MSOA-level characteristics $v_{s(d)} - v_{r(d)}$. The contribution of area effects to the standard deviation in the log-odds is estimated as $\sqrt{2} \times 1.38 = 1.95$. Thus area effects are sizeable, although variation between households in neighborhood choice is dominated by the substantial variation in household inertia. This between-MSOA variation also implies that there is a similarity between neighborhoods in close proximity (in the same MSOA) in their probability of being chosen as a place to live.

4. Discussion

This article has presented a general mixed logit model which makes use of longitudinal data to distinguish, and estimate simultaneously, the push and pull effects of multiple area attributes on residential location choice. An efficient MCMC algorithm was proposed which, together with sampling of the choice set, allows consideration of a larger set of potential destination areas, larger sample size, and longer observations period than has been possible in previous research.

Our analysis of household heterogeneity in the effects of neighborhood deprivation on out-mobility and movers' selection of a new destination suggests that the residential choices of less-advantaged households are severely constrained. We find that low income is associated with a lower probability of moving to a more advantaged neighborhood while private and social renters are less likely than owner-occupiers to move out of deprived areas. As argued by other authors (e.g., Clark, van Ham, and Coulter 2013), such constraints in the housing market lead to increasingly selective migration with disadvantaged households unable to "move up" to better-off areas, a situation which is likely to worsen with rising house prices in the United Kingdom. We also find that even for local moves within labor market areas, the influence of distance of a potential destination from the current residence depends on household characteristics.

The focus of the analysis presented here is household heterogeneity in the importance placed on two area characteristics, deprivation and distance, in location choices. Another avenue for research would be to compare the push and pull effects of multiple area attributes—such as crime, house prices, and school quality—for different types of household. It is also straightforward to extend the model to include random coefficients on interactions between household and area characteristics, $\mathbf{x}_{it-1} * \mathbf{z}_{rt-1}$ in Equation (7), thus allowing for heterogeneity in the importance placed on an area characteristic z within groups defined by $\mathbf{x}$. A consequence of this more general specification is that the covariance between individual choice

propensities given by Equations (10) and (11) would depend not only on **z**, but also on (possibly time-varying) household characteristics.

Supplementary Materials

The online supplementary materials contain further details of the MCMC algorithm and the extension to include area effects. Also provided are details of the design of the simulation study and results.

Acknowledgment

The authors are grateful to the associated editor and anonymous referees for their helpful comments and suggestions.

Funding

This work was partially supported by grants RES-062-23-2265 and RES-576-25-0032 from the Economic and Social Research Council.

References

Albert, J. H., and Chib, S. (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669–679. [1066]

Allenby, G. (1997), "An Introduction to Hierarchical Bayesian Modeling," Tutorial Notes, Advanced Research Techniques Forum, American Marketing Association. [1066]

Ben-Akiva, M., and Bowman, J. L. (1998), "Intergration of an Activity-Based Model System and a Residential Location Model," *Urban Studies*, 35, 1131–1153. [1064,1066]

Ben-Akiva, M., and Lerman, S. R. (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, Cambridge, MA: MIT Press. [1063,1066]

Bergström, L., and van Ham, M. (2010), "Understanding Neighbourhood Effects: Selection Bias and Residential Mobility," Discussion paper no. 5193, Institute for the Study of Labor (IZA). [1061]

Bhat, C. R., Govindarajan, A., and Pulugurta, V. (1998), "Disaggregate Attraction-End Choice Modeling: Formulation and Empirical Analysis," *Transportation Research Record*, 1645, 60–68. [1066]

Bhat, C. R., and Guo, J. (2004), "A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modelling," *Transportation Research Part*, B, 38, 147–168. [1062,1065]

Brooks, S. P., and Gelman, A. (1998), "General Methods for Monitoring Convergence of Iterative Simulations," *Journal of Computational and Graphical Statistics*, 7, 434–455. [1069]

Browne, W. J., and Draper, D. (2000), "Implementation Issues in the Bayesian Fitting of Multilevel Models," *Computational Statistics*, 15, 391–420. [1067]

Browne, W. J., Steele, F., Golalizadeh, M., and Green, M. (2009), "The Use of Simple Reparameterisations in MCMC Estimation of Multilevel Models With Applications to Discrete-Time Survival Models," *Journal of the Royal Statistical Society*, Series A, 172, 579–598. [1066]

Brownstone, D., Bunch, D., and Train, K. (2000), "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-Fuel Vehicles," *Transportation Research Part B*, 34, 315–338. [1066]

Bruch, E. E., and Mare, R. D. (2012), "Methodological Issues in the Analysis of Residential Preferences, Residential Mobility, and Neighborhood Change," *Sociological Methodology*, 42, 103–154. [1061,1062,1063,1066]

Charlton, C. M. J., Michaelides, D. T., Parker, R. M. A., Cameron, B., Szmaragd, C., Yang, H., Zhang, Z., Fraser, A. J., Goldstein, H., Jones, K., Leckie, G. B., Moreau, L., and Browne, W. J. (2013), "Stat-JR Version 1.0," Centre for Multilevel Modelling, University of Bristol, Bristol, UK. [1066]

Clark, W. A. V. (1992), "Residential Preferences and Residential Choices in a Multiethnic Context," *Demography*, 29, 451–466. [1061]

Clark, W. A. V., and Huang, Y. (2003), "The Life Course and Residential Mobility in British Housing Markets," *Environment and Planning A*, 35, 323–339. [1068]

Clark, W. A. V., and Ledwith, V. (2007), "How Much Does Income Matter in Neighbourhood Choice?" *Population Research and Policy Review*, 26, 145–161. [1061]

Clark, W. A. V., van Ham, M., and Coulter, R. (2013), "Spatial Mobility and Social Outcomes," *Journal of Housing and the Built Environment*. Online first. [1071,1072]

Crowder, K. D., and South, S. J. (2008), "Spatial Dynamics of White Flight: The Effects of Local and Extralocal Racial Conditional on Neighborhood Out-Migration," *American Sociological Review*, 73, 792–812. [1061]

Dibben, C., Atherton, I., Cox, M., Watson, V., Ryan, M., and Sutton, M. (2007), *Investigating the Impact of Changingthe Weights That Underpin the Index of Multiple Deprivation 2004,* London, UK: Department for Communities and Local Government. [1067]

Domanski, A., and von Haefen, R. H. (2010), *Estimation and Welfare Analysis From Mixed Logit Recreation Demand Models With Large Choice Sets*, Raleigh, NC: North Carolina State University. [1064]

Gelfand, A. E., Sahu, S. K., and Carlin, B. P. (1995), "Efficient Parameterizations for Normal Linear Mixed Models," *Biometrika*, 82, 479–488. [1066]

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004), *Bayesian Data Analysis* (2nd ed.), Boca Raton, FL: Chapman and Hall/CRC. [1069]

Gelman, A., and Hill, J. (2007), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, New York: Cambridge University Press. [1069]

Greene, W. H., and Hensher, D. A. (2003), "A Latent Class Model for Discrete Choice Analysis: Contrasts With Mixed Logit," *Transportation Research Part B: Methodological*, 37, 681–698. [1064]

Hedman, L., van Ham, M., and Manley, D. (2011), "Neighbourhood Choice and Neighbourhood Reproduction," *Environment and Planning*, A, 43, 1381–1399. [1061,1062,1071]

Hensher, D. A., and Greene, W. H. (2003), "The Mixed Logit Model: The State of Practice," *Transportation*, 30, 133–176. [1064]

Hensher, D. A., Rose, J. M., and Greene, W. H. (2005), *Applied Choice Analysis: A Primer*, New York: Cambridge University Press. [1064]

Ioannides, Y. M., and Zabel, J. E. (2008), "Interactions, Neighbourhood Selection and Housing Demand," *Journal of Urban Economics*, 63, 229–252. [1061,1071]

ISER, (2009), *British Household Panel Survey: Waves 1-17, 1991–2008* (6th ed.), University of Essex, Institute for Social and Economic Research [original data producer(s)], Colchester, Essex: UK Data Archive [distributor]. SN: 5151. [1068]

Jain, D. C., Vilcassim, N. J., and Chintagunta, P. K. (1994), "A Random-Coefficients Logit Brand-Choice Model Applied to Panel Data," *Journal of Business and Economic Statistics*, 12, 317–328. [1062,1063]

Kass, R. E., Carlin, B. P., Gelman, A., and Neal, R. (1998), "Markov Chain Monte Carlo in Practice: A Roundtable Discussion," *The American Statistician*, 52, 93–100. [1069]

Keane, M., and Wasi, N. (2013), "Comparing Alternative Models of Heterogeneity in Consumer Choice Behaviour," *Journal of Applied Econometrics*, 28, 1018–1045. [1064]

Keane, M. P., and Wasi, N. (2012), *Estimation of Discrete Choice Models With Many Alternatives Using Random Subsets of the Full Choice Set: With an Application to Demand for Frozen Pizza*, New York: University of Oxford Press. [1066]

Kearns, A., and Parkes, A. (2003), "Living in and Leaving Poor Neighbourhood Conditions in England," *Housing Studies*, 18, 827–851. [1072]

Lee, B. A., Oropesa, R. S., and Kanan, J. W. (1994), "Neighbourhood Context and Residential-Mobility," *Demography*, 31, 249–270. [1061]

Lee, B. H. Y., and Waddell, P. (2010), "Residential Mobility and Location Choice: A Nested Logit Model With Sampling of Alternatives," *Transportation*, 37, 587–601. [1062,1064]

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. J. (2000), "WinBUGS - a Bayesian Modelling Framework: Concepts, Structure, and Extensibility," *Statistics and Computing*, 10, 325–337. [1067]

Maddala, G. S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge, UK: Cambridge University Press. [1062]

McFadden, D. (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. P. Zarembka, New York: Academic Press, pp. 105–142. [1062]

McFadden, D. (1978), "Modeling the choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, eds. A. Karlqvist, F. Lundqvist, F. Snickars, and J. Weibull, Amsterdam: North-Holland, pp. 75–96. [1066]

Nerella, S., and Bhat, C. R. (2004), "Numerical Analysis of Effect of Sampling of Alternatives in Discrete Choice Models," *Transportation Research Record*, 1894, 11–19. [1066]

Rabe, B., and Taylor, M. P. (2010), "Residential Mobility, Quality of Neighbourhood and Life Course Events," *Journal of the Royal Statistical Society*, A, 173, 531–555. [1061,1067,1071]

Revelt, D., and Train, K. (1998), "Mixed Logit With Repeated Choices: Households' Choices of Appliance Efficiency Level," *Review of Economics and Statistics*, 80, 647–657. [1064,1065]

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002), "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society*, Series B, 64, 538–640. [1069]

Steele, F., Clarke, P. S., and Washbrook, E. (2013), "Modeling Household Decisions Using Longitudinal Data From Household Panel Surveys, With Applications to Residential Mobility," *Sociological Methodology*, 43, 225–276. [1068,1072]

——— (2001), "A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit," Working paper, Department of Economics, University of California, Berkeley. [1066]

Train, K. (2003), *Discrete Choice Methods With Simulation*, Cambridge, UK: Cambridge University Press. [1064,1066]

van Ham, M., and Clark, W. A. V. (2009), "Neighbourhood Mobility in Context: Household Moves and Changing Neighbourhoods in the Netherlands," *Environment and Planning A*, 41, 1442–1459. [1061]

Washbrook, E., Clarke, P. S., and Steele, F. (2014), "Investigating Non-Ignorable Dropout in Panel Studies of Residential Mobility," *Journal of the Royal Statistical Society*, Series C, 63, 239–266. [1068]