Measuring the scales of segregation: looking at the residential separation of White British and other school children in England using a multilevel index of dissimilarity

Richard Harris,

School of Geographical Sciences,

University of Bristol,

University Road,

Bristol,

UK.

rich.harris@bris.ac.uk

Measuring the scales of segregation: looking at the residential separation of White British and other school children in England using a multilevel index of dissimilarity

Abstract

Within the segregation literature there has been a movement away from measuring ethnic segregation at a single scale, using traditional indices, to instead treating segregation as a multiscale phenomenon about which measurement at a range of scales will shed knowledge. Amongst the contributions, several authors have promoted multilevel modelling as a way of looking at segregation at multiple scales of a geographical hierarchy, estimating the micro-, meso- and macro effects of segregation simultaneously. This paper takes the approach forward by outlining a multilevel index of dissimilarity that combines the advantages of using a widely-understood index with a means to identify scale effects in a way that is computationally fast to estimate and uses freely available software to do so. To demonstrate the method, a case study is made looking at the residential separation of White British pupils from six other ethnic groups in England in 2011. It examines a claim made by the Casey Review into opportunity and integration that school children are more residentially segregated than the population-at-large. The results suggest that school children were indeed more residentially divided but comparison with earlier data and the general uplift in the scales at which patterns of segregation are evident suggest a trend of decreasing segregation overall and the spreading-out of 'minority' groups.

Key words: index of dissimilarity, segregation, ethnicity, multilevel, multiscale, Casey Review

Measuring the scales of segregation: looking at the residential separation of White British and other school children in England using a multilevel index of dissimilarity

1. Introduction

This paper contributes to the literature on multi-scale measures of segregation, providing methodological innovation in the form of a multilevel index of dissimilarity that is used to look at the residential segregation of White British school pupils from other ethnic groups in England. The recently published Casey Review (Casey 2016) has rekindled the idea that social and ethnic segregation are increasing in some parts of Britain, arguing that segregation reduces how often people from different ethno-cultural groups meet and undermines social integration. It gives particular attention to the young, stating that "the school age population is even more segregated when compared to residential patterns of living" (p.11). This paper will examine that claim.

Studies of ethnic (and socio-economic) segregation continue to debate measurement issues. This is important because, as Simpson and Peach (2009, 1379) note, the act of measurement is not neutral but "depends upon what the measurer conceives segregation to be." It is also enduring because of the political and media attention that is given to ethnic segregation in the UK, as in other countries. In response to the Casey Review, the free national *Metro* newspaper published the headline, "Diverse yet divided: UK is growing apart",[1] whilst the tabloid *The Sun* wrote in characteristically measured tones, "GHETTO BLASTER

---

[1] http://metro.co.uk/2016/12/05/diverse-yet-divided-uk-is-growing-apart-casey-report-finds-6303352/

Mass immigration to Britain has changed it beyond recognition and turned communities into ghettos, reveals damning report."[2] Earlier in the same year, a Member of Parliament expressed a belief that "Britain has become more ethnically segregated with widening cracks in our communities" (The i newspaper, May 23, 2016, reporting on a speech by Chuka Umunna MP).

An important role of measurement is to challenge misperception with empirical evidence: in fact, the various studies that have compared 2011 to 2001 UK Census data have all revealed that ethnic segregation is decreasing, not increasing in the UK, as 'minority' groups spread out geographically into what are becoming increasingly mixed neighbourhoods (Johnston et al. 2013, 2014, Harris 2014, Catney 2016a, 2016b). This is not to deny the persistence of ethnic separations in the UK (or elsewhere: Nightingale 2012, Logan 2013), or to suggest that neighbourhoods all are diverse and mixed. However, it is to acknowledge that the scale of segregation in the UK is changing both numerically (decreasing in its measured quantity) and geographically (minority groups are less spatially concentrated in distinct places than they have been previously).

A recent development in the measurement of segregation is to treat it as a multiscale phenomenon, the outcome of choices, processes, structural constraints and behaviours that operate simultaneously at multiple scales – of what Galster and Shevky (2017) theorise as the differential effects of the spatial opportunity structure that create the spatial foundations of inequality. An attraction of this approach is it allows the dominant scales of segregation to be

[2] https://www.thesun.co.uk/news/2327147/british-towns-have-changed-beyond-recognition-as-mass-immigration-turns-communities-into-ghettos-as-report-raps-governments-for-failing-to-deal-with-crisis/

identified: is it something that occurs at a micro-scale, a meso-scale, a macro-scale or some combination of all three? (Manley et al. 2015).

This paper takes the idea and applies it to the well-known index of dissimilarity (Jahn et al. 1947, Duncan & Duncan, 1955). It shows how the index can be formulated within a regression framework to measure both spatial variation and scale effects thereby capturing the two principal dimensions of segregation, unevenness and clustering. A study is made of the residential segregation of White British from school children of other ethnicities in England in 2011, of the scales at which those separations occur, of how they compare with those for the whole Census population in the same year, and of whether they have changed over the decade prior. The focus on the White British is because when politicians and the media talk of ethnic segregation in the UK, they usually mean between the White British and other ethnic groups. The paper questions whether the impression of increasing segregation fostered by the Casey Review is correct.

2. Measuring segregation as a spatial and multi-scale phenomenon

Reardon et al. (2008) write that the study of racial segregation has three main analytical aims: to investigate the patterns of segregation, to investigate the causes of segregation, and to investigate the consequences of segregation. The first of these sheds light on the other two – patterns suggest processes. However, what we determine as a pattern is scale dependent. A checker or chessboard has complete segregation at the scale of the individual square that is either white or black. Yet, at the scale of the board, the black-white pattern suggests a high degree of mixing because the board is as much white as it is black, and the two groups are spread out across it equally.

The index of dissimilarity (ID) looks at whether the places where one population group is most likely to be resident are also the places where another group is most likely to reside too. If not, then it may be said that the groups have an uneven geographical distribution (relative to one another) and this is taken as evidence of segregation. Unevenness is, however, only one of the five dimensions of segregation identified by Massey and Denton (1988). The others are exposure, clustering, concentration and centralisation. Amongst them, unevenness and clustering are regarded as the most important (Reardon & O'Sullivan 1997, Oka & Wong 2014). Unevenness is a measure of spatial heterogeneity, the variation displayed across the map. Clustering measures the scale of spatial similarity, the extent to which closely located neighbourhoods are alike. Although unevenness may be related to clustering, it need not be so. Reflecting again on the case of a chess or checkers board, the standard pattern of black-white alternation can be compared with a hypothetical board for which the top half is wholly black and the bottom wholly white. A well-known problem with the ID is that these two patterns of segregation generate the same index value for the board despite the different scales of black and white clustering they present.

Because of this deficiency, several spatial alternatives haven been proposed that include creating localised values and spatial surfaces (Reardon & Sullivan 2004, O'Sullivan & Wong 2007, Lloyd 2012, 2015, Lloyd & Shuttleworth 2012), sometimes using what have been described as egocentric neighbourhoods (Lee et al. 2008, Reardon et al. 2008, 2009, Spielman & Logan 2013, Hongwei et al.

2014). These neighbourhoods are overlapping sub-spaces of the map formed by taking a location (an 'ego') and combining it with a user-defined subset of surrounding places, in much the same way as how local indicators of spatial association (Anselin 1995) or geographically weighted statistics (Brunsdon et al. 2002) are calculated. Increasing the number of locations increases the scale of aggregation and by building-up the level of aggregation, segregation profiles can be formed across a range of analytical scales (Clark et al. 2015, Fowler et al. 2016).

A limitation of this approach is that it does not measure how much of the segregation is due to a specific scale because the act of aggregation conflates but does not separate out the scales of segregation. Others have advanced a multilevel approach that does disentangle the amounts due to the various levels of a geographically hierarchical data structure such as that of the UK Census (Leckie et al. 2012, Jones et al. 2015, Leckie & Goldstein 2015, Manley et al. 2016). The multilevel approach is based on using statistical measures of variance as indicators of segregation, whereby greater variance indicates greater unevenness between places (greater spatial heterogeneity) and therefore greater segregation. Critically, the variance at any one level of the analysis is estimated net of the variances at the other levels, allowing scale differences to be identified. It also permits the variance to be apportioned between the levels of the model, allowing their separate contributions to the overall segregation to be assessed. Apportioning the variance in this way creates a measure of spatial clustering.

A problem is that the variance does not translate directly into an interpretable index of the sort offered by classic approaches (although see Johnston et al. 2017): it is not constrained to lie within a given range (but it cannot be negative) and it is dependent upon the measurement units. This is problematic if what is measured has a prevalence in the population that is susceptible to economic effects or to a conceptual redefinition, such as counts of benefit claimants per neighbourhood or the number of free school meal eligible pupils. Other problems with multilevel modelling are that it assumes the areal and hierarchical units of analysis have some meaningful correspondence with the real-world patterns of segregation studied, and it confuses levels with scale – a point that is returned to in the conclusion.

Despite these shortcomings, it is the multilevel approach that is extended in this paper. The methods outlined are motivated by the work of Jones et al. (2015) and especially of Owen (2015). However, the approach given here is different in three key ways. First, it combines the advantages of a widely-used and interpretable index (the ID) with a means to apportion the ID between the scales of analysis net of the other scales. Combining the multilevel approach with the ID allows the two dimensions of segregation, unevenness and clustering to be measured. Second, the methods are computationally fast: fitting the multilevel models to the data takes only a matter of seconds using the software R, which is free and open source – important criteria for reproducible research (Brunsdon 2016). However, the speed arises because, third, the method does not directly model the observed counts (of a population) against a random data generating process – for example, a Poisson distribution – in the way that Owen and Jones et al. outline. In some regards, what is presented is a simpler and more accessible variant of those earlier approaches. A tutorial in how to fit the models is available at https://rpubs.com/profrichharris/211772.

3. Forming a spatially disaggregated and multiscale index of dissimilarity

The Index of Dissimilarity

The index of dissimilarity (ID) is calculated as

$$\mathrm{ID}_{YX} = k \sum_{i=1}^{m} \left| \frac{n_{Yi}}{n_{Y+}} - \frac{n_{Xi}}{n_{X+}} \right| \tag{1}$$

where the calculation is based on a series of neighbourhood counts: n_{Yi} is the number of population group Y who are living in a neighbourhood, i; n_{Xi} is the number of a second, comparator group, X, living in the same; the summation is across the m neighbourhoods in the study region; n_{Y+} and n_{X+} are the total count of X and Y in the study region ($n_{Y+} = \sum_{i=1}^{m} n_{Yi}$ and $n_{X+} = \sum_{i=1}^{m} n_{Xi}$); and k is a scaling constant. Simplified, and scaled to lie in the range from 0 to 1:

$$\mathrm{ID}_{YX} = 0.5 \times \sum_{i=1}^{m} |p_{Yi} - p_{Xi}| \tag{2}$$

where p_{Yi} is the proportion of group Y that lives in neighbourhood i, and p_{Xi} is the corresponding value for X. If the two groups are distributed evenly then, for every neighbourhood, $p_{Yi} = p_{Xi}$ (so if 1 per cent of the purple population lives in neighbourhood A then 1 per cent of the orange population lives there too), and with no differences between p_{Yi} and p_{Xi}, the index will be zero, interpreted as no segregation. At the opposite extreme, if wherever the purple group is found, the orange group is not (and *vice versa*) then the index will be one, a situation of complete segregation. The index gives the proportion of either of the two groups that would have to change neighbourhoods for the index to be zero, assuming the other group remained in place.

Calculating the index within a regression framework

Since the expectation is that p_{Yi} will be equal to p_{Xi} under a condition of zero segregation so the index may be regarded as summarising the differences between a set of observed values, *y*, and a set of expected values, *x*. Writing this within a regression framework,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{3}$$

where $y_i = p_{Yi}$ and $x_i = p_{Xi}$. Fixing β_0 to zero, β_1 to one and rearranging gives,

$$y_i - x_i = \varepsilon_i \tag{4}$$

and substituting into Equation 1,

$$\mathrm{ID}_{YX} = 0.5 \times \sum_{i=1}^{m} |\varepsilon_i| \tag{5}$$

What equations (3) to (5) show is that the index can be calculated from the residuals of a regression model with dependent variable, y, no intercept and an offset equal to x.

The method is not limited to the index of dissimilarity, an index which is not invariant to changes in the group sizes when x_i and y_i sum to the total population per area of analysis (for example, when x is the White British and y is not White British). Other indices such as the Gorard index can then be used instead (Gorard 2003); doing so requires only a change to the offset. A further extension is to add weights to the models to reflect the fact that the units of measurement (the neighbourhoods) are not of equal population size. Weights proportional to the population size of each neighbourhood might be considered but are not used here to maintain comparability with the standard index of dissimilarity.

The expected value under randomisation

Although, in principle, the ID ranges from 0 to 1 (or from 0 to 100 as some authors prefer), there is a large degree of uncertainty associated with the index when small counts of population are involved (Voas & Williamson 2000). For the pupil data used in the case study that follows, there are 6.77 million pupils in total, living in 170,545 small area Output Areas (OAs), of which, for example, 105,446 (1.56 per cent) are recorded as being Asian Bangladeshi, whereas 5,001,508 (73.8 per cent) are White British. That difference in number means that it is very hard and potentially impossible for the ID to reach its theoretical value of zero – there are too few Bangladeshi pupils for their distribution across the OAs to be the same as for the White British.

The expected value for the ID, given the group sizes and given the geography of the study region can be estimated by simulation:

$$\hat{\boldsymbol{t}} = binom(N, T, \boldsymbol{p_T})$$

$$\hat{\boldsymbol{x}} = binom(N, \hat{\boldsymbol{t}}, P_x)$$

$$\hat{\boldsymbol{y}} = binom(N, \hat{\boldsymbol{t}}, P_y)$$

$$\mathrm{ID}_{YX} = \sum_{i=1}^{m} \left| \frac{n_{\hat{y}i}}{n_{\hat{y}+}} - \frac{n_{\hat{x}i}}{n_{\hat{x}+}} \right| \tag{6}$$

where *binom* indicates a random, binomial process, N is the number of areas (e.g. OAs), T is the total number of pupils (the number of trials), $\boldsymbol{p_T}$ are the proportions of the total living in each of the areas (the probabilities of success), $\hat{\boldsymbol{t}}$ is the simulated number of pupils per area, P_x is the proportion of the total population that is White British, $\hat{\boldsymbol{x}}$ is the simulated number of White British pupils per area, and P_y and $\hat{\boldsymbol{y}}$ are the corresponding values for the Bangladeshi pupils. Each of $\boldsymbol{p_T}$, $\hat{\boldsymbol{t}}$, $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ is a vector of size N. Repeating over 1000 simulations, the average is taken, $\mathrm{E(ID)} \cong \overline{\mathrm{ID}}$.

The procedure is like randomly shuffling the individual pupils around the areas, the difference being that the permutation approach (the shuffling) would hold the total population in each area at its observed value, whereas the binomial draws permit some fluctuation. The method does not incorporate an allowance for stochastic variation directly into the model of segregation in the manner of Jones et al. (2015), Owen 2015 and Johnston et al. (2017). Instead, the expected value under randomisation is modelled as a separate process.

Calculating localised measures of dissimilarity

Recall that the regression residuals are the differences between what is observed for population group Y, and what is expected from population group X, for each neighbourhood. Drawing on the spatial analysis literature, each neighbourhood-level residual can be regarded as a localised disturbance from a general trend at location (u_i, v_i), where the global value is proportional to their sum: $\mathrm{ID}_{YX} \propto \sum_{i=1}^{m} \left|\varepsilon_{(u_i,v_i)}\right|$. Simply plotting the residuals on a map would reveal any spatial variation; it would show where the proportion of group Y is greater or less than expected given the proportion of group X. The level of variation can also be gauged by the standard error of the residuals, which is itself a measure of unevenness and of segregation. In simulations, the standard error (also the variance) of the residuals was found to be correlated almost perfectly with the ID score: $r_{(ID,\sigma_\varepsilon)} \cong 0.99$ (an empirical link from which Jones et al. 2015 also benefit).

Using the residuals, localised measures of dissimilarity can be calculated using an index of the form,

$$id_{(u_i,v_i)} \propto \sum_{j=1}^{m} w_{ij} \left|\varepsilon_j\right| \tag{7}$$

where w_{ij} is a spatial weights matrix specifying which places to include in the index for neighbourhood i and its surrounding areas (see also Wong 1996, Feitosa 2007). Alternatively, we may be interested in exploring the amount by which sub-regions of the map contribute to the overall ID. If n is the number of neighbourhoods in the sub-region then the regional share of the total ID is

$$share = \frac{\sum_{i=1}^{n} |\varepsilon_i|}{\mathrm{ID}_{YX}} \tag{8}$$

From this, the parts of the map contributing most to the segregation can be identified. The problem with the measure is it increases with n. This does not matter if a series of localised values are created based on a fixed number of nearest neighbours around each location because then n would be constant. However, where census regions are considered, and where those regions are unequal in the number of neighbourhoods they contain, we should expect the share to be greater for those with more neighbourhoods, making direct comparisons difficult.

A solution is to divide $\sum_{i=1}^{n} |\varepsilon_i|$ by the expected value, which is $n\overline{|\varepsilon|}$, where $\overline{|\varepsilon|}$ is the mean absolute residual for the whole study region. In this way, the relative impact of each region upon the overall index value can be calculated,

$$impact = \frac{\sum_{i=1}^{n} |\varepsilon_i|}{n\overline{|\varepsilon|}} \times 100 \tag{9}$$

A score of 100 indicates the impact is as expected, 200 indicates it is double expectation, and 50 that it is half. The values can also be scaled to help identify impacts that are 'significant' on average:

$$\bar{z} = \frac{\sum_{i=1}^{n}|\varepsilon_i|}{n(\sigma_\varepsilon)}$$

(10)

where σ_ε is the standard error of the regression residuals.

A multilevel (multiscale) index

The methods described above are flexible in allowing the constituent parts of the ID to be 'summed up' into higher-level units. However, there is nothing specifically multilevel about them. They are based solely on adding together calculations made at a neighbourhood scale, which is useful for identifying places that contribute most to the ID score but not on separating out the scale effects due to each level. To achieve the latter, we need to handle the regression residuals in a different way using a multilevel model.

Census geographies are hierarchical. In England, the small area OAs (with a mean population size of 309 persons and 129 households), nest into what are called Lower Level Super Output Areas (LLSOAs, a mean of 1614 persons and 672 households), which themselves nest into Middle Level Super Output Areas (MLSOAs, 7806 persons and 3249 households), Local Authority Districts (LAs, 163,619 persons and 68,097 households), and governmental regions (GORs, 5,890,273 persons and 2,451,485 households).

Neighbourhood studies often focus on the OAs. There is logic in doing so. They are the smallest areas for which census data are available and will, in principle, be less susceptible to the effects of aggregation than other less detailed geographies. However, sometimes aggregation is desirable if it increases the number of observations per area unit and decreases the uncertainty associated with the estimates.

For the analysis of the pupil data that follows there is an insufficient number of pupils per OA to support calculations at that scale: for example, the mean number of Asian Bangladeshi pupils per OA is 0.618 and that figure excludes OAs where there are no pupils living there at all. At the OA scale, the ID for White British and Asian Bangladeshi pupil segregation is 0.924 but the uncertainty of the estimate is very high, with over half of it expected due to randomisation alone (the expected ID is 0.490). Aggregating the data into LLSOAs produces an ID of 0.848 against a much-reduced expected value of 0.218. LLSOAs will therefore provide the base scale for the analysis.

In any case, it is not just the small area differences that matter as these can co-exist with differences measured at coarser scales of analysis that sometimes exceed the micro-level contributions to segregation (Jones et al. 2015, Johnston et al. 2017). A multilevel model may be specified as

$$Y = \beta_0 + \beta_1 X + \lambda_i + \mu_j + \nu_k + \xi_l \tag{11}$$

where λ_i are the residuals due to the base scale (the LLSOAs), and μ_j, ν_k and ξ_l, are the residuals calculated at higher levels of the hierarchy (MLOAs, LAs and

GORs). The sums of these residuals are equal to those in the standard regression model (Equation 2); that is,

$$\varepsilon_i = \lambda_i + \mu_j + \nu_k + \xi_l \tag{12}$$

Substituting into Equation 4 gives,

$$y_i - x_i = \lambda_i + \mu_j + \nu_k + \xi_l \tag{13}$$

and therefore

$$\mathrm{ID}_{YX} = 0.5 \times \sum \left| \lambda_i + \mu_j + \nu_k + \xi_l \right| \tag{14}$$

The multilevel approach allows the influence of each level upon the overall lD to be considered with that influence measured net of the other levels of the hierarchy. There are parallels with the decomposition of Theil's entropy index of segregation (Theil 1972) into the contribution of different scales of segregation (Fischer et al. 2004, Fischer M, 2008, Lichter et al. 2015) – both it and the ID measure unevenness and both may be used to look at scale effects. However, the multilevel ID follows more closely to Leckie et al., (2012), Leckie and Goldstein (2015) and Jones et al. (2015) in that it is by estimating and comparing the variances of λ_i, μ_j, ν_k and ξ_l that the importance of each scale can be determined.

Specifically, the proportion of the total variance is calculated at each scale. For example, at the LLSOA scale that proportion is $\hat{\sigma}_{\lambda}^{2}/\left(\hat{\sigma}_{\lambda}^{2}+\hat{\sigma}_{\mu}^{2}+\hat{\sigma}_{\nu}^{2}+\hat{\sigma}_{\xi}^{2}\right)$ where the $\hat{\sigma}^{2}$ are the variance estimates at each level. The greater the proportion, the more spatially clustered the two population groups are away from each other at that level and the more it contributes to the overall ID.

In addition, a holdback approach is adopted, which is to calculate the ID with one of the levels omitted (its effect set to zero) and to consider the consequence that has on the index. For example, the percentage change in the ID of omitting level *i*, is,

$$\% \bigtriangleup_{ID(-i)} = 100\left(\frac{\sum\left|0+\mu_{j}+\nu_{k}+\xi_{l}\right|-\mathrm{ID}_{xy}}{\mathrm{ID}_{YX}}\right) \tag{15}$$

4. Case study of the residential segregation of White British pupils from other ethnic groups

Using the National Pupil Database, residential information was extracted on pupils who were in state-funded primary or secondary schools in 2011, omitting those in the upper, non-compulsory years (also those in fee-charging schools, which was 7.2 per cent of all pupils nationally).[3] Pupil counts per LLSOA were calculated using the pupils' home postcodes, matched to the census geography. From these, the multilevel index and the various measures of residential

[3] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219261/sfr10-2012nt.xls, although it varies by local authority: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219262/sfr10-2012lat.xls

segregation outlined in Section 3 were calculated for the White British against each the of six major ethnic groups included in the data: Asian Bangladeshis, Asian Indians, Asian Pakistanis, Black Africans, Black Caribbeans and those of a White but not White British ethnicity. For the analysis, the generic statistics functions in R were used, supplemented with the lme4 (Linear Mixed-Effects Models) package (Bates et al. 2015). Initial estimates were checked against those from the multilevel modelling software MLwiN and yielded the same results.

The results are in Table 1. Taking the White British – Asian Bangladeshi ID as an example, its value for the whole of England is, as previously stated, 0.848 against an expected value of 0.218. The ID value is high because the Bangladeshi group is concentrated especially in the eastern parts of London but also in parts of the urban North West such as Oldham, Burnley and Bradford, whereas the White British are more spatially dispersed across urban and rural regions.

[TABLE 1 ABOUT HERE]

Regarding the scale of their separation, the holdback scores suggest the differences between local authorities are important. To discount them has the effect of reducing the ID by 30.7 per cent, whereas discounting the MLSOA scale reduces it by 10.3, the LLSOA by 9.1 per cent, and the regions by 5.6 per cent. The estimates of the variance at each level have the same rank ordering but the differences between the LA, MLSOA and LLSOA scales are less clear: 35.3 per cent of the variance in the ID is assigned to the LA scale, 34.6 per cent to the

MLSOA scale, 29.7 per cent to the LLSOA scale, and only 0.7 per cent to the GOR scale. Looking at the local authorities, it is Tower Hamlets and Newham, both situated in the former Docklands of London, together with Oldham in the North West that make the most 'significant' contribution to the ID score, where $\bar{z} > 1.96$ (from Equation 10) and the impact scores are 2669, 969 and 679, respectively – many times greater than the 'as average' value of 100.

Whereas the variance measures indicate that the variations within local authorities, at the MLSOA scale, are almost equal to the variation between local authorities, at the LA scale, the holdback scores suggest that the effect of the LA variation is greater. This is not inconsistent because they are measuring different things. The model is additive so any uplift (or decrease) in segregation that is due to the LA scale applies to all the neighbourhoods within the LA, whereas the change due to the MLSOA is restricted to the smaller sub-group of neighbourhoods that are in that MLSOA. It is therefore entirely possible for the proportion of the variance to be small at the higher levels but for the differences between places at those levels to still have a strong cumulative effect upon all the lower levels to which they must be added. This will be picked-up on by the holdback scores.

Looking again at Table 1, although the level of residential segregation is highest from the White British for the Asian Bangladeshi pupils, it remains high for each of the other ethnic groups too. The ID usually exceeds 0.800 and is always higher than what would be expected under randomisation. Amongst the groups, the least separation is for the White British from the White Other group (ID = 0.506); the next lowest is from the Indian group (0.738). The scale of separation is usually greatest at the LA scale, although the MLSOA scale exceeds it in a few instances, most obviously for the Asian Pakistanis. This suggests that the spatial clustering of the six ethnic groups is at a scale typically above the 'micro' scale of the LLSOAs but below the macro scale of Government regions.

A lot of the LA clustering occurs within London. In fact, 28 of the 32 London Boroughs are listed as making a 'significant' contribution to the various ID scores in Table 1. For the Black African, Black Caribbean and White Other groups, none of the LAs listed is outside London. The LA for these three groups overlap – notably Brent, Hackney and Haringey – but, except for Newham, not with the LAs listed for any of the three Asian groups. Despite the spatial deconcentration and mixing of 'minority' groups in London that is evidenced by the 2001 and 2011 Censuses, there remains a geography to where the pupils of different ethnic groups are most likely to be living (see Harris 2016, Figure 11.2).

Because the model is additive, we can map how the geography unfolds cumulatively at each level. As an example, Figure 1 considers White British segregation from the three Asian groups (Bangladeshi, Indian and Pakistani) combined. It begins with the regional estimate for London (ξ_l), which is, of course, constant across the London region. It also maps the regional estimate plus the local authority estimates ($\nu_k + \xi_l$), those plus the MLOA estimates ($\mu_j + \nu_k + \xi_l$), and then the same but with the LLOA estimates included ($\lambda_i + \mu_j + \nu_k + \xi_l$). Every shaded area of the map has an 'excess share' of the Asian pupils relative to the White British ones, and the darker the shading the greater the difference is.

[FIGURE 1 ABOUT HERE]

At the regional scale, London is predicted to have a higher than expected share of the Asian pupils: its share is greater than its share of White British pupils. There are differences between the local authorities, with increased shares suggested in Redbridge, Newham, Tower Hamlets and Hounslow. However, the residential patterning does not always match local authority boundaries and more detail emerges as the MLOA effects are added in: Redbridge, Newham and Tower Hamlets do still contain clusters of Asian pupils but the cluster in Hounslow appears to be split across it and neighbouring authorities. Adding in the LLOA effects makes some but less difference to the map than those at the MLOA and local authority scales.

Comparison with the residential segregation of the Census population

Although the data are collected separately, pupils in state-funded schools in 2011 are a sub-group of the entire English population enumerated by the 2011 Census. If segregation from the White British is decreasing then we might expect the residential segregation between school pupils to be less than the ethnic segregation for the Census population as a whole because the residential locations of school children represent the younger members of society and their parents/guardians.

In fact, the reverse is true, apparently supporting what the Casey Review (discussed in the Introduction) has claimed. Under the columns headed Census 2011 and Pupils 2011, Table 2 shows that for each of the six groups, the ID value for the segregation from the White British is less for the Census population than it is for the pupils alone. Typically, the former is about 85-90 per cent of the latter. Even allowing that there is less uncertainty associated with the Census figures because they are drawn from a larger population and so the expected value under randomisation is also always less, this seems insufficient to explain the difference.

[TABLE 2 ABOUT HERE]

Looking at Table 2, it is notable that the regional differences are stronger for the Census population – the holdback scores and the variances are greater at this level than they were when calculated for the pupils. This is most evident for the Asian Indian group: the estimated 1.1 per cent of the variance due to the GOR scale in the model for the pupils now rises to 11.5 per cent for the Census population. At the opposite end of the scale, for all groups the percentage of the variance falls at the LLSOA level and for most it decreases at the MLSOA scale (the exception is the Black Caribbean group for which it rises marginally). The LA scale share of the variance rises for the Asian Bangladeshi, Asian Indian and Asian Pakistani groups but falls for the Black Africans, Black Carribeans and White Other.

The increase in the regional effect is due to London. It is evident in Figure 2, which is a caterpillar plot showing the regional level residuals and their 95 per cent confidence interval estimated for the Asian Indian – White British model of segregation. For the pupil data, although there is some evidence of an increased regional effect for London (the capital has a greater than expected share of the Asian Indian pupils), the confidence interval for the estimate overlaps with the interval for the West Midlands and is not too far apart from the other regions. London is different but not excessively so.

[FIGURE 2 ABOUT HERE]

However, for the census population, London stands clearly apart. The situation for the Asian Bangladeshi – White British model is similar but there are even starker differences between London and the other regions when the census populations of the Black African, Black Caribbean or White Other groups are considered relative to the White British. Only for the Asian Pakistani group is there little difference between London and the rest. This is the group for which the increase in the variance at the GOR scale is less than for the other groups when comparing the census population with the pupils, and, for which, the two 'significant' cases of LA-level segregation are both outside of London for the pupils (see Table 1).

To summarise, the evidence for 2011 is that school age members of 'minority' ethnic groups are unevenly distributed relative to the White British more than the all-age census population is. In addition, there is greater clustering at coarser geographical scales for the census population than for the pupils, with clear evidence of a London effect. Given children usually live with a parent of the same ethnic group, the implication is that what is observed for London is caused by residential differences between elder members of the census population. Previous studies have highlighted the reduced number of White British living in cities, notably London (Goodhart 2013, Hellen 2013, Kaufmann & Harris 2014). However, the median age of the White British is greater at 40-44 years than for other ethnic groups, especially Asian Bangladeshis, for whom the median is 20-24 years (Simpson 2013, Harris 2016b). If there is a general process of migration out from cities with aging then the White British will, on average, be ahead of other ethnic groups and more financially positioned to make the move. This would be consistent with other evidence, which has shown: (a) a pattern of spatial retrenchment and contraction of the White British out of London and other traditionally industrial cities; (b) a process of dispersion and spatial diffusion of 'minority' groups across cities as their numbers grow and they move out from their previous enclaves; (c) that the places that those groups move to have declining numbers of White British residents; but (d) that the places that the White British are moving to are gradually becoming more ethnically diverse (Harris 2014, 2015, 2016a, Johnston et al. 2014, Catney 2015b). Paradoxically, what makes London seem unusual might not be not unique at all. The reason is that a process whereby some of the 'minority' groups spread out, and some of the White British residents move out, can act either to even-up the shares of each group living in each neighbourhood or to increase the disparity between them. It all depends on the starting values and the amount of change.

Critically, what cannot be concluded is that because the school age population appears more segregated when compared to residential patterns of living so there is evidence that ethnic segregation is increasing. If the pupil data for 2011 are compared with that for 2002 (the earliest date for which comparable data are available) then there is no evidence that the ID has increased for pupils; in fact, it has decreased slightly for all groups except the Black Caribbeans, for which there has been only a marginal increase (from 0.808 to 0.820). This can be seen by looking again at Table 2 and comparing the columns headed Pupils 2011 with the columns header Pupils 2002.

Also evident from the table is a general uplift in the scale of segregation for pupils in 2011 when compared with those for 2002. This can be seen by looking at the holdback and variance measures for the two years and noting that these tend to reduce in magnitude at the lower scales and to increase at the coarser ones. The finding is again consistent with a process of spreading out of the minority groups and the spatial contraction of the White British, creating broader scale patterns of ethnic geography. These broader patterns replace some of the finer scale differences but do so against a backdrop in which the overall divisions are lessening. The evidence is of decreasing residential segregation as other studies also have shown.

Conclusion

This paper makes two connected advances. The first is a methodological development that adds to a growing interest in multiscale measures of segregation. The second is a contribution to current public and policy debates in response to the Casey Review into opportunity and integration and the perception that ethnic segregation is rising in the UK – this despite the evidence of multiple census-based studies that show it to have decreased.

The paper looks at the residential segregation of White British school pupils in comparison to other ethnic groups in England. It agrees with the Review in finding that the school age population is more segregated residentially than is the population-at-large but that difference needs to be set in the context of what is decreasing segregation for the school-age population over the decade to 2011. The general uplift in the geographical scale at which patterns of segregation can be observed is consistent with a process of the minority groups spreading out from their more traditional enclaves occurring in tandem with the spatial contraction of the White British from urban areas, especially London. However, there is no reason to conclude this is driven by or is driving a presumed lack of integration because the differential rate of movement can arise because of their different age profiles and spatial opportunity structures. The other ethnic groups are generally younger than the White British (Rees & Butt 2004, Rees et al. 2013), which means their movement into new areas to raise families can increase their number in those places and give an impression of increased segregation. However, such increases are likely to be short-lived: they are a demographic effect that will decline as those children age and themselves move.

The oobserved patterns of segregation have been studied by showing how the traditional index of dissimilarity can be disaggregated into a series of localised values, which are derived as residuals from a regression model with a zero intercept and an appropriate offset. Those residuals can also be estimated in a multilevel model that allows them to be apportioned between the various levels of a geographical hierarchy. Measures of the amount and spatial impact of the segregation at each level can be calculated, allowing for the study of scale effects. As such, the index begins to consider both the principal dimensions of segregation, unevenness and clustering.

More advanced developments would be to consider how the approach can be extended into a multi-group index (Reardon & Firebaugh, 2002) or to place the index within the framework of a hierarchical spatial autoregressive model (Dong & Harris 2015, Dong et al., 2016). Such a model provides an integration of multilevel modelling with a spatial econometric approach and would move away from the idea that patterns of segregation are contained within the discrete boundaries of the model's geographical hierarchical and can instead overspill into neighbouring locations. A weakness of the multilevel approach is it does not really look at the impact of *scale* upon segregation but upon the amount of segregation that can be attributed to the various *levels* of the model. The problem is that these levels are usually somewhat arbitrary; in fact, any one level is usually a mixture of scales because the areal units that comprise it vary in shape and size. There is potential in adopting a more consistent geography such as using the gridded data available at https://popchange.liverpool.ac.uk/ (Lloyd, 2016).

Most broadly, we need to ask what measures of segregation can actually tell us about the processes of segregation. Ultimately, the indices measure outcomes not causes, and the multilevel models reveal only (unexplained) spatial variation, not the processes by which patterns of spatial clustering and dissimilarity arose.

Although a focus on multiscale measurement is important, to better achieve its potential, a future research agenda needs to link the measurement both to the causes of spatial differentiation (demographic change, immigration, employment structures, the operation of the housing market, and so forth) and to its outcomes (for example, the impact of segregation on health, well-being or educational performance).

Acknowledgments

The research was funded partly under the ESRC's Urban Big Data Centre (http://ubdc.ac.uk/), ES/L011921/1. An early version of the paper was presented at a seminar in Ritsumeikan University at the kind invitation of Professors Yano and Nakaya. The author is extremely grateful to the constructive and helpful suggestions of the anonymous referees and the editor. Each deserves to be credited as co-authors but errors and omissions remain my own.

References

Allen R and Vignoles A 2007 What Should an Index of School Segregation Measure? *Oxford Review of Education* 33 643-668

Anselin L 1995 Local Indicators of Spatial Association - LISA *Geographical Analysis* 27 93-115

Bates D Maechler M Bolker B and Walker S 2015 Fitting Linear Mixed-Effects Models Using lme4 *Journal of Statistical Software* 67 1-48

Brunsdon C Fotheringham AS and Charlton M 2002 Geographically weighted summary statistics — a framework for localised exploratory data analysis *Computers, Environment and Urban Systems* 26 501-524

Brunsdon C 2016 Quantitative methods I. Reproducible research and quantitative geography *Progress in Human Geography* 40 687-696

Cantle T 2001 *The Cantle Report - Community Cohesion: a report of the Independent Review Team* The Home Office, London

Cantle T 2012 *Interculturalism: The New Era of Cohesion and Diversity* Palgrave Macmillan, London

Catney G 2016a Exploring a decade of small area ethnic (de-)segregation in England and Wales *Urban Studies* 5 1691-1709

Catney G 2016b The Changing Geographies of Ethnic Diversity in England and Wales, 1991-2011 *Population, Space and Place* 22 750-765

Catney G 2015 Towards an enhanced understanding of ethnic group geographies using measures of clustering and unevenness *The Geographical Journal*, in press. [Online at DOI: 10.1111/geoj.12162]

Clark WAV Anderson E Östh J and Malmerg B 2015 A Multiscalar Analysis of Neighborhood Composition in Los Angeles, 2000-2010: A Location-Based Approach to Segregation and Diversity *Annals of the Association of American Geographers* 105 1260-128

Dong G Harris R Jones K and Yu J 2015 Multilevel Modelling with Spatial Interaction Effects with Application to an Emerging Land Market in Beijing, China *Plos One* (June 18, 2015). http://dx.doi.org/10.1371/journal.pone.0130761

Dong G and Harris R 2015 Spatial Autoregressive Models for Geographically Hierarchical Data Structures *Geographical Analysis* 47 173-191

Dong, G Harris R and Mimis A 2016 HSAR: Hierarchical Spatial Autoregressive Model (HSAR). https://cran.r-project.org/web/packages/HSAR/index.html [Accessed November 26, 2016]

Duncan OD and Duncan B 1955 A Methodological Analysis of Segregation Indexes *American Sociological Review* 20 210-217

Feitosa FF Câmara G Monteiro AMV Koschitzki T and Silva MPS 2007 Global and local spatial indices of urban segregation. *International of Geographical Information Science* 21 299-323

Fischer CS Stockmayer G Stiles J and Hout M 2004 Distinguishing the Geographic Levels and Social Dimensions of US Metropolitan Segregation, 1960-2000. *Demography* 41 37-59

Fischer M 2008 Shifting Geographies: Examining the Role of Suburanization in Blacks' Declining Segregation *Urban Affairs Review* 43 475-496

Fowler CS Lee BA and Matthews SA 2016 The Contributions of Places to Metropolitan Ethnoracial Diversity and Segregation: Decomposing Change Across Space and Time *Demography* 53 1955-1977

Galster G and Shevky P 2017 Spatial Foundations of Inequality: A Conceptual Model and Theoretical Framework *The Russell Sage Journal of the Social Sciences*, 3 forthcoming

Goodhart D 2013c White flight? Britain's new problem - segregation *Prospect*, February 2013 30-31

Gorard S Taylor C and Fitz J 2003 *Schools, Markets and Choice Policies* RoutledgeFalmer, London

Harris R 2014 Evidence and trends: are we becoming more integrated, more segregated or both? in Goodhart D ed. *Mapping Integration* Demos, London 15-23. Available at http://www.demos.co.uk/project/mapping-integration/

Harris R 2016 *Quantiative Geography. The Basics* Sage, London

Hellen N 2013 Britons 'self-segregate' as white flight soars *The Sunday Times* January 27 2013 15

Hongwei X Logan JR and Short SE 2014 Integrating Space with Place in Health Research: A Multilevel Spatial Investigation Using Child Mortality in 1880 Newark, New Jersey. *Demography* 51 811-834

Jahn JA Schmid CF and Schrag C 1947 The Measurement of Ecological Segregation. *American Sociological Review* 12 293-303

Johnston R Manley D and Jones K 2017 Spatial scale and measuring segregation: illustrated by the formation of Chicago's ghetto *GeoJournal* in press

Johnston R Poulsen M and Forrest J 2013 Multiethnic residential areas in a multiethnic country? A decade of major change in England and Wales *Environment and Planning A* 45 753-759

Johnston R Poulsen M and Forrest J 2014 Increasing diversity within increasing diversity: the changing ethnic composition of London's neighbourhoods 2001–2011 *Population Space and Place* 21 38-53

Jones K Johnston R Manley D Owen D and Charlton C 2015 Ethnic Residential Segregation: A Multilevel Multigroup Multiscale Aroach Exemplified by London in 2011 *Demography* 52 1995-2019

Kaufmann E and Harris G 2014 *Changing Places: Maing the white British response to ethnic change* Demos, London Available at: http://wwwdemoscouk/project/changing-places/

Leckie G and Goldstein H 2015 A multilevel modelling aroach to measuring changing patterns of ethnic composition and segregation among London secondary schools 2001–2010 *Journal of the Royal Statistical Society Series A* 178 405-424

Leckie G Pillinger R Jones K and Goldstein H 2012 Multilevel modelling of Social Segregation *Journal of Educational and Behavioral Statistics* 37 3-30

Lee BA Reardon SF Firebaugh G Farrell CR Matthews SA and O'Sullivan D 2008 Beyond the Census Tract: Patterns and Determinants of Racial Segregation at Multiple Geographical Scales *American Sociological Review* 73 761–91

Lichter DT Parisi D and Taquino M 2015 Toward a New Macro-Segregation? Deccomposing Segregation within and between Metropolitan Cities and Suburbs *American Sociological Review* 80 843-873

Lloyd CD 2012 Analysing the spatial scale of population concentrations by religion in Northern Ireland using global and local variograms *International Journal of Geographical Information Science* 26 57-73

Lloyd C 2015 Assessing the spatial structure of population variables in England and Wales *Transactions of the Institute of British Geographers* 40 28-43

Lloyd C 2016 Spatial scale and small area population statistics for England and Wales *International Journal of Geographical Information Science* 30 1187-1206

Lloyd C and Shuttleworth I 2012 Residential Segregation in Northern Ireland in 2001: Assessing the Value of Exploring Spatial Variations *Environment and Planning A* 44 52-67

Logan JF 2013 The Persistence of Segregation in the 21st Century Metropolis 12 160-168

Manley D Johnston R Jones K and Owen D 2015 Macro- Meso- and Microscale Segregation: Modeling Changing Ethnic Residential Patterns in Auckland New Zealand 2001-2013 *Annals of the Association of American Geographers* 105 951-967

Manley D Johnston R Jones K and Owen D 2016 Occupational seggregation in London: a multilevel framework for modelling segregation in Tammaru T Marcińczak S van Ham M and Musterd S eds *Socio-Economic Segregation in European Capital Cities* Routledge, London 30-54

Massey DS and Denton NA 1988 The dimensions of residential segregation *Social Forces* 67281-315

Nightingale C 2012 *Segregation: A Global History of Divided Cities* The University of Chicago Press, Chicago

Owen D 2015 Measuring residential segregation in England and Wales: a model-based aroach Unpublished PhD thesis School of Geographical Sciences, University of Bristol

O'Sullivan D and Wong D W S 2007 A Surface-Based Aroach to Measuring Spatial Segregation *Geographical Analysis* 39 147-168

Reardon SF and Firebaugh G 2002 Measures of Multigroup Segregation *Sociological Methodology* 32 33-67

Reardon SF Farrell CR Matthews SA O'Sullivan D Bischoff K and Firebaugh G 2009 Race and Space in the 1990s: Changes in the Geographic Scale of Racial Residential Segregation 1990-2000 *Social Science Research* 38 55-70

Reardon SF Matthews SA O'Sullivan D Lee BA Firebaugh G Farrell CR and Bischoff K 2008 The Geographic Scale of Metropolitan Segregation *Demography* 45 489-514

Reardon SF and O'Sullivan D 2004 Measures of Spatial Segregation *Sociological Methodology* 34 121-162

Reardon SF Matthews SA O'Sullivan D Barrett AL Firebaugh G Farrell CR and Bischoff K 2008 The Geographic Scale of Metrpolitan Racial Segregation *Demography* 45 489-514

Rees P and Butt F 2004 Ethnic change and diversity in England 1981-2001 *Area* 36 174-186

Rees P Wohland P and Norman P 2013 The demographic drivers of future ethnic group populations for UK local area 2001-2051 *The Geographical Journal* 179 44-60

Simpson L 2013 What makes ethnic group populations grow? Age structures and immigration Dynmics of Diversity: Evidence from the 2011 Census briefing note University of Manchester: ESRC Centre on Dynamics of Ethnicity Available at:
http://wwwethnicityacuk/medialibrary/briefings/dynamicsofdiversity/what-makes-ethnic-group-populations-grow-age-structures-and-immigrationpdf

Simpson L and Peach C 2009 Measurement and Analysis of Segregation Integration and Diversity: Editorial Introduction *Journal of Ethnic and Migration Studies* 35 1377-1380

Spielman SE and Logan JR 2013 Using High-Resolution Population Data to Identify Neighbourhoods and Establish Their Boundaries *Annals of the Association of American Geographers* 103 67-84

Theil H 1972 *Statistical decomposition analysis* North-Holland Publishing Company, Amsterdam

Voas D and Williamson P 2000 The Scale of Dissimilarity: Concepts Measurement and an Alication to Socio-Economic Variation across England and Wales *Transactions of the Institute of British Geographers* 25 465-481

Wong DWS 1996 Enhancing segregation studies in GIS *Computers Environment and Urban Systems* 20 99-109

List of figures / tables

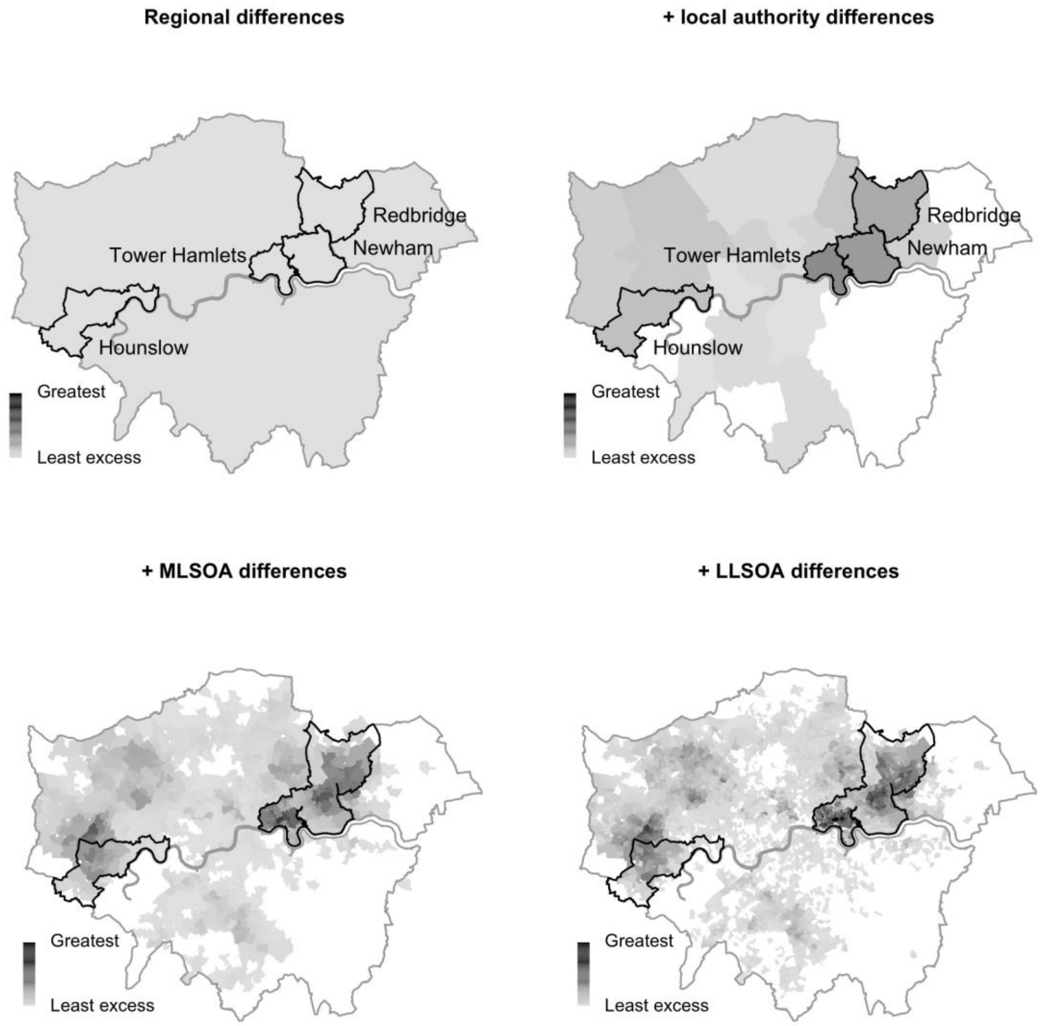

Figure 1. Showing how the 'excess share' of the combined Asian Bangladeshi, Indian and Pakistani populations unfolds at the various levels of the model for the residential locations of school pupils in London in 2011.

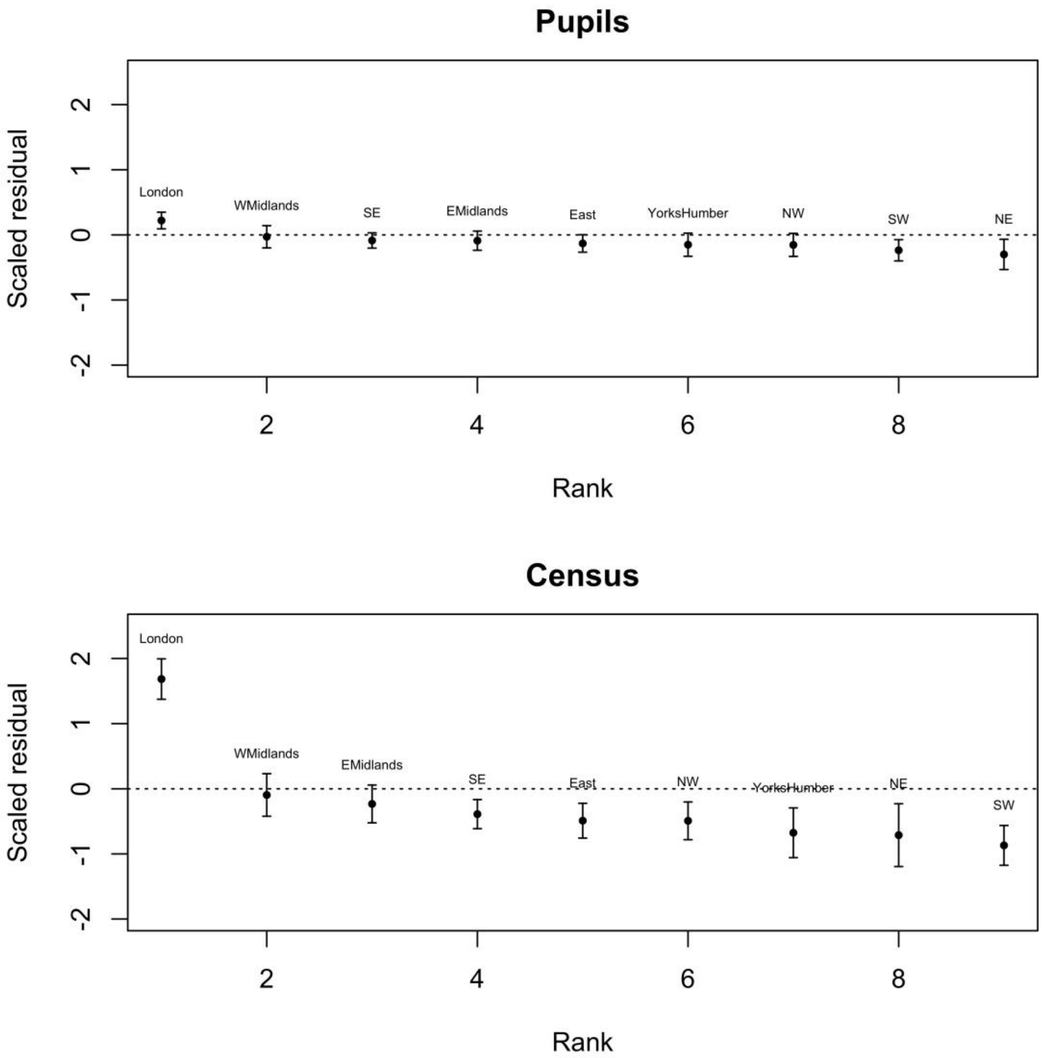

Figure 2. Caterpillar plots showing the regional level effects of Asian Indian – White British residential segregation estimated first for the pupils and then for the Census population

	Holdback score		% of variance			Holdback score		% of variance	
Asian Bangladeshi	GOR	-5.6	GOR	0.4	**Asian Pakistani**	GOR	-9.4	GOR	0.8
ID = 0.848	**LA**	**-30.7**	**LA**	**35.3**	ID = 0.842	LA	-14.3	LA	13.3
E(ID) = 0.218	MLSOA	-10.3	MLSOA	34.6	E(ID) = 0.139	**MLSOA**	**-19.4**	**MLSOA**	**61.6**
	LLSOA	-9.1	LLSOA	29.7		LLSOA	-6.8	LLSOA	24.5
Significant LAs	$\bar{z}$	Impact	% pupils Asian Bangladeshi	% pupils White British	Significant LAs	$\bar{z}$	Impact	% pupils Asian Pakistani	% pupils White British
Tower Hamlets (LDN)	8.45	2669	60.7	11.1	Bradford (Y&H)	2.79	702	33.3	50.2
Newham (LDN)	3.07	969	17.9	8.8	Pendle (NW)	2.05	515	31.0	63.5
Oldham (NW)	2.15	679	13.9	60.3					
	Holdback score		% of variance			Holdback score		% of variance	
Asian Indian	GOR	-8.5	GOR	1.1	**White Other**	GOR	-17.2	GOR	11.7
ID = 0.738	**LA**	**-25.9**	LA	28.1	ID = 0.506	**LA**	**-28.8**	**LA**	**41.7**
E(ID) = 0.171	MLSOA	-13.0	**MLSOA**	**46.5**	E(ID) = 0.130	MLSOA	-11.7	MLSOA	21.6
	LLSOA	-7.9	LLSOA	24.4		LLSOA	-11.5	LLSOA	25.1
Significant LAs	$\bar{z}$	Impact	% pupils Asian Indian	% pupils White British	Significant LAs	$\bar{z}$	Impact	% pupils White Other	% pupils White British
Leicester (EMID)	4.20	939	29.4	37.3	Enfield (LDN)	4.68	691	25.7	25.0
Blackburn (NW)	2.81	628	17.5	55.5	Haringey (LDN)	3.70	546	23.4	19.0
Harrow (LDN)	2.43	542	19.6	20.7	Waltham Forest (LDN)	2.49	368	14.4	22.0
Oadby and Wigston (EMID)	2.36	528	23.1	61.4	Hackney (LDN)	2.48	365	16.9	15.3
Hounslow (LDN)	2.26	504	17.2	29.7	Barnet (LDN)	2.39	353	17.6	35.3
Slough (SE)	2.24	500	14.8	24.7	Brent (LDN)	2.09	309	12.9	8.8
Redbridge (LDN)	2.08	465	14.4	20.5					
Black Caribbean	Holdback score		% of variance		**Black African**	Holdback score		% of variance	

ID = 0.820 E(ID) = 0.234	GOR	-19.5	GOR	5.5	ID = 0.779 E(ID) = 0.156	GOR	-23.5	GOR	8.4
	LA	**-33.6**	**LA**	**40.1**		**LA**	**-24.5**	LA	28.3
	MLSOA	-6.1	MLSOA	32.0		MLSOA	-8.0	**MLSOA**	**35.9**
	LLSOA	-4.2	LLSOA	22.4		LLSOA	-5.0	LLSOA	27.5
Significant LAs	$\bar{z}$	Impact	% pupils Black Caribbean	% pupils White British	Significant LAs	$\bar{z}$	Impact	% Black African	% pupils White British
Lewisham (LDN)	3.96	738	16.6	25.8	Barking & Dagenham (LDN)	3.43	631	22.0	39.5
Lambeth (LDN)	3.72	693	17.5	15.8	Southwark (LDN)	3.25	598	29.8	23.0
Hackney (LDN)	3.04	566	12.9	15.3	Greenwich (LDN)	2.74	504	22.5	41.6
Croydon (LDN)	3.00	559	12.1	37.1	Newham (LDN)	2.71	498	16.6	8.8
Haringey (LDN)	2.76	514	10.9	19.0	Lambeth (LDN)	2.39	441	24.0	15.8
Brent (LDN)	2.74	510	10.5	8.8	Hackney (LDN)	2.35	432	21.2	15.3
Southwark (LDN)	2.41	449	10.6	23.0	Enfield (LDN)	2.10	387	15.6	25.0
Waltham Forest (LDN)	2.39	444	8.4	22.0	Haringey (LDN)	2.08	384	17.5	19.0
					Brent (LDN)	2.07	381	16.8	8.8

Table 1. Showing the ID, expected ID, holdback scores and variance measures for the residential segregation of each of six ethnic groups from the White British at four levels of analysis (LLSOAs, MLSOAs, LAs and GORs) for pupils in state schools in England in 2011. Also shown are the LAs contributing most greatly to the ID score in each case and the regions where those LAs are located (EMID = East Midlands, LDN = London, NW = North West, Y&H = Yorkshire and the Humber, SE= South East).

	Census 2011		Pupils 2011		Pupils 2002	
Asian Bangladeshi	ID	0.772	ID	0.848	ID	0.874
	E(ID)	0.109	E(ID)	0.218	E(ID)	0.266
	% population	0.8	% population	1.6	% population	1.1
	Holdback score	% Variance	Holdback score	% Variance	Holdback score	% Variance
GOR	**-15.0**	6.4	-5.6	0.4	-2.9	0.1
LA	-4.9	**36.2**	**-30.7**	**35.3**	**-27.9**	31.5
MLSOA	-8.9	32.2	-10.3	34.6	-9.3	32.7
LLSOA	-7.3	25.2	-9.1	29.7	-11.5	**35.7**
Asian Indian	ID	0.658	ID	0.738	ID	0.772
	E(ID)	0.060	E(ID)	0.171	E(ID)	0.178
	% population	2.6	% population	2.5	% population	2.3
	Holdback score	% Variance	Holdback score	% Variance	Holdback score	% Variance
GOR	**-18.6**	11.5	-8.5	1.1	-10.2	1,0
LA	-12.1	30.0	**-25.9**	28.1	**-23.4**	23.4
MLSOA	-10.6	**40.6**	-13.0	**46.5**	-12.2	**50.1**
LLSOA	-5.2	17.9	-7.9	24.4	-6.7	25.6
Asian Pakistani	ID	0.775	ID	0.842	ID	0.852
	E(ID)	0.068	E(ID)	0.139	E(ID)	0.166
	% population	2.1	% population	3.8	% population	2.6
	Holdback score	% Variance	Holdback score	% Variance	Holdback score	% Variance
GOR	-13.9	2.8	-9.4	0.8	-9.2	0.6
LA	-12.0	15.0	-14.3	13.3	-9.1	8.4
MLSOA	**-16.8**	**61.3**	**-19.4**	**61.6**	**-20.4**	**59.6**
LLSOA	-5.5	20.9	-6.8	24.3	-9.0	31.3
Black African	ID	0.699	ID	0.779	ID	0.841
	E(ID)	0.072	E(ID)	0.156	E(ID)	0.227
	% population	1.8	% population	3.0	% population	1.4
	Holdback score	% Variance	Holdback score	% Variance	Holdback score	% Variance
GOR	**-32.9**	33.0	-23.5	8.4	-22.9	5.3
LA	-4.4	17.2	**-24.5**	28.3	**-29.9**	**34.2**
MLSOA	-8.8	**32.0**	-8.0	**35.9**	-5.0	28.9
LLSOA	-4.7	17.8	-5.0	27.5	-4.2	31.5
Black Caribbean	ID	0.717	ID	0.820	ID	0.808
	E(ID)	0.093	E(ID)	0.234	E(ID)	0.231
	% population	1.1	% population	1.3	% population	1.4
	Holdback score	% Variance	Holdback score	% Variance	Holdback score	% Variance
GOR	**-33.8**	**32.6**	-19.5	5.5	-19.1	4.4
LA	-7.5	22.4	**-33.6**	**40.1**	**-29.9**	33.2
MLSOA	-6.9	32.2	-6.1	32.0	-7.2	**37.0**
LLSOA	-2.9	12.8	-4.2	22.4	-4.5	25.4
White Other	ID	0.464	ID	0.506	ID	0.610
	E(ID)	0.041	E(ID)	0.130	E(ID)	0.165

	% population	5.7	% population	4.3	% population	2.7
	Holdback score	% Variance	Holdback score	% Variance	Holdback score	% Variance
GOR	**-34.5**	**48.5**	-17.2	11.7	-14.1	6.4
LA	-7.6	22.7	**-28.8**	**41.7**	**-29.9**	**36.4**
MLSOA	-11.3	17.4	-11.7	21.6	-10.0	31.6
LLSOA	-5.5	11.4	-11.5	25.1	-8.7	25.5

Table 2. Comparing the ID and scale of segregation from the White British for each of the six ethnic groups: for the Census population in 2011, for pupils in state schools in 2011, and for pupils in state schools in 2002, in England.