Journal of the Royal Statistical Society
SERIES A Statistics in Society

J. R. Statist. Soc. A (2010)
173, *Part* 4, *pp.* 755–774

Random-effects models for migration attractivity and retentivity: a Bayesian methodology

Peter Congdon

Queen Mary University of London, UK

[Received January 2009. Revised October 2009]

Summary. Several studies have proposed methods for deriving summary scores for describing the in-migrant attractivity of areas, as well as out-migrant push (or conversely migrant retentivity). Simple in-migration and out-migrant rates (migrant totals divided by populations) do not correct for spatial separation or the migration context of a particular area, namely the size and proximity of nearby urban areas with populations at risk of migrating to an area, or offering potential destinations for out-migrants from an area. An extended random-effects gravity model is proposed to represent the influences of attractivity and retentivity after controlling for urban structure. Whereas the existing literature is focused on fixed effects modelling (and classical estimation), the focus in this paper is on a Bayesian hierarchical random-effects approach that links estimates of pull-and-push scores across areas (i.e. allows scores to be spatially correlated) and also allows correlation between attractivity and retentivity within areas. As demonstrated by a case-study of English local authorities, a random-effects model may have a lower effective model dimension than a fixed effects model.

Keywords: Attractivity; Bayesian; Migration; Random effects; Retentivity; Spatial correlation

1. Introduction

It is important for planning resource and service provision to understand why some areas lose population through migration, whereas others are gaining. Some studies (e.g. Tobler (1979), Niedomsyl (2006) and Wall (2001)) have suggested ways of obtaining summary measures or scores for describing the in-migrant pull or ‘attractivity’ of areas. Ways of summarizing out-migrant push (or conversely migrant ‘retentivity’) have also been suggested. Related studies have been in other flow applications, e.g. recreational and travel flows (Baxter and Ewing, 1981; Lue *et al.*, 1996; Beaman *et al.*, 1997). In estimating such scores from migration interaction data, we need to correct for the spatial separation of origins and destinations, and what may be termed the ‘migration context’ of a particular area, namely the size and proximity of nearby areas with populations at risk of migrating to that area, or offering potential destinations for out-migrants from that area. Correcting for such structural effects can be regarded as a form of standardization that is analogous to demographic standardization used in comparing (say) mortality rates between areas, where one needs to correct for differing age structures (and possibly for other factors such as ethnic composition).

Such considerations mean that simple rates of in-migration and out-migration (migrant totals divided by populations) do not serve as satisfactory indices of migrant attractivity or retentivity, since for one thing they do not correct for proximity of an area to other areas with large

Address for correspondence: Peter Congdon, Centre for Statistics and Department of Geography, Queen Mary University of London, Mile End Road, London, E1 4NS, UK.
E-mail: p.congdon@qmul.ac.uk

0964–1998/10/173755

populations (Fotheringham *et al.*, 2000). As noted by Wall (2001) in an analysis of interregional migration,

> 'models that do not take into account that large regions attract more migrants bias the measure of migration rates towards large regions, with this bias becoming more severe the more diverse the regions are in size'.

Migrant flows may also be affected by the way that area boundaries are configured. For example, within the continuous urban space of London there are 33 local authorities ('areas') and what are recorded as interarea flows are in many cases short-range residential moves, so that attractivities would be apparently high for areas within London if they are measured by simple migration rates. By contrast, the attractiveness of remoter rural areas or free-standing towns that are not close to large population centres is understated by simple in-migration and out-migrant rates. Once such structural effects (of origin and destination population sizes, and of distances between origins and destinations) have been corrected for, we may distinguish remaining differentials as reflecting differential attractivities or retentivities.

We are then confronted with choosing a model for (in Bayesian terms, selecting an appropriate prior for) attractivities and retentivities. The existing literature for analysing migration interaction flows is largely focused on fixed effects modelling (and classical estimation) using variations of the gravity model, which assumes that spatial dependences are removed by including a distance effect—see Lesage *et al.* (2007) and Lesage and Pace (2008). Inclusion of other distance-based separation measures in gravity models, such as various accessibility indices (Pellegrini and Fotheringham, 2002), is also intended to account for spatial dependences. This raises a broader question of whether the spatial interaction should be appropriately modelled in the mean or whether additional structured effects are required to reflect spatial dependence (LeSage and Pace, 2009). If it were true that spatial dependences were removed simply by modelling distance decay (or effects of other accessibility indices), then a fixed effects model for attractivities and retentivities might be justified. A fixed effects model assumes independence of origin–destination effects, and that there is no gain from pooling strength under a scheme whereby unknown parameters (e.g. attractivity scores) are spatially interrelated. For example, let $s=(s_1,\ldots,s_n)$ denote attractivity parameters for n areas and $s_{[i]}=(s_1,\ldots,s_{i-1},s_{i+1},\ldots,s_n)$ denote the collection of attractivities for all areas apart from area i. Then a fixed effects model says that $s_{[i]}$ contains no information about the level of s_i (Plummer, 2008), namely that there is no benefit in looking at $(s_1,\ldots,s_n)$ as a collection of interrelated effects.

However, there is considerable evidence of spatial dependences that are not accounted for by the fixed effect gravity model. Lesage and Pace (2008) cited work by Griffith and Jones (1980) on commuting that flows from an origin are

> 'enhanced or diminished in accordance with the propensity of emissiveness of its neighboring origin locations',

and that flows that are associated with a destination are

> 'enhanced or diminished in accordance with the propensity of attractiveness of its neighboring destination locations'.

Such proximity effects on attractivities and retentivities suggest that they are likely to demonstrate strong spatial interdependences.

Accordingly, the focus in this paper is on a Bayesian hierarchical random-effects approach that links estimates of pull-and-push scores across areas. Markov chain Monte Carlo (MCMC) techniques (Gelfand and Smith, 1990), as implemented in the WinBUGS program (Lunn *et al.*, 2000), are used for parameter estimation. Such methods may be criticized (e.g. in terms of requiring specification of prior densities on parameters, or in terms of the computational demands of MCMC techniques). However, in terms of mean-squared error, such an approach provides better estimates (Haneuse and Wakefield (2004), page 270, and Lindley and Smith (1972)) and is likely to have a lower effective model dimension than a fixed effects model by virtue of 'borrowing strength' across areas (Banerjee *et al.* (2003), page 107). A lower model dimension provides a better fit when fit is measured by a likelihood criterion penalized for model complexity (Spiegelhalter *et al.*, 2002). A Bayesian approach is also likely to assist in random-effects model estimation (e.g. avoiding quadrature methods), and in assessing the distributional properties of push-and-pull indices, since a Bayesian approach provides a full posterior density for push-and-pull indices (see Hogan and Tchernis (2004)). Other inferences (e.g. on rankings of attractivities) are available under an MCMC approach. A Bayesian approach assists in hypothesis testing, e.g. in deriving posterior probabilities that the attractivity of area i is higher than that of area j, or the probability that area i has a higher attractivity than its neighbouring areas.

A further distinct feature of the current paper is the likelihood that is assumed for migration flows; these are discrete counts that are sometimes modelled as Poisson distributed (Flowerdew and Aitkin, 1982). However, in the application here the data contain a high proportion of zero flows yet still show overdispersion. A negative binomial model is accordingly adopted. A log-normal regression scheme may be appropriate when there are only a few zero migrant flows (e.g. Lesage and Pace (2009), page 223) but is problematic when there are many zero flows (Lesage and Pace (2009), page 230, and Ranjan and Tobias (2007), page 819). A negative binomial model is not the only choice but provides satisfactory fit in the case-study below; a zero-inflated negative binomial (ZINB) model may also be adopted (Yau *et al.*, 2003), and exploratory work with a ZINB model (which is not reported below) showed some improvement in fit but at the expense of greater computational complexity and slowed computing in WinBUGS. When there are excess 0s, Banerjee *et al.* (2000) similarly suggest a zero-inflated Poisson model, whereas Ranjan and Tobias (2007) relied on a tobit censoring model; other options might be a discrete mixture of Poisson or negative binomial regressions (Yau *et al.*, 2003).

So it is important that a suitable model has both an appropriate data likelihood and incorporates any latent spatial dependences. Existing studies in Britain and elsewhere suggest that migrant push-and-pull effects may be spatially correlated (e.g. places with high attractivity are disproportionately concentrated in certain regions), even though they do not typically include this feature in their model specification, meaning *inter alia* that correlations between and within areas are not represented parametrically. In particular the present study was to some degree inspired by Fotheringham *et al.* (2000), who estimated attractivities by using a fixed effect approach rather than incorporating spatial dependence. Although this study is a pioneering analysis in substantive terms, it has other methodological drawbacks such as applying a log-normal regression when there are many zero flows, and considering only attractivities (and not retentivities). It is also not clear how the method of Fotheringham *et al.* (2000) provides distributional characteristics (e.g. standard errors) for attractivities, whereas full posterior densities and other inferences (e.g. on rankings of attractivities) are available under an MCMC approach. Existing studies also suggest that push and pull are correlated within areas, and such correlation may be positive if areas tend to vary in turnover, with high turnover areas marked by both high attractivity and low retentivity (Tervo, 2001). It is straightforward to model correlation between and within areas via a random-effects approach, whereas it is not possible to model spatial correlation or within-area correlation if push or pull effects are treated as fixed effects.

The relevance of a hierarchical random-effects approach including spatial correlation and the resulting benefit to model fit is demonstrated in a case-study application. The main analysis involves all-age migrant flows between $n = 354$ English local authority areas (LAAs) in 2000–2001, with data from the 2001 census. Intra-area flows are not included in the model: only migrations involving a change in LAA of residence. However, inter-LAA flows for particular age groups (e.g. young adults or retired people) are also considered to assess differential attractivity of areas and broader regions according to age. As will be shown, migration flows of older adults (all groups over 30 years of age) show a concentration of high attractivity areas in particular English regions, particularly South West England.

2. Model for migrant flows

Consider migration flows y_{ij} (i.e. counts) from origin areas i to destination areas j ($i, j = 1, \ldots, n$; $i \neq j$) with intra-area flows y_{ii} not considered, so that there are $N = n(n-1)$ observations. For census migration data (as in the case-study) the time span over which migration is observed is most commonly within the previous year, though some census migration data relate to moves in the preceding 5 years.

Migration is typically relatively rare in relation to populations at risk P_i in origin areas, but considerable variability in rates (overdispersion) is likely. If we allow for overdispersion, a Poisson distribution for the counts y_{ij} is not appropriate, although some studies of migration have assumed a Poisson density (e.g. Flowerdew and Aitkin (1982)). Although overdispersion can be dealt with in various ways, a simple approach (based on Poisson–gamma mixing) leads to a density for y_{ij} that is marginally negative binomial (Hilbe, 2007). The appropriateness of an overdispersed model is apparent from simple summary statistics for the English LAA migration flows in the case-study, with a mean y_{ij} of 16.9 but a standard deviation of 81.

Letting X_{ij} be any known predictors, the conditional Poisson regression means $E(y_{ij}|X_{ij}) = \mu_{ij}$ for each origin–destination pair may be represented as $\mu_{ij} = P_i \nu_{ij} = P_i \exp(X_{ij}\beta)$, with P_i as an offset, and ν_{ij} as the implied annual migration rate to destination j for people living in area i. Then we may allow for heterogeneity by setting $E(y_{ij}|X_{ij}, \omega_{ij}) = \mu_{ij}\omega_{ij}$, where the multiplicative factors ω_{ij} are gamma distributed with mean 1, namely $\omega_{ij} \sim \mathrm{Ga}(\alpha, \alpha)$. These are interpretable as latent effects or frailties whereby particular origin–destination flows are at variance with the Poisson mean (e.g. Greene (2008), page 586). Integrating out the ω_{ij} leads to a marginal negative binomial density for the y_{ij}, namely

$$P(y_{ij}|X_{ij}) = \frac{\Gamma(y_{ij} + \alpha)}{\Gamma(y_{ij} + 1)\,\Gamma(\alpha)} \left(\frac{\alpha}{\mu_{ij} + \alpha}\right)^{\alpha} \left(\frac{\mu_{ij}}{\mu_{ij} + \alpha}\right)^{y_{ij}}, \tag{1}$$

and $\log(\mu_{ij})$ can be modelled as a function of attributes X_{ij} of areas i and j or their separation. Allowing latent heterogeneity induces overdispersion while preserving the conditional mean:

$$E(y_{ij}|X_{ij}) = \mu_{ij},$$
$$\mathrm{var}(y_{ij}|X_{ij}) = \mu_{ij}(1 + \kappa\mu_{ij})$$

where $\kappa = 1/\alpha = \mathrm{var}(\omega_{ij})$.

Following the principle of the well-known gravity model, we need to allow for

(a) mass effects in different origin and destination areas and
(b) distance decay (or more broadly decay according to some measure of spatial separation), namely that (*ceteris paribus*) migration declines as the distance d_{ij} between areas grows.

Mass and decay effects are both relevant to defining the 'migration context' of a particular area, namely the size and proximity of nearby areas with populations at risk of migrating to that area, or offering potential destinations for out-migrants. For example, the 33 London LAAs are close to neighbouring areas with large populations, and not allowing for this will mean that their migrant attractivity will be overstated.

To operationalize mass effects we may use area populations, or employment or housing stocks. Here populations P_i and P_j are taken as measures of mass. So, rather than taking $\log(P_i)$ as offset (with known parameter 1) in a log-link regression for the μ_{ij}, we may allow for an unknown regression effect. Thus a baseline gravity-type model would stipulate

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(P_i) + \beta_2 \log(P_j) + \gamma \log(d_{ij}), \tag{2}$$

where β_1 and β_2 are expected to be close to 1, and γ is expected to be negative by virtue of distance decay, typically taking a value between -0.5 and -2. The conventional gravity model may, however, be relatively insensitive to the spatial structure of populations at risk of migrating to area j (Fik and Mulligan, 1998), and so the gravity model is extended to include an accessibility index

$$A_j = \sum_{r \neq j} \frac{P_r}{d_{rj}}. \tag{3}$$

This is a form of competing destinations index (Fotheringham, 1983; Fik and Mulligan, 1990) and measures the proximity of destination j to alternative intervening or competing destinations. Large values of A_j indicate that destination j is close to other alternative destinations (with large populations), whereas low values indicate that destination j is spatially isolated (Pellegrini and Fotheringham, 2002). So to set the migration context (before it is sensible to represent attractivity or retentivity effects) we have

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(P_i) + \beta_2 \log(P_j) + \gamma \log(d_{ij}) + \delta \log(A_j). \tag{4}$$

However, the goal of the present paper is to derive summary indices of area-specific push-and-pull effects after controlling for migration context. Hence an extended gravity model is proposed, with origin push effects s_{1i}, and destination pull effects (attractivities) s_{2j}:

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(P_i) + \beta_2 \log(P_j) + \gamma \log(d_{ij}) + \delta \log(A_j) + s_{1i} + s_{2j}, \tag{5}$$

where the bivariate push-and-pull effects within each area $s_i = (s_{1i}, s_{2i})$ are taken as spatially dependent random effects, with zero means over all areas. Retentivity scores may be obtained simply as negative push scores, namely $r_i = -s_{1i}$. A focus on retentivity (i.e. impeded out-mobility) is in tune with several studies, e.g. on the effects of different types of housing tenure on migration (e.g. Hughes and McCormick (1981) and Boyle (1998)).

A random-effects approach to latent spatial dependences between areas was also used by Lesage *et al.* (2007) in a knowledge spillover application; they assumed a simultaneous auto-regressive dependence between area-specific attractivities, and between retentivities, but (unlike the approach of the current paper, which is described in Section 5) did not specifically allow these effects to be correlated within areas. Lesage *et al.* (2007) adopted an MCMC auxiliary mixture sampling approach in estimation (Frühwirth-Schnatter and Wagner, 2006).

3. Model variants

Should the negative binomial model not provide a satisfactory fit to a particular data set, possible alternatives are discrete mixtures of negative binomial regressions, or (when excess 0s are a particular issue) a ZINB model (Yau *et al.*, 2003). Under this model zero counts may result from either of two mechanisms: they may be true 0s with probability π, or result from a stochastic mechanism, when the process is 'active' but sometimes produces zero events. The likelihood under a ZINB model then has the form

$$P(y_{ij}) = \pi + (1-\pi)\left(\frac{\alpha}{\mu_{ij}+\alpha}\right)^{\alpha} \qquad y_{ij}=0, \tag{6a}$$

$$P(y_{ij}) = (1-\pi)\frac{\Gamma(y_{ij}+\alpha)}{\Gamma(y_{ij}+1)\,\Gamma(\alpha)}\left(\frac{\alpha}{\mu_{ij}+\alpha}\right)^{\alpha}\left(\frac{\mu_{ij}}{\mu_{ij}+\alpha}\right)^{y_{ij}} \qquad y_{ij}>0. \tag{6b}$$

There is a substantive issue here, namely whether a particular migration interaction can ever be regarded as a permanent true 0, or whether zero flows are a genuine feature related (for example) to high spatial separation, and to the rarity of migration once populations at risk are allowed for. By contrast, a threshold effect that generates a true 0 may be relevant in other economic applications, such as trade flows in particular goods (Ranjan and Tobias, 2007). Exploratory work with a ZINB model in the case-study of the current paper proved to slow computation times in WinBUGS considerably. The fit was slightly improved but inferences regarding the patterning of attractivity and retentivity were little affected.

Another possible model development involves the treatment of known influences on interarea migration. The essential aim of the methodology in the current paper is to summarize both known and unknown influences on attractivity and retentivity in generic scores, much as Fotheringham *et al.* (2000) did for attractivities. Introducing known economic predictors of migration at origin and destination (e.g. house prices or employment) into equation (5), and then in addition including random attractivity and retentivity terms, would mean that the latter are now only representing residual unknown influences. We no longer have complete measures of attractivity, but only measures in which the effects of house prices and employment opportunities (etc.) are partialled out. However, it is possible to allow a mechanism whereby explanatory variates have a role in determining attractivity, while retaining generic attractivity and retentivity terms in equation (5).

This lies in the analogy between model (5) and common spatial factor methods (e.g. Wang and Wall (2003) and Hogan and Tchernis (2004)). Thus, in the terminology of factor analysis, there are $n-1$ indicators for the attractivity score of any particular area, namely the $n-1$ flows to that area. As well as these 'multiple indicators', we can extend the spatial attractivity scheme to have 'multiple causes' in a multiple indicators, multiple causes type of factor model, with the latter meaning house prices, unemployment rates, etc. So in the spatial prior for the random attractivities (Section 5) we may account for measured causes, as Congdon (2009) illustrates in another form of application, namely an analysis of a latent small area health need index. Congdon (2009) adapts a spatial conditional auto-regressive (CAR) prior modified for known predictors, as set out, for example, by Bell and Broemeling (2000).

A further modification of equation (5) in line with a common spatial factor conceptualization is that loadings, especially on attractivities of particular destinations j, may vary according to the region R_i that the origin area i is located in. For example, we might expect that the attractivity of destination areas in South West England would be magnified for people in origins in Northern England compared with origin areas that are themselves in the South West. So for M regions we might introduce region-specific loadings (e.g. $M=8$ for the eight standard regions of England) $(\lambda_1, \ldots, \lambda_M)$ with one loading set to 1 (for identification) and the remaining loadings as unknowns. Then equation (5) is modified to become

$$\log(\mu_{ij}) = \beta_0 + \beta_1\log(P_i) + \beta_2\log(P_j) + \gamma\log(d_{ij}) + \delta\log(A_j) + s_{1i} + \lambda_{R_i}s_{2j}. \tag{7}$$

4. Substantive evidence for the prior on push-and-pull effects

On the basis of previous study evidence and other *a priori* considerations, the bivariate effects $s_i = (s_{1i}, s_{2i})$ are considered as potentially spatially correlated, and also potentially correlated with each other within areas. The basis for assuming a correlation between the two sets of effects is extensive. A negative correlation within areas between push-and-pull scores might be expected if migration plays an 'equilibrating' role in job or housing markets. The ideal equilibrating role is exemplified by job-led migration flows from high unemployment areas with restricted economic opportunities to prosperous areas, or by housing-led migration flows from high housing cost to lower cost areas.

In fact, many studies show a positive rather than negative association between in- and out-migration rates. To explain this a 'compositional' model has been proposed (Tervo, 2001), namely areas that experience much in-migration have a relatively large number of people who are more likely to move again, thus increasing out-migration. We therefore have some areas with high migrant turnover (high in- and out-migration), whereas some areas have low turnover (restricted out-mobility and limited immigration). For example, housing barriers to migration by virtue of house price differentials or immobility that is related to social rented or 'council' housing (e.g. Boyle (1998) and Hughes and McCormick (1981)) mean restricted mobility for many, even if economic opportunities are limited in the current area of residence. To quote Boyle (1993)

> 'council tenants were more likely to move within local authority areas than owner-occupiers. This implies that institutional constraints restrict migration between local authorities whose policies are designed primarily to house individuals and households from within their own jurisdictions.'

Moreover migration attractiveness, even for working age groups, is increasingly related to quality of life rather than job-led considerations (e.g. Findlay and Rogerson (1993)). Counter-urbanizing migration to less urban areas (e.g. into South West England) may actually run counter to economic opportunities. Williams (2003) and Burley (2004, 2006) showed that high immigration to Cornwall (in the extreme South West of England) is despite Cornwall's unemployment rates being far higher than elsewhere in the South West region, and earnings in Cornwall being lower than in both the South West region and the UK as a whole. Such areas often have a high turnover of seasonal workers in tourism and agriculture, and other temporary residents, so both inflows and outflows may be relatively elevated (after controlling for migration context). High migrant turnover may additionally reflect differential age-related in- and out-migration; an analysis for Cornwall (Burley, 2004) points to net out-migration among 16–29-year-olds more than offset by net in-migration among pre-retirement and retirement age groups.

Spatial clustering in attractivity or retentivity is also likely to be present. Unobserved influences on migrant attractivity or impeded out-mobility are likely to straddle arbitrary administrative boundaries. Moreover existing findings suggest spatial concentrations. The South West of England was identified by Fotheringham *et al.* (2000) as having a disproportionate number of areas with high migrant attractivity, as were other less urbanized regions such as East Anglia, and parts of Yorkshire and Humberside and the extreme North West (e.g. areas including parts of the Lake District and the Yorkshire Dales). Areas with restricted out-mobility may tend to be clustered in older industrial areas or areas with high levels of social housing.

5. Density for random push-and-pull effects

So it seems sensible to allow for the prior density for attractivity and retentivity effects to allow for spatial correlation over areas, and correlation of the two effects within areas. Thus in the extended gravity model a bivariate version of the spatial CAR prior is used for $s_i = (s_{1i}, s_{2i})$ (e.g. Song *et al.* (2006)). This is straightforward to implement in the Bayesian analysis program WinBUGS by the `mv.car` density, which represents a multivariate normal form of the spatial CAR prior.

Define the neighbourhood of area i as the set of L_i LAAs adjacent to area i. Then the conditional form of the multivariate CAR density for the case of bivariate effects is

$$s_i|s_{[i]} \sim N_2(S_i, \Lambda/L_i), \tag{8}$$

where $S_i = (S_{1i}, S_{2i})$ is a two-dimensional vector of average push and attractivity effects in the neighbourhood of area i. Λ is a 2×2 covariance matrix with diagonal elements $(\lambda_1^2, \lambda_2^2)$ representing the variances of the push-and-pull effects, and off-diagonal element $\rho\lambda_1\lambda_2$ representing covariance between push-and-pull effects. The inverse of the covariance matrix (the precision matrix) may be assigned a Wishart prior: a multivariate version of the gamma density that is often used as a prior density for inverse variances.

These assumptions regarding the proximity effect and the prior assumption on the precision matrix are standard in Bayesian spatial applications, but others are possible. The proximity effect in the CAR prior might also be defined by using functions of distance between areas (Conlon and Waller, 1998), but at the expense of at least one extra parameter to measure the distance effect within the CAR spatial prior, and so a slowing in fit via MCMC sampling. As an alternative to the Wishart prior we may adopt various forms of Cholesky decomposition, and this may provide a basis for random-effects selection (Cai and Dunson, 2006).

6. Case-study for English migration

The first part of the case-study considers all-age migrant flows y_{ij} between $n = 354$ English local authorities in 2000–2001 (from the 2001 UK census). The full model assumes bivariate random push–pull effects under a negative binomial regression model, namely

$$P(y_{ij}) = \frac{\Gamma(y_{ij}+\alpha)}{\Gamma(y_{ij}+1)\,\Gamma(\alpha)} \left(\frac{\alpha}{\mu_{ij}+\alpha}\right)^{\alpha} \left(\frac{\mu_{ij}}{\mu_{ij}+\alpha}\right)^{y_{ij}}, \tag{9a}$$

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \log(P_i) + \beta_2 \log(P_j) + \gamma \log(d_{ij}) + \delta \log(A_j) + s_{1i} + s_{2j}, \tag{9b}$$

$$s_i|s_{[i]} \sim N_2(S_i, \Lambda/L_i), \tag{9c}$$

with d_{ij} in kilometres (distances between local authority population centroids). Fotheringham *et al.* (2000), page 402, used a more sophisticated form of interarea distances, namely a

> 'migration-weighted sum of the distances between the wards or pseudo-postcode sectors nested within both districts',

though they did not show how using this measure improves on simple straight line distances between area centroids. Use of straight line distances is common in migration models, though it might be argued that other separation measures be used, for instance road distances, though this presumes that all migrations involve road transport (and that there is a uniquely defined best road route). Arguably a separation measure should be a weighted average of distances defined by relative frequencies of migration transport mode (road, train or air), which would probably involve an initial rather complex application of specialized geographical information system techniques. The effect of the distance variable on flows (and the possibility of alternative measures of distance) needs to be seen against evidence of residual spatial dependences even

Table 1. Model fit summary

Model	*Average deviance*	*Effective parameters*	*DIC value*	*log(psML)*	*Proportion of replicate data with probabilities $Q_{ij} < 0.05$ or $Q_{ij} > 0.95$†*	*Proportion of mixed replicate data with probabilities $Q_{\mathrm{mix},ij} < 0.05$ or $Q_{\mathrm{mix},ij} > 0.95$*
1, spatial push–pull (independent fixed effect push and pull scores)	726904	682	727586	−364090	0.045	—
2, spatial push–pull (independent univariate CAR)	726897	632	727529	−364028	0.044	0.035
3, spatial push–pull (bivariate CAR)	726868	531	727399	−363911	0.044	0.043

†The expected proportion combined over tails is 0.1.

Table 2. Parameter estimates under model 3

Parameter	*Posterior mean*	*Posterior standard deviation*	*Monte Carlo standard error*	*2.5 percentile*	*97.5 percentile*
ρ	0.93	0.009	0.0005	0.91	0.95
α	0.69	0.004	0.0001	0.68	0.69
β_0	−7.32	0.054	0.0063	−7.42	−7.23
β_1	0.72	0.001	0.0002	0.72	0.72
β_2	0.63	0.002	0.0002	0.62	0.63
γ	−1.41	0.003	0.0004	−1.41	−1.42
δ	0.08	0.007	0.0009	0.07	0.10

Table 3. Regional averages, attractivity and push scores

Region	*Push*	*Average attractivity*	*Average excess of attractivity over push*
East Midlands	−0.29	−0.15	0.15
East	−0.02	−0.09	−0.07
London	−0.03	−0.18	−0.15
North East	0.21	0.16	−0.05
North West	−0.14	−0.18	−0.04
South East	0.14	0.01	−0.12
South West	0.34	0.61	0.27
West Midlands	−0.31	−0.36	−0.04
Yorkshire and Humberside	0.04	0.17	0.13
England	0.00	0.00	0.00

Table 4. LAAs with highest attractivity, and highest excess of attractivity over push

Name	*Region*	*Push*	*Attractivity*	*Attractivity minus push*
Local authorities with highest attractivity				
Carrick	South West	1.16	1.70	0.54
Kerrier	South West	1.15	1.55	0.39
Plymouth	South West	1.10	1.50	0.40
North Cornwall	South West	0.97	1.50	0.53
Durham	North East	1.25	1.48	0.23
Restormel	South West	0.93	1.43	0.50
Penwith	South West	0.88	1.39	0.51
Torbay	South West	0.75	1.27	0.52
Newcastle upon Tyne	North East	1.07	1.22	0.15
Exeter	South West	0.91	1.19	0.29
South Hams	South West	0.82	1.19	0.37
Richmondshire	Yorkshire and Humberside	0.91	1.13	0.22
Norwich	East	0.75	1.05	0.30
York	Yorkshire and Humberside	0.80	1.05	0.24
West Devon	South West	0.65	1.03	0.39
Alnwick	North East	0.85	1.02	0.17
Cambridge	East	0.66	1.02	0.36
Leeds	Yorkshire and Humberside	0.65	1.01	0.36
Lancaster	North West	0.79	1.00	0.21
East Devon	South West	0.40	0.98	0.58
Local authorities with highest excess of attractivity over push				
East Devon	South West	0.40	0.98	0.58
Carrick	South West	1.16	1.70	0.54
North Cornwall	South West	0.97	1.50	0.53
Torbay	South West	0.75	1.27	0.52
Penwith	South West	0.88	1.39	0.51
Torridge	South West	0.05	0.56	0.51
Restormel	South West	0.93	1.43	0.50
North Norfolk	East	0.06	0.55	0.49
North Kesteven	East Midlands	0.32	0.81	0.49
North Devon	South West	0.43	0.88	0.45
Isle of Wight	South East	0.00	0.42	0.43
West Dorset	South West	0.39	0.81	0.43
Nottingham	East Midlands	0.45	0.87	0.42
East Lindsey	East Midlands	0.17	0.58	0.42
South Holland	East Midlands	−0.26	0.15	0.41
Plymouth	South West	1.10	1.50	0.40
Kerrier	South West	1.15	1.55	0.39
West Devon	South West	0.65	1.03	0.39
East Riding of Yorkshire	Yorkshire and Humberside	0.40	0.77	0.37
South Hams	South West	0.82	1.19	0.37

after distance decay is allowed for. As noted by Lesage and Pace (2008), the assumption that the fixed regression effect of distance in the gravity model (however measured) removes spatial dependences is simplistic.

As mentioned above, a Bayesian strategy is used for specification and estimation of the model parameters. Hence, we need to stipulate prior densities for all parameters that are involved in defining the model. Although informative priors might be adopted (e.g. constraining γ to be negative, on the basis of expected distance decay), relatively diffuse priors are adopted here. Thus

Table 5. LAAs with highest retentiveness (restricted out-mobility)

Area name	*Region*	*Retentiveness*	*Pull*
Knowsley	North West	0.76	−1.00
Walsall	West Midlands	0.72	−1.11
Cannock Chase	West Midlands	0.69	−0.99
St Helens	North West	0.68	−0.85
Erewash	East Midlands	0.68	−0.74
Rotherham	Yorkshire and Humberside	0.66	−0.59
Sandwell	West Midlands	0.65	−0.94
Dudley	West Midlands	0.65	−0.86
Oldham	North West	0.64	−0.84
Barnsley	Yorkshire and Humberside	0.63	−0.60
Corby	East Midlands	0.62	−0.67
Blaby	East Midlands	0.62	−0.68
Mansfield	East Midlands	0.60	−0.59
Gedling	East Midlands	0.60	−0.72
North Warwickshire	West Midlands	0.60	−0.85
Tamworth	West Midlands	0.59	−0.93
Wigan	North West	0.58	−0.47
Rochdale	North West	0.57	−0.72
Chesterfield	East Midlands	0.57	−0.43
Oadby and Wigston	East Midlands	0.56	−0.31

Table 6. Zero-order correlations between scores and area characteristics

Correlation between attractivity scores and	
% green space, logit (generalized land use database)	0.24
Population density 2001, logarithm	−0.31
% manufacturing employment	−0.33
% unemployment (among economically active aged 16–74 years)	−0.05
% 16–74 years economically active in hotels and catering	0.42
% 16–74 years full-time students	0.27
House prices, 2001 (logarithm) (Land Registry)	0.08
Social rented housing	−0.10
Private rented housing	0.45
Correlation between retentiveness scores and	
% green space, logit (generalized land use database)	−0.19
Population density 2001, logarithm	0.25
% manufacturing employment	0.46
% unemployment (among economically active aged 16–74 years)	0.11
% 16–74 years economically active in hotels and catering	−0.35
% 16–74 years full-time students	−0.27
House prices, 2001 (logarithm) (Land Registry)	−0.20
Social rented housing	0.11
Private rented housing	−0.41

normal $N(0, 100)$ priors are taken on the parameters $\psi = \{\beta_0, \beta_1, \beta_2, \gamma, \delta\}$, a $U(0, 1000)$ prior is adopted for α and a Wishart prior is taken on the precision matrix, namely $\Lambda^{-1} \sim W(I, 2)$, with identity scale matrix and 2 degrees of freedom.

The full model (model 3) is compared with two alternative models with less general assumptions. Thus model 1 assumes that the (s_{1i}, s_{2i}) are unrelated fixed effects, independent over areas

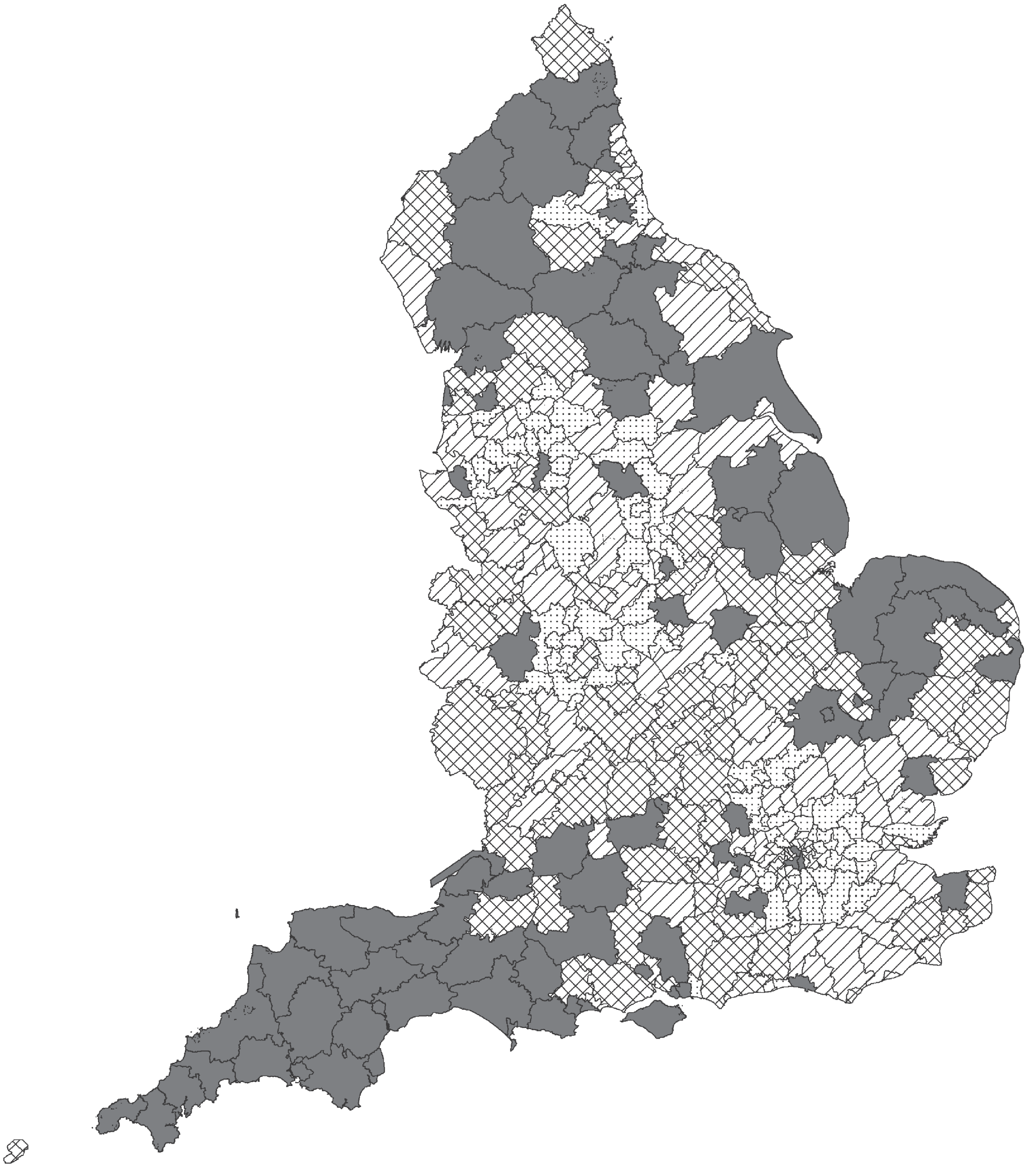

Fig. 1. Attractivity scores, model 3: ⊞, under −0.30; ▨, from −0.30 to −0.01; ⊠, from −0.01 to 0.27; ■, over 0.27

and of one another, with normal $N(0, 100)$ priors on each s_{1i} and s_{2i}. Priors on ψ and α are as in model 3. Model 2 allows for spatial correlation in push-and-pull effects, but not for their intercorrelation. It assumes separate univariate CAR priors on s_{1i} and s_{2i}.

The fit of the three models is assessed by using the deviance information criterion (DIC) of Spiegelhalter *et al.* (2002), and the logarithm of the pseudo-marginal-likelihood (log(psML)), based on Monte Carlo estimates of conditional predictive ordinates (Iyengar and Dey, 2000). The DIC procedure compares the average deviance $\bar{D}$ over MCMC chains with the deviance

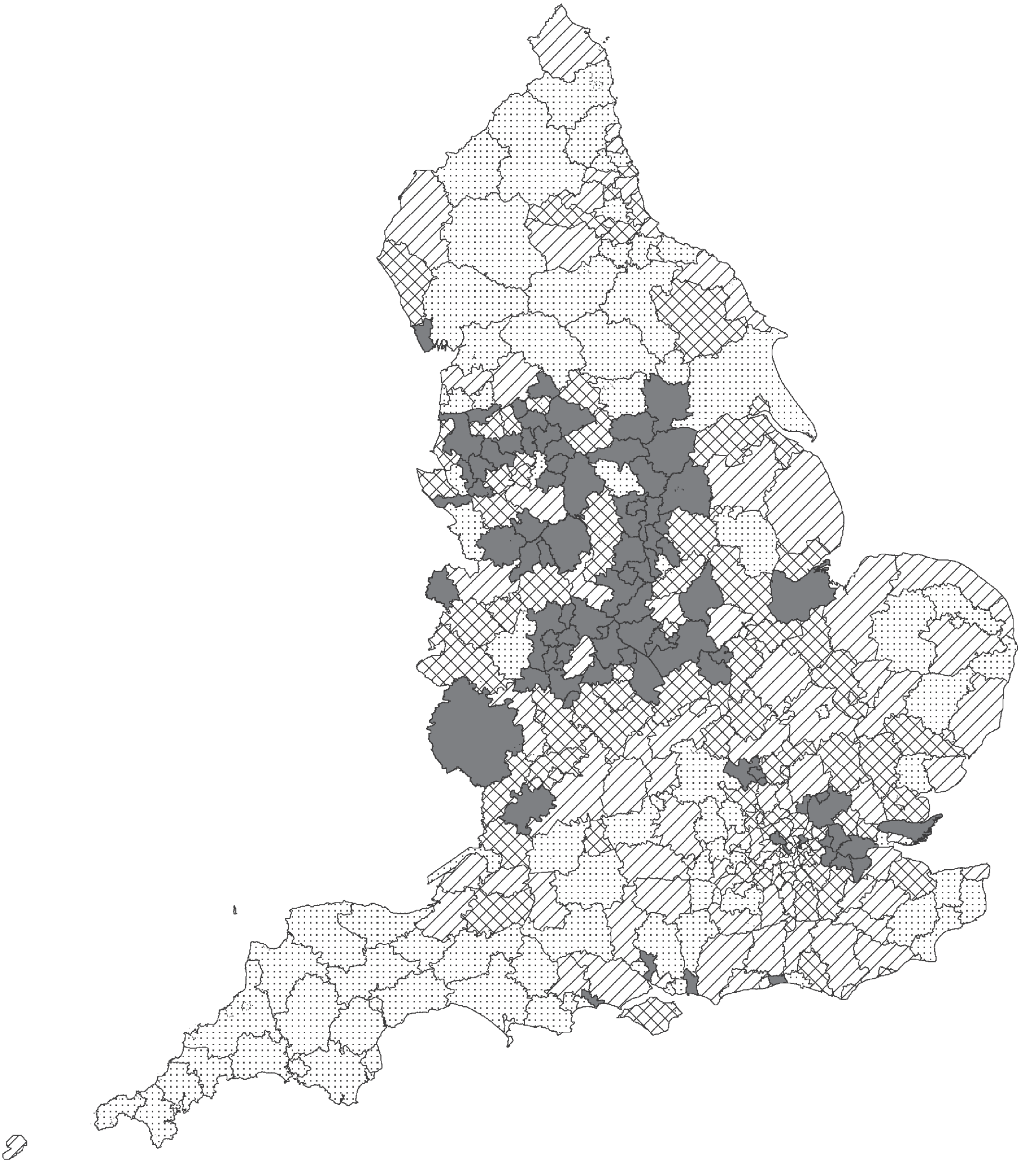

Fig. 2. Retentivity scores, model 3: ▦, under −0.16; ▨, from −0.16 to 0.02; ⊠, from 0.02 to 0.21; ■, over 0.21

$D(\bar{\theta})$ at the posterior mean $\bar{\theta}$ of the parameters to produce an estimate of effective parameters $d_e = \bar{D} - D(\bar{\theta})$, with the DIC then provided by $D(\bar{\theta}) + 2d_e$. Lower values of the DIC indicate better fit, as do higher values of log(psML). Spiegelhalter *et al.* (2002), page 613, suggested that models with DIC values within 1–2 of the 'best' deserve consideration, but that differences exceeding 3 provide less support.

Model checking involves sampling predictions $y_{\mathrm{rep},ij}$ from the posterior predictive density $p(y_{\mathrm{rep}}|y)$ which are compared with the actual migrant flows (Gelfand, 1996). Concordance with the data may be represented by probabilities $Q_{ij}=\Pr(y_{\text{rep},ij}\leqslant y_{ij}|y)$, with extreme values (e.g. under 0.05 or over 0.95) indicating origin–destination pairs where the model does not fit well (namely a tendency to overpredict and underpredict according to whether $Q_{ij}<0.05$ or $Q_{ij}>0.95$ respectively). These probabilities are estimated in practice by counting MCMC iterations r where the constraint $y^{(r)}_{\text{rep},ij}\leqslant y_{ij}$ holds, which for discrete migration flows is based on the probability (Marshall and Spiegelhalter, 2003)

$$\Pr(y_{\text{rep},ij}<y_{ij}|y)+0.5\Pr(y_{\text{rep},ij}=y_{ij}|y). \tag{10}$$

For models which take latent push-and-pull effects (s_{1i},s_{2i}) as spatial random effects, we may also obtain 'mixed replicate data' $y_{\text{mix},ij}$, leading to predictive tests which are less conservative as they are less influenced by the actual data (Marshall and Spiegelhalter, 2007). Thus replicate random effects $(s_{\text{rep}1i},s_{\text{rep}2i})$ are sampled first, and then replicate data $y_{\text{mix},ij}$ sampled by using the means

$$\log(\mu_{\text{mix},ij})=\beta_0+\beta_1\log(P_i)+\beta_2\log(P_j)+\gamma\log(d_{ij})+\delta\log(A_j)+s_{\text{rep}1i}+s_{\text{rep2i}}. \tag{11}$$

The revised way of assessing predictive concordance is based on probabilities

$$Q_{\text{mix},ij}=\Pr(y_{\text{mix},ij}<y_{ij}|y)+0.5\Pr(y_{\text{mix},ij}=y_{ij}|y). \tag{12}$$

This procedure was not feasible under the fixed effects model 1 since sampling replicates $(s_{\text{rep}1i}$, $s_{\text{rep}2i})$ from the fixed effects prior led to numerical overflow problems. For a model that satisfactorily reproduces the observations, the expected proportion of the N observations for which $Q_{\text{mix},ij}<0.05$ would then be 0.05 or lower, and similarly for observations with $Q_{\text{mix},ij}>0.95$ (Gelfand (1996), page 153).

Estimation was based on running two chains of 5000 iterations with convergence assessed according to the criteria of Brooks and Gelman (1998); in all models, convergence was present after 1000 iterations, and summaries are based on iterations 1000–5000. Table 1 summarizes the fit for the three models. Model 3 has a clearly superior fit, with fewer effective parameters than the fixed effects model, or a model not incorporating a within-area push–pull correlation. In terms of satisfactorily reproducing the data, the predictive check based on MCMC sampling of $y_{\text{rep},ij}$ from current values of μ_{ij} suggests no strong failure in reproducing the data under any of the three models, albeit under this possibly conservative full data posterior predictive test (Marshall and Spiegelhalter, 2007). The more satisfactory mixed predictive test confirms that models 2 and 3 satisfactorily reproduce the observed data.

Table 2 presents estimates (posterior means and 95% credible intervals) for ψ and α under model 3, as well as the push–pull correlation ρ within areas. Table 3 lists regional averages for push, pull and net pull (pull minus push) effects under model 3, Table 4 lists particular LAAs with high attractivity scores, and Table 5 lists areas with the highest retentivity scores (or equivalently the lowest push scores). Table 6 presents correlations between the scores and selected indicators (census and non-census) relating to the urban–rural status of the LAAs, their employment structure, their housing tenure mix and house price levels.

The highest attractivities (pull scores) in model 3 are concentrated in South West England, East Anglia and parts of the North, though some regional centres and university towns (e.g. Cambridge, York and Newcastle) also have high attractivity—as can be seen from Tables 3 and 4, and from Fig. 1. By contrast, areas with high retentivity (restricted out-mobility) tend to concentrate in highly urbanized areas in the Midlands, London and the lower North West (Table 5 and Fig. 2). Fig. 1 demonstrates a clear spatial correlation in the location of more attractive LAAs, and the same is apparent for restricted out-mobility (Fig. 2).

Table 7. Linear regression of scores on area characteristics

Predictor	*Posterior mean*	*Posterior standard deviation*	*2.5 percentile*	*97.5 percentile*
(a) Attractivity scores				
Constant	4.287	1.103	2.038	6.379
Population density (logarithm)	−0.153	0.042	−0.234	−0.067
Hotel and catering jobs	0.026	0.013	0.000	0.053
Students	0.078	0.009	0.060	0.097
Private renting	0.044	0.009	0.026	0.062
Social rented housing	−0.005	0.004	−0.014	0.004
Unemployment rate	−0.009	0.021	−0.048	0.032
House prices 2001 (logarithm)	−0.392	0.084	−0.552	−0.214
Manufacturing jobs	−0.037	0.006	−0.048	−0.025
% green space	0.003	0.003	−0.002	0.009
(b) Retentiveness scores				
Constant	−0.619	0.065	−0.748	−0.488
Unemployment rate	−0.022	0.012	−0.047	0.003
Manufacturing jobs	0.010	0.003	0.003	0.016
Social rented housing	0.037	0.003	0.030	0.043

High attractivity areas are a mix of less urbanized areas which may offer higher quality of life (e.g. in Cornwall in the extreme South West, and parts of East Anglia), and certain urban areas (university towns and coastal areas) where mobile groups (students or seasonal workers) create high migrant turnover. Typically in the former category, namely less urbanized areas with high attractivity, there is a relatively large excess of the pull index over the push index, so high attractivity in these areas is not just a matter of turnover among mobile groups but demonstrates migration gain, for instance, through quality-of-life factors.

There is a high posterior correlation (posterior mean for ρ of 0.93) between attractivity and push indices in model 3; equivalently stated, attractivity and retentivity are strongly inversely correlated. In fact there is also a high within-area correlation (0.88) between the two sets of posterior means of s_{1i} and s_{2i} in the fixed effects model 1, though correlation is not incorporated *a priori* in that model. So the compositional hypothesis, proposing a positive push–pull relationship due to flows by mobile groups that raise both inflows and outflows, receives support.

Correlations between the push-and-pull scores and selected indices of urban, employment and housing status support the relevance of counterurbanizing moves to lower density areas with more green space—as well as high attractivity due to turnover effects in university towns, and areas with a large tourist sector (Table 6). To represent the effects of such factors jointly a linear regression—involving predictors taken as substantively relevant from the literature—was carried out for attractivity and retentivity scores (Table 7). These regressions use the posterior mean scores as the dependent variable and were carried out (using default priors) in the Bayesian package BayesX (Brezger *et al.*, 2005). An analogous procedure (i.e. regressing point estimates of attractivities on relevant area characteristics) was used in the classical analysis of Fotheringham *et al.* (2000).

Table 7, part (a), shows that the main (statistically significant) influences on attractivity are lower population density, lower house prices, university town status, presence of a large private rented sector (offering short-term housing opportunities for mobile groups) and avoidance of manufacturing areas. The fact that flows tend towards lower house price areas (after controlling for other predictor variables) shows that seeking lower house prices may be a component of counterurbanizing moves. A retentivity regression including just manufacturing, social housing and unemployment shows positive effects for the first two, namely restricted out-mobility for a set of areas with above-average levels of social housing and dependence on manufacturing jobs. After controlling for these two predictors, the effect of unemployment is negative, though marginally non-significant; this is in line with higher outflows from areas with higher unemployment. So underlying equilibrating effects (movement towards lower house prices and movement away from higher unemployment) may coexist with the predominant turnover effect.

7. Analysis for particular age groups

Migration patterns for all ages combined may conflate differential patterns by age or other demographic categories. For example, London and other large UK cities attract migrants in early working life (Champion *et al.*, 2007), and this may pertain even after controlling for the effects of population size, distance decay and accessibility. Hence the second part of the case-study analysis considers age-specific migration flows: a strategy that was also followed by Fotheringham *et al.* (2000).

To assess influences on attractivity according to age, four groups were distinguished: young adults (ages 18–29 years), intermediate ages (30–44 years), middle ages (45–59 years) and retirement ages (over 60 years). Table 8 shows that for young adults there is still relatively high average attractivity for the South West but that, in terms of the net balance of attractivity over push, London has the highest average. For the older middle aged and retirees, a markedly high attractivity score for the South West is apparent (average of 0.77 for 45–59-year-olds and 0.82 for the over 60s), and heavily urbanized regions, namely London and the West Midlands, have markedly low attractivity. A similar contrast holds even for the relatively young age band of 30–44-year-olds, who are most likely to be parents for whom migration involves family units.

8. Final remarks

Census migration flows between administratively defined areas are considerably affected by the patterning of urban areas, and migration to and from a particular local authority or town will be affected by the proximity and size of other such areas. This is demonstrated in the case of England by the partitioning of continuous urban spaces (e.g. London) into several separate administrative subdivisions; by contrast, local authorities that are defined by distinct free-standing towns (e.g. medium-sized market towns) may be relatively distant from other population centres. Without controlling for the urban configuration, the migrant attractiveness of metropolitan subdivisions will be overstated, and that of relatively remote rural areas or free-standing towns will be understated (Wall, 2001; Fotheringham *et al.*, 2000).

After allowing for the urban configuration, we can assess levels of migrant attractiveness of different areas (as potential destinations) or their retentivity (as current places of residence). This raises the question of an appropriate prior for latent area effects. The existing literature for analysing migration interaction flows is largely focused on fixed effects modelling within the gravity model, which assumes that spatial dependences are removed by including distance effects or the effects of other accessibility indices in the regression mean (Lesage and Pace, 2009). If it were true that spatial dependences were removed simply by modelling distance decay (or effects of other accessibility indices), then a fixed effects model for attractivities and retentivities might be justified—though an unstructured random-effect prior is still likely to have lower complexity.

Table 8. Attractivity and push scores by region and age band†

Region and age group	*Average attractivity*	*Average push*	*Average (attractivity − push)*
Young adults			
East Midlands	−0.27	−0.24	−0.03
East	−0.09	−0.06	−0.03
London	0.37	−0.13	0.50
North East	0.05	0.24	−0.19
North West	−0.24	−0.03	−0.21
South East	0.07	0.02	0.05
South West	0.35	0.32	0.03
West Midlands	−0.35	−0.19	−0.15
Yorkshire and Humberside	0.14	0.16	−0.03
Ages 30–44 years			
East Midlands	−0.16	−0.32	0.16
East	−0.06	−0.01	−0.05
London	−0.29	0.15	−0.43
North East	0.17	0.18	−0.01
North West	−0.10	−0.16	0.06
South East	0.00	0.16	−0.15
South West	0.57	0.27	0.30
West Midlands	−0.31	−0.33	0.02
Yorkshire and Humberside	0.18	0.01	0.17
Older middle aged			
East Midlands	−0.09	−0.28	0.19
East	−0.03	0.00	−0.03
London	−0.42	−0.06	−0.36
North East	0.13	0.19	−0.06
North West	−0.16	−0.16	0.00
South East	−0.03	0.17	−0.20
South West	0.77	0.37	0.40
West Midlands	−0.41	−0.33	−0.09
Yorkshire and Humberside	0.23	−0.06	0.29
Retirement ages			
East Midlands	−0.11	−0.33	0.22
East	0.02	0.03	−0.01
London	−0.65	0.00	−0.64
North East	0.17	0.13	0.05
North West	−0.20	−0.22	0.02
South East	−0.01	0.19	−0.19
South West	0.82	0.45	0.37
West Midlands	−0.34	−0.38	0.04
Yorkshire and Humberside	0.20	−0.08	0.28

†England scores are 0.

However, findings of many studies of migration and other interaction data indicate that attractiveness and retentivity are spatially structured (Lesage *et al.*, 2007; Griffith and Jones, 1980). Such differences in attractiveness or retentivity may be related to known area attributes to some extent, but unobserved influences are likely and will tend to vary smoothly over space, regardless of arbitrary administrative boundaries. This provides a rationale for assuming spatial correlation in migrant attractivity and retentivity scores. In the model that was assumed in the analysis here, such scores are also assumed to be interrelated between areas under a random-effects density, and the case-study has demonstrated the consequent gain in fit compared with a fixed effects method. A Bayesian approach to estimation and specification is advocated and applied, and the improved fit that is obtained under a spatial random-effects approach within a negative binomial likelihood is demonstrated. Fitting such models by using other methods (e.g. quadrature) generally becomes unwieldy, as more such effects are present in a model, and will not provide the same flexibility in inferences as is available under a fully Bayes approach.

Other methods (regarding estimation, likelihood assumptions and assumptions regarding area terms) have been used to estimate migrant attractivity. Fotheringham *et al.* (2000) adopted a log-normal approximation and a fixed effects approach to estimating attractivities. The log-normal approximation assumes that flows y_{ij} are log-normal with constant residual variance; thus with $z_{ij} = \log(y_{ij} + 1)$ one assumes that $z_{ij} \sim N\{\log(\mu_{ij}), \sigma^2\}$. Drawbacks of the log-normal model were mentioned by Lesage and Pace (2009), Ranjan and Tobias (2007) and Flowerdew and Aitkin (1982), such as biased estimation of μ_{ij}, and problems in assuming log-normality when there are many zero flows. Flowerdew and Aitkin (1992) and Flowerdew (1991) argued for a Poisson regression method, though this will not be appropriate if the data are over-dispersed. Hence the current paper has opted for negative binomial regression. Wall (2001) took the response as

$$z_{ij} = \frac{y_{ij} - y_{ji}}{P_i P_j}$$

in a normal errors regression and so did not include a data-generating mechanism for the actual migration flows. Wall sought to use interregional migration (for 10 British regions, including Scotland and Wales, but with London subsumed within South East England) to estimate differences in regional quality of life, and his findings tally with those of the current paper in that the South West is estimated to have the highest quality of life.

In this paper, a negative binomial likelihood has been assumed for migration flows between the English LAA areas and was found to provide a satisfactory fit and predictive performance. With regard to modelling often encountered features of migrant interaction data such as excess 0s, it is possible that further model enhancements improve the fit yet further. However, practical computation issues are relevant here given that MCMC methods have been used. There may possibly be some further improvement in fit in adopting a more complex model (such as the ZINB model) but at the cost of greater computational demands and possibly some risk of overfitting. Exploratory work with a ZINB model in the case-study that was reported on above proved to slow computation times in WinBUGS considerably, and involved a non-standard likelihood approach (Spiegelhalter *et al.*, 2003), so replicate data cannot be obtained. The fit was slightly improved but inferences regarding the patterning of attractivity and retentivity were little affected.

The methodology of the current paper has adopted the most common strategy of not including intra-area flows. By contrast, LeSage and Pace (2008) developed a generalized latent spatial effects model that can explain both types of migration behaviour (intra-area as well as between area), while using explanatory variables that are associated with own regions as well as traditional explanatory variables for origins and destinations. Including intra-area flows raises new questions that might better be handled in an extended analysis comparing models that do and do not include intra-area flows. Including such flows raises questions regarding the arbitrary nature of administrative boundaries (a short-range migration may or may not involve a change in administrative boundaries), and about the nature of the processes being modelled: is moving within the same area essentially the same (in terms of implied area preferences) as staying in that area, especially if (as commonly observed) the move is over a short distance, or more comparable with migration out of the area in the sense that there is dissatisfaction with the former residence?

Although adopting a distinct methodology, the case-study findings in the present paper coincide in substantive terms with those of other studies. Thus the present paper tends to confirm the growing importance of quality of life in counterurbanizing migrant patterns, and it also demonstrates turnover effects (a positive correlation between inflow and outflow) linked to the composition of migrant streams. However, linear regressions involving mean area scores did show underlying equilibrating effects (e.g. moves towards lower house price areas). An age-group-specific analysis demonstrates counterurbanizing effects for all except young adults in the 16–29 years age group, and the strong migrant attractivity of South West England.

References

Banerjee, S., Carlin, B. and Gelfand, A. (2003) *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton: CRC Press.

Banerjee, S., Gelfand, A. and Polasek, W. (2000) Geostatistical modelling for spatial interaction data with application to postal service performance. *J. Statist. Plann Inf.*, **90**, 87–105.

Baxter, M. and Ewing, G. (1981) Models of recreational trip distribution. *Regl Stud.*, **15**, 327–344.

Beaman, J., Jeng, J.-M. and Fesenmaier, D. (1997) Clarification of cumulative attractivity as a concept and its measurement. *J. Travl Res.*, **36**, 74–77.

Bell, S. and Broemeling, L. (2000) A Bayesian analysis for spatial processes with application to disease mapping. *Statist. Med.*, **19**, 957–974.

Boyle, P. (1993) Modelling the relationship between tenure and migration in England and Wales. *Trans. Inst. Br. Geogr.*, **18**, 359–376.

Boyle, P. (1998) Migration and housing tenure in South East England. *Environ. Plann* A, **30**, 855–866.

Brezger, A., Kneib, T. and Lang, S. (2005) BayesX: analyzing Bayesian structured additive regression models. *J. Statist. Softwr.*, **14**, 1–22

Brooks, S. and Gelman, A. (1998) General methods for monitoring convergence of iterative simulations. *J. Computnl Graph. Statist.*, **7**, 434–445.

Burley, S. (2004) A quantitative analysis on the relationship between migration and the economy in the small areas of Cornwall. *British Society for Population Studies A. Conf., Leicester, Sept. 13th–15th.* (Available from `http://www.lse.ac.uk/collections/BSPS/annualConference/2004/migrationAndPopulationMobility.htm`.)

Burley, S. (2006) In-migration to Cornwall: stepping off the treadmill? *Conf. Periphery and Policy, Truro, Apr. 21st–22nd.*

Cai, B. and Dunson, D. (2006) Bayesian covariance selection in generalized linear mixed models. *Biometrics*, **62**, 446–457.

Champion, T., Coombes, M., Raybould, S. and Wymer, C. (2007) Migration and socioeconomic change: a 2001 Census analysis of Britain's larger cities. *Report*. Joseph Rowntree Foundation and University of Newcastle upon Tyne.

Congdon, P. (2009) A spatial structural equation model with an application to area health needs. *J. Statist. Computn Simuln*, **79**, in the press, doi 10.1080/00949650802676300.

Conlon, E. M. and Waller, L. A. (1998) Flexible neighborhood structures in hierarchical models for disease mapping. *Report 1998-018.* Division of Biostatistics, University of Minnesota, Minneapolis.

Fik, T. and Mulligan, G. (1990) Spatial flows and competing central places: towards a general theory of hierarchical interaction. *Environ. Plann* A, **22**, 527–559.

Fik, T. and Mulligan, G. (1998) Functional form and spatial interaction models. *Environ. Plann* A, **30**, 1497–1507.

Findlay, A. and Rogerson, R. (1993) Migration places and quality of life: voting with their feet? In *Population Matters: the Local Dimension* (ed. T. Champion), pp. 33–49. London: Chapman.

Flowerdew, R. (1991) Poisson regression modelling of migration. In *Migration Models: Macro and Micro Approaches* (eds J. Stillwell and P. Congdon). London: Belhaven.

Flowerdew, R. and Aitkin, M. (1982) A method for fitting the gravity model based on the Poisson distribution. *J. Regl Sci.*, **22**, 191–202.

Fotheringham, A. (1983) Some theoretical aspects of destination choice and their relevance to production-constrained gravity models. *Environ. Plann* A, **15**, 1121–1132.

Fotheringham, A., Champion, T., Wymer, C. and Coombes, M. (2000) Measuring destination attractivity: a migration example. *Int. J. Popln Geogr.*, **6**, 391–422.

Frühwirth-Schnatter, S. and Wagner, H. (2006) Auxiliary mixture sampling for parameter-driven models of time series of counts. *Biometrika*, **93**, 827–841.

Gelfand, A. (1996) Model determination using sampling based methods, In *Markov Chain Monte Carlo in Practice* (eds W. R. Gilks, S. Richardson and D. J. Spiegelhalter), ch. 9. London: Chapman and Hall.

Gelfand, A. and Smith, A. (1990) Sampling based approaches to calculate marginal densities. *J. Am. Statist. Ass.*, **85**, 398–409.

Greene, W. (2008) Functional forms for the negative binomial model for count data. *Econ. Lett.*, **99**, 585–590.

Griffith, D. and Jones, K. (1980) Explorations into the relationship between spatial structure and spatial interaction. *Environ. Planng* A, **12**, 187–201.

Haneuse, S. and Wakefield, J. (2004) Ecological inference incorporating spatial dependence. In *Ecological Inference: New Methodological Strategies* (eds G. King, O. Rosen and M. Tanner), pp. 266–300. Cambridge: Cambridge University Press.

Hilbe, J. (2007) *Negative Binomial Regression*. New York: Cambridge University Press.

Hogan, J. and Tchernis, R. (2004) Bayesian factor analysis for spatially correlated data, with application to summarizing area-level material deprivation from census data. *J. Am. Statist. Ass.*, **99**, 314–324.

Hughes, G. and McCormick, B. (1981) Do council housing policies reduce migration between regions? *Econ. J.*, **91**, 919–937.

Iyengar, M. and Dey, D. (2000) Bayesian analysis of compositional data. In *Generalized Linear Models: a Bayesian Perspective* (eds D. Dey, S. Ghosh and B. Mallick), pp. 349–364. New York: Dekker.

LeSage, J., Fischer, M. and Scherngell, T. (2007) Knowledge spillovers across Europe, evidence from a Poisson spatial interaction model with spatial effects. *Pap. Regl Sci.*, **86**, 393–421.

LeSage, J. and Pace, R. (2008) Spatial econometric modeling of origin-destination flows. *J. Regl Sci.*, **48**, 941–967.

LeSage, J. and Pace, R. (2009) *Introduction to Spatial Econometrics*. Boca Raton: Chapman and Hall–CRC.

Lindley, D. V. and Smith, A. F. M. (1972) Bayes estimates for the linear model (with discussion). *J. R. Statist. Soc.* B, **34**, 1–41.

Lue, C.-C., Crompton, J. and Stewart, W. (1996) Evidence of cumulative attraction in multidestination recreational trip decisions. *J. Travl Res.*, **35**, 41–49.

Lunn, D., Thomas, A., Best, N. and Spiegelhalter, D. (2000) WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Statist. Comput.*, **10**, 325–337.

Marshall, C. and Spiegelhalter, D. (2003) Approximate cross-validatory predictive checks in disease mapping models. *Statist. Med.*, **22**, 1649–1660.

Marshall, C. and Spiegelhalter, D. (2007) Identifying outliers in Bayesian hierarchical models: a simulation-based approach. *Baysn Anal.*, **2**, 1–33.

Niedomysl, T. (2006) Migration and place attractiveness. *Geografiska Regionstudier 68*. Department of Social and Economic Geography, Uppsala University, Uppsala.

Pellegrini, P. and Fotheringham, A. (2002) Modelling spatial choice: a review and synthesis in a migration context. *Progr. Hum. Geogr.*, **26**, 487–510.

Plummer, M. (2008) Penalized loss functions for Bayesian model comparison. *Biostatistics*, **9**, 523–539.

Ranjan, R. and Tobias, J. (2007) Bayesian inference for the gravity model. *J. Appl. Econometr.*, **22**, 817–838.

Song, J., Ghosh, M., Miaou, S. and Mallick, B. (2006) Bayesian multivariate spatial models for road-way traffic crash mapping. *J. Multiv. Anal.*, **97**, 246–273.

Spiegelhalter, D. J., Best, N. G., Carlin, P. B. and van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). *J. R. Statist. Soc.* B, **64**, 583–639.

Spiegelhalter, D., Thomas, A., Best, N. and Lunn, D. (2003) Tricks: advanced use of the BUGS language. In *WinBUGS User Manual, Version 1.4*. Cambridge: Medical Research Council Biostatistics Unit.

Tervo, H. (2001) Does the compositional effect explain the association between rates of in-migration and out-migration? *41st European Regional Science Association Congr., Zagreb, Aug. 29th–Sept. 1st*.

Tobler, W. (1979) Estimation of attractivities from interactions. *Environ. Planng* A, **11**, 121–127.

Wall, H. (2001) Voting with your feet in the United Kingdom: using cross-migration rates to estimate relative living standards. *Pap. Regl Sci.*, **80**, 1–23.

Wang, F. and Wall, M. (2003) Generalized common spatial factor model. *Biostatistics*, **4**, 569–582.

Williams, M. (2003) Why is Cornwall poor?: poverty and in-migration since the 1960s. *Contemp. Br. Hist.*, **17**, 55–70.

Yau, K., Wang, K. and Lee, A. (2003) Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. *Biometr. J.*, **45**, 437–452.