Longitudinal analysis of residential choice in England

Fiona Steele and Elizabeth Washbrook
(University of Bristol)

CASE seminar, 15th May 2013

Research aims

- Develop a method for analysing the effects of area characteristics on neighbourhood choice
- Illustrate the technique by focusing on the effect of a single characteristic - overall area deprivation (IMD)
- How far, and for whom, does area deprivation act as a
  - Push factor → increasing the probability of a move out of the current neighbourhood
  - Pull factor → increasing the probability of a move into an alternative neighbourhood

Related literatures

- Neighbourhood effects
  - Vast literature on the effects of place on behaviour and attitudes
  - Estimates of causal effects biased by neighbourhood sorting
  - Bergström and van Ham (2010) argue that understanding neighbourhood selection processes is fundamental to understanding neighbourhood effects
  - Selective migration may undermine the effectiveness of area-based initiatives (Bailey and Livingston, 2008)
- The composition of inflows and outflows from a neighbourhood affects neighbourhood change or reproduction and the degree of spatial segregation
  - Large US literature on the effects of ethnic mix on location decisions of different groups ("white flight"), e.g. South and Crowder (1998)
- The extent of, and trends in, socio-spatial mobility – is the ability to move up in terms of neighbourhoods broadly based? (Clark et al., 2011)

Previous approaches

Two-step approaches

- Model the probability of moving away from current area as function of current area characteristics
- For movers only, model the change in area characteristics (Rabe and Taylor, 2010, Clark et al., 2011)
- Only consider a single dimension of neighbourhood quality at a time

Cross-sectional models of neighbourhood choice – model current location as function of area characteristics

- Recognize neighbourhood is a “package” but no information on timing of inflows and outflows, e.g. Nechyba and Strauss (1998)

Our approach

- Longitudinal discrete choice model of neighbourhood in which to locate, within a given labour market area (Bruch and Mare, 2012)
- Every year a household decides which neighbourhood in which to live – can stay in current one or move
- Choice is fundamentally a function of neighbourhood attributes. Individual attributes do not vary within the choice set. However, individual characteristics may modify the way area attributes are valued.
- Extends the small literature using a similar approach(Hedman, van Ham and Manley, 2011; Ioannides and Zabel, 2008)

The choice set

- Choice of labour market area is treated as given.
- As in much of the residential mobility literature, our focus is on short-distance moves that adjust the nature of the dwelling to fit household characteristics (the majority of all moves)
- 71% of moves in our sample were within-labour market area, plus we incorporate a further 12% that only just crossed a boundary

This application

- Focus on the effect of area deprivation on choice of neighbourhood
- Certain groups have stronger aversion to deprivation than others – over time tend to become concentrated in low deprivation areas
- Need to be critical of the concept of “choice” in this setting as location choices reflect constraints as well as preferences (Bergström and van Ham, 2010)

Area definitions

- The choice set is restricted to the current Travel-to-Work-Area (TTWA)
- 166 TTWAs (at least partly) in England
- A neighbourhood is a lower super output area (LSOA)
- 32,482 LSOAs in England, average 1500 residents
- Number of LSOAs per TTWA ranges from 4 to 5647 (London), mean =196

Area deprivation

- Deprivation is measured by the Index of Multiple Deprivation (IMD) at the LSOA-level
- Composite measure of deprivation along multiple domains: income, employment, education, health, barriers to housing and services, living environment
- Our data cover the period 1999 to 2008, with different IMDs for 1999-2002 (IMD2004), 2003-2006 (IMD2007) and 2007-8 (IMD2010)
- Total IMD standardized on 2007 distribution for England → 1 unit = 1 SD of the national distribution

The conditional choice probability

$p_{rit|j}$ is the probability of selecting LSOA *r* in year *t*, conditional on choosing one LSOA from TTWA *j*, where *j* is the TTWA of residence in year *t*-1

Define

$$y_{it} = r \text{ if individual } i \text{ is resident in LSOA } r \text{ in year } t$$

$$C_j = \{1, ..., R_j\}, \text{ the set of LSOAs in TTWA } j$$

The conditional logit model specifies

$$p_{rit|j} = \Pr\left(y_{it} = r \mid r \in C_j\right) = \frac{\exp(\eta_{rit})}{\sum_{k=1}^{R_j} \exp(\eta_{kit})}$$

Where η_{rit} can be thought of as an expression for the expected utility derived from living in LSOA *r*

The linear predictor

In its simplest form, the linear predictor, η, depends only on z_r, a single area characteristic

$$\eta_{rit} = \gamma_{it} z_r$$

In a random utility framework, γ_{it} is the effect of z on the utility the individual derives from alternative r.

An advantage of a longitudinal model is that we can allow the valuation an individual places on z to be different for their current area of residence than other potential alternatives

Defining $D_{rit} = \mathrm{I}(y_{it-1} = r)$, we can write

$$\eta_{rit} = \alpha D_{rit} - \beta_{it} D_{rit} z_r + \gamma_{it} \left(D_{rit} - 1 \right) z_r$$

The pull effect of z: γ_{it}

Consider the choice between two neighbourhoods r and s. The log-ratio of the choice probabilities is

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

With

$$\eta_{rit} = \alpha D_{rit} - \beta_{it} D_{rit} z_r + \gamma_{it}\left(D_{rit} - 1\right) z_r$$

Suppose the individual is resident in neither r or s at t-1, so $D_{rit} = D_{sit} = 0$. Then

$$\log\left(\frac{p_{rit}}{p_{sit}}\right) = \gamma_{it}\left(z_r - z_s\right)$$

γ_{it} is the pull effect of z. It captures the effect of a unit difference in z between areas r and s on the probability of moving to r relative to s

The push effect of z: β_{it}

Now consider the decision to remain in one's current area. How does z affect the probability an individual will stay where they are or move somewhere else?

Setting D_{rit} = 1, the expected utility of staying in an area with z_r is $\eta_{rit} = \alpha - \beta_{it} z_r$, where α is the "inertia" parameter

Similarly, the expected utility of staying in an area with z_s is $\eta_{sit} = \alpha - \beta_{it} z_s$. So the log ratio of the choice probabilities is

$$\log\left(\frac{p_{rit}}{p_{sit}}\right) = -\beta_{it}\left(z_r - z_s\right)$$

β_{it} is the push effect of z. It captures the effect of a unit difference in z in one's current area on the probability of moving out (- β_{it} is the effect on staying)

Individual heterogeneity

$$\eta_{rit} = \alpha D_{rit} - \beta_{it} D_{rit} z_r + \gamma_{it} \left(D_{rit} - 1 \right) z_r$$

We can allow the push and pull effects of z to differ with individual characteristics

$$\beta_{it} = b_0 + x'_{it-1} b_x \qquad \gamma_{it} = g_0 + x'_{it-1} g_x$$

$x_{it\text{-}1}$ is a vector of observed characteristics

b_0 and g_0 are the push and pull effects respectively for some reference individual (with all elements of vector x = 0)

b_x and g_x are vectors of parameters that capture how the push and pull effects vary with observed characteristics

We can also incorporate time-invariant unobserved characteristics, but for now we cluster the standard errors at the individual level

Individual-level data

- Longitudinal data on individuals/households from the BHPS, waves 9-18
- Individuals age 18-59 at t, living in known LSOA in England at *t*-1 and *t*
- Only one record retained for couples intact between *t*-1 and *t*
- Records with a cross-TTWA move between *t*-1 and *t* dropped, *except* for moves within 45km
- 6248 individuals, 30,915 person-years; 29 million person-wave-LSOA records!

Handling large choice sets

A dataset of 29 million observations is impractical

We randomly sample neighbourhoods from within an individual's choice set with probability q_{rit}

- q_{rit} = 1 for areas of residence at t and t -1, else q_{rit} << 1
- Include –log(q_{rit}) as offset term in model
- No effect of sampling on consistency of parameter estimates from model (McFadden 1978)

Sample LSOAs with probability proportional to the square root of number of LSOAs in TTWA → 829,775 person-wave-LSOA records

Sources of variation in push and pull effects

Individual characteristics that modify the effects of area deprivation are

- Log annual income of individual/couple
- Housing tenure at t-1
- Household type
- IMD of neighbourhood at t-1 (pull only)
- Transition indicators for birth of a child and tenure changes (partnership transitions dropped as insignificant)

Results

Presented in two formats

- Summaries of estimated push and pull effects for different groups → depend on joint distribution of observed characteristics within a group.
- Model coefficients → isolate the contribution of a particular characteristic holding others constant

Compare results characteristic by characteristic, highlighting London vs the Rest of England

Note that all estimates are derived from a single model containing all conditioning variables

Overview

Summary over all person-years

	N	Mean	SD	Min	Max
Push	30,915	**0.141**	0.379	-5.071	2.278
Pull	30,915	**-0.342**	0.304	-2.244	7.743

Model coefficients for the reference category (-b_0 and g_0)

Childless couple, owners, mean log income, mean IMD at origin, no birth or tenure transition between t-1 and t

Push: 0.195* (0.050) Pull: -0.369* (0.045) * p<.01, + p<.05

Push: Effect of a 1SD ↑ in IMD in current area on moving out of that area (β_i)
Pull: Effect of a 1SD ↑ in IMD in potential area in moving into that area (γ_i)

Variation in push effect of IMD with household income

Group means of household-specific push effects, other characteristics not held constant

Model coefficients

(other characteristics held constant)

Reference group: 0.195* (0.050)

Log income × RoE: **0.010** (0.016)

Log income × London: **0.079+** (0.035)

Test of equality: p = 0.074

* p<.01, + p<.05

Variation in pull effect of IMD with household income

Group means of household-specific pull effects, other characteristics not held constant

Model coefficients

(other characteristics held constant)

Reference group: -0.369* (0.045)

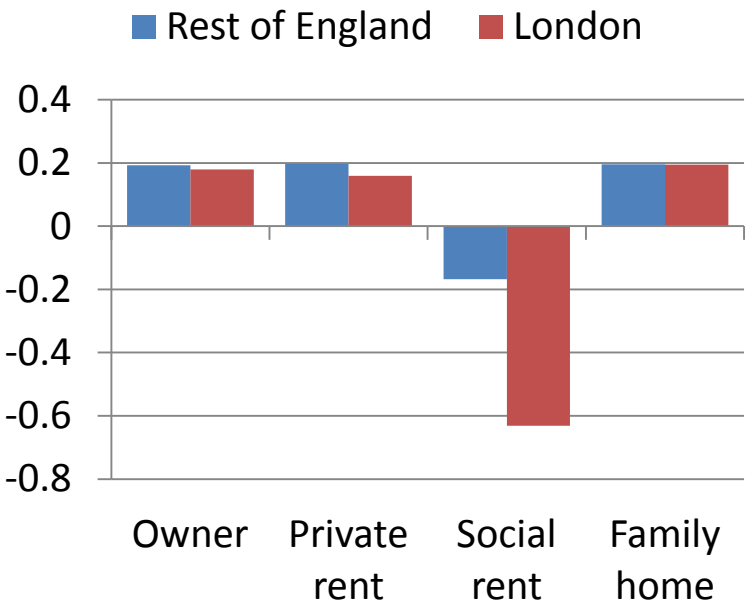
Log income × RoE: -**0.029+** (0.012)

Log income × London: -**0.090*** (0.023)

Test of equality: p = 0.018

* p<.01, + p<.05

Variation in push effect of IMD with origin housing tenure



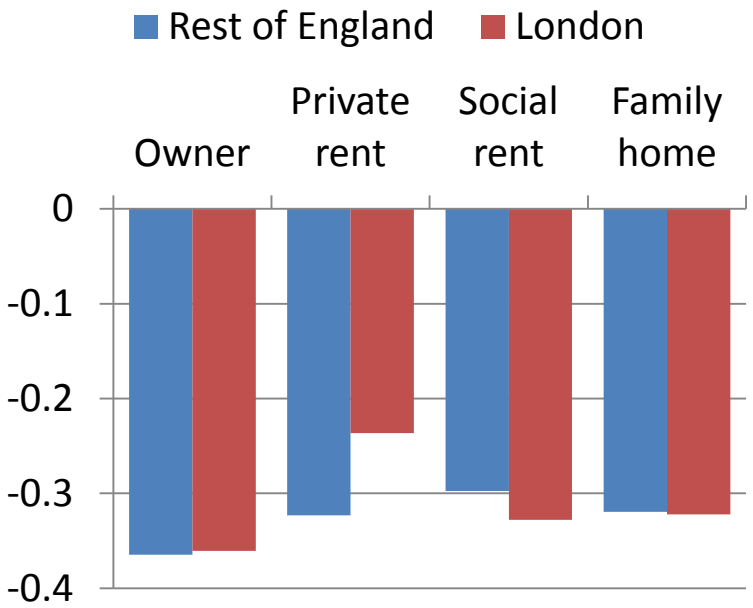
Group means of household-specific push effects, other characteristics not held constant

Model coefficients
(other characteristics held constant)

Reference group:	0.195* (0.050)
Private rent × RoE:	-**0.562*** (0.085)
Private rent × London:	**0.190** (0.192)
Test of equality: p = <0.001	
Social rent × RoE:	-**0.477*** (0.064)
Social rent × London:	**-1.048*** (0.224)
Test of equality: p = 0.012	
Family home × RoE:	-**0.333*** (0.079)
Family home × London:	-**0.038** (0.186)
Test of equality: p = 0.116	

* p<.01, + p<.05

Variation in pull effect of IMD with origin housing tenure



Group means of household-specific pull effects, other characteristics not held constant

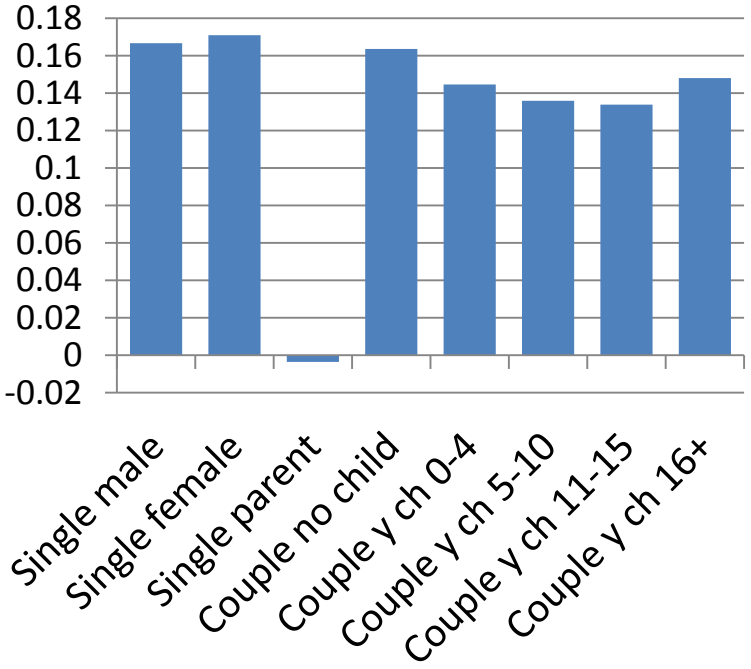
Model coefficients
(other characteristics held constant)

Reference group:	-0.369* (0.045)
Private rent × RoE:	**0.297*** (0.068)
Private rent × London:	**0.496*** (0.137)
Test of equality: p = 0.165	
Social rent × RoE:	**0.435*** (0.060)
Social rent × London:	**0.281** (0.179)
Test of equality: p = 0.402	
Family home × RoE:	**0.293*** (0.067)
Family home × London:	-**0.080** (0.153)
Test of equality: p = 0.016	

* p<.01, + p<.05

Variation in push effect of IMD with household type

Virtually no significant differences between London and RoE, so results for all England



Group means of household-specific push effects, other characteristics not held constant

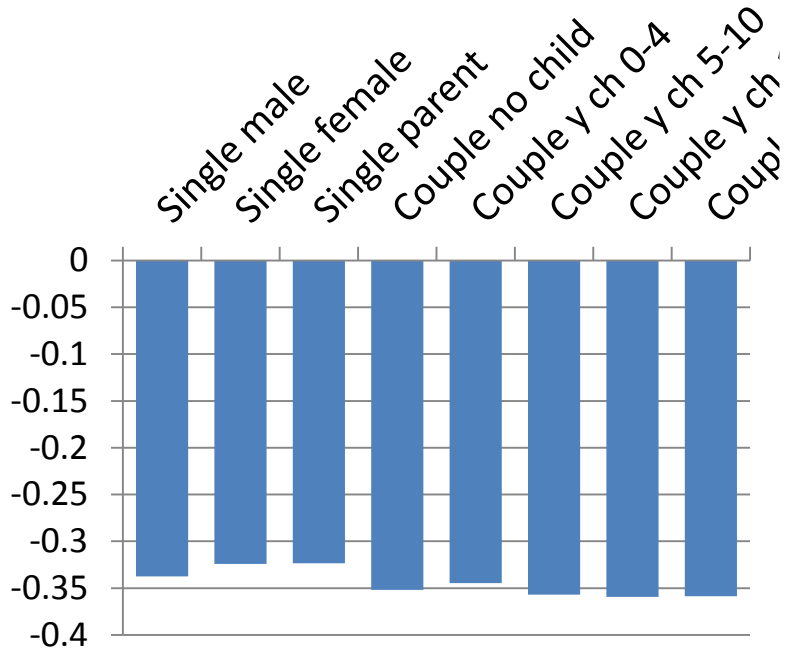
Model coefficients
(other characteristics held constant)

Reference group:	0.195* (0.050)
Single female:	-**0.026** (0.085)
Single male:	-**0.029** (0.077)
Single parent:	-**0.288*** (0.082)
Couple, y ch 0-4:	**-0.008** (0.076)
Couple, y ch 5-10:	-**0.018** (0.088)
Couple, y ch 11-15:	-**0.018** (0.097)
Couple, y ch 16+:	**0.044** (0.086)

* p<.01, + p<.05

Variation in pull effect of IMD with household type

Virtually no significant differences between London and RoE, so results for all England



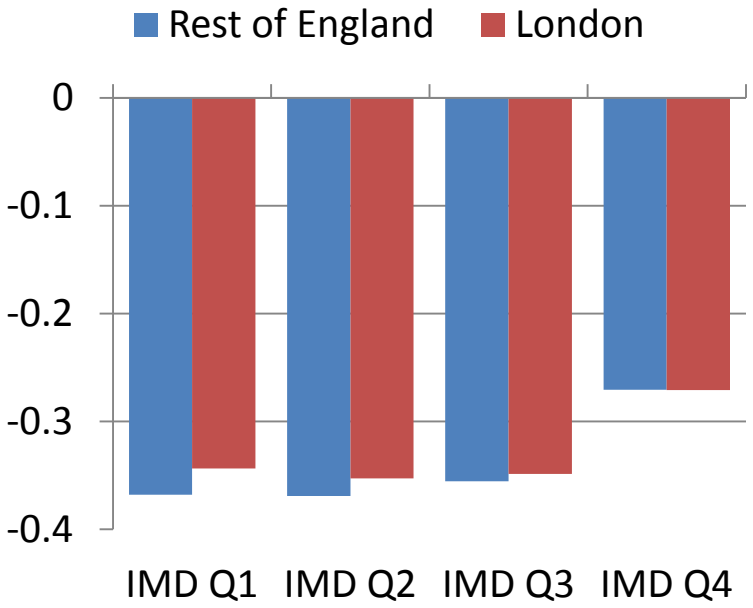
Group means of household-specific pull effects, other characteristics not held constant

Model coefficients (other characteristics held constant)

Reference group:	-0.369* (0.045)
Single female:	-**0.068** (0.074)
Single male:	**0.010** (0.066)
Single parent:	**0.059** (0.071)
Couple, y ch 0-4:	**0.004** (0.070)
Couple, y ch 5-10:	**0.047** (0.080)
Couple, y ch 11-15:	**0.091** (0.087)
Couple, y ch 16+:	**0.191+** (0.077)

* p<.01, + p<.05

Variation in pull effect of potential area IMD with origin IMD



Group means of household-specific pull effects, other characteristics not held constant

Model coefficients

(other characteristics held constant)

Reference group: -0.369* (0.045)

Origin IMD × RoE: **0.181*** (0.016)

Origin IMD × London: **0.519*** (0.063)

Test of equality: p <0.001

* p<.01, + p<.05

Variation in push and pull effects with life course transitions – birth of a child

Effect of a birth between t-1 and t

Push ↑ 0.278+ (0.115) in RoE, ↑ 0.167 (0.417) in London (p = 0.797)

Pull ↑ 0.276* (0.095) in RoE, ↓ 0.471 (0.366) in London (p = 0.047)

Births are associated with stronger aversion to current area deprivation. In London they are also associated with a stronger aversion to potential area deprivation, though the opposite is true outside London.

Variation in push and pull effects with life course transitions – buying a home

Effect of change in tenure from rent to own between t-1 and t

Push ↓ 1.011* (0.153) in RoE, ↑ 0.594 (0.375) in London ($p<0.001$)

Pull ↓ 0.028 (0.110) in RoE, ↓ 0.764+ (0.334) in London ($p = 0.036$)

Outside London, buying a home is associated with a weaker propensity to leave a deprived area and no difference in the propensity to move into an affluent area. In London, it is associated with both a greater propensity to leave a deprived area and to move into an affluent area.

Summary of findings

In general, higher levels of deprivation in one's current neighbourhood are associated with a greater propensity to move out

Home ownership and expecting a child increase the sensitivity to out-mobility to deprivation, single parenthood and social tenancy reduce it

On average, single parents are no more likely to exit deprived neighbourhoods than affluent ones

Social tenants are *more* likely to exit affluent neighbourhoods than deprived ones

Summary of findings

High area deprivation reduces the inflow rate of all groups, with relatively little variation

This aversion to deprivation when choosing a new area is weaker for those with lower incomes, renters, families with adult children, those who already live in deprived neighbourhoods and those expecting a child (outside London)

Summary of findings

Residential location decisions are different in London along a number of dimensions

- Income is more strongly associated with avoidance of deprivation than elsewhere
- The push effect of deprivation is the same for private renters as for owners
- Current deprivation is particularly strongly associated with lower exit rates among social renters
- Current deprivation has a greater negative impact on the probability of moving into a relatively affluent area
- Buying a (first) home is associated with a stronger aversion to deprivation at both origin and destination

Next steps: relaxing IIA

The conditional logit model relies on the independence of irrelevant alternatives assumption (IIA)

The choice between areas r and s depends only on characteristics of r and s, not of other areas

IIA will be invalid if there is unmeasured correlation in the utility derived from different areas

One way to relax this is to allow for unobserved individual heterogeneity in the effects of area characteristics

This involves treating b_0 , g_0 and α as random coefficients drawn from a trivariate normal distribution

MCMC estimation in Stat-JR (Browne, Charlton et al. 2013)

Other applications

More individual household characteristics – employment status, education,...

Which area characteristics matter for location choice of different groups?

- Crime, urban amenities, physical environment, social composition
- School quality
- Neighbourhood change