League Tables and Concentric Banding: How Similar are the Employment and Education Domains of the Index of Multiple Deprivation 2010?

Michael A. Nolan[1] · Michael Reynolds[2] · Stephen Trotter[1]

Abstract

This paper seeks to provide further evidence about 'league table' orderings for deprivation in local areas, especially for two domains – employment and education – which seem likely to be closely related. Rather than relying solely on administrative Local Authority areas and functional economic areas, our focus is to refine the concentric banding approach introduced in Nolan et al. (*Local Economy* 27(4): 403–418, 2012), and to apply it throughout England. We also introduce two new procedures by which areas might be systematically split into smaller components. The spatial distributions of employment and education deprivation turn out to be quite different (and distinct from the spatial distribution of overall deprivation). Throughout, there is clear evidence that it matters what size of local area is considered.

Keywords Spatial unit · Areal unit · League table · Employment deprivation · Education deprivation

JEL Classifications R12 · R23 · R58 · I2

Introduction

League tables are rather commonplace nowadays. It is unsurprising that they gain such substantial media attention, because they convey their message clearly and simply, and

✉   Michael A. Nolan
    m.a.nolan@hull.ac.uk

1   Faculty of Business, Law and Politics, University of Hull, Wharfe Building, Cottingham Road, Hull HU6 7RX, UK

2   University of Leeds, Leeds, UK

because their format has a wide familiarity (founded on the league tables associated with sporting competitions). Area-based league tables serve to paint a picture. In so doing, they raise the question of whether that picture captures the essence of reality, or is simply the equivalent of shining a rather unusual type of light. If a body of analysis creates a set of impressions of local areas which is not invariant to the choice of spatial unit, that fact must be a matter of concern in principle. There is also the issue of who is supposed to be enlightened by area-based league tables: although policy-makers from local and central government are one obvious audience, so too are employers who are considering the location of their organisation and individuals choosing where to live. Beyond this, there is the question of informing the general public. Of course, league tables based on a particular strand of performance may distort policy, or other decisions: whatever is directly measured in a league table might become the key focus, potentially to the detriment of other strands of performance of equal, or greater, importance. However, we must not ignore the fact that, in as much as an absence of league tables may also reduce available information, that absence may be even more damaging to the prospects for sound (evidence-based) decision-making.

In this paper, our spotlight falls especially on the question of how 'local areas' are to be defined: this will determine 'league membership' for local economic performance league tables. Our example is drawn from the case of England. However, in principle, similar analysis should also be undertaken for other countries. Clearly, any country can be divided up in many different ways. The basis might be administrative or political boundaries, which themselves often result from an accumulation of historical factors (including political expediency or administrative convenience). However, these spatial units – e.g., Local Authority (LA) areas in England – can have the serious disadvantage of being highly heterogeneous in size (whether that is defined by area, population or some combination of the two). Functional economic areas are another possibility (see, for example, Cörvers et al. (2009)), but these are likely to vary too, according to which policy domain is being considered. Although the definition of such spatial units is likely to be criterion-based, there may also seem to be an arbitrary element in setting the precise terms of the criteria. In terms of outcome, we need to consider whether decisions made solely on the basis of an LA ranking are likely to be optimal. It is quite possible that such decisions might have differed substantially, had they been based instead on rankings for another sort of spatial unit.

The first of this paper's two principal concerns is to develop the methodological approach proposed, and initially applied, in the work of Nolan et al. (2012). The other key element is to apply the approach to the data for a pair of specific domains from within the 2010 English Index of Multiple Deprivation (IMD2010) – rather than focusing on the overall index, as the earlier work did. This paper will provide initial empirical evidence on the extent to which local area 'league table' orderings exhibit similarities across two different – but related – policy domains. Commentary should then be possible about whether there is any substantial impact from defining spatial units in different ways – be they administrative, functional or size-based.

McLennan et al. (2011) provide a detailed review of IMD2010, which includes important discussion about technical issues associated with constructing a multiple-domain index across small geographic areas. Some of that discussion reflects a background literature which developed as successive versions of the IMD were produced (for 2000 and 2004), and their features and merits were debated – see, for

example, Deas et al. (2003) and Noble et al. (2006). Changes in IMD construction through time, as well as evolving boundaries for some administrative local areas, lead to difficulties in examining how deprivation changes and where it persists – see, for example, Ajebon and Norman (2016). Indeed, a further English Index of Deprivation was developed to enable deprivation to be examined for the period 1999–2005, as analysed in Rae (2012). Meanwhile, there is also a literature on automated zoning procedures – modifying a zoning system to adapt (on the basis of specified criteria) to changes in underpinning area characteristics that occur inevitably with the passage of time. Martin (2003) and Cockings et al. (2011) are examples, and the latter considers threshold sizes for LSOA (Lower-layer Super Output Area) population as the distribution of residential location changes over the years. Geographical differences are examined by Burke and Jones (2018), who explore the development of an index (again at the LSOA level) to reflect the particular types of deprivation experienced in rural areas.

The next section of this article looks back briefly at underlying literature on the topics of spatial resolution, aggregation and the areal unit. It then moves on with an outline of our approach to spatial unit definition. This is followed by a section which starts by briefly outlining two domains from IMD2010 – employment being the first, and education, skills and training being the other. It then proceeds to portray extracts from the league tables that emerge from the application of our approach. After some appropriate discussion, there is a brief conclusion.

Spatial Resolution, Hypotheses and Spatial Unit Definition

As mentioned in Nolan et al. (2012), the Modifiable Areal Unit Problem (MAUP) was identified by Openshaw (1984), who noted that areal (spatial) units may be defined arbitrarily, and maybe in accordance with the idiosyncratic views of an individual undertaking a geographic study. More specifically, Huby et al. (2009) found that an increase in the size of base spatial unit will generate a reduction in inequality. The relevance of this finding to the current paper is substantially bolstered by the fact that these earlier authors were considering the aggregation LSOAs – which are the same baseline spatial unit that we will be using. Looking at Montreal, Séguin et al. (2012) found greater concentration of poverty for the smallest areal unit type – the dissemination area. The work of Jones (2000), which provides a way to estimate the unemployment rate at a spatially disaggregated level beyond the scope of the data then available, is also potentially of particular interest in the context of the employment domain of IMD2010 we consider here. Commuting patterns offer a ready explanation of the divergence of his unemployment rates, for Cardiff, from their published counterparts. However, it should be noted that the IMD2010 employment domain is rather broader – taking in substantial numbers of individuals that would be defined as economically inactive, in addition to the (claimant) unemployed.

Given the MAUP, it seems reasonable to put forward an initial outline null hypothesis (H_{01}) that changes in the resolution and construction criteria of the spatial units used (typically below LA area resolution) should make no fundamental difference to the composition of a league table for a deprivation index or measure of local economic performance. A supplementary null hypothesis (H_{02}) might be that changes in spatial unit resolution and construction criteria should not systematically influence regression

results, for a given underlying specification. A final null hypothesis (H_{03}) could be that the impact of spatial unit resolution and construction criteria changes are similar for two distinct deprivation index domains. Findings with regard to each of these hypotheses offer potentially valuable insights – for academics, local policy-makers, employers considering locational decisions, or individuals making choices about where to work and/or reside – in seeking to interpret league tables released in the media, and regression-based results from research articles.

Our approach to the definition of spatial units is based heavily on the concentric analysis laid out in Nolan et al. (2012). However, in this paper (in contrast to its predecessor), we utilise the definitions of LA areas in force after the April 2009 boundary changes (which cut the number of LA areas in England from 354 to 326). The first step of the process of specifying spatial units is to define the centre of each LA area. Each of England's 32,482 LSOAs is defined as being centred at its population-weighted centroid (as available via the Office for National Statistics (ONS) Geoportal web pages). In a modification to the next step, the LA centre is given as the population-weighted mean[1] of all the LSOA centres within a given LA. Once the LA centre has been determined, the geometric distance can be calculated between it and any LSOA centre. For any given LA area, all of England's LSOAs can be placed in ascending order of distance from that particular LA centre. The closest LSOAs to a particular LA centre will almost always be interior to that LA, while the furthest will be exterior to it. In between, the transition from interior LSOAs to exterior LSOAs may be smooth in some cases, but not others. New spatial units can be defined and generated through the systematic addition of LSOAs according to their ranking on distance between LSOA centre and the chosen LA centre. Later, we consider spatial units constructed with reference to cut-offs at integer numbers of kilometres of distance from the LA centre. Although this gives an underlying basis of a series of concentric circles, actual spatial units will be irregularly shaped. Moreover, as distance from LA centre is allowed to increase, the newly created spatial units will increasingly overlap.

Figure 1 below shows LSOAs for the case of Kingston upon Hull, for distances of 3 km (solid blue), 6 km (horizontal stripes), 12 km (solid green) and 24 km (diagonal stripes) from the LA centre. All of the LSOAs that are within 3 km of the LA centre are within the Hull LA area, but that is not true of some of the LSOAs within 6 km of the Hull LA centre, most of the LSOAs in the 6–12 km range and all of those in the 12–24 km range.

This paper also investigates two further, related, approaches: they are novel and we name them Procedure 1 and Procedure 2. Both of them split each LA area into newly-defined spatial units of what we will term 'similar population'. First, a benchmark (Y) is chosen for this 'similar population' (in thousands) – for example, $Y = 50$. If the LA area i has a population of X_i, then it is split into k_i parts, where we define $k_i = \max(1, X_i / Y)$ with the result rounded to the nearest integer. Once k_i has been determined, the construction of sets of newly-defined spatial units can commence. As before, newly-created spatial units are formed from the aggregation of LSOAs, making use of the centre of each LSOA and the centre of each LA – in each case, using population

---

[1] For the employment domain, we eventually decided not to use full population weights – replacing these with weighting by the working-age population. As noted in Nolan et al. (2012), we could alternatively have used some sort of official or pseudo-official measurement of the relevant city centre.
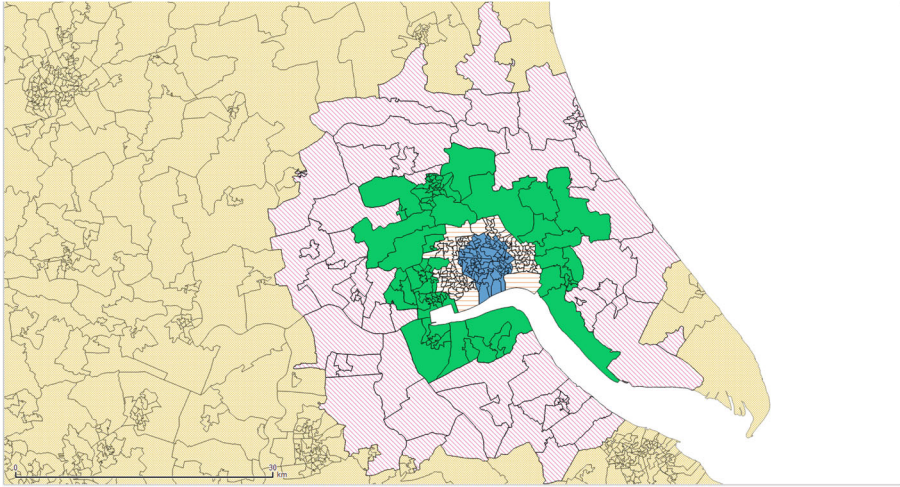
Fig. 1 Concentric banding (for 3 km, 6 km, 12 km and 24 km), for Kingston upon Hull

weighted centroids. However, under Procedures 1 and 2, no given spatial unit is permitted to cross the boundary of a single LA area (so that a particular set of newly-created spatial units has no overlaps).

Under Procedure 1, LSOAs are added to a newly-defined spatial unit, in ascending order of distance of LSOA centre from the centre of the LA area, until the threshold population (X_i / k_i) is first reached or breached. Once that first spatial unit has been established, a second spatial unit is constructed – again by proceeding through LSOAs in ascending order of distance from the LA centre, this time until the threshold ($2X_i / k_i$) is reached or breached. The process then continues, being undertaken k_i times for that LA; and it is repeated for each and every LA area in England. The example shown in Fig. 2, for Kingston upon Hull (163 LSOAs, with population 261,098), splits the LA area (using $Y = 50$) into five parts (labelled A_1-E_1) on the basis of distance from the LA population weighted centroid.[2]

Procedure 2 proceeds in a similar fashion to Procedure 1, with the one distinction being that (under Procedure 2) LSOAs are added in ascending order of the *angle of the LSOA centre* from the centre of the LA area. We define the angles in such a way that the 0° / 360° boundary is placed where an LSOA centre is directly due west from the centre of the LA area.[3] Another important point to note, for both Procedure 1 and Procedure 2, is that the same process can be replicated for any chosen benchmark 'similar population' (Y) – provided that Y is not smaller than the population size of at least the vast majority of the LSOAs. Of course, the number of newly-defined spatial units created will depend (negatively) upon the chosen value of Y. The example shown below in

---

[2] For example, the first threshold (261,098 / 5 = 52,219.6) is breached by the inclusion of the 33rd LSOA, which takes the population of area A_1 up to 52,868. It is visually evident that these units are not necessarily contiguous.

[3] It might be helpful to imagine the drawing of a circle around the LA area's population weighted centroids – starting on the middle left and proceeding around in a clockwise direction. This paper utilised Stata's atan2 function, undertaking the appropriate simple transformations of the output (including that of angles from radians to degrees).
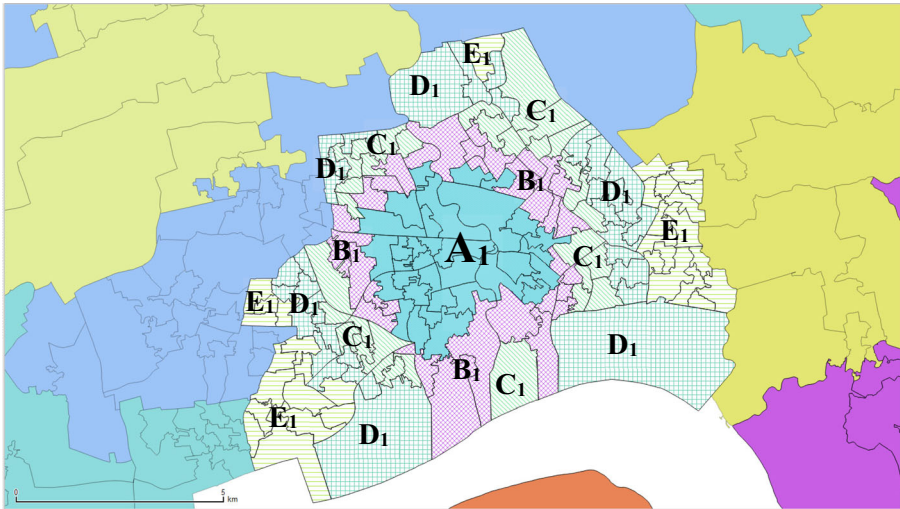
Fig. 2 Procedure 1, with $Y = 50$, using Kingston upon Hull as an example

Fig. 3, for Kingston upon Hull, splits the LA area (using $Y = 50$) into five parts (labelled A_2-E_2) on the basis of angle of LSOA population weighted centroids from the LA population weighted centroid.[4] As Figs. 2 and 3 illustrate, Procedure 1 has the advantage of each part (e.g. the five areas marked C_1 in Fig. 2) being roughly equidistant from the LA centre, but the disadvantage that they may well not be contiguous. Procedure 2's angles-based approach much more readily achieves contiguity (see area C_2 in Fig. 3) but at the cost of sacrificing equal raw distance. Thus, neither procedure obviously dominates the other. Many articles which analyse the IMD (or its components) focus on the grouping of socio-economically similar spatial units, regardless of whether they are contiguous – or, indeed, even spatially close to each other.[5] However, a key intended purpose of this research is to inform policy regarding, and perceptions of, local area economic performance – where the contiguity and proximity of spatial units (of population well above the LSOA norm) is more likely to be of practical relevance.

English IMD2010 Employment and Education Domains

The details of the Index of Multiple Deprivation 2010 (IMD2010) for England are laid out in DCLG (2011). As stated there, the main focus of analysis of the IMD2010 – to avoid the risk of loss of information – should be on England's 32,482 LSOAs. These were designed to have similar populations of around 1500, and they nest within LA areas. The Appendix reveals an average LSOA population (in 2008) of 1584, while 95% of LSOAs range in population from 1165 to 2281. Splitting England's land area of around 13 million hectares, the LSOA average area (Table 16 in the Appendix) is 401 ha, but 85% of LSOAs are smaller than that (and half of those are less than 40 ha) due to substantial positive skewness. Although the overall IMD2010 is a weighted

---

[4] Contiguity of these spatial units, under Procedure 2, is much more likely than was the case under Procedure 1.
[5] We are grateful to a referee for suggesting this clarifying observation.
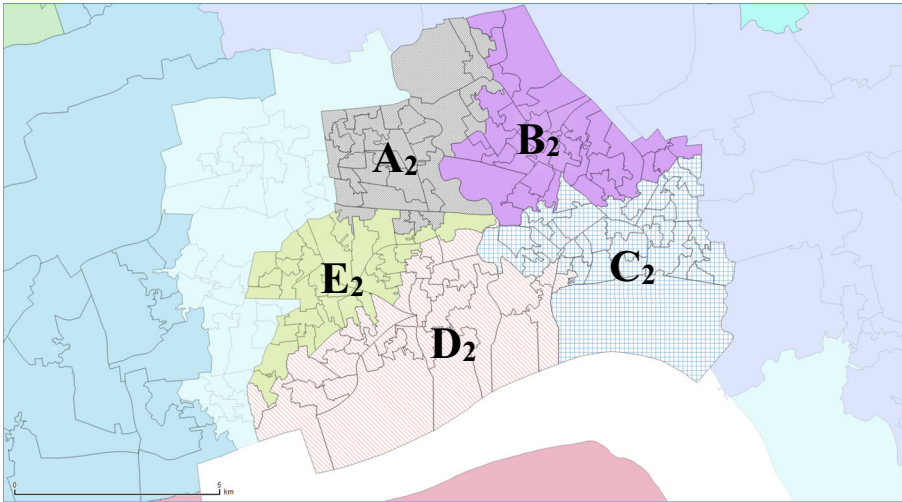
Fig. 3 Procedure 2, with $Y = 50$, using Kingston upon Hull as an example

amalgamation of seven policy domains, our focus in this paper concerns just two of those domains.[6] It is perhaps worth noting that the first, employment deprivation, is given a 22.5% weighting in the overall index (equal top billing, along with income deprivation); whereas the other (education, skills and training deprivation) is offered a more limited weighting of 13.5%. For each of these two policy domains, the domain score is itself the result of the aggregation of several elements. For example, the employment domain score includes not only working-age Job Seekers Allowance (JSA) claimants (a fairly narrow definition of unemployment); but also Incapacity Benefit (IB) claimants and Severe Disablement Allowance (SDA) claimants of working age (and both of these groupings include substantial components of economic inactivity, even if they are partly argued to include elements of 'hidden' unemployment). There are also four other components in the employment score – all based on working-age population (18–64 for men and 18–59 for women in that era) – details of these being shown on pages 23–24 of DCLG (2011). In other words, the employment domain score is defined – in principle – as the total number of employment-deprived persons divided by the total working-age population.[7]

The education, training and skills domain is split into a pair of sub-domains: one is focused on children and young people, while the other relates to adult skills. The first sub-domain includes six elements: the first three capture attainment at each of the Key Stages 2–4 (for pupils aged 11, 14 and 16, and importantly based on LSOA of pupil

---

[6] On the other hand, Nolan et al. (2012) considered the overall index – and for a narrower subset of England's LA areas.

[7] In fact, the IMD2010 data include an employment score (as a decimal, rather than a percentage) for each LSOA area. We have weighted these LSOA employment scores by the working age population of that LSOA area, in order to construct employment scores – converted to values in the 0–100 range – for each of our newly-defined spatial units (each aggregated across a number of LSOAs). The working age population data are drawn from LSOA-based mid-year 2008 population estimates. It should also be noted that we have not undertaken a further application of the 'shrinkage technique' which is applied to the LSOA employment deprivation scores, as outlined in detail in Annex E of DCLG (2011).

residence, rather than educational establishment), while the others relate to absenteeism from school, entry to Further Education and entry to Higher Education. The weightings of the elements were determined by factor analysis, as indicated on page 37 of DCLG (2011). The second sub-domain looks at the proportion of low-skilled prime age (25–54) adults in Census 2001 (see ONS (2001)). In stark contrast to the case of the employment domain, even some of the different elements in the first sub-domain are fractional rates that use different denominators (for example, populations for different particular age ranges). This makes a straightforward detailed interpretation of the magnitude of the education domain score problematic[8] (within the IMD2010 output, it is reported on a 0–100 scale, and we also follow that approach).

Employment Deprivation Domain

Initially (in Table 1), we list the 30 most deprived LA areas in England for the employment domain, based on weighted average of the LSOA employment scores (proportions) – the weightings being based on LSOA working-age population. Of these LA areas, 17 would be in the corresponding league table for the overall deprivation IMD2010 index scores.

By contrast, IMD2010 employment scale values are reported for LA areas in DCLG (2011); they are higher (ceteris paribus) for areas of greater population, because employment scale measures the raw number of employment-deprived individuals (just over three million across England, out of a working age population of a little over thirty million). Table 1 is an incomplete league table, but it also includes – for comparison – the LA areas at the approximate quartiles (82 and 245) and median (164) of the employment deprivation score distribution, and the least employment-deprived LA (326).

Beyond a raw league table, we need to explore the extent to which some key explanatory variables can account for variation in LA area employment scores. Table 2 below shows results for a basic linear regression of the natural logarithm of LA employment score[9] on the natural logarithm of LA land area (in hectares), the natural logarithm of 2008 mid-year LA population and a set of eight region dummies (London is the base region). A few basic summary statistics for the LA areas can be seen at the start of Table 16 in the Appendix.

Since the linear regression is a double-log specification for the land area and population regressors, it should be noted that the first two estimated slope coefficients can be interpreted as elasticities. Of England's 326 LA areas, 33 are in London, and the sample mean of the dependent variable for those cases is 2.177. This places London less favourably (i.e., more employment deprived) than the East Midlands, East (Anglia), the South East and the South West – whereas only the South East has a negative regression estimate, and that is by a narrow magnitude. Of course, the regression results control for land area and population – and London LAs average six times less land area

---

[8] Basically, the education score is a weighted average of various proportions, each of which is relevant to the measurement of education deprivation. For each LSOA, IMD2010 reports a value in the 0–100 range. For our newly-defined spatial units (aggregated across LSOAs), we used total population data drawn from LSOA-based mid-year 2008 population estimates. We did not undertake any further application of the 'shrinkage technique' (empirical Bayesian estimation) detailed in Noble et al. (2006) and Annex E of DCLG (2011), although it had been applied to all but one of the constituent indicators within the education score.

[9] The employment scores as percentages (see Table 1) were used, rather than raw scores, prior to the taking of logs.

Table 1 Thirty most employment-deprived LA areas (with rank, and a weighted score (as a %))

1	Knowsley	19.85	16	St. Helens	15.94
2	Liverpool	19.50	17	Wirral	15.90
3	Hartlepool	19.04	18	Wolverhampton	15.82
4	Blackpool	18.55	19	Redcar and Cleveland	15.45
5	Middlesbrough	17.14	20	Salford	15.43
6	Stoke-on-Trent	17.01	21	Kingston upon Hull	15.22
7	Burnley	16.84	22	Sandwell	15.18
8	South Tyneside	16.74	23	Manchester	15.18
9	Halton	16.62	24	Mansfield	15.15
10	Blackburn with Darwen	16.23	25	Hyndburn	15.11
11	Rochdale	16.17	26	Durham County	15.11
12	Sunderland	16.07	27	Birmingham	14.97
13	Barnsley	16.06	28	Gateshead	14.96
14	Barrow-in-Furness	16.01	29	Wigan	14.90
15	Hastings	15.96	30	Bolsover	14.74
82	Telford and Wrekin	11.43	164	Redbridge	8.40
245	Wiltshire	6.19	326	Isles of Scilly	2.17

than those in any other region and higher population than any other region except West Midlands (lower land area and higher population are both associated with higher employment deprivation).

Table 2 Natural logarithm of LA employment score, regression results

Variable	Estimate	Std. Err.	t ratio	P value	Sample means	
					Regressor	Dependent variable
ln (LA land area)	−0.1119	0.0147	−7.64	0.000	9.897	
ln (LA population, 2008)	0.2496	0.0291	8.57	0.000	11.800	
North East	0.7273	0.0985	7.38	0.000	0.037	2.684
North West	0.6086	0.0724	8.41	0.000	0.120	2.505
Yorkshire & The Humber	0.3976	0.0882	4.51	0.000	0.064	2.246
East Midlands	0.3676	0.0761	4.83	0.000	0.123	2.145
West Midlands	0.3659	0.0771	4.74	0.000	0.092	2.245
East	0.1631	0.0726	2.25	0.025	0.144	1.980
South East	−0.0184	0.0671	−0.27	0.784	0.206	1.839
South West	0.2907	0.0783	3.71	0.000	0.113	2.050
Constant	0.0350	0.3671	0.10	0.924	1.000	2.129
Number of observations	326	$F_{10,315} =$	33.10	0.000		
Sum of squared residuals	25.068	$R^2 =$	0.5124	Adjusted	$R^2 = 0.4969$	
Test for region controls		$F_{8,315} =$	24.40	0.000		

The above regression could be set up equivalently (using Y_1 for LA employment score and X_1 for LA land area) with the natural log of LA population *density* ($\ln(X_3)$) in place of the log of LA population ($\ln(X_2)$). Then we can write $\ln(X_3) = \ln(X_2/X_1) = \ln(X_2) - \ln(X_1)$, which is equivalent to $\ln(X_2) = \ln(X_1) + \ln(X_3)$, and to $\ln(X_1) = \ln(X_2) - \ln(X_3)$. With predicted $\ln(Y_1)$ denoted as $\hat{Z}_1$, there are three alternative valid expressions that can be written for the London region:

A. Using Table 2, $\hat{Z}_1 = 0.0350 - 0.1119\ln(X_1) + 0.2496\ln(X_2)$.
B. $\hat{Z}_1 = 0.0350 + 0.1377\ln(X_1) + 0.2496\ln(X_3)$.
C. $\hat{Z}_1 = 0.0350 + 0.1377\ln(X_2) + 0.1119\ln(X_3)$.

All of the estimates attached to $\ln(X_1)$, $\ln(X_2)$ and $\ln(X_3)$ in these three equivalent expressions are statistically significant at the 1% significance level. The impact of LA land area on employment deprivation score is negative (-0.1119) having controlled for population as well as region, but positive ($+0.1377$) after controlling instead for population density.

More urban areas typically have lower land area and higher population. They therefore tend to have higher population density. For the LA data, the respective bivariate correlations of $\ln(X_1)$, $\ln(X_2)$ and $\ln(X_3)$ with $\ln(Y_1)$ are -0.224, 0.403 and 0.374: this confirms that LA employment deprivation scores tend to be higher in more urban areas. In turn, that raises the issue of whether LA-based league tables of employment deprivation score – ranging across both urban and rural LA areas – can be expected to naturally favour less densely populated LA areas. Going beyond that, there is the issue of the extent to which a particular LA's employment deprivation score can be considered to be a natural consequence of the relevant set of its underlying attributes, rather than underperformance of the local economy. Benchmarking actual scores, using predicted values from an appropriate regression, might be an interesting approach. Although Table 2 uses a very basic regression which is too straightforward to be likely to be genuinely appropriate, it is interesting to consider the 'benchmarked' league table it generates, based on land area, population and region (see Table 17 in the Appendix). There, only 12 of the LA areas seen in Table 1 now occupy top 30 positions (for having most employment deprivation over and above their 'benchmark level').

Some further basic spatial analysis of LA area employment deprivation scores can be undertaken by using Moran's I, Geary's c and the Getis and Ord statistic. Moran's I was introduced by Moran (1950) to measure global spatial autocorrelation, and – for the case of 326 LA areas – it can be defined as follows[10]:

$$I = \frac{326}{\sum_i^{326}\sum_j^{326} w_{ij}} \frac{\sum_i^{326}\sum_j^{326} w_{ij}\left(Y_{1i}-\overline{Y_1}\right)\left(Y_{1j}-\overline{Y_1}\right)}{\sum_i^{326}\left(Y_{1i}-\overline{Y_1}\right)^2}, \tag{1}$$

---

[10] For LA areas, this was implemented via the spatwmat and spatgsa procedures, using Stata software (see Pisati (2001)). However, although it was also possible to use these procedures for some of the later work (for the range $Y = 60$ to $Y = 30$), memory limitations then have an impact on speed and feasibility. As a consequence, Microsoft Excel was used to generate the spatial analysis statistics for $Y = 25$. In turn, however, this is also very unwieldy. Eventually, mata (Stata's matrix language) was used to confirm the spatial analysis statistics for $Y = 25$, calculate their statistical significance, and undertake similar analysis for all the smaller values of Y, as well as for the data on individual LSOAs.

where the weights w_{ij} can be given by the reciprocal of the distance (in kilometres) between the centre of LA_i and the centre of LA_j, and those weights may then be 'row standardised' to sum to unity; or they can be defined in a binary format to be one for 'near neighbour LA areas' with a centre within a certain threshold distance and zero for other LA areas, beyond that threshold distance. The value of I itself lies between minus one and plus one. For LA area employment deprivation, Moran's I takes the value 0.331 – that is, moderate positive spatial autocorrelation – with the weighting matrix standardised. With a binary weighting matrix, the value of Moran's I depends on the cut-off point for 'near neighbours', but it is at least 0.245. The positive sign of Moran's I (and statistical significance in both cases above its expected value $(-1/(326-1)))$ indicates that, unsurprisingly, LA areas of similar employment deprivation tend to be found in spatial proximity to one another.

On the other hand, Geary's c (see Geary (1954)) has more sensitivity to local spatial autocorrelation. For the case of 326 LA areas, it is given by the following expression:

$$c = \frac{325}{2\sum_i^{326}\sum_j^{326}w_{ij}} \frac{\sum_i^{326}\sum_j^{326}w_{ij}\left(Y_{1i}-Y_{1j}\right)^2}{\sum_i^{326}\left(Y_{1i}-\overline{Y_1}\right)^2}. \tag{2}$$

This statistic lies between zero and two, with a value of one indicating no spatial autocorrelation. For LA area employment deprivation, Geary's c is 0.672. Since it is less than unity (and statistically significantly so), positive spatial autocorrelation is again indicated.

The Getis and Ord statistic (see Getis and Ord (1992, 1995)) also investigates local spatial autocorrelation. For the case of 326 LA areas, it is given by the following expression:

$$G(d) = \frac{\sum_i^{326}\sum_j^{326}w_{ij}(d)Y_{1i}Y_{1j}}{\sum_i^{326}\sum_j^{326}Y_{1i}Y_{1j}}, \tag{3}$$

where $i \neq j$ and the weights are in binary format. The value of $G(d)$ depends on the threshold distance, d, for 'near neighbours'. It must be compared against its mathematical expectation, which (for 326 LA areas) is given by:

$$E[G(d)] = \frac{\sum_i^{326}\sum_j^{326}w_{ij}(d)}{325*326}.$$

When we choose a large enough value of d (120 km) to ensure that even the Isles of Scilly LA has at least one 'near neighbour', the expected value of the Getis and Ord statistic is 0.305; $G(120) = 0.280$ for our data lies significantly beneath. This indicates 'cold spot' clustering of areas of low employment deprivation. However, when a smaller value of d (such as 40 km, still giving each LA area an average of nearly 20 'near neighbours') is used, $E[G(40)] = 0.0605$ and $G(40) = 0.058$: this offers no evidence of LA employment deprivation clustering.

LSOA-based employment deprivation scores inevitably exhibit much more variation than their LA counterparts (since England has nearly 100 times as many of these

smaller spatial units, compared to the number of LA areas) – see Table 16 in the Appendix. Since LSOA scores are typically based on much smaller populations, they can also be expected to be less reliable on average than those for LA areas. However, we still undertake an LSOA-based log employment score regression corresponding to our earlier LA-based one – although without showing the coefficients on the region dummies in Table 3 (below). In addition, we test for the joint significance of a full set of (325) LA controls, in place of the eight region dummies. While LA area populations range from just over two thousand to a little beyond one million, LSOA populations are rather more similar to each other (from about five hundred and fifty to roughly twenty times that figure – see Table 16 in the Appendix[11]). The most prominent difference between Table 3 and Table 2 relates to the estimated coefficient attached to log population: for the LSOA-based regression, this elasticity has a negative sign – opposite to the sign from the LA case. Such a disparity indicates that the way in which LSOAs are aggregated into LA areas is, in itself, important: more highly populated LSOAs are predicted to exhibit a **lower** rate of employment deprivation. This is initial evidence against our null hypothesis H_{02}, that spatial resolution should not systematically influence regression results.

If the log of LSOA population density is inserted in place of the log of LSOA population, it has the same negative estimated elasticity attached as is reported in Table 3 for log population. With LA controls, the estimate attached to the 'log of LSOA land area' regressor is -0.2497 ($= -0.1328 - 0.1168$). The impact of LSOA land area on employment deprivation score is negative having controlled for population (-0.1168), and even more negative (-0.2497) after controlling instead for population density. The lower R^2 value for the LSOA regression may bear out the point above about LSOA scores having lower reliability.[12]

Calculation of Moran's I, Geary's c and the Getis and Ord statistic for the LSOA data is less straightforward in practical terms. This is because the 'spatgsa' procedure for the Stata software package does not work when the number of observations is so large. Fortunately, the mata matrix language offers an alternative approach (whilst still using the Stata software). Expressions [1] and [2] are readily implemented, but confirmation of statistical significance requires the accompanying variances – provided, for example, in the mathematical appendix of Sokal et al. (1998). The appropriate value of Moran's I is 0.189, and that of Geary's c is 0.829. Each is comprehensively statistically significant, and indicative of positive spatial autocorrelation. Meanwhile, expression [3] yields a value for the Getis and Ord statistic ($G(90)$) of 0.206. Using the variance from Getis and Ord (1992), that can be seen as statistically significant for 'cold spot' clustering of low employment deprivation.

Table 4 shows the 20 most deprived LA areas on the basis of the LSOAs centred within (respectively) 3 km, 6 km, 12 km and 24 km of their LA centre (as illustrated in Fig. 1 for the case of Kingston-upon-Hull).

It is worth remarking even on some rather obvious points. Firstly, the considerable majority of these employment-deprived LA centres can be found in northern England –

---

[11] Also note that only seven LSOAs have a population of less than 750, and only two have a population above 7500.

[12] This is true even with 325 control dummies for the 326 LA areas. Note that the absence of discussion here about Moran's I, Geary's c, and the Getis and Ord statistic can be explained by a practical difficulty: the weighting matrices in the LSOA case have about 10,000 times more cells than in the LA case.

Table 3 Natural logarithm of LSOA employment score, two regression specifications

Variable	Estimate	Std. Err.	t ratio	P value	Sample Mean
Region controls					
ln (LSOA land area)	−0.1360	0.0022	−62.00	0.000	4.320
ln (LSOA population, 2008)	−0.1715	0.0189	−9.07	0.000	7.353
Constant	3.7941	0.1389	27.32	0.000	1.000
Number of observations	32,482	$F_{10,32,471} =$	991.61	0.000	[Region controls]
Sum of squared residuals	9945.96	$R^2 =$	0.2339	Adjusted	$R^2 = 0.2337$
LA controls					
Test for region controls		$F_{8,32,471} =$	662.28	0.000	
ln (LSOA land area)	−0.1168	0.0023	−50.67	0.000	4.320
ln (LSOA population, 2008)	−0.1328	0.0173	−7.70	0.000	7.353
Constant	3.8981	0.1397	27.90	0.000	1.000
Number of observations	32,482	$F_{327,32,154} =$	69.38	0.000	[LA controls]
Sum of squared residuals	7612.05	$R^2 =$	0.4137	Adjusted	$R^2 = 0.4077$
Test for LA versus region controls		$F_{317,32,154} =$	31.10	0.000	

with almost all the exceptions to that rule (Birmingham, Sandwell, Staffordshire Moorlands, Stoke-on-Trent, Walsall and Wolverhampton) being situated in the West Midlands. That just leaves Bolsover (East Midlands) and Great Yarmouth (East of England). Secondly, there are some LA centres for which employment deprivation persists across all these chosen radii: notably Hartlepool, Wirral and Knowsley; and, to a somewhat lesser extent, Liverpool and St. Helens. Unsurprisingly, all of the corresponding LA areas are in the top 20 most deprived, well within the bounds of Table 1. Thirdly, there is some evidence against our null hypothesis H_{01}. Two LA centres appear only once in this table (of four rankings) – namely Bolsover and Durham County (newly created in the April 2009 boundary changes).[13] However, both of the corresponding LA areas appear in Table 1 – unlike several of the other names from Table 4. Two of the most notable cases are Wyre and Staffordshire Moorlands, since their LA centres are well within the top 20 for a 12 km radius, but their underlying LA areas lie outside the top 100 for employment deprivation. The other case worth highlighting is Bury, whose LA centre is named for the 12 km radius and the 24 km radius in Table 4, while having an underlying LA area which is only the 72nd most employment deprived. Finally, we can focus on the LA areas within Table 1 which are not to be found within Table 4: Hastings LA moves from being within the top 15 to being outside the top 25 based on radius from centre. Similarly, Salford, Kingston upon Hull, Manchester and Mansfield are each within the top 25 most employment-deprived as LA areas, but would never be higher than 30th (and usually outside the top 40) if Table 4 were to be extended.

To give a somewhat clearer impression of the relationship between employment deprivation rankings based on LA areas, and each of the sets of rankings based on our

---

[13] It is not necessarily surprising that there would be differences between the four ranking lists in Table 2. Since the radius doubles from one column to the next, the area of the corresponding circles are in the ratios 1: 4: 16: 64.

Table 4 20 most employment-deprived LA centres at 3 km, 6 km, 12 km and 24 km radius

	3 km radius	6 km radius	12 km radius	24 km radius
1	Hartlepool	Liverpool	Knowsley	Wirral
2	Blackpool	Hartlepool	Liverpool	Sefton
3	Wirral	Knowsley	Wirral	Liverpool
4	Great Yarmouth	Bolsover	Hartlepool	Hartlepool
5	Knowsley	Middlesbrough	Redcar and Cleveland	Halton
6	Halton	Wirral	Sefton	Knowsley
7	St. Helens	Blackpool	St. Helens	St. Helens
8	Rochdale	South Tyneside	Sunderland	Durham County
9	Birmingham	Sunderland	Oldham	Stockton-on-Tees
10	Blackburn with Darwen	Burnley	Halton	Redcar and Cleveland
11	Stoke-on-Trent	Great Yarmouth	Wyre	Middlesbrough
12	Walsall	Barnsley	Rochdale	West Lancashire
13	Burnley	Barrow-in-Furness	Hyndburn	Rossendale
14	Liverpool	Birmingham	Bury	Sunderland
15	Barnsley	Hyndburn	Burnley	North Tyneside
16	Wolverhampton	Halton	Stockton-on-Tees	South Tyneside
17	Barrow-in-Furness	St. Helens	Staffs. Moorlands	Doncaster
18	Wigan	Blackburn with Darwen	Barnsley	Newcastle upon Tyne
19	South Tyneside	Stoke-on-Trent	South Tyneside	Barrow-in-Furness
20	Oldham	Gateshead	Sandwell	Bury

There are 14 of England's 326 LA areas that are smaller than a circle of 3 km radius. For circles of 6 km, 12 km and 24 km, the respective numbers of LA areas are 121, 229 and 318. England's largest LA area is just fractionally smaller than a circle of 40 km radius

concentric banding methodology (with cut-offs at integer numbers of kilometres from the LA centre), we can calculate Spearman's rank correlation coefficient (r_s). Values of r_s are reported in Table 5 below, for each integer concentric banding radius (in km):

Firstly, it should be noted that, for the 1–4 km radii, some LA centres have no LSOA centre within that distance – and thus no ranking to compare against the LA area ranking. For a given radius, we exclude such LA areas – and this may help to explain the initial increase in r_s (as radius increases) – although it is also likely to be linked to the fact that more than 80% of actual LA areas in England are larger in size than a circle of radius 4 km. However, beyond a 6 km radius, r_s begins to fall back – although it is still as high as 0.88 at a radius consistent with mean LA land area, and 0.778 for a radius of 24 km (the largest in Table 4). Thus, evidence against null hypothesis H_{01} here is rather modest.

The plot below (Fig. 4) illustrates employment deprivation for some LA centres highlighted above:

The pronounced peak for Bolsover at a 5 km radius indicates why it appears only once in Table 4. A rather inexorable rise for Durham County leads to its entry into Table 4 at a 24 km radius.

Table 5 Values of Spearman's rank correlation coefficient (in each case, the LA area ranking is compared to a ranking for LA centres based on concentric banding at a particular radius)

Radius (km)	r_s	Radius (km)	r_s	Radius (km)	r_s	Radius (km)	r_s
1	0.8139	11	0.8865	21	0.7998	31	0.7152
2	0.8454	12	0.8769	22	0.7916	32	0.7046
3	0.8727	13	0.8676	23	0.7854	33	0.6944
4	0.8877	14	0.8568	24	0.7780	34	0.6863
5	0.9047	15	0.8493	25	0.7706	35	0.6776
6	0.9071	16	0.8389	26	0.7597	36	0.6711
7	0.9032	17	0.8301	27	0.7531	37	0.6676
8	0.9004	18	0.8232	28	0.7443	38	0.6633
9	0.8972	19	0.8144	29	0.7337	39	0.6563
10	0.8934	20	0.8068	30	0.7237	40	0.6483

Now using Procedure 1 instead, to split each LA area into subsets of its constituent LSOAs, we can obtain the two sets of results shown in Table 6, below[14] – the upper half of the rows involve specifications which use region controls (eight dummies), whilst the lower half of the rows utilise LA controls (325 dummies). The key difference between the contents of Table 2 and Table 3 was with respect to the estimates attached to the population regressor. However, whilst the estimates attached to the population regressor in Table 6's upper rows (region controls) share the positive sign and statistical significance seen initially in Table 2 (for the regression based on LA area data), those in Table 6's lower rows (LA controls) are statistically insignificant at the 5% level until Y drops below 20. This offers some evidence against null hypothesis H_{02} as resolution is altered, for construction criterion Procedure 1. It is also worth considering our measures of spatial autocorrelation for these values of Y, and their statistical significance. The information is displayed in the first half of Table 19 in the Appendix, and shows analogous results to the earlier LA and LSOA cases.

Using Procedure 2, we can obtain the two sets of results shown in Table 7, below (with the same region controls in the upper half, and LA controls for the lower half). The pattern of the estimates attached to the population regressor, across different values of Y, shows less fluctuation than under Procedure 1. With region controls, the estimate shrinks substantially only when Y is reduced to 7.5, but the picture is less clear when LA controls are included. However, the regression results differ for Procedures 1 and 2, and this offers some further evidence against null H_{02}.

Procedure 1 and Procedure 2 group together LSOAs according to quite different rules (which have quite different grouping outcomes). Thus, it is interesting that the respective first halves of Tables 6 and 7 are characterised chiefly by similarities. The similarity extends to the measures of spatial autocorrelation shown in the second half of Table 19 in the Appendix.

---

[14] Statistical significance at the 1%, 5% and 10% significance levels is denoted – respectively – by the symbols "*", "+" and "#". These markings are used in Tables 7, 13 and 14 (in addition to Table 6).
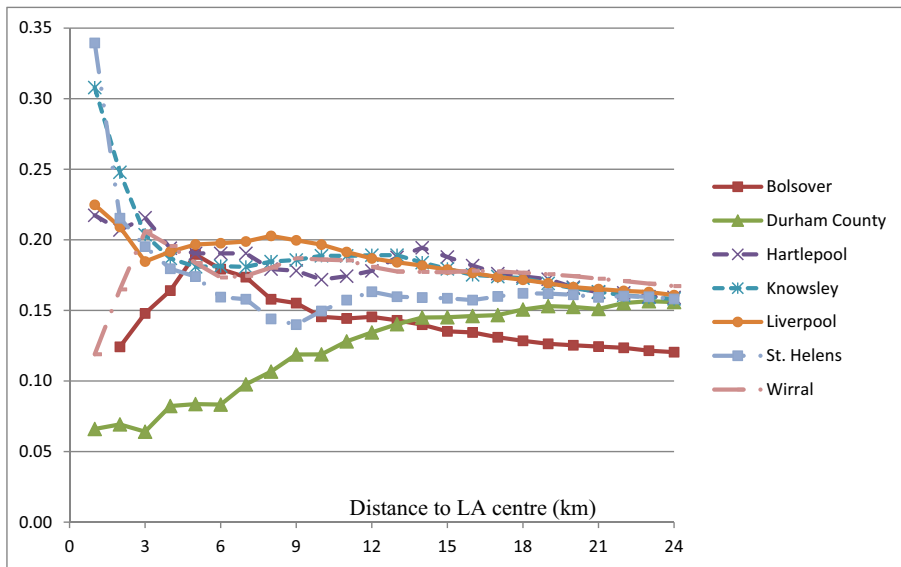
Fig. 4 IMD2010 employment deprivation domain score for seven selected LA centres

However, although regression results do not appear to be markedly different across the total of 35 cases captured across Tables 2, 3, 6 and 7, the character of league tables may be less similar. Recall that we have effectively created 16 separate new employment deprivation league tables: this does not include Table 1 (which shows the top 10% or so LA areas for employment deprivation) or the corresponding league table for LSOA employment deprivation. Beyond this, Table 4 offers four 'top-20 most deprived' league tables – one for each of four chosen radii, for potentially overlapping areas, each of which is centred at the population-weighted centroid of the base LA area. There is a separate (basically unreported) league table underlying each column of

Table 6 Regression estimates across Y values, dependent variable ln(LSOA employment score)

Procedure 1	Estimates, based on different values of Y (in thousands)							
Variable	60	50	40	30	25	20	12.5	7.5
Region controls								
ln (land area)	−0.135*	−0.137*	−0.139*	−0.139*	−0.141*	−0.139*	−0.140*	−0.135*
ln (population, 2008)	0.326*	0.383*	0.318*	0.444*	0.261*	0.423*	0.294*	0.091+
Constant	−0.452	−1.019	−0.266	−1.517#	0.406	−1.189	0.164	2.017*
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.492	0.486	0.485	0.476	0.466	0.457	0.430	0.382
LA controls								
ln (land area)	−0.133*	−0.134*	−0.135*	−0.135*	−0.136*	−0.129*	−0.125*	−0.111*
ln (population, 2008)	−0.151	0.030	0.367	0.204	0.031	0.134	0.117+	0.067+
Constant	3.408	1.729	−1.420	0.106	1.726	0.724	0.867	1.175*
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.817	0.818	0.809	0.797	0.788	0.771	0.729	0.663

Table 7 Regression estimates across Y values, dependent variable ln(LSOA employment score)

Procedure 2	Estimates, based on different values of Y (in thousands)							
Variable	60	50	40	30	25	20	12.5	7.5
Region controls								
ln (land area)	−0.138*	−0.140*	−0.142*	−0.142*	−0.144*	−0.143*	−0.142*	−0.137*
ln (population, 2008)	0.319*	0.397*	0.320*	0.441*	0.284*	0.398*	0.352*	0.175*
Constant	−0.358	−1.160	−0.271	−1.477#	0.169	−0.938	−0.393	1.258*
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.469	0.453	0.452	0.433	0.424	0.414	0.381	0.342
LA controls								
ln (land area)	−0.145*	−0.140*	−0.147*	−0.140*	−0.139*	−0.139*	−0.130*	−0.116*
ln (population, 2008)	0.207	0.293	0.184	0.097	0.174	0.019	0.196*	0.127*
Constant	0.134	−0.699	0.362	1.132	0.404	1.852	0.152	0.676#
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.760	0.740	0.737	0.710	0.704	0.691	0.643	0.593

Table 6 and each column of Table 7. As an illustration, in the case where Y is 60, one part of the Oldham LA area appears within the top 1% of the 858 observations for employment deprivation score under Procedure 1, whilst one part of each of the Walsall and Bolton LA areas appears around the top 2%. For Procedure 2, those LA names are not evident, but Sefton (top 1%) and Bradford (top 2%) are. None of these LA areas appear in Table 1 (showing the top 10% or so LA areas for employment deprivation), and Bolton and Bradford do not appear in Table 4 either. This is further exemplar evidence against null H_{01}, and we might conclude that those considering the employment deprivation (or otherwise) of Bolton and Bradford would be well advised to go beyond an examination of a basic LA area league table and beyond straightforward league tables based on concentric distance from LA centre. Instead, league tables from our two new procedures should be considered.

Education, Skills and Training Deprivation Domain

We turn now to a parallel analysis for the education domain. Initially (in Table 8), we list the 30 most deprived LA areas based on weighted average of the LSOA education scores – the weightings in this case being based on LSOA overall population. However, this choice is not clear-cut – especially given the fact that the two component education sub-domains are based on two different groups (respectively, those who have not reached adulthood, and those who have). Table 8 also names the LA areas that occupy the upper and lower quartiles of the education deprivation distribution, as well as the median, and the least education-deprived LA area.

Only 14 LA areas are common to the top 30 in Table 1 and Table 8. Although strong in principle, the link between LA area employment deprivation and education deprivation is limited in reality. This fact may be relevant background when our null hypothesis H_{03} is being considered later.

Results shown below in Table 9 are for a basic linear regression of the natural logarithm of LA education score on the natural logarithm of LA land area (in hectares),

Table 8 Thirty most education-deprived LA areas (with rank, and a weighted score)

1	Kingston upon Hull	47.53	16	Bradford	35.76
2	Sandwell	42.42	17	Great Yarmouth	35.59
3	Stoke-on-Trent	41.12	18	North East Lincolnshire	35.49
4	Knowsley	39.78	19	Norwich	35.33
5	Walsall	38.95	20	Manchester	34.88
6	Mansfield	38.49	21	Wolverhampton	34.85
7	Barnsley	37.77	22	Blackpool	34.36
8	Nottingham	37.67	23	Rotherham	34.29
9	Leicester	37.53	24	Swale	34.20
10	Middlesbrough	37.41	25	Blackburn with Darwen	34.04
11	Bolsover	37.25	26	Wakefield	33.66
12	Corby	36.60	27	Oldham	33.46
13	Doncaster	36.59	28	Liverpool	33.39
14	Ashfield	36.50	29	Pendle	33.09
15	Burnley	36.10	30	Barking and Dagenham	32.95
82	Lincoln	26.71	164	Cheshire West & Chester	18.38
245	Aylesbury	13.41	326	Isles of Scilly	4.02

the natural logarithm of 2008 mid-year LA population and a set of eight region dummies. In this case, a simple ordering of the regional sample mean LA log education scores places London most favourably (i.e., with the least education deprivation, at 2.532) – in keeping with all other regions having positive regression estimates. Using

Table 9 Natural logarithm of LA education score, regression results

Variable	Estimate	Std. Err.	t ratio	P value	Sample means	
					Regressor	Dependent variable
ln (LA land area)	−0.1308	0.0198	−6.61	0.000	9.897	
ln (LA population, 2008)	0.2757	0.0393	7.02	0.000	11.800	
North East	1.0608	0.1329	7.98	0.000	0.037	3.337
North West	0.8927	0.0977	9.13	0.000	0.120	3.103
Yorkshire & The Humber	0.9962	0.1191	8.37	0.000	0.064	3.146
East Midlands	0.9657	0.1027	9.40	0.000	0.123	3.043
West Midlands	0.8846	0.1041	8.50	0.000	0.092	3.074
East	0.7838	0.0980	8.00	0.000	0.144	2.905
South East	0.5295	0.0905	5.85	0.000	0.206	2.697
South West	0.7488	0.1057	7.08	0.000	0.113	2.804
Constant	0.2249	0.4955	0.45	0.650	1.000	2.901
Number of observations	326	$F_{10,315} =$	18.33	0.000		
Sum of squared residuals	45.670	$R^2 =$	0.3678	Adjusted	$R^2 = 0.3478$	
Test for region controls		$F_{8,315} =$	18.41	0.000		

Y_2 as LA education score, and with predicted ln (Y_2) denoted as $\hat{Z}_2$ (with X_1, X_2 and X_3 again defined (respectively) as LA land area, population and population density), there are three equivalents for the London region:

    A. From Table 9, $\hat{Z}_2 = 0.2249 - 0.1308\ln(X_1) + 0.2757\ln(X_2)$.

    B. Equivalently, $\hat{Z}_2 = 0.2249 + 0.1449\ln(X_1) + 0.2757\ln(X_3)$.

    C. Alternatively, $\hat{Z}_2 = 0.2249 + 0.1449\ln(X_2) + 0.1308\ln(X_3)$.

The estimates attached to $\ln(X_1)$, $\ln(X_2)$ and $\ln(X_3)$ in the above three expressions are all statistically significant at the 1% significance level. The impact of LA land area on education deprivation score is negative (-0.1308) when population is controlled for, but positive ($+0.1449$) after controlling instead for population density. We can conclude that LA education deprivation scores also tend to be higher in more urban areas (which usually have lower land area, higher population and higher population density). For the LA data, the respective bivariate correlations of $\ln(X_1)$, $\ln(X_2)$ and $\ln(X_3)$ with the log of education deprivation score are -0.082, 0.257 and 0.184: this suggests less education deprivation in the less urban LA areas. If so, LA-based league tables on education deprivation score might be expected to favour less densely populated LA areas, just like their counterparts for employment deprivation.

Using, as 'benchmarks', predicted values (for education deprivation score) generated from the basic regression that produced Table 9, a 'benchmarked' league table can be constructed (see Table 18 in the Appendix). This time, 16 of the LA areas seen in Table 8 occupy top 30 positions (for having most education deprivation over and above their 'benchmark level').

Basic spatial analysis of LA area education deprivation scores, analogous to that undertaken previously for the employment deprivation case, again indicates evidence of positive spatial autocorrelation. Moran's I takes the values 0.236 (standardised weighting matrix) or 0.193 (binary), and the combination of a positive sign and strong statistical significance indicates positive global spatial autocorrelation. Geary's c is 0.776, and being statistically significantly below unity, strongly indicates positive local spatial autocorrelation. For the Getis and Ord statistic, $G(120) = 0.284$ and $G(40) = 0.051$: given that these values lie significantly below their respective mathematical expectations (0.305 and 0.0605), the indication is that there is 'cold spot' clustering of areas of low education deprivation. In this instance, the result does not appear to depend critically on the threshold distance for 'near neighbours'.

An LSOA-based regression for log education deprivation scores corresponds to the LA-based regression with results in Table 9 above. In parallel to Table 3, Table 10's LSOA-based results are shown initially for eight region controls, and then for a set of 325 LA controls.

Comparison of LA and LSOA results (Tables 9 and 10) shows no change of sign for the estimate on the log of population (although it loses its significance for Table 10's region controls specification), so there is no evidence here against null H_{02}. A re-specification of the regression with log of LSOA population density in place of the log of LSOA population would yield a negative estimate attached to the separate 'log of LSOA land area' regressor for either set of controls. For LA controls, $-0.0777 = -0.2144 - (-0.1367)$. The impact of LSOA land area on education deprivation score is negative: it remains so after controlling for population density, although such control shrinks the estimate, rather than expanding it – as was the case for employment

Table 10 Natural logarithm of LSOA EDUCATION score, regression results

Variable	Estimate	Std. Err.	t ratio	P value	Sample mean
Region controls					
ln (LSOA land area)	−0.2118	0.0040	−52.40	0.000	4.320
ln (LSOA population, 2008)	0.0404	0.0349	1.16	0.246	7.353
Constant	2.6027	0.2559	10.17	0.000	1.000
Number of observations	32,482	$F_{10,32,471} =$	472.78	0.000	[Region controls]
Sum of squared residuals	33,758.65	$R^2 =$	0.1271	Adjusted	$R^2 = 0.1268$
Test for region controls		$F_{8,32,471} =$	373.9	0.000	
LA controls					
ln (LSOA land area)	−0.2144	0.0043	−49.44	0.000	4.320
ln (LSOA population, 2008)	0.1367	0.0325	4.21	0.000	7.353
Constant	0.1555	0.4787	0.32	0.745	1.000
Number of observations	32,482	$F_{327,32,154} =$	42.86	0.000	[LA controls]
Sum of squared residuals	26,934.71	$R^2 =$	0.3035	Adjusted	$R^2 = 0.2965$
Test for LA versus region controls		$F_{317,32,154} =$	25.70	0.000	

deprivation. Thus, we have some initial evidence against null hypothesis H_{03}. In this instance, Moran's I is 0.138, and Geary's c is 0.876. Once again, each is statistically significant, and demonstrates positive spatial autocorrelation. Meanwhile, the Getis and Ord statistic ($G(90)$) is 0.202, so education deprivation also displays statistically significant 'cold spot' clustering.

Taken together, Tables 1, 4 and 8 plus Table 11 (below) demonstrate that there is a rather different spatial mix of education deprivation, compared with employment deprivation. From the south, the London Borough of Barking and Dagenham appears in Table 8, whilst Basildon – which is less than 35 km from the edge of Greater London – can be seen in the first column of Table 11, along with Sedgemoor and Swale (contrast just Hastings in Table 1). There is also a change in the distribution of LA centres on the northern side of the north-south divide – with significantly more representation from the Yorkshire and The Humber region, and also from the East Midlands (whereas Bolsover was alone in Table 4).

Based on the LSOAs centred within (respectively) 3 km, 6 km, 12 km and 24 km of their LA centre, Table 11 shows the 20 most deprived LA areas, for the education deprivation domain. It is also very striking that only two LA centres appear within the top 20 for education deprivation for all the four chosen radii, and they are Barnsley (Yorkshire and The Humber) and Walsall (West Midlands). Those that can be seen for three out of the four are Kingston upon Hull, King's Lynn and West Norfolk, Ashfield, Dudley, Sandwell and South Staffordshire. Of these eight named LA areas, five appear within the top 15 of Table 8. However, the other three are some way outside the top 30 – specifically King's Lynn and West Norfolk (58th), Dudley (60th) and South Staffordshire[15] (214th). This is further evidence against null H_{01}. Since some parallel cases

---

[15] The very particular shape and population density of the South Staffordshire LA area, and its proximity to relatively deprived LA areas such as Wolverhampton and Walsall, combine to give much more adverse rankings on the basis of concentric bandings (rather than the simple LA area itself). We would argue strongly that, for at least some league table purposes, the measurement of deprivation in close proximity (rather than simply within a given administrative area) is informative.

Table 11 Top 20 most education-deprived LA areas at 3 km, 6 km, 12 km and 24 km radius

	3 km radius	6 km radius	12 km radius	24 km radius
1	Great Yarmouth	Kingston upon Hull	Rotherham	Doncaster
2	Walsall	Bolsover	Wolverhampton	Barnsley
3	Sandwell	Great Yarmouth	Mansfield	Lichfield
4	NE Lincolnshire	Walsall	Bolsover	Rotherham
5	Basildon	Sedgemoor	South Staffordshire	Bassetlaw
6	Stoke-on-Trent	Sandwell	Kingston upon Hull	Wakefield
7	Oldham	Middlesbrough	Ashfield	King's Lynn & W. Norfolk
8	King's Lynn & W. Norfolk	Barnsley	Walsall	Sheffield
9	Birmingham	Rotherham	Doncaster	Kingston upon Hull
10	Barnsley	Bradford	Barnsley	Bolsover
11	Blackburn with Darwen	Mansfield	Sandwell	Chesterfield
12	Bradford	Burnley	Swale	NE Derbyshire
13	Dudley	Tendring	Staffs. Moorlands	Mansfield
14	North Lincolnshire	King's Lynn & W. Norfolk	Wakefield	East Yorkshire
15	Burnley	Stoke-on-Trent	Redcar & Cleveland	Walsall
16	Corby	Ashfield	Dudley	Wolverhampton
17	Ashfield	Corby	NE Lincolnshire	Tamworth
18	Blackpool	Gateshead	Great Yarmouth	Hartlepool
19	South Staffordshire	Hyndburn	Oldham	Dudley
20	Middlesbrough	Leeds	Tendring	South Staffordshire

were found previously for employment deprivation, this might also be at least soft evidence that is broadly consistent with null H_{03}.

As we did for employment deprivation, we now examine the relationship between education deprivation rankings based on LA areas, and each of the sets of rankings based on our concentric banding methodology (with cut-offs at integer numbers of kilometres from the LA centre), using Spearman's rank correlation coefficient (r_s). This is shown in Table 12 below.

Again the initial increase in r_s (as radius increases) is probably likely to be linked to the fact that more than 80% of actual LA areas in England are larger in size than a circle of radius 4 km. Just as was true for employment deprivation, r_s begins to fall back beyond a 6 km radius. The values of r_s at a given radius are always lower for education deprivation – although 0.695 for a radius of 24 km is still a substantial rank correlation. As with Table 5, the evidence here against null hypothesis H_{01} is rather modest. There is also at least an indication of evidence in line with null H_{03}.

Figure 5 below illustrates education deprivation for some LA centres highlighted in the earlier discusion of Table 11.

Note that all these LA centres appear in the top 20 for education deprivation in at least three of our four ranking lists (see Table 11): it would be interesting to get a wider

Table 12 Values of Spearman's rank correlation coefficient (in each case, the LA area ranking is compared to a ranking for LA centres based on concentric banding at a particular radius)

Radius (km)	r_s	Radius (km)	r_s	Radius (km)	r_s	Radius (km)	r_s
1	0.7110	11	0.8337	21	0.7070	31	0.6452
2	0.7763	12	0.8166	22	0.7040	32	0.6340
3	0.8122	13	0.8072	23	0.7018	33	0.6222
4	0.8446	14	0.7899	24	0.6950	34	0.6140
5	0.8635	15	0.7757	25	0.6854	35	0.6078
6	0.8737	16	0.7590	26	0.6764	36	0.6064
7	0.8655	17	0.7444	27	0.6700	37	0.6058
8	0.8584	18	0.7339	28	0.6641	38	0.6048
9	0.8569	19	0.7238	29	0.6604	39	0.6017
10	0.8460	20	0.7148	30	0.6540	40	0.5939

perspective by viewing some of these cases benchmarked against LA centres with much less education deprivation.

Using Procedure 1, we can obtain regression results for education deprivation. With LA controls (lower half of Table 13), the estimates on population are rather different from either the employment deprivation results, or Tables 9 and 10 for education deprivation – being negative for higher values of Y, and statistically significant at the 10% level for $Y = 50$ and $Y = 60$.

Thus, there is some evidence here against two null hypotheses – H_{02} (regarding spatial resolution) and H_{03}. However, our measures of spatial autocorrelation – in the first half of Table 20 in the Appendix – yield similar conclusions to their counterparts from the employment deprivation case.

Using Procedure 2, we can obtain the set of results shown in Table 14, below.

With LA controls (lower half of Table 14), the pattern of the estimates attached to the population regressor, across different values of Y, this time exhibits greater volatility than (the trend) under Procedure 1. It is interesting that the relationship between the
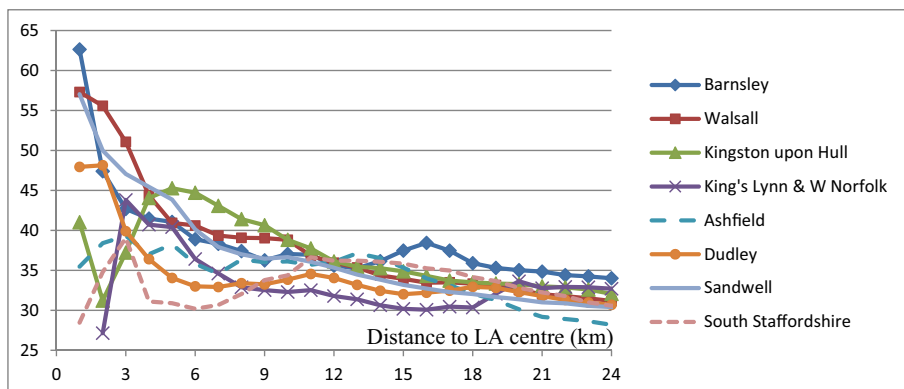

Fig. 5 IMD2010 education deprivation domain score for eight selected LA centres

Table 13 Regression estimates across *Y* values, dependent variable ln(LSOA education score)

Procedure 1	Estimates, based on different values of *Y* (in thousands)							
Variable	60	50	40	30	25	20	12.5	7.5
Region controls								
ln (land area)	−0.159*	−0.166*	−0.169*	−0.171*	−0.173*	−0.172*	−0.179*	−0.178*
ln (population, 2008)	0.373*	0.425*	0.280*	0.362*	0.249+	0.558*	0.258+	0.160+
Constant	−0.464	−0.941	0.670	−0.134	1.050	−2.011	1.026	1.908*
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.370	0.366	0.357	0.341	0.338	0.322	0.299	0.254
LA controls								
ln (land area)	−0.179*	−0.184*	−0.188*	−0.183*	−0.183*	−0.173*	−0.178*	−0.166*
ln (population, 2008)	−1.112#	−0.849#	−0.289	−0.141	−0.284	0.181	0.029	0.148*
Constant	13.029+	10.603#	5.392	3.987	5.318+	0.917	2.369+	1.072#
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.712	0.708	0.698	0.682	0.678	0.652	0.611	0.535

results in Tables 13 and 14 appears somewhat different to that seen previously for employment deprivation (in Tables 6 and 7). Thus, there is further evidence against null hypothesis H$_{02}$ for spatial resolution and spatial unit construction criteria, and some evidence against null H$_{03}$ across spatial unit construction criteria. The second half of Table 20 (in the Appendix) is, however, broadly similar to the second half of Table 19.

As with Tables 6 and 7, there is a separate (basically unreported) league table underlying each column of Table 13 and each column of Table 14. However, as before, the *Y* = 60 case (for example) offers evidence against null hypothesis H$_{01}$, with some concentration of education deprivation in LA areas not shown in the raw LA-based league table (Table 8). In this instance, part of Southampton lies well within the top 1% of the 858 observations for Procedure 1 – while a part of Birmingham is in the top 2%,

Table 14 Regression estimates across *Y* values, dependent variable ln(LSOA education score)

Procedure 2	Estimates, based on different values of *Y* (in thousands)							
Variable	60	50	40	30	25	20	12.5	7.5
Region controls								
ln (land area)	−0.157*	−0.163*	−0.162*	−0.167*	−0.167*	−0.168*	−0.177*	−0.178*
ln (population, 2008)	0.370*	0.449*	0.310*	0.419*	0.306+	0.573*	0.430*	0.327*
Constant	−0.463	−1.257	0.274	−0.795	0.385	−2.222	−0.657	0.366
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.321	0.304	0.289	0.267	0.260	0.250	0.228	0.200
LA controls								
ln (land area)	−0.183*	−0.181*	−0.180*	−0.182*	−0.179*	−0.184*	−0.186*	−0.176*
ln (population, 2008)	0.677	0.596	0.851#	0.515	0.156	0.293	0.293+	0.310*
Constant	−3.654	−2.907	−5.293	−2.145	1.183	−0.069	−0.050	−0.286
Observations	858	1036	1285	1724	2060	2580	4109	6869
Adjusted R^2	0.583	0.556	0.555	0.528	0.523	0.514	0.476	0.436

Table 15 Values of Spearman's rank correlation coefficient (in each case, the employment deprivation ranking is compared to education deprivation for LA centres based on concentric banding at a particular radius)

Radius (km)	r_s	Radius (km)	r_s	Radius (km)	r_s	Radius (km)	r_s
1	0.8077	11	0.8022	21	0.7758	31	0.7569
2	0.8174	12	0.8012	22	0.7750	32	0.7600
3	0.8058	13	0.7955	23	0.7714	33	0.7623
4	0.8057	14	0.7925	24	0.7708	34	0.7620
5	0.7961	15	0.7923	25	0.7690	35	0.7648
6	0.7923	16	0.7935	26	0.7678	36	0.7669
7	0.7922	17	0.7913	27	0.7660	37	0.7698
8	0.7969	18	0.7882	28	0.7583	38	0.7747
9	0.7961	19	0.7862	29	0.7539	39	0.7792
10	0.8015	20	0.7811	30	0.7540	40	0.7819

and a part of Bristol is just outside that range. For Procedure 2, a part of Leeds and a part of Bristol are well within the top 1%, with a part of Sheffield just outside and two parts of Birmingham within the top 2%. Of these, only Southampton and Bristol are also missing from radius-based Table 11. The impact of changes in spatial resolution, and of changes in spatial unit construction criteria, are quite similar to those for employment deprivation – evidence in line with null H_{03}.

The Two Domains Compared

Another issue to consider is the strength of the correlation between employment deprivation ranking and education deprivation ranking. On the basis of England's
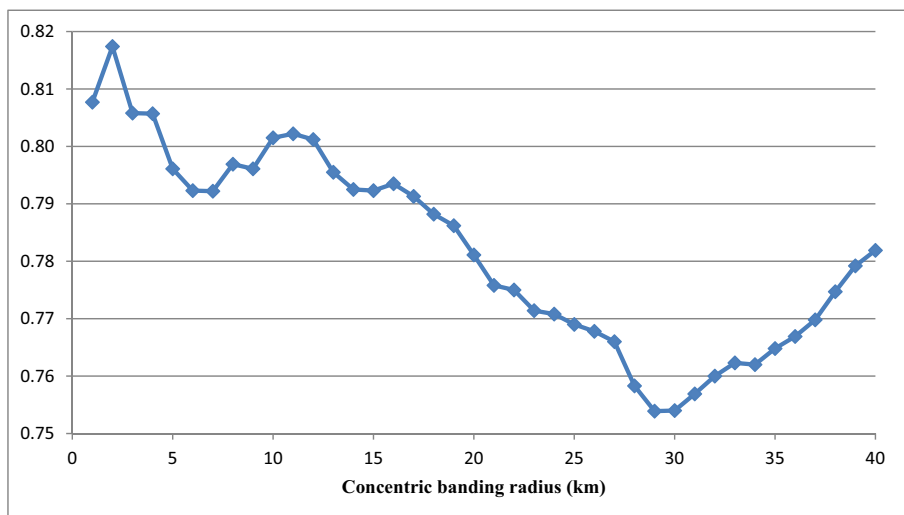

Fig. 6 Spearman's rank correlation coefficient, comparing employment deprivation and education deprivation rankings for LA centres, using concentric banding for a range of radii

326 LA areas themselves, Spearman's rank correlation coefficient for this relationship is 0.8086. This compares with a corresponding value of 0.7898 through the 32,482 observations at LSOA level – and seems to be in keeping with the findings of Gehlke and Biehl (1934), who reported a tendency for the (Pearson) correlation coefficient to increase when (US) census tract data were grouped together in contiguous groupings.[16]

Table 15 and Fig. 6 report and illustrate the rank correlations for employment deprivation and education deprivation based instead upon our concentric banding methodology.

Although, on the face of it, this remains a strong positive rank correlation throughout – as we might expect – it should be noted that r_s exhibits several turning points within the first three columns (up to 30 km radius). This indicates that, for England, the relationship between the location of local employment deprivation and that of local education deprivation is not entirely straightforward. That much was apparent from observation of the different regional spreads for league tables of local deprivation shown above in Tables 1, 4, 8 and 11 – which revealed indications that null hypothesis H_{01} does not really hold, although there is little clear evidence to contradict null hypothesis H_{03}.

Conclusions

In addition to comparing and contrasting regression estimates for administrative units such as LA areas and LSOAs, and for systematically constructed collections of LSOAs, it is worthwhile to produce and compare league tables (given their popularity for communication with the general public). We have demonstrated some rather large league table differences between which of England's LA areas exhibit employment deprivation and which demonstrate deprivation for the education, skills and training domain. Concentrations of the former type of deprivation are particularly evident in northern England (although not much in Yorkshire and The Humber) and the West Midlands. On the other hand, education deprivation has a considerable focus in Yorkshire and The Humber, as well as the West Midlands; while there is also some spread into the East Midlands, the East of England and even Greater London.

Rather unsurprisingly, we find evidence against our first null hypothesis. League table positions, based on either of our measures of deprivation, can depend crucially on the size category of spatial unit being considered, and the criteria by which those spatial units are constructed. Traditional simple league table construction based solely on defined LA areas is unlikely to give a sufficiently clear picture. Precisely how far to investigate beyond the initial LA-based league table depends upon the circumstances. We have given examples, for employment deprivation in Bolton and Bradford, and for education deprivation in Bristol and Southampton, where a concentration of deprivation across an area with a population of the order of 60,000 people becomes evident for league tables based on our two procedures. With regard to league tables, this evidence is also in line with our final null hypothesis that the impact of spatial resolution and

---

[16] Corresponding values for the Pearson correlation coefficient in our case are 0.7985 (LA) and 0.7822 (LSOA).

spatial unit construction criteria on different deprivation domains may be similar. From a practical perspective, fuller investigation offers an opportunity for a deeper understanding of the concentration of deprivation – and thus to target local economic assistance measures more efficiently where they are most urgently needed. Simultaneously, the identification of an area as being less deprived than it initially appears might alter perception favourably and change locational decisions regarding the siting of organisations or inward mobility for work or residence.

As basic regression analysis demonstrates, both employment and education deprivation scores are likely to be higher (more adverse) for more urban LA areas with higher population totals, higher population density and lower land area – but physically larger LA areas are more deprived, once population density has been controlled for. However, the picture is less clear when LSOA area data are subjected to similar analysis – with larger LSOA areas tending to be less deprived (and, in the case of employment deprivation, more so after controlling for population density). Additional regression analysis, performed after having introduced a couple of distinct procedures for splitting up LA areas into similarly populated subsets of their constituent LSOAs, yields somewhat mixed results. Overall, regression results are mildly dependent on area unit definitions, with some evidence of different outcomes from concentric and angular approaches, and also across our two deprivation measures. This constitutes at least limited evidence against our second and third null hypotheses. Meanwhile, our measures of spatial autocorrelation are broadly stable across both areal specification and the two deprivation domains.

Future work should consider other data, concerning local economic performance, that are available at similar levels of spatial disaggregation. Such analysis could be augmented by comparisons with results generated for such functional areas as Travel To Work Areas (and also Primary Urban Areas, if the precise definitions of these becomes publicly accessible – see DCLG (2010), and also Smith et al. (2010) for some discussion). There is also scope for further investigation of the links between rankings on education deprivation and employment deprivation, as well as other deprivation domains, and various measures of local economic performance. Ideally, similar analysis should also be attempted for other countries, since there is no guarantee that identical findings will be obtained in every nation's case: it is of interest to discover whether variations are systematic and, if so, in what ways. As indicated by He et al. (2018), the exploration of changes in deprivation over time would need to keep in mind the impact of changes in the definition of administrative boundaries – in our case, some LSOAs were redefined in the 2011 Census.

Compliance with ethical standards

Conflict of Interest   The authors declare that they have no conflict of interest.

Appendix

Table 16 Summary statistics

Variable	Mean	Std. Dev.	Minimum	Maximum
Split by LA				
Land area	40,010.76	55,322.75	290.39	502,608.6
Population, 2008	157,867	107,101	2261	1,019,228
Population density	15.85	22.41	0.24	141.09
Employment score	9.08	3.55	2.17	19.85
Education score	20.14	8.66	4.02	47.53
Split by LSOA				
Land area	401.08	1137.24	2	67,282
Population, 2008	1584.41	309.23	548	11,525
Population density	40.92	40.52	0.026	751.5
Employment score	10.05	6.53	0.250	75.45
Education score	21.69	18.81	0.012	99.34
Split using Procedure 1				
$Y = 60$ land area	15,184.11	25,362.29	290	286,260
$Y = 60$ population, 2008	59,982.1	8457.23	2261	89,637
$Y = 60$ population density	24.02	28.08	0.214	173.94
$Y = 60$ employment score	9.85	3.93	2.170	21.74
$Y = 60$ education score	21.38	10.29	3.217	56.84
Split using Procedure 2				
$Y = 60$ land area	15,184.11	25,361.11	290	281,940
$Y = 60$ population, 2008	59,982.1	8448.03	2261	89,637
$Y = 60$ population density	23.00	27.72	0.229	188.22
$Y = 60$ employment score	9.85	4.02	2.170	28.66
$Y = 60$ education score	21.39	10.85	3.336	65.28
Split using Procedure 1				
$Y = 50$ land area	12,575.26	21,699.15	275	276,909
$Y = 50$ population, 2008	49,676.3	5520.55	2261	74,832
$Y = 50$ population density	24.23	28.12	0.183	175.36
$Y = 50$ employment score	9.85	4.00	2.170	24.89
$Y = 50$ education score	21.42	10.50	3.282	56.92
Split using Procedure 2				
$Y = 50$ land area	12,575.26	21,782.06	246	274,585
$Y = 50$ population, 2008	49,676.3	5520.55	2261	74,832
$Y = 50$ population density	23.36	27.85	0.195	203.82
$Y = 50$ employment score	9.86	4.14	2.170	28.65
$Y = 50$ education score	21.43	11.22	2.292	65.91
Split using Procedure 1				
$Y = 40$ land area	10,138.5	18,074.98	195	214,234
$Y = 40$ population, 2008	40,050.31	3738.31	2261	59,088
$Y = 40$ population density	24.63	28.29	0.238	201.73

Table 16 (continued)

Variable	Mean	Std. Dev.	Minimum	Maximum
$Y = 40$ employment score	9.87	4.06	2.170	24.92
$Y = 40$ education score	21.45	10.67	2.910	59.57
Split using Procedure 2				
$Y = 40$ land area	10,138.5	18,491.29	213	250,613
$Y = 40$ population, 2008	40,050.31	3753.88	2261	59,088
$Y = 40$ population density	24.02	28.08	0.162	198.20
$Y = 40$ employment score	9.88	4.22	2.170	28.35
$Y = 40$ education score	21.46	11.49	2.409	65.26
Split using Procedure 1				
$Y = 30$ land area	7556.83	13,819.53	121	144,219
$Y = 30$ population, 2008	29,851.88	2194.37	2261	38,121
$Y = 30$ population density	25.04	28.71	0.170	254.13
$Y = 30$ employment score	9.85	4.13	2.170	26.49
$Y = 30$ education score	21.46	10.96	2.806	64.26
Split using Procedure 2				
$Y = 30$ land area	7556.83	14,085.19	146	239,493
$Y = 30$ population, 2008	29,851.88	2194.35	2261	40,466
$Y = 30$ population density	24.58	28.72	0.135	191.48
$Y = 30$ employment score	9.86	4.34	2.170	33.13
$Y = 30$ education score	21.47	11.98	1.342	64.19
Split using Procedure 1				
$Y = 25$ land area	6324.26	12,261.21	96	196,929
$Y = 25$ population, 2008	24,982.84	1730.13	2261	35,575
$Y = 25$ population density	25.61	28.93	0.131	238.18
$Y = 25$ employment score	9.88	4.18	2.170	28.25
$Y = 25$ education score	21.48	11.08	2.847	65.35
Split using Procedure 2				
$Y = 25$ land area	6324.26	12,301.34	120	222,301
$Y = 25$ population, 2008	24,982.84	1751.10	2261	35,575
$Y = 25$ population density	25.22	29.09	0.117	212.24
$Y = 25$ employment score	9.89	4.43	2.170	34.46
$Y = 25$ education score	21.49	12.23	1.435	67.73
Split using Procedure 1				
$Y = 20$ land area	5049.60	9778.25	96	122,976
$Y = 20$ population, 2008	19,947.54	1236.55	2261	25,766
$Y = 20$ population density	25.90	29.08	0.140	225.93
$Y = 20$ employment score	9.85	4.24	2.170	29.57
$Y = 20$ education score	21.43	11.32	2.274	67.34
Split using Procedure 2				
$Y = 20$ land area	5049.60	9927.78	89	205,532
$Y = 20$ population, 2008	19,947.54	1224.42	2261	26,136
$Y = 20$ population density	25.73	29.32	0.101	215.93
$Y = 20$ employment score	9.87	4.50	2.170	36.06
$Y = 20$ education score	21.45	12.51	1.261	69.23
Split using Procedure 1				
$Y = 12.5$ land area	3170.59	6659.34	42	125,582

Table 16 (continued)

Variable	Mean	Std. Dev.	Minimum	Maximum
$Y = 12.5$ population, 2008	12,524.86	848.69	2261	16,675
$Y = 12.5$ population density	27.33	30.24	0.107	271.79
$Y = 12.5$ employment score	9.87	4.40	1.966	30.82
$Y = 12.5$ education score	21.45	11.86	0.622	71.81
Split using Procedure 2				
$Y = 12.5$ land area	3170.59	7069.25	48	201,875
$Y = 12.5$ population, 2008	12,524.86	887.25	2261	19,090
$Y = 12.5$ population density	27.42	30.52	0.069	240.69
$Y = 12.5$ employment score	9.88	4.70	0.799	37.70
$Y = 12.5$ education score	21.46	13.17	0.592	74.44
Split using Procedure 1				
$Y = 7.5$ land area	1896.63	4318.95	20	90,240
$Y = 7.5$ population, 2008	7492.31	798.43	2261	17,301
$Y = 7.5$ population density	29.52	31.82	0.082	314.12
$Y = 7.5$ employment score	9.89	4.67	1.664	31.67
$Y = 7.5$ education score	21.48	12.79	0.642	84.39
Split using Procedure 2				
$Y = 7.5$ land area	1896.63	4623.28	21	172,814
$Y = 7.5$ population, 2008	7492.31	781.45	1599	15,800
$Y = 7.5$ population density	29.80	32.08	0.049	315.86
$Y = 7.5$ employment score	9.89	4.93	0.632	39.80
$Y = 7.5$ education score	21.48	13.99	0.430	81.67

Table 17 LA league table on 'excess employment deprivation above benchmark' (with rank)

1	Hastings	8.87	16	Hackney	5.04
2	Thanet	7.59	17	Dover	4.95
3	Great Yarmouth	7.46	18	Copeland	4.71
4	Bolsover	6.42	19	Corby	4.68
5	Knowsley	6.34	20	Barrow-in-Furness	4.67
6	Shepway	5.83	21	Fenland	4.60
7	Hartlepool	5.73	22	Waveney	4.53
8	Barnsley	5.63	23	Rother	4.50
9	East Lindsey	5.63	24	Eastbourne	4.49
10	Mansfield	5.43	25	Boston	4.49
11	Tendring	5.40	26	Barking and Dagenham	4.38
12	Burnley	5.39	27	Bassetlaw	4.36
13	Isle of Wight	5.35	28	Swale	4.34
14	Stoke-on-Trent	5.20	29	King's Lynn and West Norfolk	4.27
15	West Somerset	5.10	30	Scarborough	4.21

Table 18 LA league table on 'excess education deprivation above benchmark' (with rank)

1	Barking and Dagenham	21.05	16	Pendle	13.27
2	Swale	20.39	17	Stoke-on-Trent	13.22
3	Great Yarmouth	17.52	18	Adur	13.09
4	Fenland	17.22	19	Sandwell	12.69
5	Bolsover	16.83	20	Thanet	12.52
6	Corby	16.08	21	Waveney	12.49
7	Burnley	15.35	22	Ashfield	12.24
8	Knowsley	14.81	23	Bassetlaw	12.00
9	Kingston upon Hull	14.44	24	Carlisle	11.97
10	Boston	14.43	25	Barnsley	11.75
11	Copeland	14.39	26	Shepway	11.64
12	Hastings	14.19	27	Norwich	11.07
13	Mansfield	14.09	28	Thurrock	11.05
14	King's Lynn and West Norfolk	14.01	29	Blackburn with Darwen	10.98
15	Tendring	13.97	30	Havant	10.98

Table 19 Spatial autocorrelation tests, log LSOA employment score, various values of Y

	Different values of Y (in thousands)							
	60	50	40	30	25	20	12.5	7.5
Procedure 1								
Moran's I	0.460	0.455	0.454	0.446	0.447	0.431	0.400	0.354
E[I]	−0.0012	−0.0010	−0.0008	−0.0006	−0.0005	−0.0004	−0.0002	−0.0001
z stat for Moran's I	35.31	41.41	48.74	62.63	73.42	86.59	126.12	193.60
Geary's c	0.559	0.558	0.556	0.564	0.567	0.581	0.613	0.662
E[c]	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
z stat for Geary's c	−29.11	−33.68	−39.15	−48.25	−54.14	−61.43	−80.47	−103.28
G[90] (Getis & Ord)	0.206	0.209	0.210	0.205	0.206	0.205	0.206	0.206
E[G[90]]	0.215	0.220	0.221	0.214	0.215	0.215	0.215	0.215
z stat for G	−3.41	−4.28	−4.48	−4.58	−5.18	−5.72	−6.99	−8.38
Procedure 2								
Moran's I	0.346	0.346	0.342	0.330	0.326	0.320	0.306	0.294
z stat for Moran's I	55.40	65.20	77.21	95.92	108.09	126.89	173.03	245.47
Geary's c	0.665	0.666	0.670	0.683	0.685	0.693	0.706	0.720
z stat for Geary's c	−37.46	−43.02	−48.29	−54.93	−58.45	−65.11	−79.69	−100.67
G[90] (Getis & Ord)	0.205	0.203	0.204	0.204	0.205	0.205	0.205	0.205
E[G[90]]	0.214	0.214	0.214	0.213	0.215	0.214	0.215	0.215
z stat for G	−3.39	−4.09	−4.18	−4.41	−4.95	−5.45	−6.64	−8.00
Observations	858	1036	1285	1724	2060	2580	4109	6869

Table 20 Spatial autocorrelation tests, log LSOA education score, various values of Y

	Different values of Y (in thousands)							
	60	50	40	30	25	20	12.5	7.5
Procedure 1								
Moran's I	0.391	0.394	0.395	0.378	0.380	0.363	0.329	0.284
E[I]	−0.0012	−0.0010	−0.0008	−0.0006	−0.0005	−0.0004	−0.0002	−0.0001
z stat for Moran's I	30.06	35.89	42.41	53.18	62.41	73.05	103.87	155.47
Geary's c	0.648	0.641	0.638	0.650	0.649	0.664	0.695	0.737
E[c]	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
z stat for Geary's c	−23.24	−27.54	−32.16	−39.36	−45.10	−50.55	−66.04	−83.87
G[90] (Getis & Ord)	0.202	0.206	0.207	0.201	0.202	0.202	0.202	0.202
E[G[90]]	0.215	0.220	0.221	0.214	0.215	0.215	0.215	0.215
z stat for G	−3.86	−4.58	−4.96	−5.27	−5.85	−6.36	−7.74	−9.22
Procedure 2								
Moran's I	0.278	0.273	0.267	0.256	0.252	0.248	0.238	0.226
z stat for Moran's I	44.52	51.54	60.30	74.32	83.56	98.46	134.82	189.01
Geary's c	0.757	0.760	0.766	0.775	0.776	0.779	0.786	0.796
z stat for Geary's c	−27.04	−31.22	−35.03	−40.89	−44.89	−50.78	−63.04	−78.25
G[90] (Getis & Ord)	0.201	0.199	0.200	0.200	0.202	0.201	0.201	0.201
E[G[90]]	0.214	0.214	0.214	0.213	0.215	0.214	0.215	0.215
z stat for G	−3.69	−4.45	−4.69	−4.86	−5.34	−5.78	−7.09	−8.51
Observations	858	1036	1285	1724	2060	2580	4109	6869

References

Ajebon, M. O., & Norman, P. (2016). Beyond the census: a spatial analysis of health and deprivation in England. *GeoJournal, 81*, 395–410.

Burke, A., & Jones, A. (2018). The development of an index of rural deprivation: a case study of Norfolk, England. *Social Science and Medicine* in press and available online 15 September 2018.

Cockings, S., Harfoot, A., Martin, D., & Hornby, D. (2011). Maintaining existing zoning systems using automated zone-design techniques: methods for creating the 2011 census output geographies for England and Wales. *Environment and Planning A, 43*, 2399–2418.

Cörvers, F., Hensen, M., & Bongaerts, D. (2009). Delimitation and coherence of functional and administrative regions. *Regional Studies, 43*(1), 19–31.

Deas, I., Robson, B., Wong, C., & Bradford, M. (2003). Measuring neighbourhood deprivation: a critique of the index of multiple deprivation. *Environment and Planning C: Government and Policy, 21*, 883–903.

Department for Communities and Local Government. (2010). *primary urban areas and travel to work area indicators: Updating the evidence base on cities* http://webarchive.nationalarchives.gov.uk/20120919132719/http://www.communities.gov.uk/publications/corporate/statistics/primaryurbanareas042010. Accessed 1 Sep 2018

Department for Communities and Local Government. (2011). *The English indices of deprivation 2010*. See URL (valid on 29/01/2018): https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6871/1871208.pdf. Accessed 1 Sep 2018

Geary, R. C. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician, 5*(3), 115–145.

Gehlke, C. E., & Biehl, K. (1934). Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association, 29*(185, supplement), 169–170.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis, 24*(3), 189–206.

Getis, A., & Ord, J. K. (1995). Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis, 27*(4), 286–306.

He, S., King, C., Chung, L., Bayrak, M. M., & Wang, W. (2018). Administrative boundary changes and regional inequality in provincial China. *Applied Spatial Analysis and Policy, 11*, 103–120.

Huby, M., Cinderby, S., White, P., & de Bruin, A. (2009). Measuring inequality in rural England: the effects of changing spatial resolution. *Environment and Planning A, 41*, 3023–3037.

Jones, C. (2000). The estimation of rates of unemployment for small spatial units. *Local Economy, 15*(2), 144–158.

Martin, D. (2003). Extending the automated zoning procedure to reconcile incompatible zoning systems. *International Journal of Geographical Information Science, 17*(2), 181–196. https://doi.org/10.1080/713811750.

McLennan, D., Barnes, H., Noble, M., Davies, J., Garratt, E. and Dibben, C. (2011): *The English indices of deprivation 2010: Technical Report*, DCLG. See URL (valid on 29/01/2018): https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/6320/1870718.pdf

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*(1–2), 17–23.

Noble, M., Wright, G., Smith, G., & Dibben, C. (2006). Measuring multiple deprivation at the small-area level. *Environment and Planning A, 38*, 169–185.

Nolan, M. A., Reynolds, M. M., & Trotter, S. D. (2012). Does the local economic performance league table lie? Concentric banding and the index of multiple deprivation 2010. *Local Economy, 27*(4), 403–418.

Office for National Statistics (2001): *Census 2001*. See URL (valid on 29/01/2018): https://www.ons.gov.uk/census/2001censusandearlier/aboutcensus2001.

Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geo Books.

Pisati, M. (2001). Tools for spatial data analysis. *Stata Technical Bulletin*, (60), 21–37. See the following URL (valid on 09/01/2018): http://www.stata.com/products/stb/journals/stb60.pdf.

Rae, A. (2012). Spatially concentrated deprivation in England: an empirical assessment. *Regional Studies, 46*(9), 1183–1199. https://doi.org/10.1080/00343404.2011.565321.

Séguin, A.-M., Apparicio, P., & Riva, M. (2012). The impact of geographical scale in identifying areas as possible sites for area-based interventions to tackle poverty: the case of Montréal. *Applied Spatial Analysis and Policy, 5*, 231–251.

Smith, T., Noble, S., Avenell, D., & Lally, G. (2010). *Technical feasibility for improving the state of the cities database*, Department for Communities and Local Government.

Sokal, R. R., Oden, N. L., & Thomson, B. A. (1998). Local spatial autocorrelation in a biological model. *Geographical Analysis, 30*(4), 331–354.