THE CENTRE FOR MARKET AND PUBLIC ORGANISATION

Using Local Statistics to Portray Ethnic Residential Segregation in London

Michael Poulsen, Ron Johnston
and James Forrest

March 2009

Working Paper No. 09/213

Centre for Market and Public Organisation
Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol BS8 1TX
http://www.bristol.ac.uk/cmpo/

Tel: (0117) 33 10799
Fax: (0117) 33 10705
E-mail: cmpo-office@bristol.ac.uk

The Centre for Market and Public Organisation (CMPO) is a leading research centre, combining expertise in economics, geography and law. Our objective is to study the intersection between the public and private sectors of the economy, and in particular to understand the right way to organise and deliver public services. The Centre aims to develop research, contribute to the public debate and inform policy-making.

CMPO, now an ESRC Research Centre was established in 1998 with two large grants from The Leverhulme Trust. In 2004 we were awarded ESRC Research Centre status, and CMPO now combines core funding from both the ESRC and the Trust.

ISSN 1473-625X

CMPO Working Paper Series No. 09/213

Using Local Statistics to Portray Ethnic Residential Segregation in London

Michael Poulsen[1], Ron Johnston[2]
and James Forrest[1]

[1]Macquarie University
[2] CMPO, University of Bristol

March 2009

Abstract

Much has been written about ethnic residential segregation in urban areas, almost all of it deploying single-index numbers to measure the degree of segregation. These give very little detailed appreciation of the extent to which different ethnic groups live apart from each other, and where. This paper suggests that a combination of measures derived from local spatial statistics, which identify the geography of clustering, and a typology of residential areas, which describes the population composition of each area, provides much greater insight into the nature and extent of segregation. Data for London in 2001 illustrate the potential of this approach.

Keywords: segregation, ethnicity, London, local statistics

JEL Classification: R1

Electronic version: **www.bristol.ac.uk/cmpo/publications/papers/2009/wp213.pdf**

Address for Correspondence
CMPO, Bristol Institute of Public Affairs
University of Bristol
2 Priory Road
Bristol
BS8 1TX
R.Johnston@bristol.ac.uk
www.bristol.ac.uk/cmpo/

CMPO is jointly funded by the Leverhulme Trust and the ESRC

1. Introduction

The links between immigration, multicultural policies, residential segregation, and social cohesion are much debated in contemporary societies. Within such debates, there is a widespread belief that segregation – whether in residential neighbourhoods, schools, or a variety of other formal and informal social arena – is not conducive to the development of an open and tolerant society. Instead, it is argued that ethnic groups living apart tends to embed and increasingly exacerbate problems of difference within an increasingly diverse cultural milieu. Identifying the extent and nature of ethnic residential segregation, as a preliminary to policy development, is therefore a fundamental task for social scientists.

There has been much debate over the measurement of residential segregation: how can the degree to which various cultural groups live apart be portrayed in ways that fully reflect the complex reality? Most work in this area has deployed a number of single-number indices, as illustrated by Simpson's (2007) recent work on segregation levels in British cities. These suffer from a number of substantial weaknesses and we suggest and explore an alternative approach, using geostatistical procedures with data for London drawn from the 2001 Census of England and Wales as the exemplar. The approach combines an analysis of residential clustering – which takes a relative perspective on segregation – with one that focuses on the ethnic composition of individual areas – the latter introducing an absolute measurement of segregation. Together, the two offer a comprehensive overview of the geography of residential segregation in a multi-ethnic city.

2. On segregation

The concept of segregation is used in much social science and general writing to refer to both pattern and/or process (Johnston et al., 2005). As pattern, it refers to the extent to which members of different groups live apart from each other within the urban fabric – i.e. the degree to which they live in relatively exclusive, separate neighbourhoods, either through (cultural) choice or because of disadvantage (if not discrimination) in labour and housing markets. As process, it refers to the degree to which the geography is moving towards or away from that end state – complete separation in distinct areas – before addressing questions as to the causes stimulating such movement. The approach here focuses entirely on pattern-description, recognising that it could readily be adapted to address process-analysis as well.

Most students of ethnic residential segregation have used single-index measures to summarise the spatial pattern, recognising – following Massey and Denton (1988) – that five separate dimensions to the degree of segregation can be identified: unevenness, isolation, centralisation, concentration and clustering. Most focus on the first two of these – especially the first. All suffer from the same basic problem. As single-number indices, they identify the average situation without any intimation of the degree of variation about that figure. Thus, for example, an index of isolation of 0.4 for ethnic group *x* indicates the probability that any member of group *x* being selected at random from within the city will encounter another member of the same group living in the same area (i.e. the areal unit deployed for the index measurement

procedure, such as a local government ward) also selected at random.[1] But it says nothing about the variation about that average: does the probability of 0.4 apply to all members of *x*, wherever they live in the city, or are there some members with very different probabilities – those living in areas almost exclusively populated by members of *x* could have a probability of 0.95 whereas others, living in areas almost exclusively housing members of *non-x*, could have a probability of 0.005? Knowledge of such variation is crucial given the definition of segregation above: if there is almost complete segregation this will be shown by an index close to 1.0 and an almost complete lack of segregation will produce an index close to 0.0, but an index of 0.4 on its own tells us nothing about the proportion (if any) of *x* living largely apart from members of *non-x* in separate, exclusive (or relatively so) residential areas.

A further difficulty with the commonly-deployed measures of unevenness and isolation is that although geographical units – census tracts, wards etc. – are fundamental to their calculation nevertheless they are in a very important sense ageographical. Thus if all members of *x* lived in spatial units that were exclusive to them but those areas were randomly distributed across the urban area you would get the same indices of segregation (one of the commonly-used measures of unevenness) as would be produced if all of those areas were clustered into one part of the urban area only;[2] an index of 0.4, for example, would tell you that 40 per cent of ethnic group *x* would have to be redistributed across the areal units for it to be distributed in the same proportion in each unit as the remainder of the population, but not whether the areas where members of *x* were relatively concentrated formed a single territorial block rather than being widely distributed across the city. Because of this, often termed the checkerboard problem, analyses need to take account of the geography of the segregation as well as its intensity. Some have essayed this by combining indices of, say, unevenness and clustering (see Reardon, 2006; Reardon and O'Sullivan, 2004), but these pose considerable problems of interpretation and are, of necessity, also single-index measures without any representation of variation: is there a single

---

[1] The formula for the index of isolation is:

$$_{i}II_{i} = \sum_{j}^{n} (x_{ij} / X_{i}) * (x_{ij} / t_{j})$$

where
x_{ij} is the number of members of group i living in area j;
X_i is the total number of members of group i living in the city;
t_j is the total population of area j;
summation is over all n areas into which the city is divided; and
$_{i}II_{i}$ is the index of isolation for group i.

[2] The formula for the index of segregation is:

$$IS_{i} = [\sum_{j}^{n} \left| x_{ij} - x_{rj} \right|] / 2$$

where
x_{ij} is the proportion of the city's total population of group i living in area j;
x_{rj} is the proportion of the city's total population excluding group i living in area j;
summation is over all n areas into which the city is divided; and
IS_i the index of segregation for group i.

cluster, or are there several? do most members of *x* live in relatively exclusive areas, or in mixed areas, or…?

3. Introducing local statistics

Although index-based measures of segregation can usefully summarise a general pattern, therefore, they fail to illustrate many features of a spatial distribution and to address the question of how apart different groups live from each other in a comprehensive way. To provide the needed answers, we introduce methods of local statistics developed for the analysis of complex mapped distributions. Having introduced these, and illustrated their positive qualities, we then combine them with a further approach, thereby making much greater use of the available census data than the single-index numbers.

Spatial analysts have long employed general measures of spatial clustering and/or autocorrelation. Of these, Moran's I is probably most widely used (Moran, 1950). It is the ratio of the cross-products of values of the variable in question – such as the percentage of the local population who claim Bangladeshi ethnicity – for spatially-adjacent observations to the cross-products of all possible values, and can be interpreted in the same way as a correlation coefficient. An associated *Z*-value indicates the probability that this ratio is significantly different from 0.0. The formula is :

$$I = [[\sum_{i}^{n} \sum_{j}^{n} w_{ij} * \{(x_i - \overline{X}) * (x_j - \overline{X})\}] / \sum(x_i - \overline{X})^2] * n / S_o$$

where
x_i and x_j are the percentages of the population of areas i and j respectively in ethnic group x;
$\overline{X}$ is the mean percentage of the population of all areas in ethnic group x;
w_{ij} is the spatial proximity weight for areas i and j, coded 1 if they are adjacent and 0 otherwise;
n is the number of areas into which the city is divided;
S_o is the sum of all w_{ij} across all *n* areas; and
I is the value of Moran's *I*.

Like other indices of clustering, Moran's *I* is a system-wide average which provides no indication of the degree of variation. More information is provided by local measures of clustering, developed on the same principles as Moran's I but which focus on variations across the map rather than its overall pattern. One such measure, whose use is explored here, is Getis-Ord's *G** (Getis and Ord, 1992; Ord and Getis, 1995, 2001).

$$G_i^* = [\sum_{j}^{n} w_{ij}(x_j - \overline{X}) \sum_{j}^{n} w_{ij}] / [S \sqrt{\{n \sum_{j}^{n} w_{ij}^2 - (\sum_{j}^{n} w_{ij})^2\} / (n-1)}]$$

where
x_j is the percentage of the population of area j in ethnic group *x*;
$\overline{X}$ is the mean percentage of the population of all areas in ethnic group x;
w_{ij} is the spatial proximity weight for areas i and j, coded 1 if j is within *d* metres of i, and 0 otherwise;

n is the number of areas into which the city is divided;
S is

$$\sqrt{\ [\{\sum_{j}^{n} \quad x_j^2\}/n\ ] - \ \overline{X}^2}$$

d in this case is 1000 metres; and
and
G_i^* is the value (distributed as *Z*) for area i.

This takes the separate areas into which the map has been divided and for each area *i* calculates a ratio between the cross-products of all pairs of *x* with areas *j* within a fixed distance of *i* (*z* meters between their centroids, say) in their value of the variable of interest, *x*, and, as with Moran's *I*, the cross-products of all possible pairs. This is derived as a *Z*-Value, thereby indicating the degree to which each area *i* is significantly clustered with neighbouring areas having a similar value – either above or below the overall average – on the variable under consideration. A significant positive *Z*-value indicates an areal unit with a higher than average value on variable *x* whose neighbouring areas have a similar higher than average values. An insignificant *Z*-value indicates that neighbouring areas do not have similar relatively high or low values, and a significant negative *Z* indicates a cluster of neighbouring areas with similarly lower-than-average values.

Clearly, the detailed results of such an analysis will depend on the size of the neighbouring area incorporated in the distance band. The selection of the fixed distance threshold needs to reflect both the average areal extent of the spatial units being analysed and the anticipated size of local residential communities. Ideally, it should be large enough to ensure that at least four surrounding areal units are involved in the local averaging process but the entire area should be on average no more than about two kilometres radius from the observation unit's centroid. For the analyses of Greater London reported here the average areal extent of an Output Area (OA: the census areal unit employed) was 0.67 hectares and with a 1000-metre fixed distance threshold on average there were 47.5 OAs in each calculation.[3]

4. Ethnic residential segregation in London, 2001

We have data for the six main ethnic groups in London according to the typology of self-assessed ethnicity used in the 2001 Census – Bangladeshi, Indian, Pakistani, Black African, Black Caribbean, and White – at the Output Area scale. Output Areas (OAs) are small bespoke areal units defined for reporting the census data. They were defined to produce relatively homogeneous areas on two variables – housing type and housing tenure – with a contiguity constraint, nested within local authority electoral wards (Martin, 2001, 2002). There were 24,140 OAs for Greater London, with an average population of 297 (standard deviation, 66).

[3] There is clearly considerable potential for experimentation with different thresholds to identify scale effects in the pattern of segregation, as undertaken in a different context by Reardon et al. (2008, 2009)

The values of Moran's *I* for the percentages of each ethnic group across those OAs are in Table 1.[4] These show very high levels of spatial autocorrelation, with associated *Z*-values indicating very high levels of statistical significance. All six groups are strongly clustered into certain parts of the urban area, therefore, but the nature of the clusters – where are they? does each group have only one cluster? what proportion of the group's members live in those clusters? do the clusters comprise exclusive residential areas for the respective groups? etc. – cannot be derived from those single indices. We know there is substantial segregation, but little else; another example of the single-index number problem.

The clusters

For each ethnic group, *G** analysis produces a vector of 24,140 *Z*-values, whose geographies are best appreciated through mapping. Using a *Z*-value of 2.58 as the basic threshold (i.e. the cut-off point for statistically-significant differences at the 0.01 level), we divide London into the areas where each group is significantly clustered in either above-average or below-average proportions (positive and negative *Z*-values greater than 2.58 respectively), and where there is no such clustering (insignificant *Z*-values).

Each map shows a particular geography of clustering with its own features that single-index numbers could not identify. That for Bangladeshis is the simplest (Figure 1): there is significant clustering of OAs with above-average Bangladeshi percentages in eastern London – most of them in the adjacent boroughs of Newham and Tower Hamlets, with a second concentration in Camden – and a small area of significant negative *Z*-values (in Hammersmith & Fulham and neighbouring Kensington & Chelsea). The remainder of Greater London has insignificant *Z*-values, indicating an absence of significant clustering of areas with either above- or below-average concentrations of Bangladeshis.

In contrast to the relatively simple map for Bangladeshis, those for Indians and Pakistanis (Figures 2-3) show that each has a major concentration in both the city's northwest and northeast – more extensive in the former case for Indians and in the latter case for Pakistanis – as well as a smaller one in the inner southwest. In addition, and in clear contrast to the situation for Bangladeshis, there is also a substantial swathe of inner London with negative *Z*-values, indicating large continuous tracts where neighbouring OAs all have lower-than-average percentages of these groups living there. The maps for Black Africans and Black Caribbeans (Figures 4-5) have a considerable amount in common: main clusters to the south and north of the Thames and a further cluster – more marked for Black Caribbeans – in the inner northwest. Much of the rest of London has extensive continuous areas of negative *Z*-values, indicating the relative absence of these two groups from much of the city's outer suburbs.

The final map – for those claiming White ethnicity – is very different again (Figure 6). Not surprisingly, it is to a considerable extent the inverse of the previous five: the areas where the main ethnic minority groups are clustered – in the east, west and inner

[4] All of the analyses reported here using *I* and *G** were undertaken within the ArcGIS© software.

south – have negative *Z*-values, indicating the relative absence of Whites from large continuous tracts. Most of the outer suburbs, on the other hand, plus favoured inner city areas (along both banks of the river upstream from Westminster and around Hampstead) have positive Z-values – areas where Whites are clustered to the relative exclusion of members of the other five groups.

One feature not displayed in Figures 1-6 – which cannot readily be done because of the complexity involved – is any difference across the six groups in the degree of concentration within the areas where each is significantly clustered. A *Z*-value of 2.58 has been chosen as the threshold for cluster definition because it represents a very significant variation from a random distribution, but many observed values are much larger. For example, whereas the maximum observed *Z*-value for an OA was only 9.7 for Whites, it was 29.6 and 38.8 for Black Caribbeans and Black Africans respectively, 42.5 and 48.6 for Indians and Pakistanis and 80.6 for Bangladeshis. There are thus clear differences in the intensity of the clustering – of high above-average values across blocks of neighbouring OAs. These are emphasised in Table 2, the first block of which shows the number of OAs with a range of positive *Z*-values for each group. These suggest a clear patterning of extreme clustering – of neighbouring areas all with similarly high above-average percentages of the relevant group compared to an even distribution across the whole of Greater London. The most intensively clustered are Bangladeshis, followed by the other two south Asian groups and the two Black groups, with Whites having the least intensive clustering. For each of the groups, therefore, there is a spatial ordering of their major clusters: cores with very intensive clustering surrounded by peripheries where it is less so, although still above the 2.58 threshold.

The lower block of Table 2 shows no comparable intensity in the clustering of areas which have lower-than-average percentages of the relevant groups. For the five minority groups, only six OAs have a negative *Z*-value in excess of 10.32 and only in the case of the majority Whites is there intense clustering of areas from which they are relatively absent. Only for the majority population, therefore, are there substantial tracts where neighbouring OAs have very low percentages – areas where non-Whites predominate.

Cluster populations

The *Z*-values associated with the *G** statistics indicate the where and the intensity of ethnic clustering within the city, but leave unaddressed the absolute concentration of groups into those clusters and their exclusivity as separate residential areas. Different groups may be relatively concentrated in different parts of the city, but still not be highly segregated there.

Table 3 shows the percentage of each of the six groups living in the relevant three main cluster areas according to the key thresholds of *Z*=+/-2.58, plus a subdivision of the positive clusters according to the relative size of the *Z*-values. There is a clear distinction between the five minority groups, on one hand, and the majority Whites, on the other. Whereas 60-68 per cent of the former live in their group's main clusters, the comparable figure is less than 50 for Whites. The latter, in turn and not surprisingly given their relative importance within London's population, are much more likely to live in the areas where they are relatively absent; members of all

groups are about equally likely to live in the more mixed areas (insignificant *Z*-values: note that there may be individual OAs in these areas where one group predominates, for example, but that this characteristic is not shared with adjacent areas).

There are also differences among the five minority groups, which parallel their degree of clustering. Bangladeshis and, to a slightly lesser extent, Indians and Pakistanis are more concentrated in their respective areas of extreme clustering than are Black Africans and Caribbeans: some 40 per cent of Bangladeshis are in OAs with Z-values in excess of 20.65, for example, as are 21 and 25 per cent of Indians and Pakistanis respectively, but only 9 and 8 per cent of the two Black minority groups.

Clustering and residential exclusivity

These maps and the associated tables give a very clear impression of London's ethnic geography, highlighting those parts of the urban fabric in which each group is concentrated and others from which it is relatively absent. They provide greater insight into a complex geography than single index values such as Moran's *I* could deliver. But they do not address the issue of what proportion, if any, of group members are living in relatively exclusive residential enclaves. The *G** values indicate relative clustering only – they identify the blocks of OAs where the relevant group is on average either over- or under-represented relative to its global percentage of the metropolitan population. Having identified those areas where the groups are clustered, therefore, it is then necessary – given the definition of segregation as the degree to which groups live apart from others in relatively exclusive areas – to inquire into the ethnic composition of OAs (their degree of homogeneity) within the clusters, adding an absolute measure of segregation to the relative one provided by *G**.

Tables 4 and 5 address this issue. Table 4 presents the distribution of OAs within the major clusters, according to the percentage that the relevant ethnic group contributes to the OA population. Thus, for example, within the major Bangladeshi clusters (shown in Figure 1), 79 per cent of those 2256 OAs (Table 2) have a Bangladeshi percentage of less than 20 and in only 3 per cent do Bangladeshis comprise as much as 60 per cent of the OA total. Bangladeshis are clustered together into certain parts of London, therefore, but do not dominate the local population there save in a small proportion of the constituent Output Areas. A very similar picture emerges for the other four minority ethnic groups: indeed, with the partial exception of the Indians, they are even less dominant in the areas where they are clustered than are Bangladeshis. For Pakistanis, Black Africans and Black Caribbeans, there are virtually no OAs within their significant clusters where they form even 40 per cent of the local population. There is separation and living in general proximity to co-ethnics, but not segregation, except to some extent from the majority White population.

The opposite is the case with the majority White population and the 10,048 OAs (41 per cent of London's total) where it is significantly clustered. Here Whites form at least 60 per cent of the total population in virtually every OA, and over 80 per cent in the vast majority. Thus whereas in most of those parts of London where members of the minority ethnic groups are significantly clustered their members are in a minority in nearly all of those areas' local populations, where Whites are clustered they are invariably in a substantial majority. In terms of living in relatively exclusive

residential areas, therefore, this is a characteristic of London's majority population, not its ethnic minorities who live in much more mixed contexts.

This conclusion is reinforced by the data in Table 5, which show the percentage distribution of each of the six groups according to the clusters (negative, insignificant, positive) and the population composition of the OAs in each. Thus, for example, only one per cent of all London's Bangladeshis live in the small number of OAs from which they are relatively absent and which are clustered together (the western negative cluster in Figure 1); not surprisingly, Bangladeshis form less than 20 per cent of the population there. Just under one-third of Bangladeshis live in the relatively mixed areas where they are neither positively nor negatively significantly concentrated; again, all of them are in areas where Bangladeshis form less than 20 per cent of the population. The remaining 68 per cent of all of London's Bangladeshis live in the positive clusters, with 15 per cent in OAs where Bangladeshis are in a clear majority.

Around one-tenth of each the other four minority groups live in OAs where they are negatively clustered and 20-30 per cent in the areas where there is no significant clustering: as with the Bangladeshis, almost all of these are OAs where the relevant group forms less than one-fifth of the total population. Within the areas where they are significantly clustered, most of the members of three groups live in OAs where they form less than one-fifth of the population; they cluster together, but at relatively low densities and they do not dominate the local population. The partial exception is with the Indians, around 40 per cent of whom live in OAs within their clusters where they form a substantial portion – though rarely a majority – of the population.

One clear conclusion from these two tables, therefore, is that although each of London's five main ethnic minority groups is concentrated into clearly-defined territorial clusters, those ethnic enclaves are not exclusive residential areas where the group predominates, or even dominates. Indeed, in most parts of those clusters they form not only a minority but a relatively small minority of the total population.

This is not the case with the White majority, however, who are the most segregated of the city's ethnic groups. Even in the clusters where they are significantly under-represented (where the relevant *Z*-values are less than -2.58), they form a substantial proportion of the population in most of the constituent OAs. The same applies in the parts of the city where there is no clustering, and in those areas where Whites are significantly clustered; unlike the other groups, Whites are in the majority in most of those OAs (with over half of all of London's Whites living in OAs within those clusters where they form at least 80 per cent of the local population).

Overlapping clustering

So far we have treated each ethnic group separately and identified that whereas the majority of London's Whites live in areas where Whites are clustered and predominate, the majority of the members of the five minority ethnic groups live in clusters where their group is over-represented but nevertheless does not form a majority of the local population (in many cases, not even a substantial minority). But do the group clusters overlap to any extent? To answer this, we can cross-classify the cluster maps.

Table 6 categorises each OA separately for each of the six groups into one of three types: negative clusters (i.e. in an area where the group is significantly under-represented in neighbouring OAs: $Z<-2.58$); mixed areas, where there is no significantly clustering, either positive or negative ($Z>-2.58$ & $Z<2.58$); and positive clusters ($Z>2.58$). Each pair of groups is the subject of a separate 3 x 3 table, with every cell having two percentages reported: the first is the percentage of the row total and the second the percentage of the column total. Thus in the Bangladeshi-Indian comparison, 94 per cent of all OAs which are in Bangladeshi negative clusters are also in Indian negative clusters, whereas only 7 per cent of all OAs in Indian negative clusters are also in Bangladeshi negative clusters.

Of the nine cells in each of these matrices, those in the bottom-right corner carry most information regarding cluster overlap; they show the extent to which the areas where one group is positively clustered are shared with another group also positively clustered there. One of the largest pair of such values is for Indians and Pakistanis: 61 per cent of the OAs in the Indian positive clusters are also included in the Pakistani positive clusters, and the respective percentage for the other comparison is 64. The two maps overlap very substantially, therefore, as further exemplified by the top-left cell in that matrix: 77 per cent of the OAs in the Indian negative clusters are also in the Pakistani negative clusters, and the comparable Pakistani-Indian figure is 88 per cent. There is a great deal of sharing residential space involving London's Indian and Pakistani populations, which also share space – to a lesser extent – with the other three minority groups.

The other large pair of cell values is in the matrix for Black Africans and Black Caribbeans: 78 per cent of the OAs in the former group's positive clusters are also in the latter's, with a reverse figure of 71 per cent. Where there is clustering of one of London's Black ethnic minority groups, therefore, in most cases a significant clustering of the other group can also be found.

The clear exception to this overlapping of the cluster maps is shown by the five comparisons involving London's White population. Indeed, there is virtually no overlapping at all, with at most only one per cent of the OAs in the White positive clusters also being in one of the other group's positive clusters. Where Whites are clustered, the other groups are absent.

Ethnic mixing

These findings suggest a marked difference between two main types of segregated residential areas in Greater London, therefore: on the one hand there are the, mainly suburban, areas where Whites are clustered into exclusive residential areas; and on the other are the areas where members of the main minority groups are clustered but the populations are nevertheless ethnically diverse. To capture that difference, and in particular the diversity of the latter type, we use a classification of residential areas developed for comparative purposes (Poulsen et al., 2001; Johnston et al., 2007).[5]

[5] An alternative approach, developing on the typology but not addressing the checkerboard problem, is Brimicombe (2007).

This empirical typology – constructed after evaluation of a range of options as providing a valuable overview of structural differences – places each OA into one of five categories according to its population composition:

I – the White population exceeds 80 per cent of the total;
II – the White population is between 50 and 80 per cent of the population, inclusive;
III – the White population is less than 50 and more than 30 per cent of the population;
IV – the White population is 30 per cent or less of the population;
V – the White population is 30 per cent or less of the population, and one of the minority ethnic groups is at least two-thirds of the total non-White population there.

The first two types are thus areas where the White population either predominates or dominates respectively; the next two have White minorities. Types IV-V both have non-Whites predominating: in Type V one of those minority groups predominates within the non-White total, thus distinguishing such areas from those in Type IV which are characterised by a mixed ethnic population. OAs in Type V are separately identified according to which non-White group predominates.

Tables 7 and 8 illustrate two ways in which this typology can be deployed. The first looks at the 2256 OAs in the Bangladeshi positive cluster (i.e. the areas identified in Figure 1), and shows the distribution of all of the residents there according to the type of OA in which they lived. Thus, for example, one-quarter of the 104,073 Bangladeshis living in those clusters were in Type V areas where they predominated – compared to very small percentages of members of each of the other groups. Of the other Bangladeshis living in the cluster, nearly 40 per cent lived in OAs with a non-White majority, but 36 were in areas with a White majority. A small part of the cluster comprised relatively exclusive Bangladeshi neighbourhoods, therefore, but much of it was made up of relatively mixed areas ethnically – containing around three-quarters of all of the Indians and Pakistanis living there. Most of the 358,000 Whites found in those clusters where Bangladeshis were concentrated lived in areas with White majorities – one-third of them in OAs where Whites predominated. The Bangladeshi clusters are far from extensive tracts of exclusive Bangladeshi residential areas, therefore; most of their constituent OAs have a mixed population ethnically.

Table 8 allows this conclusion to be compared with the situation in all six sets of positive clusters, showing the distributions for the total population there only (thus the column for the Bangladeshi clusters is the same as the final column in Table 7). One clear conclusion to be drawn is the very small proportion of London's population that was living in areas where one of the minority ethnic groups predominated at the time of the 2001 Census (just 6 per cent of the 695,000 living in the Bangladeshi clusters, 5 per cent of the 1.3 and 1.2 million living in the Indian and Pakistani clusters respectively – all of them in areas where Indians predominated; there were no Type V OAs with Pakistani predominance, and 1 per cent of the nearly 2 million living in the Black African clusters: there were no Type V areas with Black Caribbean predominance). Contrast this situation with that for the areas where Whites were significantly clustered; 95 per cent of the nearly 3 million people living there were in areas where Whites predominated.

5. Conclusions

Many different ways have been suggested for depicting the degree of segregation of ethnic groups across an urban area's residential fabric. Most of them rely on single indices which, although constructed using small geographical units, nevertheless are ageographical. Thus, for example, indices of segregation of 0.71 and 0.40 respectively for London's Bangladeshis and Whites tell us that 71 and 40 per cent of those groups would have to be redistributed across the city's OAs to a different configuration to achieve an even distribution relative to the rest of the population (i.e. so that Whites form 71 per cent of every OA's population and Bangladeshis 1.9 per cent). But this index of unevenness tells us very little about either the ethnic composition of the neighbourhoods in which members of those groups live or the degree to which the areas where they are relatively concentrated are clustered together within the urban fabric. Similarly, indices of isolation of 0.36 and 0.77 for Bangladeshis and Whites respectively suggest that the latter are much more segregated than the former, but again tell us very little about the varying neighbourhood contexts in which individual Bangladeshis and Whites live.[6]

Single-number indices reveal a little about the potentially complex geography of segregation in a multi-cultural city, therefore, but conceal a great deal more. They take no account of two extremely salient features of that geography – the extent to which any ethnic group's members are spatially concentrated into particular parts of the city and the degree to which they share residential spaces with members of different groups. Addressing those features calls for different approaches, more nuanced than alternative single-number indices – such as Moran's *I* – which do look at such issues as clustering, can provide.

This paper has proposed and reported on an initial exploration using an alternative approach to identifying the nature of London's ethnic segregation based on a combination of geostatistical measures and a typology of areas according to their ethnic composition. Its conclusions highlight three important features of that segregation:

- The urban area is divided into three main segments. The first comprises those parts of London in which members of the White majority are clustered, and where they predominate in the local population. The second consists of portions of the residential fabric where there is no significant clustering of any ethnic group, no tracts of territory where neighbouring areas have similar concentrations of one or more of the groups. Finally, there are sections of London where members of the main ethnic minorities are clustered. In total, 41 per cent of Londoners lived in the first segment in 2001, 9 per cent in the second, and exactly half in the third.
- The first and third of those segments – London's ethnically segregated residential areas, comprising over 90 per cent of the total – were very different in their character. The former, the clusters where Whites formed above-average proportions of the population, were almost exclusively White, indicating a high degree of residential segregation for half of London's White

[6] Modified to take account of the size of each group, following Cutler et al. (1999), the indices for the two are .24 for Bangladeshis and .20 for Whites, suggesting that each group is 20-25 per cent more spatially segregated than if they were evenly distributed across all OAs.

population. In contrast, very few of the OAs within the segment of London where the various non-White groups were clustered comprised exclusive residential areas where one of those groups predominated: multi-ethnic mix, with a substantial White component, was the norm there.

- The geography of those segments shows a complex patterning of the areas where the various ethnic groups are clustered. For the Whites it involves a series of, mainly suburban, blocks of territory where they predominate in the local demographic structure. For the five non-White groups, each has a number of clusters where they form above-average proportions of the total population; none is concentrated into a single cluster, however.

Overall, therefore, the city is divided into some exclusive White suburbs surrounding a series of enclaves characterised more by their multi-ethnic diversity than by the dominance of a particular group. Each of the five has parts of the city where it is more prevalent than the others, but in none are they as segregated as the majority population.

These conclusions draw on a combination of two types of analysis, local statistics and an are typology, which avoid the major problem of single-number indices of segregation – of unevenness, isolation and clustering; the lack of any indication of either variation around the average situation or the degree to which various groups live in exclusive residential areas. Bringing together a spatial analysis of each ethnic group's relative clustering with a portrayal of the ethnic composition of areas both within and outwith those clusters provides the insights that the over-simplified approaches using index-numbers conceal. The contours of London's multi-ethnic geography are revealed in ways that address the fundamental question regarding residential segregation – to what extent do various groups live apart from others in relatively exclusive areas, and where?

The detailed answers to that question provided here are contingent on a number of decisions – for example, to use a distance band of 1000 metres around each OA when undertaking the G* analyses, and a *Z*-score of 2.58 as the threshold for determining the territories where groups are clustered. Different distance bands (and distance functions; instead of giving all OAs within the set distance equal weight in the calculations, distance could itself be weighted, giving greater emphasis to the most proximate OAs within the chosen band) and different thresholds would lead to a different configuration of clusters being identified, with implications for some of the derived statistics. It is very unlikely that the overall patterns identified here would be contradicted, however – their lineaments are very clear and the differences would be in the detail only. Further investigations are needed to explore their nature and extend the utility of this potentially valuable complementary approach to the study of residential segregation.

Public debate about the emerging nature of Britain's multi-cultural, multi-ethnic society has been characterised in recent years by scare stories about the country 'sleepwalking towards segregation' with the portent of emerging ghettos. These expectations have never had any basis in academic analyses (Johnston and Poulsen, 2006), and comparisons with the situation elsewhere, especially the United States, show very clearly that the extremes identified there are not a feature of British cities too. Nevertheless, continuous monitoring of the situation is desirable, using methods

such as those outlined here which can provide a comprehensive overview rather than a single-figure, often misleading, average.

References

Brimicombe A. (2007) Ethnicity, religion and residential segregation in London: evidence from a computational typology of minority communities. *Environment and Planning B: Planning and Design*, 34, 884-904

Cutler, D. M., Glaeser, E. L. and Vigdor, J. L. (1999) The rise and decline of the American ghetto. *Journal of Political Economy*, 107, 456-506.

Getis, A. and Ord, J. K. (1992) The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24, 189-206.

Johnston, R. J. and Poulsen, M. F. (2006) Ethnic residential segregation in England: getting the right message across. *Environment and Planning A*, 38, 2195-2199.

Johnston, R. J., Poulsen, M. F. and Forrest, J. (2005) On the measurement and meaning of segregation: a response to Simpson. *Urban Studies*, 42, 1221-1227.

Johnston, R. J., Poulsen, M. F. and Forrest, J. (2007) The geography of ethnic residential segregation: a comparative study of five countries. *Annals of the Association of American* Geographers, 97, 713-738.

Martin, D. J. (2001) Optimising census geography: the separation of collection and output geographies. *International Journal of Geographic Information Science*, 12, 673-685.

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

Moran, P. A. P. (1950) Notes on continuous stochastic phenomena. *Biometrika*, 37, 17-33.

Massey, D.S. and Denton, N. S. (1988) The dimensions of residential segregation. *Social Forces*, 67, 281-315.

Ord, J. K. and Getis, A. (1995) Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis*, 27, 286-306.

Ord, J. K. and Getis, A. (2001) Testing for local spatial autocorrelation in the presence of global autocorrelation. *Journal of Regional Science*, 41, 411-432.

Poulsen, M. F., Johnston, R. J. and Forrest, J. (2001) Intraurban ethnic enclaves: introducing a knowledge-based classification method. *Environment and Planning A* 33: 2071-2082.

Reardon, S. F. (2006) A conceptual framework for measuring segregation and its association with population outcomes. In *Methods in Social Epidemiology* (eds. J. M. Oakes and J. S. Kaufman). San Francisco: Jossey-Bass, 169-192.

Reardon, S. F., Farrell, C. R., Matthews, S. A., 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, S. F., Matthews, S. A., O'Sullivan, D., Lee, B. A., Firebaugh, G., and Farrell, C. R. (2008) The geographic scale of metropolitan racial segregation. *Demography*, 45, 489-514.

Reardon, S. F. and O'Sullivan, D. (2004) Measures of spatial segregation. *Sociological Methodology*, 34, 121-62.

Simpson, L. (2007) Ghettos of the mind: the empirical behaviour of indices of segregation and diversity. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170, 405-424.

Table 1. The values of Moran's *I* for London's six ethnic groups, 2001

	B	I	P	BA	BC	W
I	0.86	0.57	0.67	0.60	0.73	0.65
Z	818.6	548.5	638.2	576.3	695.4	617.5

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

Table 2. The distribution of Output Areas where each group is either positively or negatively clustered

Z-value	B	I	P	BA	BC	W
Positive						
2.58 – 5.16	613	982	1073	2220	1517	4774
5.17 – 10.32	498	1089	1116	1943	2350	5274
10.33 – 20.64	591	1229	812	1693	2604	0
20.65 – 41.28	257	586	688	528	481	0
41.29 – 82.56	297	5	61	0	0	0
TOTAL	2256	3891	3750	6384	6952	10048
Negative						
2.58 – 5.16	868	9498	9892	8321	7556	2247
5.17 – 10.32	0	2144	333	1353	3584	3387
10.33 – 20.64	0	0	0	0	6	2237
20.65 – 41.28	0	0	0	0	0	243
TOTAL	868	11642	10225	9674	11146	8114

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

Table 3. The distribution of the members of each ethnic group according to the Z-value for their Output Areas.

Z-value	B	I	P	BA	BC	W
-2.58 –	0.5	12.7	9.6	11.1	12.0	25.7
-2.58 – 2.58	31.8	26.0	28.7	29.1	21.8	24.5
2.58 – 5.16	5.6	6.9	8.0	13.9	9.4	22.8
5.17 – 10.32	6.1	11.0	13.3	16.2	18.1	27.0
10.33 – 20.64	15.0	22.6	15.6	20.5	30.5	0
20.65 – 41.28	12.4	20.6	21.9	9.2	8.2	0
41.29 –	28.6	0.2	2.9	0	0	0
Total Z>2.58	*67.7*	*61.3*	*61.7*	*59.8*	*66.2*	*48.8*

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

Table 4. The distribution of Output Areas within the major clusters for each ethnic group (i.e. $Z>2.58$), according to the percentage the relevant group contributes to the OA's total population (per cent of the Output Areas in the clusters)

Group %	B	I	P	BA	BC	W
0 – 19	79	62	96	92	89	0
20 – 39	13	29	4	8	10	0
40 – 59	5	8	0	0	1	1
60 – 79	3	1	0	0	0	15
80-	0	0	0	0	0	85

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

Table 5. The distribution of each ethnic group according to the *Z*-value for each Output Area and the percentage that group contributes to the OA's total population (percentage of each group's total)

Group %	B	I	P	BA	BC	W
Z<-2.58						
0 – 19	1	13	10	11	12	0
20 – 39	0	0	0	0	0	3
40 – 59	0	0	0	0	0	11
60 – 79	0	0	0	0	0	10
80-	0	0	0	0	0	1
Z>-2.58 & Z<2.58						
0 – 19	31	25	29	29	21	0
20 – 39	0	1	0	0	1	0
40 – 59	0	0	0	0	0	2
60 – 79	0	0	0	0	0	14
80-	0	0	0	0	0	8
Z<2.58						
0 – 19	22	19	52	54	57	1
20 – 39	17	26	10	6	9	3
40 – 59	13	13	0	0	0	13
60 – 79	13	3	0	0	0	30
80-	2	0	0	0	0	53

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

Table 6. The overlapping of clusters: comparing each pair of ethnic groups. (The figures in each cell are, respectively, the percentage of all OAs classified according to the row category for the first-named group and the percentage of all OAs classified according to the column category for the second-named group.)

B/I	-	0	+	B/P	-	0	+	B/BA	-	0	+
-	94/7	6/1	0/0		82/7	18/2	0/0		71/6	29/3	1/0
0	45/80	39/96	16/88		39/80	47/96	15/82		42/91	34/87	25/81
+	65/13	14/4	21/12		59/13	10/2	31/18		12/3	35/10	53/19
B/BC	-	0	+	B/W	-	0	+	I/P	-	0	+
-	63/5	31/5	6/1		0/0	6/1	93/8		77/88	20/24	2/7
0	46/87	25/86	29/88		31/80	26/90	44/91		14/12	73/62	13/30
+	42/8	24/9	34/11		73/20	24/9	3/1		0/0	38/15	61/64
I/BA	-	0	+	I/PC	-	0	+	I/W	-	0	+
-	39/48	30/43	31/56		47/49	20/39	33/55		27/39	23/45	50/58
0	49/44	30/32	20/27		53/41	24/37	23/28		23/24	28/41	49/42
+	20/8	53/25	27/17		29/10	41/27	30/17		77/37	23/15	0/0
P/BA	-	0	+	P/BC	-	0	+	P/W	-	0	+
-	43/46	27/34	30/48		53/49	17/29	30/43		26/33	21/37	53/54
0	48/50	35/44	17/27		52/47	30/51	18/26		22/27	34/57	45/45
+	11/4	47/22	42/25		12/4	32/20	56/30		88/41	9/6	3/1
BA/BC	-	0	+	BA/W	-	0	+	BC/W	-	0	+
-	89/76	12/18	0/0		6/7	12/20	82/78		8/11	19/36	73/81
0	29/21	46/634	24/29		31/31	42/58	27/22		34/25	36/36	30/18
+	5/3	18/19	78/71		78/62	21/23	1/0		75/64	24/28	2/1

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

- $Z<-2.58$; 0 – $Z>-2.58$ & $Z<2.58$; + $Z>2.58$

Table 7. The percentage distribution of each ethnic group across different types of residential area (defined in the text) within the Bangladeshi positive clusters identified in Figure 1.

Type	B	I	P	BA	BC	W	∑
I	4	7	2	8	8	32	21
II	32	19	18	48	48	47	42
III	29	26	30	30	28	15	21
IV	10	47	48	12	13	4	11
V (B)	25	1	1	2	2	2	6
V (I)	0	0	0	0	0	0	0
V (BA)	0	0	0	0	0	0	0
TOTAL	104,073		28,415		35,342		695,240
		43,193		58,386		358,591	

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

Table 8. The percentage distribution of the total population across different types of residential area (defined in the text) within each ethnic group's positive clusters identified in Figures 1-6.

Type	B	I	P	BA	BC	W
I	21	14	9	18	20	95
II	42	48	49	60	60	5
III	21	25	28	17	16	0
IV	11	8	8	5	4	0
V (I)	0	5	5	0	0	0
V (B)	6	0	0	0	0	0
V (BA)	0	0	0	1	0	0
TOTAL	695,240		1,231,284		2,080,745	
		1,301,192		1,908,144		2,906,697

Key: B – Bangladeshis; I – Indians; P – Pakistanis; BA – Black Africans; BC – Black Caribbeans; W – Whites.

Captions

Figure 1. The clustering of those claiming Bangaldeshi ethnicity in London, 2001: the areas with *Z*-values of >2.58 have above-average proportions and those with *Z*-values of <-2.58 have below-average proportions of Bangaldeshis in the relevant OA and its near-neighbours.

Figure 2. The clustering of those claiming Indian ethnicity in London, 2001: the areas with *Z*-values of >2.58 have above-average proportions and those with *Z*-values of <-2.58 have below-average proportions of Indians in the relevant OA and its near-neighbours.

Figure 3. The clustering of those claiming Pakistani ethnicity in London, 2001: the areas with *Z*-values of >2.58 have above-average proportions and those with *Z*-values of <-2.58 have below-average proportions of Pakistanis in the relevant OA and its near-neighbours.

Figure 4. The clustering of those claiming Black African ethnicity in London, 2001: the areas with *Z*-values of >2.58 have above-average proportions and those with *Z*-values of <-2.58 have below-average proportions of Black Africans in the relevant OA and its near-neighbours..

Figure 5. The clustering of those claiming Black Caribbean ethnicity in London, 2001: the areas with *Z*-values of >2.58 have above-average proportions and those with *Z*-values of <-2.58 have below-average proportions of Black Caribbeans in the relevant OA and its near-neighbours..

Figure 6. The clustering of those claiming White ethnicity in London, 2001: the areas with *Z*-values of >2.58 have above-average proportions and those with *Z*-values of <-2.58 have below-average proportions of Whites in the relevant OA and its near-neighbours..

Figure 1

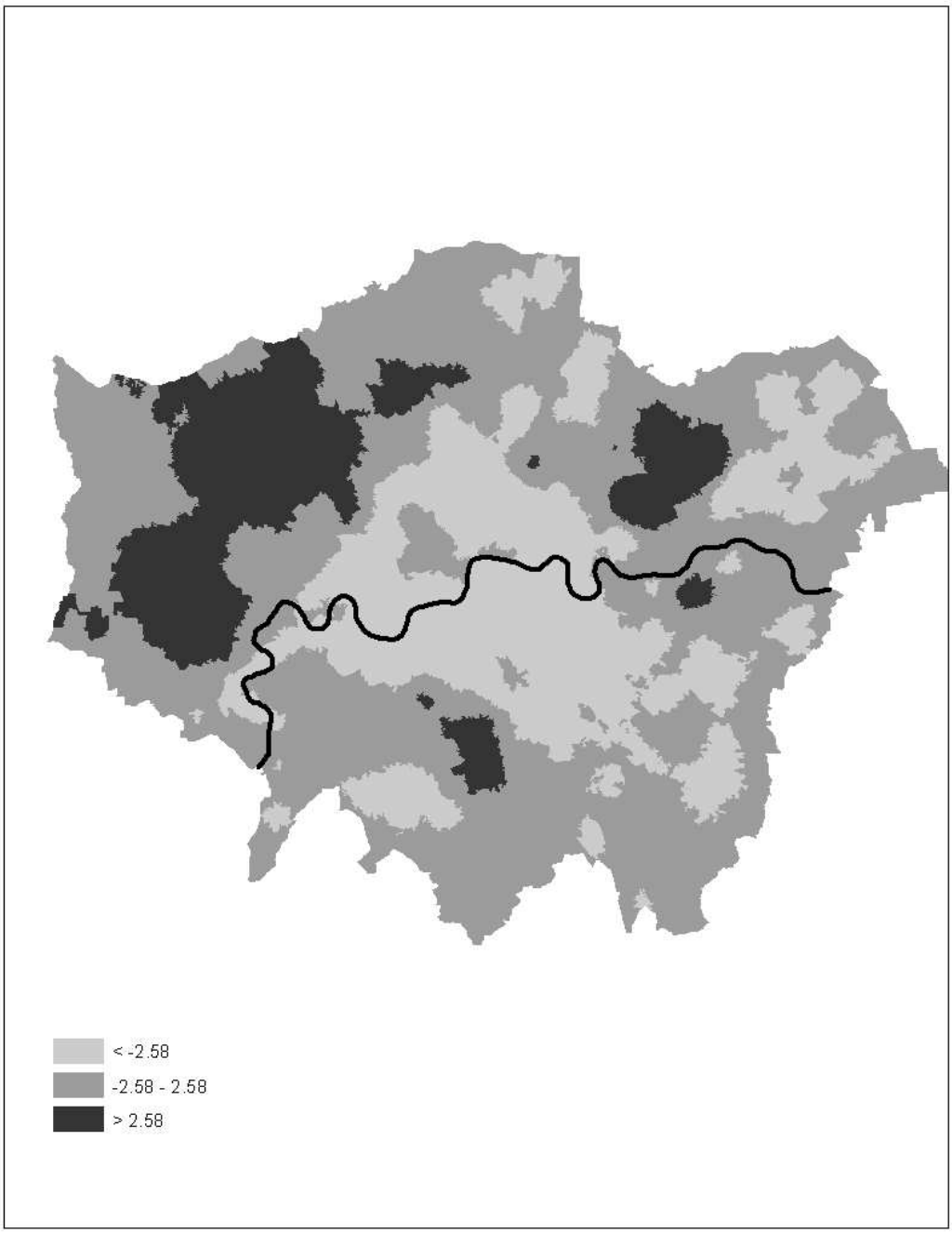

Figure 2

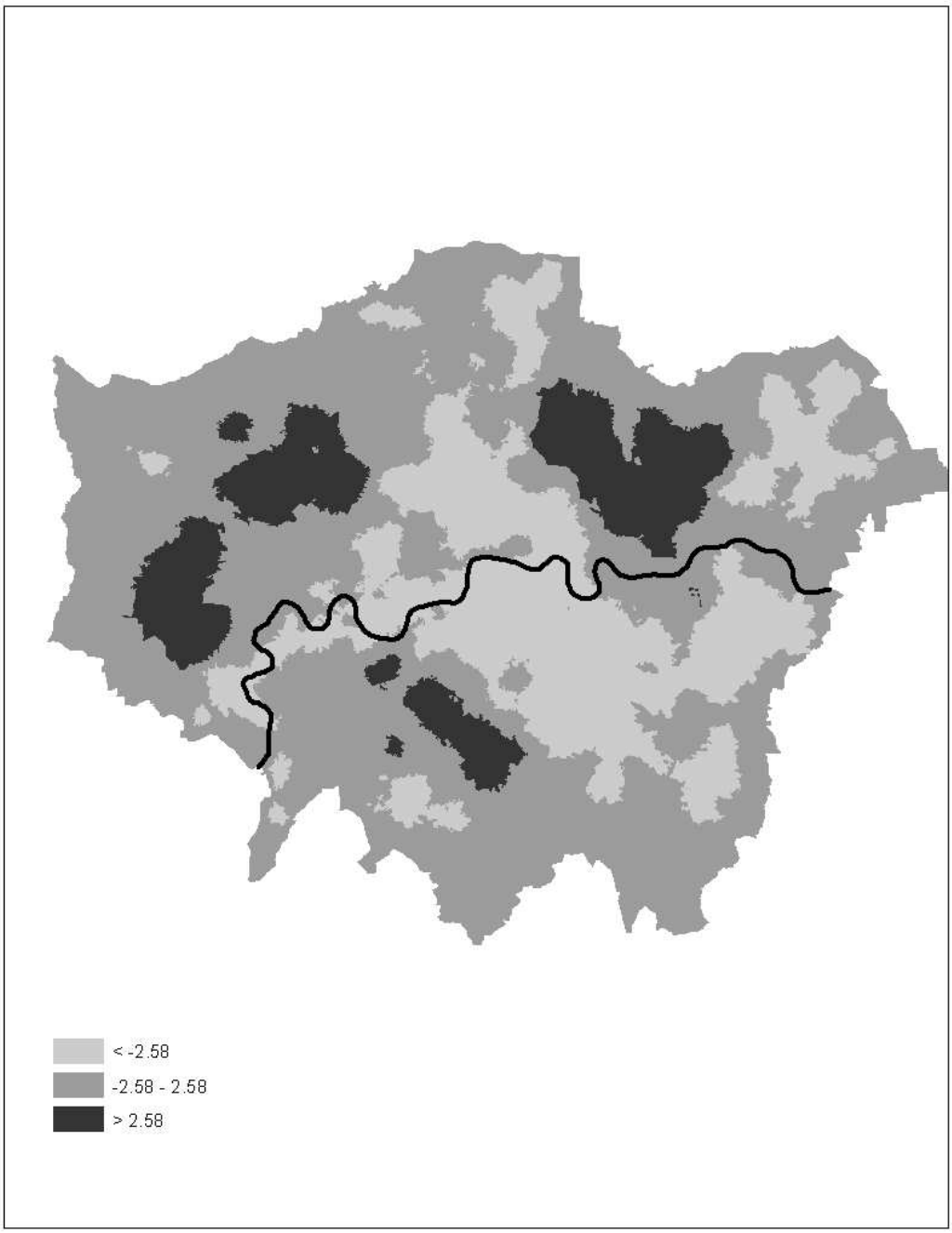

Figure 3

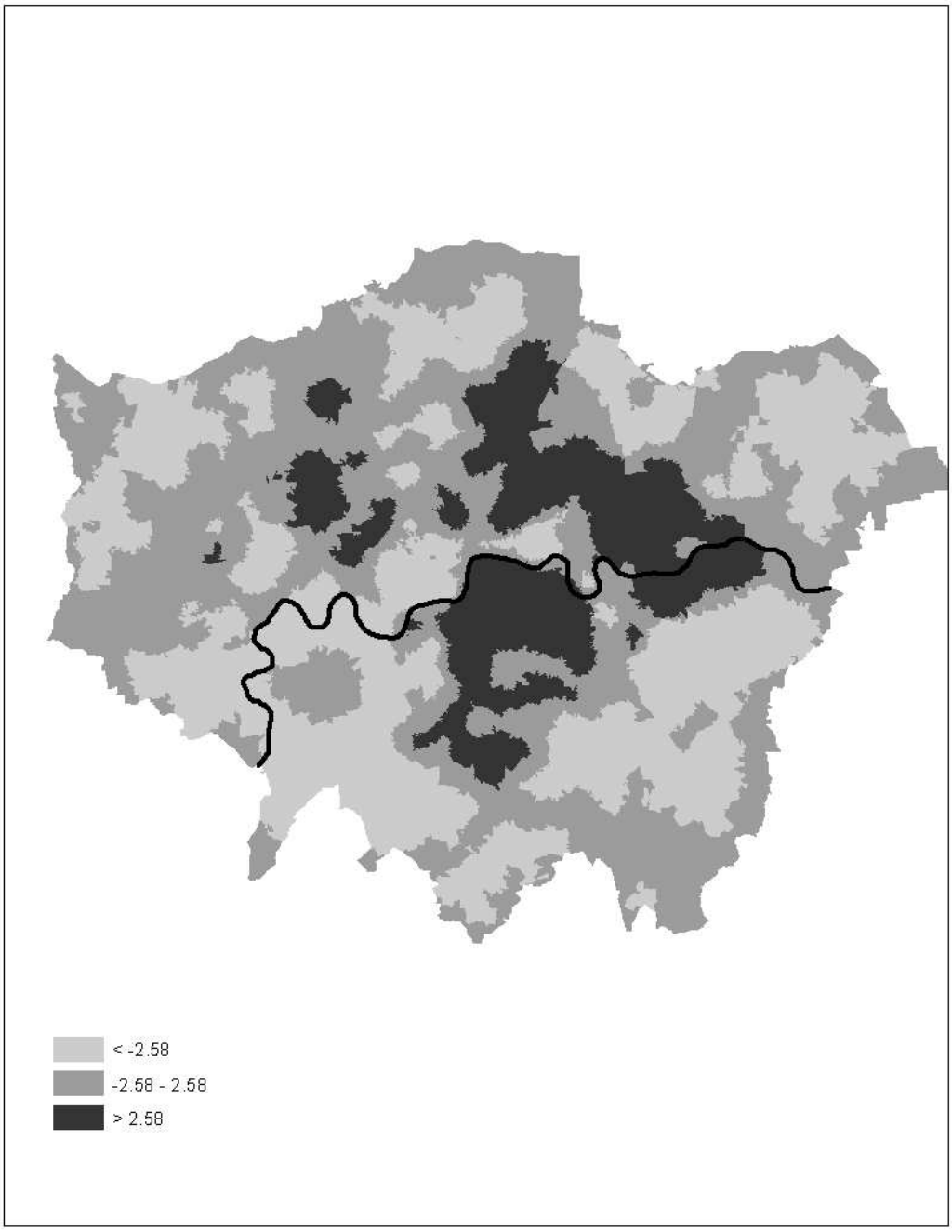

Figure 4

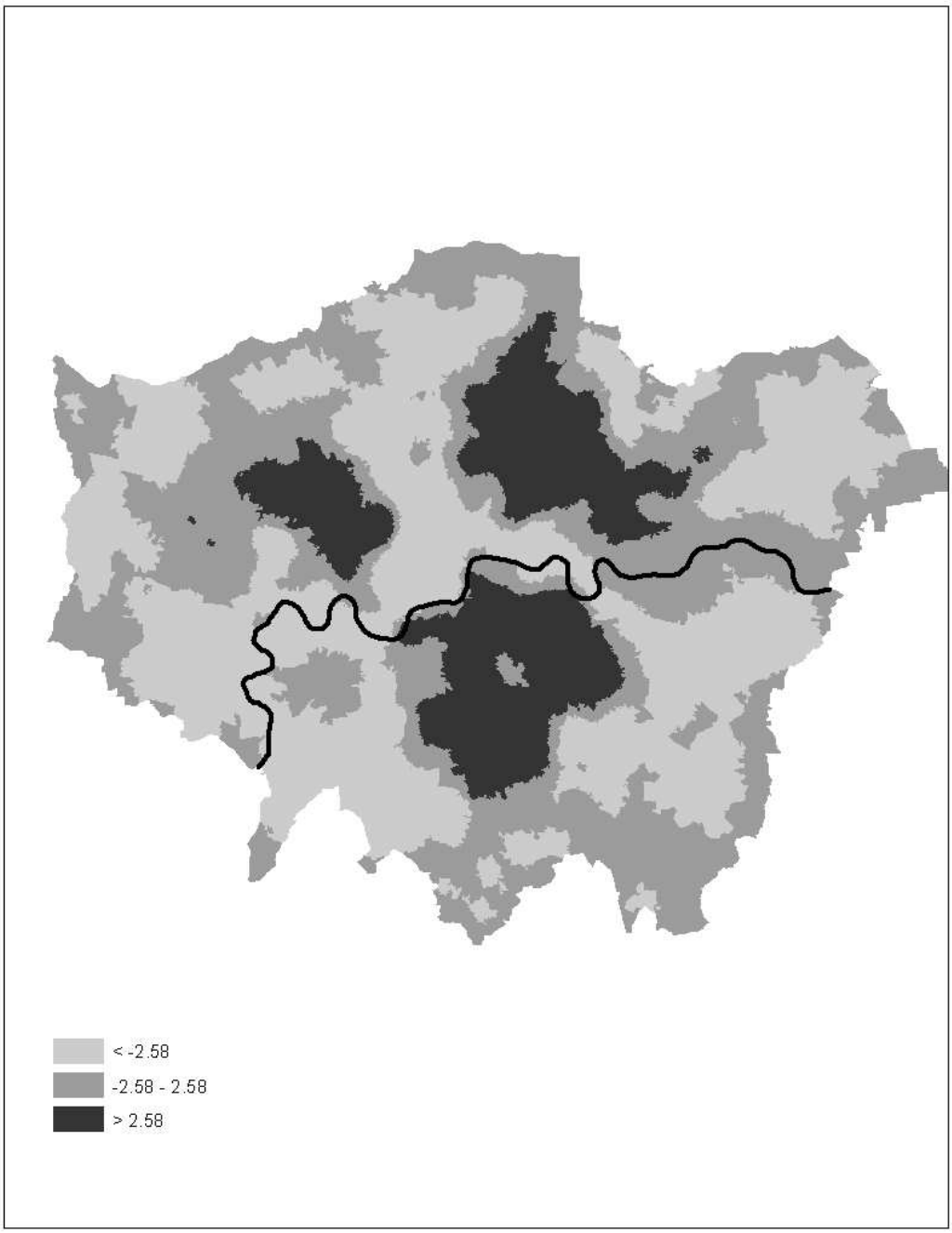

Figure 5

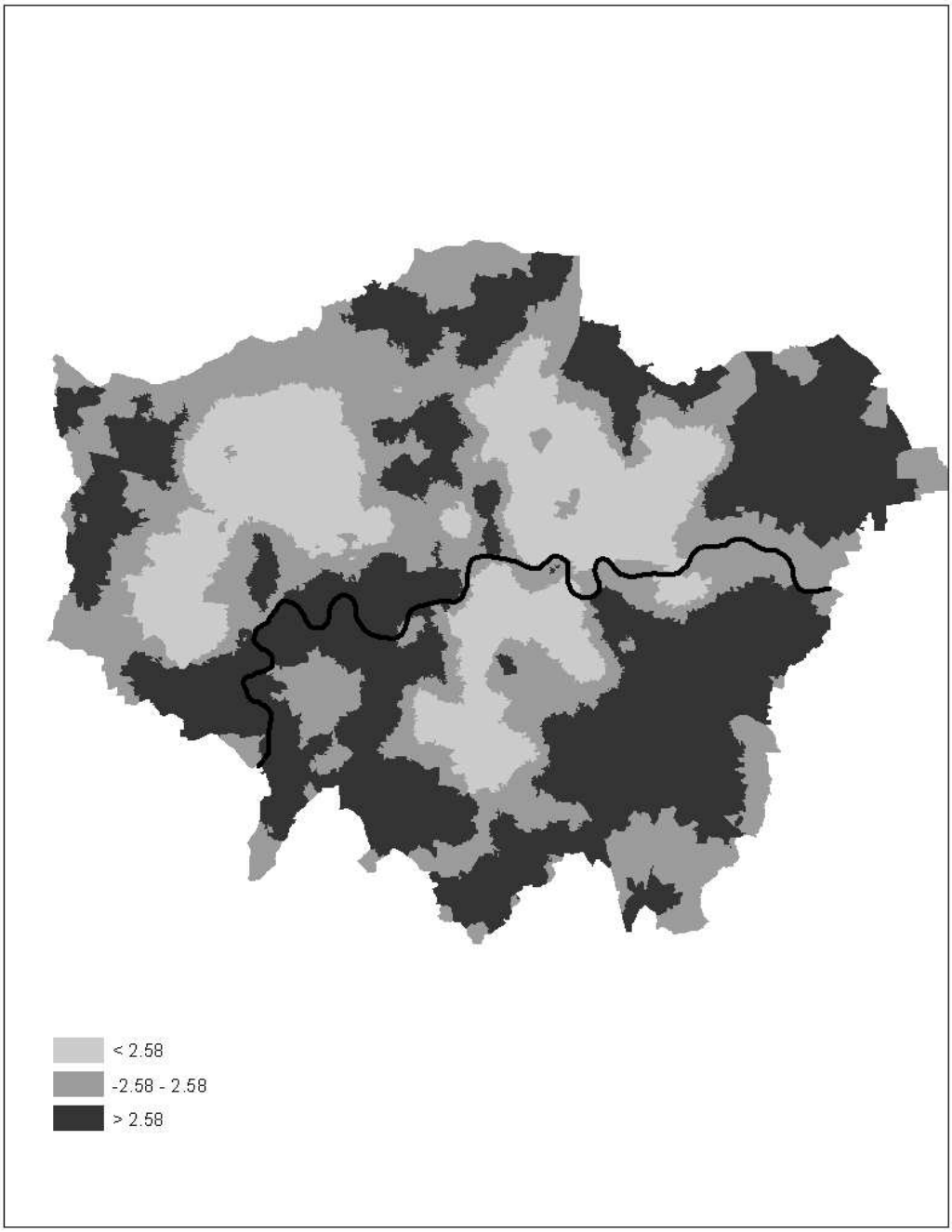

Figure 6.