# Modelling Spatial Variation in the Determinants of Neighbourhood Family Migration in England with Geographically Weighted Regression

**Stephen Jivraj · Mark Brown · Nissa Finney**

**Abstract**  There have been many studies which have modelled internal migration in the UK. However, most of these have used data at geographical scales that conceal the majority of migration flows between neighbourhoods. They have also tended to use Ordinary Least Squares (OLS) regression or spatial interaction models. The latter are computationally unfeasible for migration flows between a large number of neighbourhoods. This paper uses a spatial modelling technique called Geographically Weighted Regression (GWR) to model family out-migration from neighbourhoods in England. GWR can take account of the spatial variation in the relationship between migration and its associated factors which are not accounted for using OLS. The variables included in the model are derived from theory and empirical research and include housing, socio-economic, and environmental factors. The results show that the proportion of private renting, terraced housing, worklessness and non-domestic building land space in a neighbourhood each affect out-migration at varying levels across the country. For example, the effect of worklessness on out-migration is much stronger in neighbourhoods in the South East than the North of England. Therefore, all other things held constant, a successful intervention to reduce worklessness, initiated to discourage out-migration, would have a greater effect on out-migration in neighbourhoods in the South East compared with neighbourhoods in the North.

## Introduction

Encouraging residents not to leave deprived neighbourhoods was a core theme of the Labour Government's 2001 National Strategy for Neighbourhood Renewal (NSNR).

S. Jivraj (✉) · M. Brown · N. Finney
Cathie Marsh Centre for Census and Survey Research, School of Social Sciences,
University of Manchester, Humanities Bridgeford Street, Manchester M13 9PL, UK
e-mail: stephen.jivraj@manchester.ac.uk

The flagship programme of the NSNR was New Deal for Communities (NDC) which involved 39 neighbourhood partnerships across England each with £50 million to spend over 10 years. The NDC partnerships aimed to narrow the gap in rates of worklessness, crime, health, qualifications, and poor housing and physical environment between the most deprived neighbourhoods and the national average. One major long-term vision of the programme was "not to have neighbourhoods where so many people's number one priority is to move out" (SEU 2001, p24). The NSNR was concerned that "if people start to leave an area, high turnover can destabilise communities, and if properties are not filled, desertion creates a deteriorating physical environment and more opportunity for crime" (p. 17).

The evaluation of the NDC programme suggests there was wide variation across partnerships in relation to the impact of internal migration (Beatty and Cole 2009; Pearson and Lawless 2012). NDCs with a large student population or concentration of private rented housing have seen the greatest residential turnover and most of these are located within London. Moreover, those neighbourhoods with the highest internal migration rates have made the least gain towards meeting the overarching objectives of the programme (Beatty et al. 2009). However, this is contingent upon local circumstances of the wider area rather than a result of internal migration (Pearson and Lawless 2012). Robson et al. (2008) have suggested that migration is likely to affect neighbourhoods in different ways depending on the context a particular neighbourhood is located within as well as the composition of the neighbourhood. For example, a deprived neighbourhood which is isolated within a wider affluent area might lose its higher socioeconomic residents to other neighbourhoods whereas a deprived neighbourhood within a wider depressed area is less likely to see such an adverse effect because most migration is made over short distances.

This paper aims to determine what drives out-migration from neighbourhoods and whether determinants of neighbourhood family migration vary across the country. For example: is the effect of neighbourhood worklessness on out-migration greater in the South East where the labour markets are more buoyant and enable families to move away from areas of high worklessness more easily? Is the effect of terraced housing on out-migration greater in the North and Midlands where large terraced housing estates are losing population? These types of effects across different areas are referred to as spatial non-stationarity effects (Fotheringham et al. 2002). This paper takes account of the spatial nature of migration rates of families in neighbourhoods in England by using Geographically Weighted Regression. In doing so, the paper bridges a gap between known associates of migration at the district level and unknown associates at the neighbourhood level.

Most of the previous research that has modelled internal migration rates on predictor variables has tended to summarise relationships between migration rates and explanatory variables using global statistical methods or spatial interaction models. The latter require enormous data structures which are not feasible for modelling neighbourhood migration flows (Congdon 2010; Fotheringham et al. 2000). A migration matrix between all neighbourhoods in England could create over one thousand million cases of migration flows to be modelled. Empirical studies using global linear regression have crudely taken account of the spatial structure of data using categorisations of areas within higher geographical units, for example, dummy variables for the region a smaller area is located in (Fotheringham et al.

2004). However, this imposes an a priori spatial structure on the effect of explanatory variables on the outcome which may or may not reflect reality.

This paper compares modelling results of family out-migration rates in England from standard Ordinary Least Squares (OLS) regression and Geographically Weighted Regression (GWR). The application of GWR takes account of spatial non-stationarity. Should the effect of the determinants of migration simply reflect a wider regional housing or labour market, this could be taken into accounting using an OLS model. The spatial dependency of migration and its relationship with explanatory variables might simply reflect regional or local housing markets. In this case a fixed effect in a global model would be appropriate if the local boundaries represented the same space over which the explanatory variables operated. However, GWR allows an analyst not to have to set a priori boundaries. It therefore treats space as a continuous variable rather than categorical.

The paper is structured as follows: The next section explores the determinants of out-migration using existing empirical evidence. Section three briefly describes the study area. Section four discusses the GWR modelling approach. Section five explores the data and measures used in the study. Section six presents results of comparable OLS and GWR models. Section seven discusses the implications of the substantive results for policy and theory related to internal migration.

## Determinants of Neighbourhood Out-Migration

Champion et al. (1998) provide a review of the determinants of migration and suggest that family migration at the district level in England is driven most strongly by labour market factors and environmental factors, including the differences between urban and rural areas. Fotheringham et al. (2004) have used a linear regression model to predict out-migration from district-sized areas in England and Wales. They find that for children and child-rearing-age adults, out-migration is greater from areas with more non-white residents, greater employment growth, lower housing prices, more terraced housing, fewer dwellings with no central heating, fewer vacant houses, more urban surroundings and poorer air quality.

Housing factors are likely to be important determinants of neighbourhood out-migration, especially for families. A large concentration of rented housing in a neighbourhood will mean that many of the households will be able to move quite freely, should they wish, which will lead to higher out-migration. One might expect this to have differing effect in parts of the country where housing and labour markets are more buoyant which, in turn, encourage out-migration. A higher proportion of social rented housing, however, may not be as strongly associated with out-migration because housing in nearby neighbourhoods will be relatively more expensive, which will impede families moving away from neighbourhoods with a high proportion of social housing.

The nature of the housing stock in a neighbourhood is also likely to be related to neighbourhood out-migration. Detached housing is often favoured and this is reflected by its cost, independent of house size (Adair et al. 1996; Bramley and Pawson 2002). Detached housing provides desirable features, often preferable to families, including private green space, and, therefore, they are

likely to be less willing to move away from these neighbourhoods. Terraced housing, on the other hand, is often the least desirable housing type even when it is privately owned. Areas of low housing demand include large proportions of terraced housing particularly in parts of the North and Midlands of England (Bramley and Pawson 2002). Bramley and Pawson (2002) point out that contemporary preferences have moved away from terraced developments in favour of suburban housing with gardens. However, there are neighbourhoods in large metropolitan areas, most notably London, where terraced housing is desirable and expensive (Hamnett 2009) which may result in the effect of terraced housing on out-migration varying across space.

While migration flows between neighbourhoods may be driven primarily by housing choices, the socioeconomic profile of a neighbourhood is also expected to be a factor influencing out-migration. South & Crowder (1997) suggest that high unemployment is a likely factor that will encourage out-migration as it reflects limited employment opportunities in the neighbourhood. Clearly, local labour markets extend beyond a resident's own neighbourhood but deprived neighbourhoods with high unemployment tend to be situated within wider areas of low employment opportunity (Conley and Topa 2002). Therefore, it can be anticipated that there will be a varying effect over space of high unemployment on out-migration.

The physical environment of a neighbourhood may also be linked to its pattern of migration. In general, migrants are attracted to areas with low population density with the exception of young adults who move to urban areas for education and access to high paid entry-level employment (Champion et al. 2007). There have been a number of studies which have documented the process of counterurbanisation from UK cities which some argue has been motivated by a desire to move away from densely populated urban settlements (Champion 1989; Halfacree 2008). Boyle et al. (1998) state that non-urban residential areas offer quality of life attractions, notably in terms of open space and better housing. There are a number of physical features which previous authors have identified as related to out-migration. Congdon (2010) and Li et al. (2009) find that as the proportion of green space in an area increases, so do levels of out-migration. Fotheringham et al. (2000) find that more urban districts experience greater out-migration.

## Study Area

The neighbourhoods used in this paper are Lower Super Output Areas (LSOA) in England. There are 32,482 LSOAs in England, shown in Fig. 1 by their population weighted centroids. LSOAs were designed for the reporting of neighbourhood data using 2001 Census data. They are broadly evenly sized in terms of population, containing, on average, 1,500 people (Bailey 2012). The population weighted centroids of LSOA neighbourhoods, which act as summary points of local population distribution (Martin 2002), are used as the data points in the GWR model. There are dense clusters of LSOA neighbourhood centroids, shown by the overlapping of points on the map, within the major population centres including Greater London,

Fig. 1 Population weighted centroids for Lower Super Output Area (LSOA) neighbourhoods in England

West Midlands, Greater Manchester, West Yorkshire, Merseyside, South Yorkshire, and Tyne and Wear.

## Geographically Weighted Regression Model Specification

Geographically Weighted Regression (GWR) is a technique specifically designed to analyse spatial relationships that are not accounted for by standard OLS regression (Fotheringham et al. 2002). GWR is a form of 'local' modelling where the relationship between the dependent variable and explanatory variables is allowed to vary from place to place compared with a 'global' OLS model where a single parameter estimate is produced. GWR produces local parameter estimates for each regression point that are weighted by a function of distance from every other data point. Observations that are located closer to one another are weighted more heavily than those further away on the

assumption that nearby places are more likely to be similar than those further apart.

The GWR model can be specified as follows:

$$Y^i = \beta_0(u_i, v_i) + \sum_{j,k} \beta_k^i(u_i, v_i)X_k^j + \varepsilon^i$$

Where $Y$ is the dependent variable, $X_k^j$ is a function of a range of covariates, $\beta_0$ and $\beta_k$ are the parameters to be estimated, $\varepsilon$ is the random residual term and $(u_i, v_i)$ denotes the location of the $i$th neighbourhood in geographical space. To estimate $\beta_k(u_i, v_i)$ at the location of the $i$th neighbourhood a weighted regression is used where each neighbourhood is given a weight $w_{ij}$. An important decision when fitting a GWR model is the choice of the weighting function. The weighting function is used to estimate the regression parameter for each data point, in this study a neighbourhood, in the data. The weighting ($w_{ij}$) of a data point $i$ on a regression parameter for a particular location (x) becomes smaller the further away in distance ($d_{ij}$) that data point is located from $j$. This can be referred to as a distance-decay weighting function. The weighting function is determined by the shape of the spatial kernel and the size of its bandwidth. Bandwidth refers to the threshold distance or set of neighbours beyond which the influence of one area on another is zero. Spatial kernels with a small bandwidth will have a steeper weighting function than spatial kernels with a large bandwidth (Fotheringham et al. 2002).

In practice, the results obtained through GWR tend not to be sensitive to the choice of weighting function, but they are sensitive to the determination of bandwidth. Consequently, when estimating a model it is necessary to determine the optimum bandwidth (Pineda Jaimes et al. 2010). For a fixed weighting function the bandwidth will be a distance which is the same for each neighbourhood. If an adaptive weighting function is selected the bandwidth will be a count of the number of nearest neighbourhoods to include in the weighting function (Charlton and Fotheringham 2009). The bandwidth could be applied directly to the model if it is known a priori. However, this is unlikely and in most instances a cross-validation estimate or Akaike Information Criterion (AIC) of minimum discrepancy estimation is used. The AIC estimation has the advantage of being more generally applicable, as it can be applied to linear models. It can also be used to compare whether the results from GWR present a better fit than the global model, taking both models' degrees of freedom into consideration (Fotheringham et al. 2002).

The optimum weighting function and bandwidth can be determined by examining the model diagnostics, including AIC and R-squared values. An adaptive weighting function and AIC bandwidth estimation are found to be preferable for the results of the GWR model in this study. The GWR algorithm produces an adaptive bandwidth parameter of 233 nearest neighbourhoods to determine the regression parameters for each local regression point in the model for out-migration. This means that the regression parameters for each neighbourhood are estimated using a distance decay weighting function from 0.7 % or 233 of the surrounding neighbourhoods in England. The effect of a larger bandwidth would be smoother regression results across space that would appear to be closer to the estimates from a global OLS

regression. This is because each regression equation produced for each neighbour-hood would take account of a greater number of other neighbourhoods in the study area. The models in this paper are fitted using ArcGIS 10.0.

## Data and Measures

This paper uses data from the School Census in England collected during 2002–2007 to calculate neighbourhood family out-migration. The School Census records details of all state school pupils in England and is derived from an electronic administrative form completed by each school (Simpson et al. 2010). It is collated in England by the Department for Education (DfE) through Local Education Authorities (LEA). The School Census includes information about individual compulsory school age pupils including their home postcode. The inclusion of a unique pupil number in the School Census allows data to be matched between academic years and a change in the postcode of a pupil's matched record can indicate whether they have moved during a given time frame. The data can provide more up-to-date migration information than the decennial census of population which is the only other national dataset that can measure internal migration accurately at the neighbourhood level in England. Although migration of pupils is measured it is considered to reflect family migration and, to a lesser extent, the migration of all people apart from young adults. Jivraj (2011) has shown that the prevalence and pattern of school pupil migration between and within districts is very similar for all other age groups with the exception of those aged 16–29 when compared with 2001 Census migration data.

The School Census can be used to calculate out-migration rates for neighbour-hoods, which are used as the dependent variable in the local (GWR) and global (OLS) models. The models control for in-migration rates to take account of the strong association between out and in-migration and that some explanatory factors are related to gross migration rather than out-migration independently. The aim of the paper is to determine what drives outward movement from neighbourhoods and not the net effect of migration on population change. Out-migration rates are calculated using the average number of pupils moving out of a neighbourhood between 2002–03 and 2006–07 divided by the average pupil population between 2002 and 2006, expressed as a percentage. The data are pooled over time to provide a more accurate estimate of out-migration in neighbourhoods.

Four variables are used to explain variation in out-migration rates in neighbour-hoods in England (all expressed as a percentage for the neighbourhood): private-rented housing, terraced housing, worklessness and land space covered by non-domestic buildings. These variables are the most powerful predictors of out migration rates when entered into a global OLS model, including a range of factors known to be associated with neighbourhood migration discussed in Section 2 (results available from the authors). The four variables are entered using a stepwise approach in order of predictive power until the GWR model selection criteria suggests poorer model fit, or the model does not converge due to multicollinearity (Wheeler and Tiefelsdorf 2005).

The percentage of households living in private rented housing and terraced housing are measured from the 2001 Census. The percentage of working age

population claiming an out-of-work benefit (worklessness) is measured from Department for Work and Pensions Client Group data for 2001 (DWP n.d.). Worklessness is preferred to a census measure of unemployment because the latter excludes a large share of the working-age population who are not in work but are claiming an out-of-work benefit. The percentage of land hectares covered by non-domestic buildings (i.e. commercial premises) is measured from the Generalised Land Use Database produced by the Department for Communities and Local Government in 2001 (DCLG 2007).

It is expected that the degree of variation in the effect of each predictor varies across neighbourhoods in England. The effect of worklessness might operate at larger spatial scales providing smoother results from the GWR model because neighbourhoods are likely to be similar within sub-regional labour markets. However, the effects of the other variables are expected to be more localised. The effect for the proportion of non-domestic building land space in the neighbourhood on out-migration rates is expected to be separated zonally around urban centres. Moreover, the effect of private renting and terraced housing might be even more localised because neighbourhoods located close to one another might have differing housing profiles. Nonetheless, the assumption remains for all variables that neighbourhoods located close to one another will have more in common than neighbourhoods located further apart.

## Results of Geographically Weighted Regression Model

The results from the local GWR model and its global OLS equivalent model are shown in Table 1. They include the R-squared and AICc[1] values from the GWR and OLS models as well as parameter estimates. The GWR model compared with the OLS model improves the amount of variance explained by the predictor variables in the out-migration rates of English neighbourhoods by raising the adjusted R-squared value from 0.599 to 0.667. A preferred indication of goodness of fit is the AICc value, which if lower in a nested model indicates better model fit (Charlton and Fotheringham 2009). The AICc is 4,467 lower in the local model indicating better model fit.

This suggests that there is a varying effect of predictor variables across space that has not been accounted for in the OLS model and that not taking account of this variation could lead to a wrongly specified model. This is because it is assumed that the relationship between the explanatory variables (private rented housing, terraced housing, worklessness and land space covered by non-domestic buildings) and the dependent variable (out-migration rate) is constant over space in an OLS model. The interquartile range of the parameter estimates from the GWR model provides a rough indication of the spatial non-stationarity for each variable when compared with the standard error of the OLS global mean. If the interquartile range is greater than twice the standard error derived from the OLS estimate, the relationship appears to exhibit a relatively high degree of spatial non-stationarity i.e. the effect of the predictor varies across neighbourhoods in England (Nelson 2008). This is because 50 % of the local parameter values lie within the interquartile range and approximately 68 % of the

---

[1] AICc is adjusted for sample size.

**Table 1** Comparison of parameter estimates from global (OLS) and local (GWR) models of LSOA neighbourhood out-migration rates

| | Estimate | Standard error × 2 | p-value | VIF |
|---|---|---|---|---|
| a) Global model (Ordinary Least Squares regression) | | | | |
| Intercept | 1.137 | 0.095 | 0.000 | – |
| Private renting | 0.136 | 0.004 | 0.000 | 1.320 |
| Terraced housing | 0.030 | 0.002 | 0.000 | 1.310 |
| Worklessness | 0.086 | 0.003 | 0.000 | 1.290 |
| Non-domestic buildings | 0.097 | 0.007 | 0.000 | 1.180 |
| In-migration | 0.503 | 0.009 | 0.000 | 1.160 |
| Adjusted R-square | 0.599 | | | |
| AICc | 154164 | | | |

| | Mean | Min | Lower quartile | Median | Upper quartile | Max | IQ range[a] |
|---|---|---|---|---|---|---|---|
| b) Local model (Geographically Weighted Regression) | | | | | | | |
| Intercept | 1.496 | −5.119 | 0.641 | 1.508 | 2.357 | 10.954 | 1.716 |
| Private renting | 0.148 | −0.073 | 0.093 | 0.142 | 0.200 | 0.503 | 0.106 |
| Terraced housing | 0.031 | −0.076 | 0.016 | 0.031 | 0.046 | 0.152 | 0.030 |
| Worklessness | 0.118 | −0.134 | 0.076 | 0.110 | 0.154 | 0.383 | 0.078 |
| Non-domestic buildings | 0.103 | −0.296 | 0.036 | 0.092 | 0.165 | 0.671 | 0.129 |
| In-migration | 0.431 | −0.052 | 0.337 | 0.428 | 0.520 | 0.938 | 0.183 |
| Adjusted R-square | 0.667 | | | | | | |
| AICc | 149697 | | | | | | |

Each explanatory variable is expressed as a percentage

[a] Interquartile range.

values in the normal distribution will fall within +/− 1 standard deviation of the mean OLS coefficient, and therefore, if the former is greater, this suggests the relationship might be non-stationary.

The estimates from the GWR model for each of the predictor variables indicate the presence of a non-stationary effect on the out-migration rates across neighbourhoods in England. The global OLS estimate for private renting is positive which suggests that in neighbourhoods with a higher percentage of private renting there will be a higher out-migration rate, when controlling for the other variables in the model. More specifically, a one percentage point increase in private renting in a neighbourhood is associated with a 0.14 percentage point increase in the out-migration rate. This is not surprising as people living in private rented accommodation tend to be highly mobile even when controlling for other individual characteristics (Champion et al. 1998; Finney and Simpson 2008). This is because they have more residential flexibility than social renting tenants and owner occupiers.

The local estimates for the effect of the percentage of private renting on the out-migration rate in a neighbourhood ranged from −0.073 to 0.503. The estimate is greater than 0.2 in three-quarters of neighbourhoods, shown by the value of the lower quartile. This suggests that in a small number of neighbourhoods there is a small negative effect of private renting on the out-migration rate (i.e. higher proportions of private renting are associated with lower out-migration rates) whereas in most neighbourhoods there is a positive effect. Therefore, if a policy maker was advised to reduce the number of private rented houses in a neighbourhood to reduce the out-migration rate they would find that this intervention would have a different result in different parts of the country. For example, if the two neighbourhoods at the extremes of the GWR model estimate range had 10 % of households living in private rented housing the predicted difference between the out-migration rates would be 6 % when all other effects are held constant. Although this is an unlikely scenario because planners will not intervene in the private rental market to lower out-migration, it highlights the point that the effect of private renting on out-migration varies over space.

Figures 2, 3, 4, and 5 show the distribution of each local estimate from the GWR model using a natural break classification in ArcGIS 10.0 (ESRI 2008). Figure 2 displays the patterns of spatial variation in the effect of private renting on out-migration from the local model. Neighbourhoods with an insignificant local estimate are highlighted in the lightest shade. The positive effect of private renting on out-migration is strongest, shown by the darkest shade, in parts of East London, Essex and Kent. It is not clear why a higher percentage of private rented households would be more strongly associated with higher out-migration rates in this selection of neighbourhoods. This requires localised analysis which is outside the bounds of this paper.

The global effect of the percentage of terraced housing in a neighbourhood is also positive which suggests the greater its value the higher the out-migration rate in a neighbourhood. A one percentage point increase in the terraced housing in a neighbourhood is associated with a 0.03 percentage point increase in the out-migration rate. It has been suggested that contemporary preferences have moved away from terraced housing in favour of suburban types of housing with gardens, which means this effect is not unexpected (Bramley and Pawson 2002). However, in some neighbourhoods terraced housing can be expensive and desirable (Hamnett 2009).
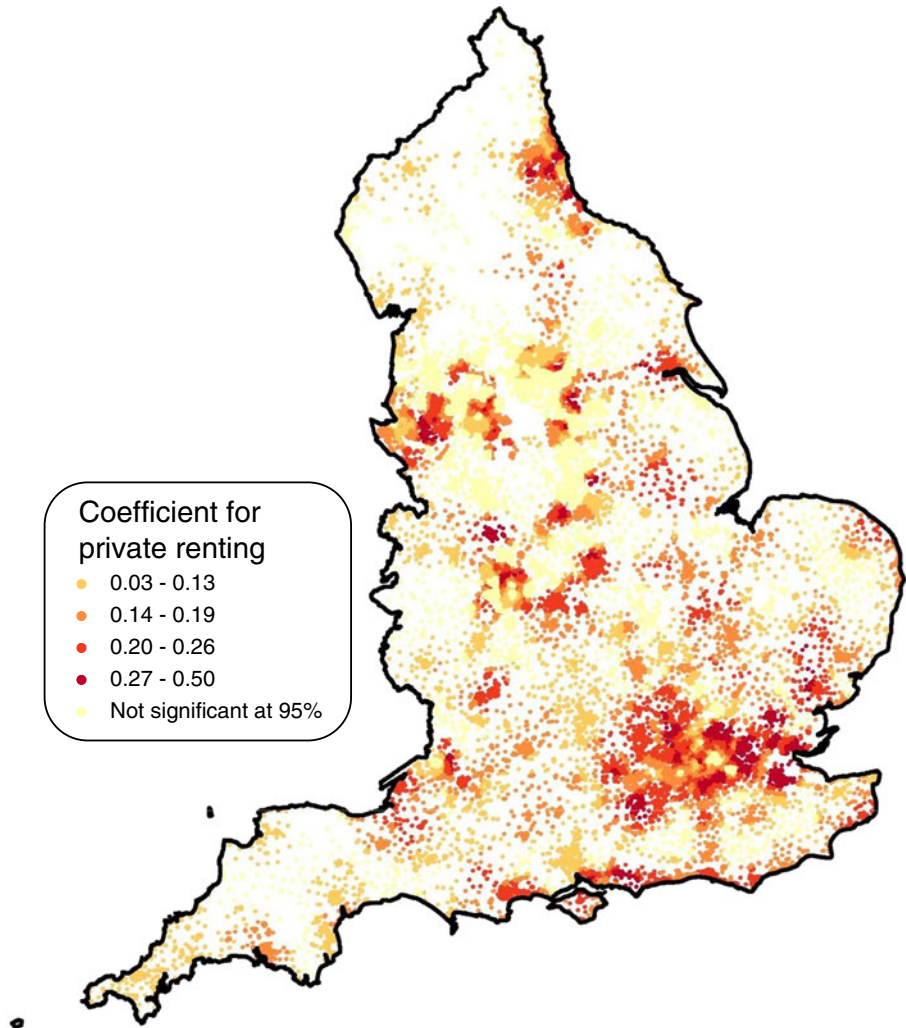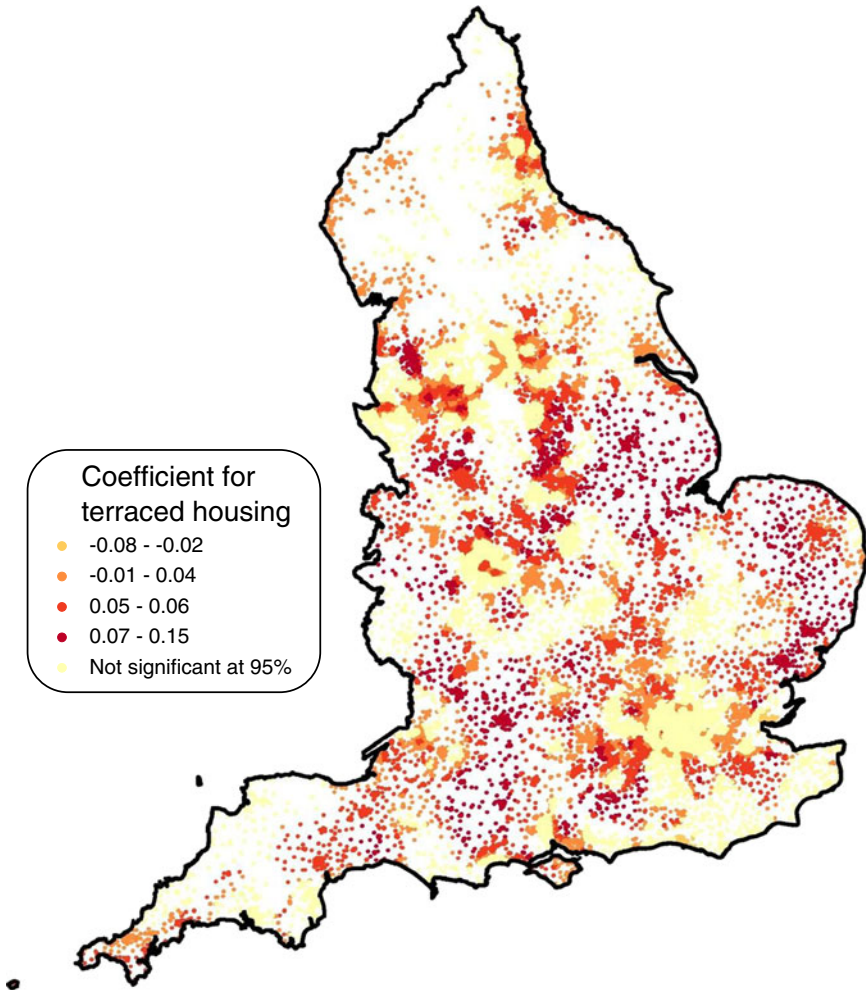
Note: darkest shade indicates neighbourhoods with the strongest positive effect (i.e. higher proportion of private renting is associated with a higher out-migration rate for a neighbourhood).

**Fig. 2** Spatial variation in the effect of private renting on out-migration rate: GWR coefficients for LSOA neighbourhoods in England

Fotheringham et al. (2002) suggest that within London terraced houses are viewed differently across the capital which has a non-stationary effect on house prices.

The parameter estimates from the local model show that the effect of the percentage of terraced housing in a neighbourhood on the out-migration rate varies from −0.076 to 0.152, with 50 % of the estimates between 0.016 and 0.046. This suggests the effect of terraced housing on out-migration is very weak. Figure 3 shows that the positive effect of terraced housing on out-migration is strongest in neighbourhoods across the North and Midlands of England where demand for housing is low. This is seen, in particular, in neighbourhoods located alongside the M1 motorway from Loughborough in the East

Note: darkest shade indicates neighbourhoods with the strongest positive effect (i.e. higher proportion of terraced housing is associated with a higher out-migration rate for a neighbourhood).

**Fig. 3** Spatial variation in the effect of terraced housing on out-migration rate: GWR coefficients for LSOA neighbourhoods in England

Midlands to Rotherham in South Yorkshire. Bramley and Pawson (2002) have identified this part of the country as containing former industrial and mining areas which include neighbourhoods of low housing demand where terraced housing is prevalent and of poor quality by modern standards. In contrast, the local coefficients for the effect of terraced housing show that in the majority of neighbourhoods in the South of England, particularly along the south coast where housing demand is high, the effect of terraced housing is marginal (positive or negative) or insignificant. This suggests that families might be more willing to stay in neighbourhoods with a high proportion of terraced housing in neighbourhoods where housing demand is high.

The effect of worklessness on out-migration is positive when controlling for other variables in the global model. A one percentage point increase in the worklessness in a

neighbourhood is associated with a 0.086 percentage point increase in the out-migration rate. Therefore, a higher percentage of people claiming an out of work benefit in the neighbourhood is related to greater out-migration. This conforms to migration theory that suggests people will move away from areas where economic opportunity is low (Champion et al. 1998). Nonetheless, empirical studies that have tested this effect have not always shown consistent results. Some find that a higher percentage of people out of work can be associated with lower migration from the area (Fotheringham et al. 2004). This is qualified by the assertion that worklessness or unemployment is associated with greater economic deprivation which may limit individuals' ability to move. One reason for this conflicting finding across previous studies might be a result of the non-stationary effect of worklessness on migration. For example, its effect might differ across broad labour markets which are more and less flexible.

The local parameter estimates for the effect of worklessness on out-migration rates vary from −0.134 to 0.383. The vast majority of coefficients are above 0.076 shown by the value of the lower quartile. This means that the effect is mildly negative in a small number of areas and fairly strongly positive in other areas. Figure 4 shows that the effect is strongest in the South East of England, particularly on the outskirts of Greater London. The effect is much less strong in neighbourhoods located in the North of England, including Merseyside and Greater Manchester. In isolated parts of Central London the effect is negative. These findings may reflect the ability of some families to move away from neighbourhoods of high worklessness more freely when the local labour market conditions are more buoyant. This is likely to be the case in the South East rather than the North of England and may confirm a hypothesis that higher levels of worklessness can both foster and hinder out-migration.

The variable for the percentage of land space that is covered by non-domestic buildings is intended to give an indication of the urbanness of the neighbourhood. It provides a measure of the built-up environment rather than the residential population density. It has been suggested that people favour suburban areas which may not have low population density but are not considered as urban (Boyle 1994). The global parameter estimate for the effect of the percentage of non-domestic building land space in the neighbourhood on the out-migration rate is positive. A one percentage point increase in non-domestic building land space in a neighbourhood is associated with a 0.097 percentage point increase in the out-migration rate. This suggests that, in general, families are moving out of neighbourhoods with a higher proportion of non-domestic land use. This should be interpreted as a small effect for most areas because in the majority of neighbourhoods the percentage of land space covered by non-domestic buildings is less than 2 %.

The local effect of non-domestic building land space on the out-migration rate varies considerably across neighbourhoods in England. In the GWR model, this parameter has the largest interquartile range for its local estimates suggesting a stronger non-stationary effect than the other explanatory variables. Although the effect is negative in some neighbourhoods, it is positive in three-quarters of neighbourhoods. Figure 5 shows that the effect is mainly insignificant in the largest urban city-regions in England where the proportion of non-domestic building land space is greatest, for example, Greater London, West Midlands, Greater Manchester, Merseyside, West Yorkshire and Tyne and Wear. This suggests that the proportion of non-domestic building land space does not motivate out-migration from these large

urban centres. This might reflect the fact that non-domestic building land space is much more likely to include amenities in urban areas, including shopping centres.

The GWR model shows that families are more likely to move away from concentrations of non-domestic land space in less urban areas where non-domestic building land space is more likely to include industrial sites. A strong positive effect of non-domestic building land space is found in neighbourhoods located in selected town and rural districts. This suggests families are more selective about this characteristic in these areas. It might reflect the fact that if a family has chosen to live in a more rural area they will not want to compromise on large proportions of non-domestic building land space in the neighbourhood. In neighbourhoods located in Kent,
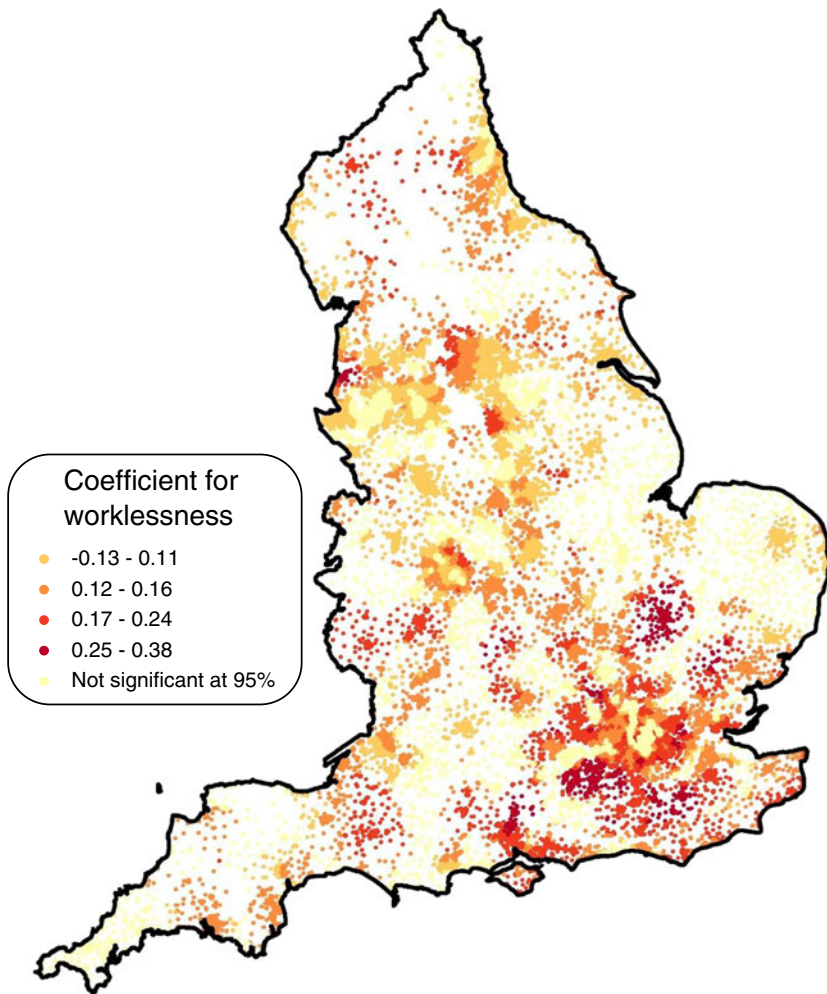


Note: darkest shade indicates neighbourhoods with the strongest positive effect (i.e. higher proportion of worklessness is associated with a higher out-migration rate for a neighbourhood).

**Fig. 4** Spatial variation in the effect of worklessness on out-migration rate: GWR coefficients for LSOA neighbourhoods in England

Note: darkest shade indicates neighbourhoods with the strongest positive effect (i.e. higher proportion of non-domestic building space is associated with a higher out-migration rate for a neighbourhood).
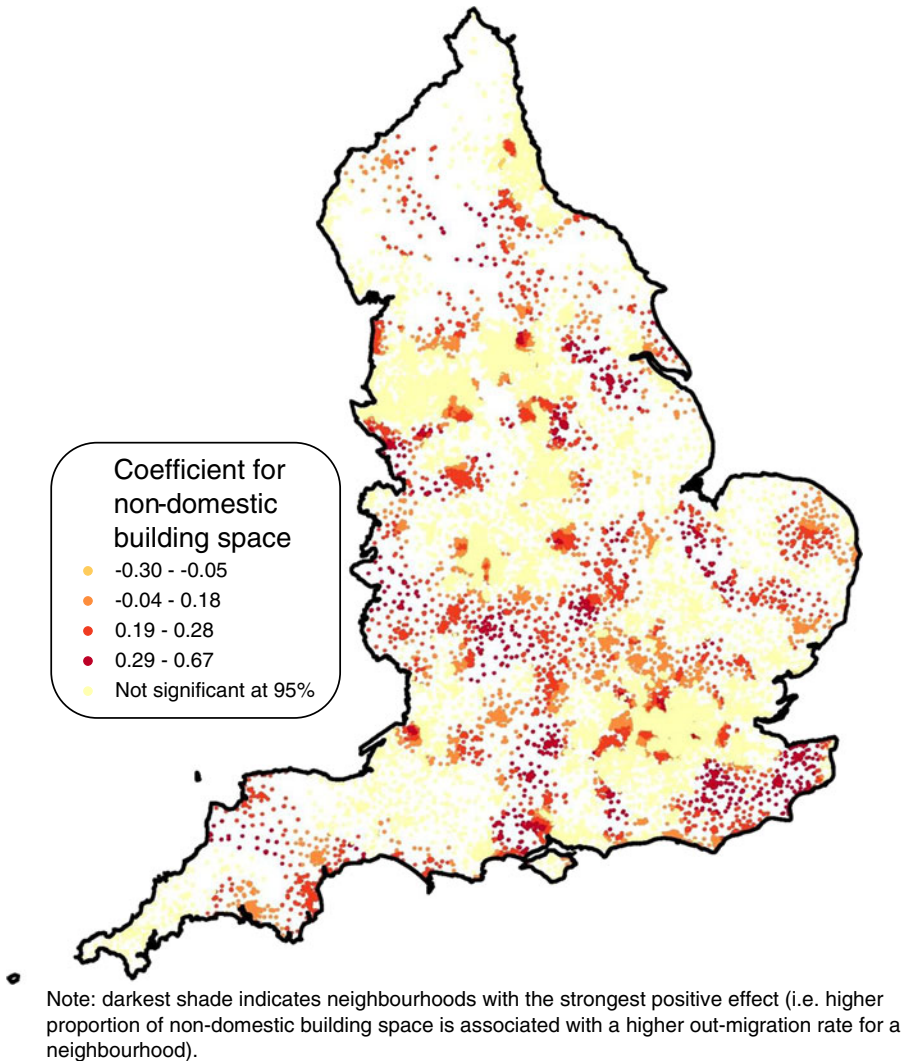
**Fig. 5** Spatial variation in the effect of non-domestic building space on out-migration rate: GWR coefficients for LSOA neighbourhoods in England

Hampshire, Devon and the Cotswolds, for example, a high percentage of non-domestic building land space is associated with higher out-migration when controlling for other variables in the model.

The GWR model can produce local R-squared values as well as local parameter estimates. These are shown in Fig. 6 for each neighbourhood in England. The R-squared values provide a measure of the amount of variance explained in the outcome variable for the regression equation fitted for each neighbourhood. Low values of the R-squared statistic are clustered in Central London. This means that the explanatory variables explain the least amount of variation in neighbourhood out-migration rates in the capital. The low R-squared values might reflect the unique housing and labour
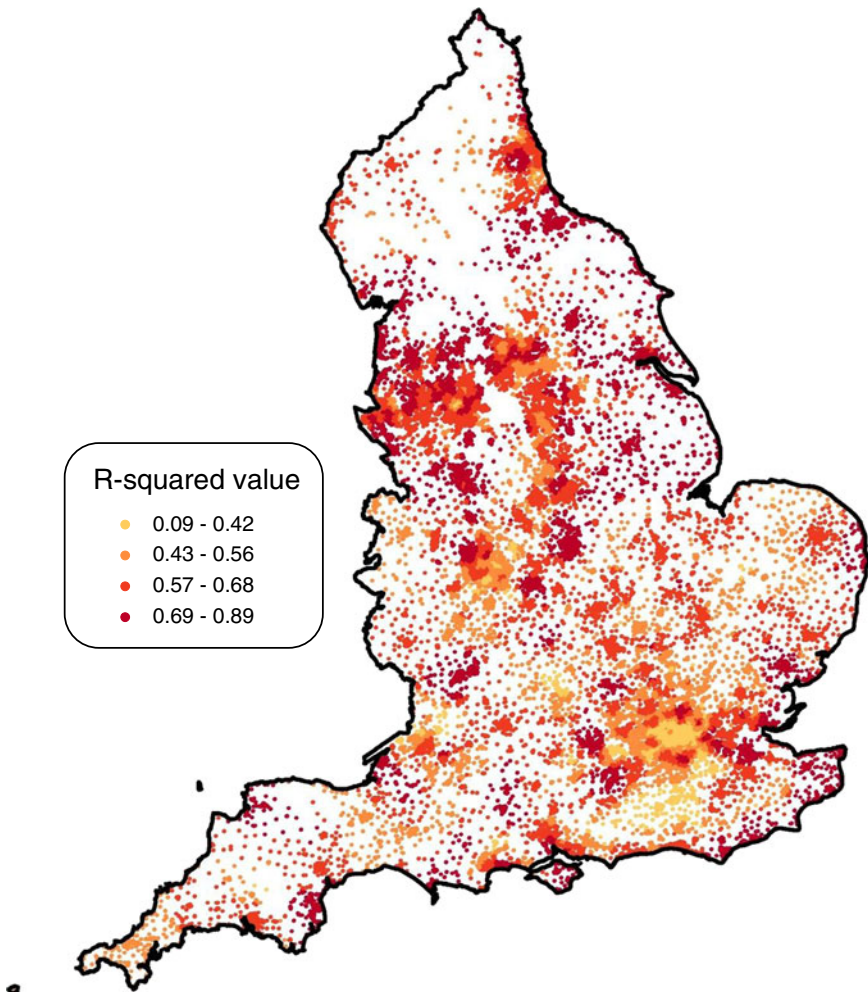
Note: darkest shade indicates greatest amount of variance in out-migration rate explained by explanatory variables for a neighbourhood.

**Fig. 6** Spatial variation in GWR model fit: local R-squared values for out-migration rate for LSOA neighbourhoods in England

markets that exist in London. Average house prices and earnings tend to be higher in the capital than other parts of the country (Duranton and Monastiriotis 2002; Holly et al. 2011). This means that the average family will not be able to be as selective regarding the characteristics of the area they reside in compared with neighbourhoods outside London where housing costs are considerably lower. For example, a family might choose to live in a neighbourhood in London with a high level of private rented housing, terraced housing, worklessness and non-domestic building land space so they can live in London. However, in neighbourhoods in Kent and Lancashire, for example, families appear to be more selective with regard to private rented housing, terraced housing, worklessness and non-domestic building land space in their area, shown by the higher R-squared values in these areas in the local model.

## Discussion

This paper has explored the spatial variation in the effect of private renting, terraced housing, worklessness and non-domestic building land space on out-migration rates of families from neighbourhoods in England using a standard Ordinary Least Squares regression model and a Geographically Weighted Regression model. The GWR results show that the effect of the percentage of private renting, terraced housing, worklessness and non-domestic building land space on out-migration rates all vary across neighbourhoods in England. This indicates how the standard linear regression model can be wrongly specified, giving misleading conclusions in some areas. If standard global models were used as the basis for neighbourhood policies, these might not have the desired effect in all neighbourhoods across the country. For example, in neighbourhoods in the South East compared with other areas the level of people out of work and claiming benefits is more strongly associated with out-migration. This is compared with parts of the North of England where the effect is positive but weaker. Therefore, all other variables held constant, a successful intervention to reduce the worklessness in a neighbourhood which was initiated to discourage out-migration will have a greater effect in neighbourhoods in the South East compared with neighbourhoods in the North. The spatially varying effect of worklessness on out-migration might reflect the higher levels of economic dynamism in the South East where labour markets are more buoyant (Boheim and Taylor 2002; Faggian and McCann 2009; McCormick 1997).

There are a number of limitations to the current analysis. Geographically Weighted Regression is described as an exploratory technique for identifying spatially varying relationships between outcome and explanatory variables and its parametric qualities as a statistical method have been questioned (Wheeler and Tiefelsdorf 2005; Wheeler 2007). The results are highly dependent on the selection of the bandwidth for the weighting function. The bandwidth is the distance from a regression point where other data points will affect that estimate. An adaptive bandwidth was chosen for the current analysis because some neighbourhoods are spatially concentrated (see Fig. 1) whereas others are fairly isolated in England. The larger the bandwidth the smoother the results will be. In the current analysis the weighting function for each local estimate is based on 233 or 0.7 % of the nearest neighbours.

Moreover, the analysis in this paper is based on aggregate data and therefore it does not enable inference to individual migratory flows. The effect of worklessness on out-migration might reflect people who are already employed or who are returning to employment moving away from neighbourhoods where a high percentage of people are out of work rather than an effect of people moving to find work. This might explain why there is the spatially varying relationship between these variables in the South East compared with neighbourhoods in the North of England. There is a higher percentage of people in work in the South East compared with the North.

To determine whether individuals with particular characteristics are more likely to move and whether this varies across the country one would require individual level geocoded data. Individual geocoded data is available from the School Census. However, there are few contextual variables known to be associated with migration that could be used to predict whether a family moves or not. The specificity of the School Census migration data should also be acknowledged. Jivraj (2011) has shown

that the propensity and pattern of migration is similar for school age children and the general population with the exception of young adults. However, young adults account for a disproportionate proportion of migration relative to the number of people they account for in the population. Therefore, the findings in this paper would change if the out-migration rates for the entire population were used.

Another limitation of the current analysis is that it treats England as a closed migration system and the estimates for neighbourhoods that border Wales and Scotland are only influenced by other neighbourhoods in England. In the UK, national boundaries might represent a real barrier to migration particularly within the social housing sector. However, within the private housing market there is little that would restrict a family moving across these boundaries. The inclusion of data for movement between neighbourhoods in Wales and Scotland might alter the findings in the current analysis especially for neighbourhoods on the borders.

In spite of these limitations, this paper has shown how GWR can provide a more nuanced empirical test of the correlates of internal migration. The analysis in this paper adds to the small but growing migration literature which uses the technique, and other forms of local modelling. The presentation of the results through maps, which is possible when results are generated for each neighbourhood, is an advantage for disseminating findings to non-technical audiences including policy makers. Figures 2, 3, 4, 5, and 6 provide a clear representation of where certain factors are more important drivers of out-migration of families. Although it is not within the bounds of this paper to explore every local variation, the detailed results might interest local policy makers by helping them to explain why their area is different from others, and thereby to plan more effective local interventions.

# References

Adair, A. S., Berry, J. N., & McGreal, W. S. (1996). Hedonic modelling, housing submarkets and residential valuation. *Journal of Property Research, 13*, 67–83.

Bailey, N. (2012). How spatial segregation changes over time: sorting out the sorting processes. *Environment and Planning A, 44*, 705–722.

Beatty, C., & Cole, I. (2009). Stability, residential satisfaction and the dynamics of change: evidence from the New Deal for Communities programme in England. *Journal of Urban Regeneration and Renewal, 3*, 141–153.

Beatty, C., Lawless, P., Pearson, S., & Wilson, I. (2009). *Residential mobility and outcome change in deprived areas: Evidence from the New Deal for Communities Programme*. London: Department for Communities and Local Government.

Boheim, R., & Taylor, M. P. (2002). Tied down or room to move? Investigating the relationships between housing tenure, employment status and residential mobility in Britain. *Scottish Journal of Political Economy, 49*, 369–392.

Boyle, P. (1994). Metropolitan out-migration in England and Wales, 1980–81. *Urban Studies, 31*, 1707–1722.

Boyle, P., Halfacree, K., & Robinson, V. (1998). *Exploring contemporary migration*. Harlow: Longman.

Bramley, G., & Pawson, H. (2002). Low demand for housing: Incidence, causes and UK national policy implications. *Urban Studies, 39*, 393–422.

Champion, A. G. (Ed.). (1989). *Counterurbanization: the changing pace and nature of population deconcentration*. London: Edward Arnold.

Champion, T., Fotheringham, A. S., Rees, P., Boyle, P., & Stillwell, J. (1998). *The determinants of migration flows in England: a review of existing data and evidence a report prepared for the department of the environment, transport and the regions*. Newcastle upon Tyne: Department of Geography, University of Newcastle upon Tyne.

Champion, T., Coombes, M., Raybould, S., & Wymer, C. (2007). *Migration and socioeconomic change: a 2001 Census analysis of Britain's larger cities*. Bristol: Policy Press.

Charlton, M., & Fotheringham, A. S. (2009). Geographically weighted regression: a tutorial on using GWR in ArcGIS 9.3 [online] http://ncg.nuim.ie/ncg/GWR/GWR_Tutorial.pdf (accessed 23 May 2012).

Congdon, P. (2010). Random-effects models for migration attractivity and retentivity: a Bayesian methodology. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 173*, 755–774.

Conley, T. G., & Topa, G. (2002). Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics, 17*, 303–327.

DCLG (2007). Generalised land use database statistics for England 2005. Department for Communities and 563 Local Government, [online] http://www.communities.gov.uk/documents/planningandbuilding/pdf/564154941.pdf (accessed 23 May 2012).

Duranton, G., & Monastiriotis, V. (2002). Mind the gaps: the evolution of regional earnings inequalities in the UK, 1982–1997. *Journal of Regional Science, 42*, 219–256.

DWP (n.d.). Client group analysis statistics. Department for Work and Pensions, [online] http://research.dwp.gov.uk/asd/index.php?page=cga (accessed 23 May 2012).

ESRI (2008). Natural breaks (Jenks) [online] http://webhelp.esri.com/arcgiSDEsktop/9.3/index.cfm?TopicName=Natural_breaks_%28Jenks%29 (accessed 24 June 2011).

Faggian, A., & McCann, P. (2009). Human capital, graduate migration and innovation in British regions. *Cambridge Journal of Economics, 33*, 317–333.

Finney, N., & Simpson, L. (2008). Internal migration and ethnic groups: evidence for Britain from the 2001 Census. *Population Space and Place, 14*, 63–83.

Fotheringham, A. S., Champion, T., Wymer, C., & Coombes, M. (2000). Measuring destination attractivity: a migration example. *International Journal of Population Geography, 6*, 391–421.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester: Wiley.

Fotheringham, A. S., Rees, P., Champion, T., Kalogirou, S., & Tremayne, A. R. (2004). The development of a migration model for England and Wales: overview and modelling out-migration. *Environment and Planning A, 36*, 1633–1672.

Halfacree, K. (2008). To revitalise counterurbanisation research? Recognising an international and fuller picture. *Population, Space and Place, 14*, 479–495.

Hamnett, C. (2009). Spatially displaced demand and the changing geography of house prices in London, 1995–2006. *Housing Studies, 24*, 301–320.

Holly, S., Pesaran, M. H., & Yamagata, T. (2011). The spatial and temporal diffusion of house prices in the UK. *Journal of Urban Economics, 69*, 2–23.

Jivraj, S. (2011). *The effect of internal migration on the socioeconomic composition of neighbourhoods in England*, Ph.D. dissertation. University of Manchester.

Li, W. J., Holm, E., & Lindgren, U. (2009). Attractive vicinities. *Population Space and Place, 15*, 1–18.

Martin, D. (2002). Geography for the 2001 Census in England and Wales. *Population Trends, 108*, 7–15.

McCormick, B. (1997). Regional unemployment and labour mobility in the UK. *European Economic Review, 41*, 581–589.

Nelson, P. B. (2008). Life-course influences on nonearnings income migration in the United States. *Environment and Planning A, 40*, 2149–2168.

Pearson, S., & Lawless, P. (2012). Population mobility in regeneration areas: trends, drivers, and implications; evidence from England's New Deal for Communities Programme. *Environment and Planning A, 44*, 2023–2039.

Pineda Jaimes, N. B., Bosque Sendra, J., Gómez Delgado, M., & Franco Plata, R. (2010). Exploring the driving forces behind deforestation in the state of Mexico (Mexico) using geographically weighted regression. *Applied Geography, 30*, 576–591.

Robson, B., Lymperopoulou, K., & Rae, A. (2008). People on the move: exploring the functional roles of deprived neighbourhoods. *Environment and Planning A, 40*, 2693–2714.

SEU. (2001). *A new commitment to neighbourhood renewal: National Strategy Action Plan Social Exclusion Unit*. London: Cabinet Office.

Simpson, L., Marquis, N., & Jivraj, S. (2010). International and internal migration measured from the School Census in England. *Population Trends, 140*, 106–124.

South, S. J., & Crowder, K. D. (1997). Residential mobility between cities and suburbs: race, suburbanization, and back-to-the-city moves. *Demography, 34*, 525–538.

Wheeler, D. C. (2007). Diagnostic tools and a remedial method for collinearity in geographically weighted regression. *Environment and Planning A, 39*, 2464–2481.

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems, 7*, 161–187.