Geography and economic performance: Exploratory Spatial Data Analysis for Great Britain

Patacchini, Eleonora; Rice, Patricia

Postprint / Postprint
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:
Patacchini, E., & Rice, P. (2007). Geography and economic performance: Exploratory Spatial Data Analysis for Great Britain. *Regional Studies*, *41*(4), 489-508. https://doi.org/10.1080/00343400600928384

Geography and economic performance: Exploratory Spatial Data Analysis for Great Britain

Journal:	*Regional Studies*
Manuscript ID:	CRES-2005-0247.R1
Manuscript Type:	Main Section
JEL codes:	O4 - Economic Growth and Aggregate Productivity < O - Economic Development, Technological Change, and Growth, R11 - Regional Economic Activity: Growth, Development, and Changes < R1 - General Regional Economics < R - Urban, Rural, and Regional Economics, R12 - Size and Spatial Distributions of Regional Economic Activity < R1 - General Regional Economics < R - Urban, Rural, and Regional Economics, O18 - Regional, Urban, and Rural Analyses < O1 - Economic Development < O - Economic Development, Technological Change, and Growth
Keywords:	regional disparities,, income per worker, , productivity,, occupational composition, , spatial autocorrelation

Geography and Economic Performance: Exploratory Spatial Data Analysis for Great Britain

Eleonora Patacchini*
University of Rome "La Sapienza"

Patricia Rice
University of Southampton

Abstract

This paper uses the techniques of exploratory spatial data analysis to analyse patterns of spatial association for different indicators of economic performance, and in so doing identify and describe the spatial structure of economic performance for Great Britain. This approach enables us to identify a number of significant local regimes – clusters of areas in which income per worker differs significantly from the global average – and investigate whether these come about primarily through spatial association in occupational composition or in productivity. Our results show that the contributions of occupational composition and productivity vary significantly across local regimes. The 'winner's circle' of areas in the south and east of England benefits from both above average levels of productivity and better than average occupational composition, while the low income regime in the north of England suffers particularly from poor occupational composition.

Key words: regional disparities, income per worker, productivity, occupational composition, spatial autocorrelation

JEL Classification: O18, O4, R11, R12

We would like to thank Henry Overman and Anthony Venables for valuable comments. The research is funded by the Evidence Based Policy Fund , HM Treasury, Department of Trade and Industry and the Office of the Deputy Prime Minister of the UK as part of the project "Regional inequalities in the UK: productivity, earnings and skills".

*Corresponding author: Eleonora Patacchini, University of Rome "La Sapienza", Faculty of Statistics, P.le A. Moro, 5, 00185, Rome, Italy. Email: eleonora.patacchini@uniroma1.it

1. Introduction

Recent research in economic geography has drawn attention to the potential for positive externalities arising from agglomeration of economic activity (Fujita, Krugman and Venables, 1999, Ottaviano and Puga, 1997). The benefits to firms and workers being located close to each other in space may come from a variety of sources: knowledge spillovers, thick market effects in the labour market, proximity to consumers and to specialist input suppliers in markets with trade costs and increasing returns.[1] As a result of these developments, economists have started to pay closer attention to the spatial configuration of economic data for evidence of significant spatial clustering.

The visualisation and exploration of spatial data can provide valuable insights into the nature and extent of spatial clustering in economic variables (Dall'Erba, 2005; Lopez-Bazo et al. 1999). However, much of the empirical work undertaken to date has tended to focus on identifying the spatial properties of a single economic variable – usually GDP per capita or its growth rate (see for example, Rey and Montouri, 1999; Ertur and Le Gallo, 2003; Roberts, 2004). In this paper, we use the techniques of exploratory spatial data analysis to compare and contrast patterns of spatial association in related measures of economic performance. More specifically, we decompose sub-regional income per worker into a productivity component and an occupational composition component, and analyse the spatial structure of each of these variables. This approach offers valuable insights into the sources of spatial dependence and spatial heterogeneity in income per worker. This is very distinct from the information that may be gained using spatial regression methods which focus on identifying and estimating average effects across space.[2]

Deleted: M

Deleted: , however,

Deleted: ed

The focus of our analysis is the significant disparities in economic performance that persist across the sub-regions of the UK. These are well documented, most recently in the Treasury report "Productivity in the UK: the Local Dimension" (July 2003). However, views differ as to whether these disparities represent a significant divide between an impoverished 'north' and an affluent 'south'; or whether the picture is more diffuse with intra-regional differences in economic outcomes as significant as those between the major regions of the UK. (Adams and Robinson, 2002, HM

Treasury, 2003). Data for 2001 shows that income per capita in London is 154 percent of the national average, as compared with just 73 percent of the national average in the North East region and 86 percent in Yorkshire and Humberside. That said, the cities of Leeds and York are both within the upper quartile of the UK distribution of income per head, while areas of Outer London fall in the lower quartile.

We start by examining income per worker in the NUTS3 sub-regions of Great Britain and address the following questions[3]. What is the relationship between the economic performance of one area and that of its neighbours and over what range does this relationship persist? Is there evidence of spatial clustering with areas of high (low) income surrounded by 'neighbours' with similar levels of income? Or are high performing areas observed as atypical areas of high productivity surrounded by lower performing neighbours? These questions are addressed using exploratory spatial data analysis techniques to characterise the relationship between the value of an economic variable in one region and that of its neighbours, and thereby detect patterns of spatial association, spatial clusters, and atypical localisations.

Deleted: spatial outliers

Deleted: –

Deleted: or spatial outliers (Haining, 1990; Anselin, 1988).

The analysis finds strong evidence of a positive spatial association in income per worker at the sub-regional (NUTS3) level in Great Britain. In other words, areas of relatively high (low) income tend to be located 'close to' other areas of high (low) income. The results show that for these purposes 'close' is within an estimated travel time of some 90 minutes. At distances beyond this, the evidence of positive spatial association persists but is weakening. Within this global structure, one can identify significant local regimes – clusters of areas in which the value of income per worker differs significantly from the average for the UK as a whole. Thus, there is strong evidence of a 'winner's circle' in the south and east of England – a cluster of areas with income per worker significantly higher than the global average. There is evidence also – albeit less strong – of two further regimes, both of a low-income type. The larger of these is located in the north west centred around the metropolitan areas of Liverpool and Manchester; while the second smaller cluster is in the south-west of England. Within each regime, there are atypical areas with dissimilar

Deleted: spatial outliers -

values to their neighbours. For example, within the high-income regime in the south and east, metropolitan areas such as Brighton and Hove and Portsmouth are significantly underperforming relative to surrounding areas.

Having identified the spatial structure of income per worker, we examine whether this derives primarily from spatial dependence in the types of jobs available or in productivity in a given job. A location may derive high income per worker from having a high concentration of good quality (i.e. well paid) jobs. Or, it may that for some reason – possibly related to the agglomeration effects identified in the economic geography literature – output per worker within a given occupation is higher here than elsewhere. As far as the high-income regime in the south and east of Great Britain is concerned, the cluster benefits from both above average job quality and higher than average worker productivity in given jobs. The picture within the low-income regimes is more mixed. For the north west, the evidence suggests that occupational composition plays the bigger role in shaping the spatial structure of income, while in the south west, low worker productivity rather than poor quality jobs appears to be the issue.

The paper is structured as follows. In Section 2, we describe the data used in this study and examine the basic descriptive statistics relating to the levels and dispersion of the variables across the NUTS3 areas of Great Britain. Section 3 presents the evidence relating to the spatial distribution of income per worker across the sub-regions of Great Britain. Section 4 compares these results with those for the data relating to the occupational composition and productivity. Section 5 concludes.

2. Income, Earnings and Productivity: Data and Descriptive Statistics

Our analysis is based on data for the sub-regional NUTS3 spatial units of Great Britain. There are 126 NUTS3 administrative areas in Great Britain but, in order to compile a consistent dataset, a number of these are combined to give a sample of 119 sub-regional units (that we will term 'areas'). The data series relate to the period 1998 to 2001 and the four years of data are averaged in order to remove some of the short-run volatility. Full details of the sample and the data used are provided in the Data Appendix to this paper (Appendix 1).

Deleted: 2

Several different types of income data are available.[4] Estimates of workplace-based gross value-added at the NUTS3 level are calculated according to the income approach by the Office of National Statistics (ONS, 2003). We construct a measure of GVA per hour worked by employees, taking as the denominator an estimate of the total hours worked by employees in the area. A limitation of GVA as a measure of income is that it is sensitive to the assumptions made in allocating profits and other non-wage income across the NUTS 3 areas (see ONS 2003 for further discussion). An alternative measure that avoids this problem focuses on income from employment only and for this we use data for average hourly earnings from the New Earnings Surveys for the relevant years. In so far as the measurement errors in the income variables are temporary, they are mitigated by averaging the data over the four year period.

Deleted: methods used to allocate

(insert Table 1 here)

Table 1 gives summary statistics for each of these measures and the relationship between them. The numbers in brackets are the same statistics with Inner London (East and West) excluded from the sample. Correlation coefficients between each of the variables are reported in the lower part of the table. GVA per hour worked (g_i) varies across the 119 NUTS3 regions from 14.79 in Stoke on Trent to 25.20 in Inner London (West) with a mean of 18.67 and a coefficient of variation of 0.10. Average hourly earnings (e_i) displays an even greater degree of spatial variation although this falls significantly with the exclusion of Inner London from the sample. As one would expect, the two income series are highly correlated with a correlation coefficient of 0.76. However, there are some major outliers, most notably the two Inner London areas where average earnings are high relative to GVA per hour worked. In general, the areas with a high ratio of average earnings to GVA per hour worked tend to be the metropolitan areas including East Merseyside, Solihull, Brighton and Hove and Liverpool, as well as Inner London. By contrast, average hourly earnings tend to be low relative to GVA per hour worked in more rural areas – Torbay, West Lothian, East Ayrshire, South West Wales.

Spatial variation in average earnings derives from two sources – differences in the wage rates paid to workers in a given occupation, and differences in the occupational composition of employment. These two contributions to the spatial structure of average earnings can be separated out as follows. Let w_i^k and l_i^k denote the wage and level of employment in occupation k and area i. Total employment in area i is $L_i = \Sigma_k l_i^k$, and the share of occupation k in employment in this area is $\lambda_i^k = l_i^k / L_i$. The average wage of occupation k in the economy as a whole (i.e. aggregating across all i) is given by $\overline{w}^k = \Sigma_i l_i^k w_i^k / \Sigma_i l_i^k$, while $\overline{\lambda}^k = \Sigma_i l_i^k / \Sigma_i L_i$ is the share of occupation k in total employment for the economy as a whole. It follows that average earnings in area i, e_i, may be decomposed as follows:

$$e_i \equiv \Sigma_k w_i^k \lambda_i^k = \Sigma_k w_i^k \overline{\lambda}^k + \Sigma_k \overline{w}^k \lambda_i^k + \Sigma_k (w_i^k - \overline{w}^k)(\lambda_i^k - \overline{\lambda}^k) - \Sigma_k \overline{w}^k \overline{\lambda}^k. \qquad (1)$$

The first term on the right-hand side of (1) is the average level of earnings at location i conditional on the occupational composition being the same as for the economy as a whole; it will be denoted $q_i = \Sigma_k w_i^k \overline{\lambda}^k$. q_i measures the spatial variation in wages while controlling for occupational structure, and as such reflects spatial differences in productivity.[5] We will refer to it as the productivity index. The second term on the right-hand side measures the average level of earnings of area i given its specific occupational composition but assuming that the wage rate for each occupation is equal to the UK average in that occupation. It will be denoted $c_i = \Sigma_k \overline{w}^k \lambda_i^k$ and referred to as the occupational composition index. The remaining terms in (1) measure the covariance in earnings and composition across occupations in area i and will be denoted by r_i. Before proceeding it is important to note that (1) is an arithmetic decomposition of the data and does not depend on any particular model of the determinants of productivity or of occupational composition, or of the relationship between them. The value of the decomposition lies in allowing us to identify ex post the contribution of the spatial variation in productivity and in occupational composition to the overall spatial structure of income per worker. In practise the quality of the decomposition depends on the level of occupational disaggregation that is feasible given available data. Ideally, the level of occupational disaggregation should be such that the occupational categories are relatively homogenous, but in practise sample sizes restrict the level of disaggregation that is practical

Deleted: allo

Sub-regional data on earnings by occupation from the New Earnings Survey and on employment shares by occupation taken from the Labour Force Survey are used to compute the productivity index and the occupational composition index for each of the NUTS3 areas of Great Britain. The productivity index, q_i, is constructed from data on earnings by occupation for each of 38 minor occupational groups, using as weights the share of each occupation in the total employment of Great Britain as a whole. The composition index, c_i, requires data on employment shares by occupation at the level of the NUTS3 area, which is available from the Labour Force Survey but in this case, reliable estimates are available only for the 9 major occupational groups.

Deleted: Ideally, the level of occupational disaggregation should be such that the occupational categories are relatively homogenous, but in practise sample sizes restrict the level of disaggregation that is practical

Summary statistics for these indices are reported in Table 1, columns 3 to 5. First, note that the sample properties of the productivity index do not vary significantly with the level of occupational disaggregation. As we would expect, the more disaggregated index (i.e. the one computed for 38 distinct occupational categories) displays a little less spatial variation. However, the two indices are very highly correlated (0.987) and their relationship with the other variables appears very similar. As one might expect, the occupational composition index and the productivity index are positively correlated so that areas with high productivity tend to have a good occupational composition also, although the correlation at approx. 0.66 is far from perfect. Variance in the productivity index accounts for some 60% of the overall variance in average hourly earnings.[6] The remaining 40 percent is attributable to variance in the composition index and the covariance term.

3. Spatial Structure of Income

In this section of the paper, we examine the spatial structure of income per worker across the UK. Is it appropriate to characterise the outcome as a 'north-south' divide between the affluence of the south of England and the impoverishment of the regions of the north (IPPR 2003)? At first sight, the maps of the NUTS3 regions of Great Britain designated according to the quintiles of the income distribution in Figure 1 would appear to support this view. In terms of GVA per hour worked, the south and east of England has a preponderance of NUTS3 regions in the top 40 percent of the distribution, while the regions in the lowest quintile tend to be located in the north of the country. The picture for average hourly earnings is, however, less clear cut, with areas of relatively high (low) average earnings appearing more spatially dispersed. Do the groupings of high and low values apparent in Figure 1 represent a statistically significant departure from spatial randomness? To answer that question, we use the methods of exploratory spatial data analysis to describe and formally test the global and local spatial properties of the two income measures – GVA per hour worked and average hourly earnings. (For details on these methods see Haining, 1990, Anselin, 1995a and 1995b, Getis and Ord, 1992, Ord and Getis, 1995.)

Deleted:

Deleted:

Deleted: a

Deleted: on

Deleted: The principles and the methods of exploratory spatial data analysis used in this paper are reviewed in detail in Appendix 1 to this paper.

A basic characteristic that distinguishes spatial data from other types of cross-section data is the spatial arrangement of the n observations. For purposes of exploratory data analysis, the spatial linkages or proximity of the units of observations are summarised by defining a $n \times n$ spatial weight matrix, $W=\{W_{ij}\}$ where $W_{ij} = 1$ if sites i and j are designated as neighbours, and $W_{ij} = 0$ otherwise. A number of alternative criteria can be used for the specification of the neighbourhood set. A standard approach is to define proximity in terms of contiguity i.e. areas are designated as neighbours if they share a common boundary. However, where the basic units are defined by administrative boundaries, as in this case, this approach can give rise to neighbourhoods that vary greatly in terms of both the number of linkages and the geographical area covered. A more economically meaningful measure of proximity may be obtained by considering travel times between the units so that areas are neighbours if they are within a specified travel time d of each other. In the analysis that follows, spatial proximity is measured in terms of the average road journey time between the population centres of NUTS3 areas.[7] The estimated road journey time between a pair of NUTS3 areas in the sample varies between 21 minutes and 748 minutes, with a mean journey time of 237.5 minutes and a median journey time of approximately 220 minutes. The potential interactions between locations are summarised by the matrix $W_d = \{W_{ij,d}\}$ where $W_{ij,d} = 1$ if the spatial units i and j are within d minutes of each other and $W_{ij,d} = 0$ otherwise, where initially values of d = {30, 60, 90, 120, 150, 180} are considered.

3.1 Global Spatial Properties

Under the assumption of spatial randomness, any grouping of high or low values of the variable in space is totally spurious. The existence of a spatial structure is detected by the presence of spatial correlation that can be defined as the "coincidence of value similarity with locational similarity" (Anselin, 2001). There is positive spatial autocorrelation when high or low values of a random variable tend to cluster in space and there is negative spatial autocorrelation when geographical areas tend to be surrounded by neighbours with very dissimilar values. To investigate the global properties of the data, we consider two measure of global spatial autocorrelation, the Moran's I and the Geary's c statistics. (Cliff and Ord, 1981).

Deleted: (spatial clustering)

Deleted: (spatial outliers)

(insert Table 2 here)

Table 2 reports the Moran's I statistic and the Geary's c statistic for the two alternative income measures - GVA per hour worked and average hourly earnings – based on (row standardised) spatial proximity matrices corresponding to estimated road journey times of up to 30, 60, 90, 120, 150 and 180 minutes. Along with the test statistics, we report the standardized z-value for each statistic, the associated significance level, $p1$, assuming the (asymptotic) distributions of I and c are approximately normal, and an alternative indicator of statistical significance, $p2$, based on the conditional randomisation approach with 10000 permutations.

These results provide strong evidence of positive spatial autocorrelation in income across the NUTS3 regions of Great Britain. NUTS3 regions with relatively high (low) income are located close to other sub-regions with relatively high (low) income more often than would be observed if their locations were purely random. Both the Moran's I and Geary's c statistics are highly significant irrespective of the chosen inference strategy (i.e. both $p1$ and $p2$ are always close to 0) at distances of up to 120 minutes travel time. The values of the standardised test statistics, $z(I)$and $z(c)$, suggest that the spatial linkages are strongest at distances of up to 90 minutes travel time.[8] Moreover, these findings hold irrespective of whether we measure income in the region by GVA per hour worked or by average hourly earnings.

The Moran scatterplot shows the relationship between the value of the variable of interest for a given area i and the average value for the areas in its neighbourhood set. The Moran's I measure of global spatial autocorrelation is formally equivalent to the OLS estimate of the slope coefficient of the line fitted to the Moran scatterplot and hence standard regression diagnostics may be used to detect outliers and to identify individual areas that exert strong influence on the global Moran's I statistic (Anselin, 1996). Figure 2 depicts the Moran scatterplots for the two income variables based on the spatial weight matrix for d=90 minutes. Simple visual inspection of Figure 2 identifies no potential outliers as far as the GVA data is concerned, but Inner London (West) does appear to have a very large residual in respect of average hourly earnings. Formal statistical tests confirm that Inner London (West) is a significant outlier in this case.[9] However, dropping Inner London (West) observation from the average hourly earnings series has no significant impact on the measures of global spatial autocorrelation reported in Table 2, and there is nothing to suggest that the finding of strong positive spatial correlation in average earnings is being driven by this outlier (see Table A1 in Appendix 2).

Deleted: f

The four quadrants of the so-called Moran scatterplot correspond to the four types of local spatial association between a location and its neighbours: HH (upper right), contains areas with a high value surrounded by areas with high values, HL (lower right) consists of high value areas with relatively low value neighbours; LL (lower left) consists of low value areas surrounded by other areas with low values; LH (upper left) contains low value areas with high value neighbours. In the case of GVA per hour worked, 65% of the NUTS3 regions of Great Britain display association of similar values; 28% of the HH type and 37% of the LL type. For average hourly earnings, the proportion is even higher with 80% of NUTS3 regions characterised by positive spatial association, of which 36% are of the HH type and 44% of the LL type. The Moran scatterplot may be used also to identify atypical areas, i.e. areas deviating from the global pattern of positive spatial autocorrelation. These correspond to sample points in the HL (lower right) quadrant or the LH (upper left) quadrant of the scatterplot. For example, the NUTS3 regions of Cheshire, Derby and Edinburgh appear to be areas of relatively high income with low income neighbours on the basis of both GVA per hour worked and average hourly earnings. By contrast, Leicester, Dudley and Sandwell, Southend-on Sea appear to be areas of relatively low income surrounded by high income neighbours. The significance of these apparent local patterns of spatial association are explored in greater detail in the next section of the paper

3.2 Local Spatial Regimes

Different statistics of local spatial correlation have been developed to assess spatial dependence in a particular sub-region of the sample. These statistics describe the association between the value of the variable at a given site and that of its neighbours, and between the value within the neighbourhood set and that for the sample as a whole. The most widely used are the Getis-Ord's G statistic and the local Moran's *I*. The Getis-Ord's G statistic (Getis and Ord, 1992), is based on a comparison of the average value within a given neighbourhood set and the global average, and as such may be used to identify local regimes of relatively high or relatively low values of a variable. The local Moran's *I* statistic (Anselin, 1995a) measures the correlation between the value for a given area and that for its neighbours, and may be used to identify atypical localisations as well as clusters of high or low values.

Deleted: Getis-Ord

Deleted: Getis-Ord

A number of complications arise in assessing the significance of local indicators of spatial association. First, the distribution of both the local Moran's I and the Getis-Ord's G statistics are affected by the presence of global spatial association, and hence inference based on the normal approximation is likely to be misleading (Anselin, 1995a) Given the evidence of global spatial autocorrelation in Table 2, the nonparametric approach of conditional randomisation provides a more reliable basis for inference in this case. The conditional randomisation method yields the same empirical reference distribution for both the local Moran's *I* and the Getis-Ord's G statistics and so inference based on this nonparametric approach gives identical results for the two statistics (Anselin, 1995a).

A second complicating factor is that the local statistics for any pair of locations, *i* and *j*, are correlated whenever their neighbourhood sets contain common elements (Ord and Getis, 1995). Given this, Ord and Getis suggest using a Bonferroni bounds procedure to assess significance such that for an overall significance level of α, the individual significance level for each observation is taken as α/n where n is the number of observations in the sample. In this particular study with a sample of 119 observations, an overall significance level of 0.05 implies an individual significance level for each observation of just 0.00043. However, in practise for any given location, the number of other locations in the sample with correlated local statistics is likely to be considerably small than *n*, and so this procedure is expected to be overly conservative.[10]

For present purposes, we compute both the Getis-Ord's G and the local Moran's I statistic for each NUTS3 area using the spatial weights matrix for d=90 minutes, the distance at which spatial linkages seem to be strongest (see Table 2).[11] The values of the local statistics together with the associated significance levels based on the normal approximation and that derived using the conditional randomisation method are reported in full in Table A2 of Appendix 2.[12] Inference based on the normal approximation and that based on the conditional permutation approach give very similar results in the case of the Getis-Ord's G statistic, but not for the local Moran's I statistic. This is not surprising, given that it is now common knowledge that the reference distribution of the local Moran's I statistic deviates substantially from the normal distribution (see, e.g., Tiefelsdorf, 2002). Figure 3 shows the NUTS3 regions for which the local statistics are significant at the 0.05 level based on the conditional permutation approach. By considering the two extreme significance levels

Deleted: (

Deleted:)

Deleted: local Moran

Deleted: Getis-Ord

Deleted: (

Deleted:)

Deleted: (

Deleted: 4)

Deleted:

Deleted: Getis-Ord

Deleted: A3

Deleted: the

Deleted: to the paper

Deleted: Getis-Ord

Deleted: local Moran

Deleted: This is consistent with the finding of earlier studies that suggest that the Getis-Ord statistic tends to be more robust to departures from the assumption of normality and no global spatial autocorrelation.

of 0.005 and 0.00043 – the first applying if the individual local statistics are independent, while the second is appropriate if all the local statistics are correlated – we are able to place bounds on the extent of local spatial regimes in the data.

In the case of the GVA per hour worked data, two spatial regimes are clearly identified. The largest is a 'high-income' regime – a cluster of NUTS3 areas in which GVA per hour worked is significantly higher than the average for the sample as a whole – located in the south and east of England. This regime includes the majority of NUTS3 regions within the East, London and South East regions, extending to Warwickshire, Northamptonshire and Cambridgeshire to the north, Oxfordshire and Wiltshire to the west, Essex and Kent to the east. Adopting the more conservative criteria implied by using the Bonferroni procedure leaves this regime largely unchanged - only Warwickshire to the north and Wiltshire to the west are excluded. Within this high-income regime, the local Moran's *I* statistic identifies a number of atypical areas i.e. areas that exhibit negative spatial correlation with their neighbours. Brighton and Hove and Portsmouth in the south, Medway and Southend-on-Sea in the east, Peterborough and Suffolk to the north are identified as atypical low income areas on the periphery of the high income regime.

The second significant regime is a 'low income' type – a cluster of areas in which average GVA per hour worked is low relative to the global average. This regime covers the North West and North East regions of the UK extending south to South Yorkshire and Stoke-on-Trent and west into Wales. In this case, adopting the tighter criteria for statistical significance shrinks the regime considerably such that it is centred on the metropolitan areas of Manchester, Merseyside and west Yorkshire in the north-west of England. Here too the local Moran's *I* statistic identifies a number of atypical areas within the low income cluster. Cheshire, York and North Yorkshire are all areas with significantly higher GVA per hour worked than the average for their neighbours.

How does this picture change if we focus on employment income only? As far as the 'high income' regime in the south of the UK is concerned, the answer is very little. The main difference is that some areas that appeared to be underperforming relative to their neighbour in terms of GVA are

Deleted:

Deleted: local Moran

Deleted: local Moran

not similarly placed with respect to earnings, and vice versa. For example, the cities of Brighton and Hove and Portsmouth are significantly worse off than their neighbouring areas in terms of GVA per hour worked but not average hourly earnings. However, the evidence for a ‘low income’ regime in the north of Great Britain in terms of the earnings data is less strong. Applying the more conservative criteria, there are no neighbourhoods for which the average for hourly earnings is significantly below the global average. Relaxing the criteria and applying a significance level of 0.05 for each observation, a ‘low income’ regime emerges which includes Northumberland and Durham in the north east, Cumbria in the north west and areas around Manchester and the west of Yorkshire but this is far less extensive than is apparent in the GVA data. Furthermore, at this significance level, we identify a second low earnings cluster centred around Plymouth and Devon in the south west of England.

The question arises as to the sensitivity of these findings to the particular choice of spatial weight matrix. Is it the case that changing the cut-off point in terms of travel times results in significant changes in the statistical significance of the local indicators of spatial association? Following the approach used in Ertur and Le Gallo (2003), our results suggest not. We find that all areas with local statistics significant at the 0.05 percent level with a cut-off point of 90 minutes in terms of travel time are still associated with a significant value using a cut-off of 120 minutes, and the same is true when we compare 120 minutes with 150 minutes. As the maximum travel time is extended and more neighbours are included in the structure of proximity, so the proportion of the areas with a significant value for the local statistic tends to increase. Moreover, considering the transitions of the areas between spatial regimes, we find that the tables are largely diagonal implying that an area tends to remain in the same spatial regime irrespective of the choice of the spatial weight matrix.[13] On the basis of these findings we conclude that the identification of spatial regimes, local clusters and atypical areas documented in Table A2 is robust to the spatial scale of the weight matrix.

Deleted:

Deleted: O

Deleted: (see Table A4 of Appendix 3)

Deleted: spatial outliers

Deleted: escribed in Figures 3 and 4

To sum up, across Great Britain as a whole, regions of high/low income tend to be located close in terms of travel time to other regions of high/low income. This result holds whether we measure income broadly in terms of GVA per hour worked or, more narrowly, in terms of average hourly earnings. Does the evidence endorse the stereotypical view of a north-south divide in prosperity? To some extent, although is more appropriate to characterise the divide as one between a 'winner's circle' in the south east corner of England and the remainder of the country. The NUTS3 regions in the south and east of Great Britain are strongly identified as a high income regime – a cluster of areas for which average income within the cluster is significantly above the average for Great Britain as a whole. The evidence that areas in the north of England constitute a distinct 'low income' regime is more tentative. The data for GVA per hour worked does point to a cluster of areas centred around the metropolitan areas of Manchester, Merseyside and west Yorkshire with average income significantly lower than the global average for Great Britain as a whole. However, if we consider income from employment only, the evidence for a 'low income' regime is much weaker.

4. Spatial Structure of Earnings: Productivity v. Occupational Composition

Having identified the spatial structure of average hourly earnings in Great Britain, we turn now to the question of whether this derives primarily from spatial variation in types of jobs available or from in worker productivity in given jobs. More specifically, is the high income regime in the south and east of Great Britain due to the fact that productivity, and hence wages, in a given occupation is higher in these areas, or is it because these areas have better jobs than elsewhere? For this purpose, we analyse the global and local spatial properties of the productivity and occupational composition indices described in Section 2.

(insert Table 3 here)

The global test statistics reported in Table 3 point to positive spatial autocorrelation in both the productivity index and the occupational composition. The results for the productivity index are very close to those obtained for average hourly earnings, with both the Moran's *I* and the Geary's *c* measures statistically significant at distances up to 120 minutes travel time. Areas with relatively high (low) productivity tend to be located close to other areas with relatively high (low) productivity. By comparison, the degree of positive spatial association displayed by the occupational composition index is less strong, and the Geary's *c* statistics fails to reject the null hypothesis of no spatial association at travel times within 30 minutes and beyond 120 minutes. For both occupational composition and the productivity index, the spatial linkages appear strongest over distances of up to 90 minutes travel time as before.[14] Thus at the global level, positive spatial association in worker productivity appears to play a major role in driving positive spatial association in income per worker.

Table A3 (in Appendix 2) reports the results of the local indicators of spatial association – Getis Ord and local Moran's *I* statistics – in respect of productivity and occupational composition. Figure 4 shows the NUTS3 regions for which the local indicators of spatial association – Getis Ord and local Moran's *I* statistics – in respect of productivity and occupational composition are statistically significant at the 5 percent level.[15] What do these tell us about the factors giving rise to the local spatial regimes in income per worker described in Section 3? The first point to note is that the high-income regime identified in the earnings data is reproduced here in both the productivity index and the occupational composition index. The cluster of high income areas in the south and east of Great Britain derives its relative prosperity from both higher than average productivity (and hence a higher than average level of wages in any given occupation), and better than average occupational composition (i.e. a higher than average proportion of better paid jobs).

Deleted: 5

Deleted: (see Table A6 in Appendix 3)

The local Moran's *I* provides further insights into the presence of atypical areas within this cluster – those NUTS3 regions which are underperforming relative to their neighbours in terms of average hourly earnings. Of the four significant atypical areas in the high earnings cluster, three regions – Southend-on-Sea, East Sussex and Wiltshire – are found to be atypical areas in respect of the productivity index but not the occupational composition. In other words, in these areas, the low level of earnings relative to their neighbours is primarily the result of low wages for a given set of

Deleted: spatial outliers

Deleted: spatial outliers

Deleted: spatial outliers

occupations, rather than due to a poor occupational composition. Only for the Medway region would it appear that occupational composition is mainly responsible for the poor performance in terms of average hourly earnings.

What of the low earnings regime identified in the north-west of Great Britain? This regime is largely reproduced in the data for occupational composition. The NUTS3 areas which make-up the low earnings regime are in the majority of cases, areas with a poor occupational structure relative to the global average. Some of these areas – Northumberland and Tyneside, Bradford and Calderdale - are underperforming relative to the global average with respect to the productivity index also. Likewise, the atypical areas within this low income regime are identified as atypical areas with respect to occupational composition. These results suggest that the spatial structure of occupational composition is playing the more influential role in shaping the low earnings regime in the north of England. The situation appears rather different in the south west of England where there is weak evidence of a low earnings regime also. Here the cluster is associated with low productivity and hence low wages relative to the global average, rather than with poor occupational composition.

Deleted: spatial outliers

Deleted: outliers

5. Concluding Remarks

This paper uses the techniques of exploratory spatial data analysis to investigate the contribution of differences in types of jobs and differences in productivity for a given job to the spatial variation in income per worker across Great Britain. This approach not only identifies global patterns of spatial association but also highlights the roles played by occupational composition and productivity in observed spatial heterogeneity. In this respect, the methods are complementary to the techniques of spatial regression analysis which focus on the identification and estimation of average effects across a given space.

The formal statistical analysis confirms that the spatial distribution of income per worker across the sub-regions of Great Britain is not random. Whether we consider average hourly earnings or the more broadly defined GVA per hour worked, there is strong evidence of positive spatial association in income per worker – areas of relatively high (low) income tend to be located 'close to' other areas of high (low) income, where for these purposes 'close' is within an estimated road journey time of some 90 minutes. Moreover, while both occupational composition and worker productivity display positive spatial autocorrelation, the degree of spatial association is far stronger for productivity. Controlling for occupational composition, areas with high (low) wages tend to be located close to other areas of high (low) wages; a picture consistent with the hypothesis of significant returns to agglomeration.

Within this overall global pattern, we are able to identify significant local regimes – spatial clusters of areas in which the measures of economic performance diverge significantly from the average for the economy as a whole. The most clear cut of these is the high income regime in the south and east of England – a large cluster of areas for which both GVA per hour worked and average hourly earnings are significantly higher than the global average. Elsewhere in Great Britain, we find weak evidence of low income regimes – one in the north of England, and a second smaller regime in the south west of the country. One interesting observation to emerge from this investigation is that the evidence of underperformance in areas in the north of England is much stronger in the data for GVA per hour worked than in that for average hourly earnings. This evidence suggests that the these areas are at a particular disadvantage with respect to the distribution of non employment income - profits, trading surpluses, rents etc. However, questions remain over the reliability of the methods used to allocation these types of income across NUTS 3 areas when compiling the GVA estimates, and further investigation is needed.

While this analysis is largely descriptive, it does provide some insights into the economic factors underlying these local spatial patterns, and hence inform policy. The high earnings regime in the south and east of Great Britain is reproduced in the data for both productivity and occupational composition. The locations in this 'winners circle' benefit from both above average level of productivity and a better than average occupational composition. As far as the low earnings areas of the north west of Great Britain is concerned, occupational composition rather than productivity emerges as the more significant factor. – poorer quality jobs than the average for the country as a whole. By contrast, in the south west of England, below average earnings is the result of poor productivity rather than worse than average occupational composition. These findings highlight the spatial heterogeneity of economic performance across the UK, and the need for detailed analysis of local outcomes in order to identify the appropriate policy tools.

References

ADAMS, J. and P. ROBINSON (2002) A New Regional Policy for the United Kingdom, London, Institute for Public Policy Research.

ANSELIN, L. (1988) Spatial Econometrics; Methods and Models, Dordrecht, Kluwer.

ANSELIN, L. (1995a) Local indicators of spatial association-LISA, *Geographical Analysis* 27, 93-115.

ANSELIN, L. (1995b) *Spacestat Version 1.80 User's Guide*, Regional Research Institute, West Virginia University, Morgantown WV.

ANSELIN, L. (1996) The Moran scatterplot as an ESDA tool to assess local instability in spatial Association, in Fisher M, Scholten HJ, Unwin D (eds), *Spatial Analytical Perspectives on GIS*, London: Taylor and Francis.

ANSELIN, L. (2001) Spatial econometrics, in Baltagi B (eds) *Companion to Econometrics*, Oxford: Basil Blackwell.

CLIFF, A. and J.K. ORD (1981) Spatial Processes, Models and Applications, London, Pion.

DALL'ERBA, S. (2005) Distribution of regional income and regional funds in Europe 1989-1999: An exploratory data analysis, *The Annals of Regional Science* 39, 121-148.

ERTUR, C. and J. LE GALLO, (2003) Exploratory spatial data analysis of the distribution of regional per capital GDP in Europe 1980-1995, *Papers in Regional Science* 82, 175-201.

FINGLETON, B. (2001) Equilibrium and economic growth: spatial econometric models and simulations, *Journal of Regional Science* 41, 117-147.

Deleted: pp.

Deleted: pp.

Deleted: pp.

FINGLETON, B. (2003) Increasing returns: evidence from local wage rates in Great Britain, *Oxford Economic Papers* 55, 716-739.

Deleted: pp

FUJITA, M., P. KRUGMAN and A.J.VENABLES (1999) The Spatial Economy: Cities, Regions and International Trade, Cambridge (MA), MIT Press.

FUJITA, M. and J.THISSE (2002) The Economics of Agglomeration, Cambridge, Cambridge University Press.

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

ORD, J.K. and A. GETIS (1995) Local spatial autocorrelation statistics; distributional issues and an application, *Geographical Analysis* 27, 286-305.

Deleted: , A. and J.K. Ord

HM Treasury (2003) *Productivity in the UK – The Local Dimension*, London: HMT(July).

HAINING, R. (1990) Spatial Data Analysis in the Social and Environmental Sciences, Cambridge: Cambridge University Press.

LOPEZ-BAZO E, VAYÀ E, MORA A.J and J SURIÑACH (1999) Regional economic dynamics and convergence in the European Union, *Annals of Regional Science* 33, 343-370.

Deleted: ¶

Office of National Statistics (2003) *NUTS 3 Gross Value Added: Methods and Background*, London: ONS (December).

OTTAVIANO, G. and D. PUGA (1997) Agglomeration in the Global Economy: A Survey of the 'New Economic Geography, Centre for Economic Policy Discussion Paper 1699, London:CEPR.

REY, S.J. and B.D. MONTOURI (1999) US regional income convergence: a spatial econometric perspective, *Regional Studies* 33, 143-156.

Deleted: pp

RICE, P. and A.J. VENABLES (2004) Spatial determinants of productivity: analysis for the regions of Great Britain, Centre for Economic Performance Paper, CEPDP0642, London, Centre for Economic Performance, LSE.

ROBERTS, M. (2004) The growth performance of GB counties: Some new empirical evidence for 1977-1993, *Regional Studies* 38, 149-165.

Deleted: pp

TIEFELSDORF, M. (2002) The saddlepoint approximation of Moran's I and Local Moran's Ii's reference distributions and their numerical evaluation, *Geographical Analysis* 34, 187-206.

Table 1: Income, Earnings and Productivity - Summary Statistics

(Bracketed term: excluding Inner London – East and West)

	GVA per (employee) hour worked g_i	Average hourly earnings e_i	Composition index, c_i (9 major groups)	Productivity index, q_i (9 major groups	Productivity index, q_i (38 groups)
Mean (£)	18.66 (18.58)	9.82 (9.71)	10.17 (10.15)	9.93 (9.86)	9.57 (9.51)
Variance	3.71 (3.28)	1.66 (0.97)	0.22 (0.19)	0.76 (0.47)	0.62 (0.40)
Coefficient of variation	0.1032 (0.0974)	0.1314 (0.1016)	0.0460 (0.0420)	0.0878 (0.0697)	0.0823 (0.0667)
Minimum	14.79	7.79	9.12	8.47	8.31
Maximum	25.20 (24.18)	17.54 (13.16)	12.03 (11.35)	14.53 (11.90)	13.52 (11.45)
Correlation coefficients					
GVA, g_i	1.00	0.7610 (0.7414)	0.6695 (0.6148)	0.7207 (0.6812)	0.7217 (0.6798)
Earnings, e_i		1.00	0.8202 (0.8077)	0.9638 (0.9450)	0.9569 (0.9387)
Composition index, c_i			1.00	0.6573 (0.5801)	0.6767 (0.6077)
Productivity index, q_i (9)				1.00	0.9875 (0.9807)

Table 2: Measures of Global Spatial Autocorrelation

	GVA per (employee) hour worked				Average hourly earnings			
Moran's I test - Spatial weight matrix: estimated travel time ≤ d minutes								
d	I	z(I)	p1	p2	I	z(I)	p1	p2
30	0.6034	7.8324	0.0000	0.0001	0.5949	7.7233	0.0000	0.0001
60	0.3496	7.3830	0.0000	0.0001	0.4208	8.8515	0.0000	0.0001
90	0.2609	8.0003	0.0000	0.0001	0.3224	9.8283	0.0000	0.0001
120	0.1718	7.3745	0.0000	0.0001	0.2113	8.9931	0.0000	0.0001
150	0.0920	5.1837	0.0000	0.0003	0.1241	6.8397	0.0000	0.0004
180	0.0556	3.8419	0.0001	0.0036	0.0693	4.6616	0.0000	0.0017
Geary's c test - Spatial weight matrix: estimated travel time ≤ d minutes								
d	c	z(c)	p1	p2	c	z(c)	p1	p2
30	0.5208	-4.3783	0.0000	0.0001	0.6048	-3.6106	0.0003	0.0374
60	0.6739	-5.6034	0.0000	0.0001	0.6636	-5.7802	0.0000	0.0057
90	0.7614	-5.7480	0.0000	0.0001	0.7361	-6.3579	0.0000	0.0082
120	0.8507	-4.3763	0.0000	0.0007	0.8350	-4.8359	0.0000	0.0145
150	0.9349	-2.1293	0.0332	0.0413	0.9247	-2.4658	0.0137	0.1064
180	0.9786	-0.8444	0.3984	0.2209	0.9591	-1.6158	0.1061	0.2177

Table 3: Measures of Global Spatial Autocorrelation

	Productivity index				Occupational composition index			
Moran's I test - Spatial weight matrix: estimated travel time ≤ d minutes								
d	I	z(I)	p1	p2	I	z(I)	p1	p2
30	0.6171	8.0081	0.0000	0.0001	0.3451	4.5258	0.0000	0.0001
60	0.4386	9.2183	0.0000	0.0001	0.2655	5.6493	0.0000	0.0001
90	0.3410	10.3798	0.0000	0.0001	0.2276	7.0128	0.0000	0.0001
120	0.2323	9.8524	0.0000	0.0001	0.1454	6.2976	0.0000	0.0001
150	0.1331	7.3044	0.0000	0.0002	0.0963	5.4050	0.0000	0.0004
180	0.0692	4.6564	0.0000	0.0013	0.0513	3.5827	0.0003	0.0043
Geary's c test - Spatial weight matrix: estimated travel time ≤ d minutes								
d	c	z(c)	p1	p2	C	z(c)	p1	p2
30	0.5293	-4.3008	0.0000	0.0020	0.7992	-1.8350	0.0665	0.0623
60	0.6226	-6.4845	0.0000	0.0004	0.7932	-3.5527	0.0004	0.0024
90	0.6985	-7.2623	0.0000	0.0004	0.8080	-4.6254	0.0000	0.0007
120	0.7945	-6.0223	0.0000	0.0048	0.8822	-3.4534	0.0005	0.0067
150	0.8911	-3.5661	0.0004	0.0282	0.9401	-1.9611	0.0499	0.0619
180	0.9410	-2.3329	0.0197	0.0883	0.9692	-1.2167	0.2237	0.1568

[1] For further discussion see Fujita and Thisse (2002).

[2] Examples of this approach applied to UK data can be found in Fingleton (2001) and (2003).

[3] The Nomenclature of Territorial Units for Statistics (NUTS) was established by the Statistical Office of the European Communities (Eurostat) to provide a single uniform breakdown of territorial units for the production of EU regional statistics. Great Britain is divided into 10 regions at the NUTS1 level, and 126 areas at the NUTS3 level. For example the area of Greater London is made up of 5 NUTS3 areas.

[4] In the absence of sub-regional price deflators, we are only able to look at *nominal* income.

Deleted:

[5] For further discussion of the theoretical basis for this assertion see Rice and Venables. (2004), pp. 8-10.

[6] Given the decomposition $e_i = q_i + c_i + r_i$, the contribution of the productivity index (q_i) to the spatial variation in earnings (e_i) is measured by [var(q_i)+cov(q_i, $c_i + r_i$)]/var(e_i) (i.e. the share of the variance of q_i plus its covariance in the total variance of e_i) and is equal in value to the slope coefficient of the simple regression of the productivity index (q_i) on earnings (e_i).

[7] Travel times between the NUTS 3 areas are estimated using Microsoft Autoroute 2002. The Microsoft Autoroute software computes the driving time between two locations on the basis of the most efficient route given the road network in 2002, and allowing for different average speeds of travel depending on the type of road.

[8] The persistence of the correlation over several lags may be indicative of non-stationarity in the area data; e.g. the presence of a simple trend in space analogous to a time trend in time-series data

[9] The results of the outliers analysis are available upon request. For an application of these techniques in a similar context see Ertur and Le Gallo, 2003.

Deleted: For the statistical test

Deleted: see

Deleted: Table A1 in Appendix 3)

Deleted: (

Deleted:)

[10] An alternative approach of dividing the p-value by the average number of neighbours (i.e.,20.15) produces similar evidence as using the Bonferroni bound.

Deleted: leads in our case roughly to the same evidence

[11] The only three unconnected observations, namely Aberdeen, Cumbria and Highlands, have been linked to the relative closest area in terms of travel time (i.e. Durham, Angus - Dundee City , Perth – Kinross - Stirling respectively). The average number of links is 20.15, the percentage of non-zero weights is 17.08.

[12] As noted above, the pseudo-significance levels obtained by the conditional randomisation approach are identical for the Getis-Ord's G and local Moran's I statistics. Figures showing the NUTS3 areas for which the local statistics are significant using the far more conservative criteria of 0.00043 implied by the Bonferroni procedure are available upon request.

Deleted: Getis-Ord

[13] The results of the robustness analysis using transition matrices are available upon request.

[14] Once again Inner London (west) emerges as a possible outlier in respect of the productivity index but there is no evidence that this observation is unduly influencing the results (results available upon request).

Deleted: see Table A5 in Appendix 3

[15] These findings are robust with respect to the choice of the spatial weight matrix (results available upon request).

Deleted: see Table A7 in Appendix 3

Appendix 1: Data Appendix

All data is at the level of the 126 NUTS 3 areas of Great Britain. Because of missing or inadequate information on our target variables, the following NUTS 3 areas are aggregated: East Cumbria and West Cumbria; South and West Derbyshire and East Derbyshire; North Nottinghamshire and South Nottinghamshire; Isle of Anglesey and Gwynedd; Caithness, Sutherland and Ross & Cromarty, Inverness and Nairn & Moray, Badenoch & Strathspey, Lochaber, Sky, Lochalsh & Argyll and the Islands. The Western Isles, Orkney Islands and Shetland Islands are excluded from the sample. Unless otherwise stated, the data relate to the period 1998 to 2001, and the four years of data are averaged to remove short-run volatility.

GVA per (employee) hour worked (g_i): Estimates of workplace-based gross value added at basic prices are from the Office of National Statistics (2003). ONS estimates of GVA are computed using the income approach. Estimates of the main components of income – wages and salaries for employees, self-employment income and gross trading profits – based on the location of the workplace are derived from a range of sources including the Annual Business Inquiry, New Earnings Survey and the Inland Revenue Survey of Personal Income. The remaining components such as rental income are apportioned to a given area using a wages and salaries indicator. For further details see Office of National Statistics (2003).

Total hours worked by employees is computed from data on the numbers of full-time employees and of part-time employees and the average weekly hours worked by each group taken from the Annual Business Inquiry.

Average hourly earnings (e_i): Estimates of the average hourly earnings of all full-time employees whose pay was not affected by absence at the NUTS 3 level based on the location of workplace are taken from the New Earnings Surveys for the appropriate years.

Productivity index ($q_i = \Sigma_k w_i^k \overline{\lambda}^k$): Weighted sum of the average earnings of each occupational group in area *i*, with weights equal to the share of the occupational group in total GB employment. Data on average hourly earnings of full-time employees at the level of the occupational major group and at the two digit occupational level from the New Earnings Survey. The weights are computed from data on the share of 2-digit occupations in total GB employment from the Labour Force Survey 2001.

Composition index ($c_i = \Sigma_k \overline{w}^k \lambda_i^k$): Weighted sum of the shares of each occupational major group in employment in area *i*, with weights equal to the GB average earnings of the occupational major group. Estimates of occupational shares in employment are derived from the Labour Force Survey for the appropriate years. Labour Force Survey data is residence-based, rather than workplace

Deleted: Appendix 1: Methods of Exploratory Spatial Data Analysis¶

¶

Global spatial autocorrelation¶

When the variable under investigation is measured on a continuous scale, the measurement of global spatial autocorrelation is usually based on Moran's *I* and Geary's *c* statistics (Cliff and Ord, 1981). ¶

Moran's *I* is defined as ¶

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j -}{\sum_i (x_i - \bar{x})^2}$$

(1)¶

where *n* is the number of observations, x_i denotes the observation at site *i* for the variable of interest *X*, and w_{ij} denotes the elements of the spatial weights matrix. S_0 is a scaling factor equal to the sum of all the elements in the weight matrix. The spatial weight matrix may be row-standardised such that the elements $\tilde{w}_{ij}$ in each row sum to 1 in order to normalise the size of the neighbourhood set for each site. In this case S_0 = n and the expression (1) simplifies to a ratio of the spatial cross-product to the variance. Moran's *I* is a cross product coefficient scaled to be less than one in value, with an expected value $E(I) = -1/(n-1) \approx 0$ for *n* sufficiently large. Values for Moran's *I* larger (smaller) than the expected value indicate positive (negative) spatial correlation.¶

An alternative measure of global spatial autocorrelation is given by Geary's *c* coefficient which is based on squared deviations. Geary's *c* is defined as ¶

$$c = \frac{(n-1)}{2S_0} \frac{\sum_i \sum_j w_{ij}(x_i - x_j)}{\sum_i (x_i - \bar{x})^2}$$

(2)¶

The expected value for Geary's c is 1. Values of Geary's *c* less than one indicate positive spatial correlation, while values larger than one suggest negative spatial correlation. ¶

. Inference is typically based on a standardised z-value of the statistic computed by subtracting the expected value and dividing by the standard deviation in the usual way. Assuming that the variable of interest is normally distributed, the z-value follows a standard normal distribution, and the significance of the test statistic may be judged by comparing the computed z-value with its probability in the standard normal tables (for a theoretical discussion and detailed expressions for the moments of the (asymptotic) distributions of *I* and *c* under various assumptions, see Cli ... [1]

Deleted: 2

Deleted: To achieve a consistent data set based, and this, coupled with the fact that the data is available only for the major occupational groups, limits comparability with the productivity index. Weights are computed from data on the GB average hourly earnings by major occupational group from the New Earnings Survey.

Travel times: Driving times between the population centres of NUTS 3 areas, are estimated using Microsoft Autoroute 2002. The Microsoft Autoroute software computes the estimated driving time between two locations on the basis of the most efficient route given the road network in 2002, and allowing for different average speeds of travel depending on the type of road. In general the average speeds are set at the upper limit for road type, namely 70mph for motorways, 60 mph for other highways; 40 mph for major roads, 30mph for minor roads and 18mph for streets within urban areas.

Appendix 2: Additional Tables

Deleted: 3

Deleted: TABLE A1 : Moran Scatterplot – Tests for Outliers¶ ... [2]

Deleted: 2

TABLE A1: Measures of global spatial autocorrelation for average hourly earnings (Inner-London-West excluded from sample)

Moran's I test – spatial weight matrix: estimated travel time < *d* minutes				
d	*I*	*z(I)*	*p1*	*p2*
30	0.5501	7.0313	0.0000	0.0001
60	0.4321	8.9909	0.0000	0.0001
90	0.3494	10.5236	0.0000	0.0001
120	0.2293	9.6311	0.0000	0.0001
150	0.1467	7.9229	0.0000	0.0002
180	0.0767	5.0609	0.0000	0.0012
Geary's c test - spatial weight matrix: estimated travel time < *d* minutes				
d	*c*	*z(c)*	*p1*	*p2*
30	0.5379	-4.1575	0.0000	0.0008
60	0.6307	-6.2815	0.0000	0.0057
90	0.6899	-7.3973	0.0000	0.0082
120	0.8060	-5.6378	0.0000	0.0009
150	0.8820	-3.8230	0.0001	0.0147
180	0.9417	-2.2780	0.0227	0.0595

TABLE A2: Local Spatial Correlation Statistics*

Deleted: 3

	GVA per (employee) hour worked					*Average hourly earnings*				
	Getis Ord		*Local Moran's I*			*Getis Ord*		*Local Moran I*		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
NORTH EAST										
Hartlepool and Stockton-on-Tees	-1.9743	0.0483	0.4825	0.6295	**0.0220**	-2.1016	0.0356	0.7615	0.4464	0.0519
South Teesside	-2.0240	0.0430	1.4228	0.1548	**0.0390**	-2.2653	0.0235	0.4317	0.6660	0.0709
Darlington	-1.9720	0.0486	1.8954	0.0580	**0.0230**	-2.0517	0.0402	1.4743	0.1404	0.0549
Durham CC	-2.0201	0.0434	0.9113	0.3621	**0.0285**	-2.1054	0.0305	0.8437	0.3988	**0.0330**
Northumberland	-2.1464	0.0318	1.5079	0.1316	**0.0070**	-2.4646	0.0137	1.5508	0.1209	**0.0060**
Tyneside	-2.0810	0.0374	1.4865	0.1371	**0.0320**	-2.3194	0.0204	0.6688	0.5036	**0.0290**
Sunderland	-2.2398	0.0251	1.4474	0.1478	**0.0488**	-1.9743	0.0483	0.8985	0.3689	0.0859
NORTH WEST										
Cumbria	-2.5453	0.0109	1.6031	0.1089	**0.0030**	-2.2323	0.0256	0.8099	0.4180	**0.0470**
Halton and Warrington	-3.3348	0.0009	1.4824	0.1382	***0.0001***	-0.9093	0.3632	-0.1167	0.9071	0.1628
Cheshire CC	-3.4567	0.0005	-4.4379	0.0000	***0.0001***	-0.8130	0.4162	-0.8086	0.4187	**0.0338**
Greater Manchester South	-3.2489	0.0012	0.1868	0.8518	**0.0020**	-1.9894	0.0467	-0.7185	0.4724	**0.0470**
Greater Manchester North	-3.6121	0.0003	2.7621	0.0057	***0.0001***	-2.0660	0.0388	0.9307	0.3520	**0.0150**
Blackburn with Darwen	-3.6540	0.0003	4.5314	0.0000	***0.0001***	-2.2050	0.0275	1.1428	0.2531	**0.0410**
Blackpool	-2.8554	0.0043	4.6547	0.0000	**0.0030**	-0.8837	0.3768	0.9380	0.3483	0.1888
Lancashire CC	-2.9256	0.0034	0.9580	0.3381	***0.0001***	-1.8575	0.0632	0.0039	0.9969	0.0749
East Merseyside	-3.3615	0.0008	3.6716	0.0002	***0.0001***	-1.1063	0.2686	-0.2086	0.8348	0.1219
Liverpool	-3.4931	0.0005	2.7573	0.0058	***0.0001***	-1.1185	0.2634	-0.1134	0.9097	0.1259
Sefton	-2.8885	0.0039	3.6875	0.0002	***0.0001***	-1.0236	0.3060	0.4678	0.6400	0.1379
Wirral	-3.7059	0.0002	3.8252	0.0001	***0.0001***	-0.9571	0.3385	0.1340	0.8934	0.1658
YORKSHIRE AND HUMBERSIDE										
Kingston upon Hull, City of	-0.6084	0.5429	0.6029	0.5466	0.2687	-1.1948	0.2322	1.3093	0.1904	0.0909
East Riding of Yorkshire	-1.0709	0.2842	-1.1031	0.2700	0.1259	-1.3370	0.1812	0.6090	0.5425	0.0589
North and N E Lincolnshire	-1.2171	0.2236	-0.2430	0.8080	0.1009	-1.5537	0.1203	1.3015	0.1931	0.0760
York	-2.2397	0.0251	-0.0571	0.9544	**0.0120**	-2.0240	0.0430	-0.1461	0.8838	**0.0280**
North Yorkshire CC	-2.7163	0.0066	-0.0352	0.9720	**0.0020**	-2.3194	0.0204	-0.3151	0.7527	**0.0200**
Barnsley, Doncaster, Rotherham	-1.7866	0.0740	2.4241	0.0153	**0.0300**	-1.0226	0.3065	1.0724	0.2835	0.1558
Sheffield	-1.6065	0.1082	1.1692	0.2423	0.0519	-0.8094	0.4183	0.0139	0.9889	0.2078
Bradford	-2.7535	0.0059	2.2371	0.0253	***0.0001***	-1.9720	0.0486	1.0713	0.2840	**0.0270**
Leeds	-3.0842	0.0020	0.6545	0.5128	***0.0001***	-2.0133	0.0441	-0.1071	0.9147	**0.0120**
Calderdale, Kirklees.Wakefield	-3.0971	0.0020	1.3543	0.1756	***0.0001***	-2.1233	0.0337	0.2559	0.7981	**0.0101**

TABLE A2: Local Spatial Correlation Statistics

Deleted: 3

	GVA per (employee) hour worked					*Average hourly earnings*				
	Getis Ord		*Local Moran's I*			*Getis Ord*		*Local Moran I*		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
EAST MIDLANDS										
Derby	-1.4595	0.1444	-1.8269	0.0677	0.0599	-0.4502	0.6525	-0.1350	0.8926	0.3487
Derbyshire	-1.7355	0.0827	0.7561	0.4496	0.0530	-0.7223	0.4701	0.2744	0.7838	0.2368
Nottingham	-1.1985	0.2307	0.1858	0.8526	0.1069	-0.6652	0.5059	0.3749	0.7077	0.2527
Nottinghamshire	-1.4116	0.1581	-0.5943	0.5523	0.0789	-1.1820	0.2372	0.2625	0.7929	0.1169
Leicester	1.0880	0.2766	-0.8740	0.3821	0.1429	-1.9508	0.0511	-2.1468	0.0318	**0.0160**
Leicestershire CC and Rutland	-0.7468	0.4552	-0.3132	0.7541	0.2248	-0.2487	0.8036	0.0314	0.9750	0.4236
Northamptonshire	3.5503	0.0004	1.5492	0.1213	***0.0001***	4.6307	0.0000	0.1091	0.9132	***0.0001***
Lincolnshire	-0.4116	0.6806	0.1651	0.8689	0.3277	-1.3850	0.1661	1.2779	0.2013	0.0870
WEST MIDLANDS										
Herefordshire, County of	-0.6231	0.5332	0.3372	0.7360	0.2727	-0.3140	0.7535	0.3208	0.7484	0.4236
Worcestershire	1.2271	0.2198	-0.9134	0.3611	0.1319	1.4855	0.1374	-0.3422	0.7322	0.0859
Warwickshire	2.8078	0.0050	1.6238	0.1044	**0.0070**	3.5460	0.0004	2.1943	0.0282	***0.0001***
Telford and Wrekin	-1.6480	0.0994	2.2125	0.0269	0.0550	-0.1822	0.8554	0.1537	0.8778	0.4486
Shropshire CC	-1.1869	0.2353	0.9590	0.3376	0.0979	-0.6436	0.5198	0.2280	0.8197	0.2737
Stoke-on-Trent	-2.6670	0.0077	5.1602	0.0000	**0.0020**	-0.6644	0.5065	0.9062	0.3648	0.2488
Staffordshire CC	-1.7538	0.0795	1.4644	0.1431	0.0570	-0.0555	0.9558	0.0719	0.9427	0.4815
Birmingham	0.7039	0.4815	-0.0687	0.9453	0.2527	2.3396	0.0193	1.5699	0.1164	**0.0110**
Solihull	1.7905	0.0734	0.0587	0.9532	**0.0390**	3.1917	0.0014	3.7016	0.0002	***0.0001***
Coventry	1.7812	0.0749	1.3013	0.1931	**0.0310**	3.1283	0.0018	1.8684	0.0617	***0.0001***
Dudley and Sandwell	0.3673	0.7134	-0.4382	0.6613	0.4086	1.0447	0.2962	-0.6027	0.5467	0.1628
Walsall and Wolverhampton	-1.4061	0.1597	1.4126	0.1578	0.0629	0.1398	0.8888	-0.0551	0.9561	0.4406
EAST										
Peterborough	2.3719	0.0177	-1.9094	0.0562	**0.0110**	2.0517	0.0402	0.2061	0.8367	0.0759
Cambridgeshire CC	4.2114	0.0000	3.3262	0.0009	***0.0001***	4.5379	0.0000	4.7235	0.0000	***0.0001***
Norfolk	-0.8324	0.4052	0.2376	0.8122	0.2138	-0.3295	0.7418	0.1802	0.8570	0.4186
Suffolk	1.8479	0.0646	-0.6707	0.5024	**0.0410**	1.4929	0.1355	-0.8515	0.3945	0.0909
Luton	4.6471	0.0000	4.7175	0.0000	***0.0001***	5.9200	0.0000	2.7687	0.0056	***0.0001***
Bedfordshire CC	4.6879	0.0000	4.3188	0.0000	***0.0001***	5.6646	0.0000	4.0528	0.0001	***0.0001***
Hertfordshire	4.9516	0.0000	9.5087	0.0000	***0.0001***	6.2600	0.0000	10.4383	0.0000	***0.0001***
Southend-on-Sea	5.4217	0.0000	-0.8032	0.4219	***0.0001***	6.0924	0.0000	-0.6357	0.5250	***0.0001***
Thurrock	5.4694	0.0000	9.0834	0.0000	***0.0001***	6.2459	0.0000	2.8865	0.0039	***0.0001***
Essex CC	5.7496	0.0000	3.4163	0.0006	***0.0001***	6.4316	0.0000	3.9092	0.0001	***0.0001***

TABLE A2: Local Spatial Correlation Statistics

Deleted: 3

	GVA per (employee) hour worked					*Average hourly earnings*				
	Getis Ord		*Local Moran's I*			*Getis Ord*		*Local Moran I*		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
LONDON										
Inner London - West	5.5528	0.0000	17.5800	0.0000	***0.0001***	6.8816	0.0000	33.9408	0.0000	***0.0001***
Inner London - East	5.8535	0.0000	11.4038	0.0000	***0.0001***	6.3619	0.0000	22.3287	0.0000	***0.0001***
Outer London - East and NE	5.5813	0.0000	4.0448	0.0001	***0.0001***	6.2284	0.0000	3.6520	0.0003	***0.0001***
Outer London – South	5.8925	0.0000	4.1391	0.0000	***0.0001***	6.4050	0.0000	8.2618	0.0000	***0.0001***
Outer London - West and NW	5.1358	0.0000	12.7564	0.0000	***0.0001***	6.0581	0.0000	13.6382	0.0000	***0.0001***
SOUTH EAST										
Berkshire	5.5693	0.0000	15.1784	0.0000	***0.0001***	6.2977	0.0000	14.5421	0.0000	***0.0001***
Milton Keynes	4.0335	0.0001	3.3265	0.0009	***0.0001***	5.2221	0.0000	4.9983	0.0000	***0.0001***
Buckinghamshire CC	5.5158	0.0000	13.3499	0.0000	***0.0001***	6.5328	0.0000	13.3738	0.0000	***0.0001***
Oxfordshire	5.1405	0.0000	5.7871	0.0000	***0.0001***	5.7668	0.0000	6.8417	0.0000	***0.0001***
Brighton and Hove	5.2022	0.0000	-1.9719	0.0486	***0.0001***	5.8707	0.0000	3.6247	0.0003	***0.0001***
East Sussex CC	4.1851	0.0000	0.7904	0.4293	***0.0001***	5.6519	0.0000	-3.0916	0.0020	***0.0001***
Surrey	5.7886	0.0000	12.5129	0.0000	***0.0001***	6.2644	0.0000	16.2595	0.0000	***0.0001***
West Sussex	4.3061	0.0000	2.7113	0.0067	***0.0001***	5.4583	0.0000	3.5740	0.0004	***0.0001***
Portsmouth	4.7795	0.0000	-2.2289	0.0258	***0.0001***	5.8527	0.0000	0.7856	0.4321	***0.0001***
Southampton	5.6638	0.0000	0.1889	0.8502	***0.0001***	6.3029	0.0000	1.2213	0.2220	***0.0001***
Hampshire CC	5.5521	0.0000	3.8396	0.0001	***0.0001***	6.0120	0.0000	5.0708	0.0000	***0.0001***
Isle of Wight	0.1268	0.8991	-0.1691	0.8657	0.4246	0.5679	0.5701	-0.3751	0.7076	0.2148
Medway	6.1721	0.0000	-2.9074	0.0036	***0.0001***	6.6653	0.0000	-1.0532	0.2922	***0.0001***
Kent CC	5.2560	0.0000	1.4774	0.1396	***0.0001***	5.9797	0.0000	1.3738	0.1695	***0.0001***
SOUTH WEST										
Bristol, City of	1.7671	0.0772	0.8981	0.3691	**0.0400**	0.7132	0.4757	0.3466	0.7289	0.2208
North and North East Somerset, South Gloucestershire	0.9128	0.3613	1.1927	0.2330	0.2008	0.4723	0.6367	0.2567	0.7974	0.3077
Gloucestershire	-0.0406	0.9676	-0.0012	0.9991	0.4725	0.0222	0.9822	0.0549	0.9562	0.4895
Swindon	4.4627	0.0000	5.0513	0.0000	***0.0001***	5.2742	0.0000	4.3853	0.0000	***0.0001***
Wiltshire CC	*2.4036*	*0.0162*	0.9761	0.3290	**0.0130**	1.9537	0.0507	-0.3465	0.7290	**0.0290**
Bournemouth and Poole	0.8573	0.3913	0.0721	0.9425	0.1988	0.9697	0.3322	0.0671	0.9465	0.1618
Dorset CC	0.7111	0.4770	-0.1497	0.8810	0.2338	0.2406	0.8098	-0.1087	0.9135	0.3427
Somerset	1.1018	0.2706	-0.1380	0.8902	0.1608	-0.5052	0.6134	0.4238	0.6717	0.3027
Cornwall and Isles of Scilly	0.0029	0.9977	-0.0150	0.9881	0.4456	-1.0298	0.3031	1.3713	0.1703	0.0839
Plymouth	-0.4501	0.6526	-0.0186	0.9852	0.3247	*-1.9894*	*0.0467*	1.9823	0.0474	**0.0040**
Torbay	0.1894	0.8498	0.1414	0.8875	0.4166	-0.9291	0.3528	1.2461	0.2127	0.1429
Devon CC	-0.1301	0.8964	0.0262	0.9791	0.4466	-1.7700	0.0767	0.8670	0.3860	0.0601

TABLE A2: Local Spatial Correlation Statistics

Deleted: 3

	GVA per (employee) hour worked					Average hourly earnings				
	Getis Ord		*Local Moran's I*			*Getis Ord*		*Local Moran I*		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
WALES										
Gwynedd+Anglesey	0.1194	0.9050	-0.0396	0.9684	0.4406	-0.6745	0.5000	0.4225	0.6727	0.2168
Conwy and Denbeighshire	-2.1689	0.0301	1.0819	0.2793	**0.0110**	-0.6093	0.5423	0.4113	0.6808	0.2897
South West Wales	-0.2931	0.7695	-0.1769	0.8596	0.3896	-0.9763	0.3289	0.7547	0.4504	0.1538
Central Valleys	0.4876	0.6259	-0.2187	0.8269	0.3247	-0.5898	0.5553	0.6091	0.5425	0.2637
Gwent Valleys	0.5120	0.6086	-0.2825	0.7776	0.3127	-0.6326	0.5270	0.5525	0.5806	0.2737
Bridgend and Neath Port Talbot	0.2895	0.7722	0.1094	0.9129	0.3766	-0.7217	0.4705	0.3800	0.7039	0.2527
Swansea	-0.0431	0.9657	0.0376	0.9700	0.4905	-0.8587	0.3905	0.2708	0.7865	0.1758
Monmouthshire and Newport	1.3349	0.1819	-0.8973	0.3696	0.1029	0.6919	0.4890	-0.0540	0.9569	0.2488
Cardiff and Vale of Glamaorgan	0.8101	0.4179	0.2645	0.7914	0.2328	-0.1652	0.8688	-0.0571	0.9545	0.4535
Flintshire and Wrexham	-3.7802	0.0002	-2.4876	0.0129	***0.0001***	-1.2089	0.2267	0.4652	0.6418	0.1039
Powys	-0.3595	0.7192	-0.0979	0.9220	0.3746	-1.1541	0.2484	1.0736	0.2830	0.0539
SCOTLAND										
Aberdeen City, Aberdeenshire and North East Moray	-0.1501	0.8807	-0.2206	0.8254	0.4735	-0.5909	0.5546	-0.2980	0.7657	0.2617
Angus and Dundee City	0.5564	0.5779	-0.0668	0.9467	0.2957	-0.4891	0.6248	0.3258	0.7446	0.3566
Clackmannanshire and Fife	-0.2863	0.7747	0.3418	0.7325	0.3606	-1.0936	0.2741	0.6017	0.5474	0.0939
East Lothian and Midlothian	-0.7398	0.4594	-0.3548	0.7227	0.2318	-1.0206	0.3074	0.8588	0.3905	0.1479
Scottish Borders, The	-0.2589	0.7957	0.3662	0.7142	0.3926	-0.7209	0.4710	0.9835	0.3254	0.2408
Edinburgh, City of	-0.9112	0.3622	-1.0676	0.2857	0.1528	-1.5221	0.1280	-1.8096	0.0704	0.0513
Falkirk	-0.7125	0.4762	-0.2627	0.7927	0.2468	-1.1833	0.2367	0.2198	0.8260	0.1019
Perth and Kinross and Stirling	-0.0928	0.9261	0.0867	0.9309	0.4585	-0.4739	0.6356	0.2849	0.7757	0.3636
West Lothian	-0.7405	0.4590	-0.9818	0.3262	0.2428	-1.1038	0.2697	0.5241	0.6002	0.1219
East and West Dunbartonshire, Helensburgh and Lomond	0.1467	0.8833	-0.0117	0.9907	0.4366	-0.8266	0.4085	-0.3380	0.7354	0.1978
Dumfries and Galloway	0.4678	0.6399	-0.6478	0.5171	0.3127	-0.6493	0.5161	0.5368	0.5914	0.2607
East Ayrshire and North Ayrshire Mainland	-0.2237	0.8230	-0.0910	0.9275	0.3906	-0.5493	0.5828	0.5186	0.6041	0.3247
Glasgow City	-0.2227	0.8238	0.1591	0.8736	0.3966	-1.1249	0.2606	-0.1981	0.8430	0.0959
Inverclyde, East Renfrewshire and Renfrewshire	0.0310	0.9753	0.0353	0.9719	0.4775	-0.6898	0.4903	0.0938	0.9253	0.2428
North Lanarkshire	-0.1333	0.8939	0.1430	0.8863	0.4436	-1.0777	0.2812	0.6300	0.5287	0.1299
South Ayrshire	-0.6068	0.5440	-0.3618	0.7175	0.2867	-0.8034	0.4217	0.1031	0.9179	0.2218
South Lanarkshire	-0.6801	0.4965	0.2263	0.8210	0.2557	-1.2271	0.2198	0.1945	0.8458	0.0849
Highlands	-1.6165	0.1068	0.1034	0.9176	0.2150	-0.8587	0.3905	0.1544	0.8774	0.1270

* p-values significant at the 0.05 level and using the Bonferroni correction (i.e. 0.00043) are marked in bold and bold-italics respectively.

Deleted: ¶
TABLE A4 Robustness analysis for local indicators of spatial associa ... [3]

TABLE A3: Local Spatial Correlation Statistics*

Deleted: 6

	Index of occupational composition					Index of productivity				
	Getis Ord		Local Moran's I			Getis Ord		Local Moran I		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
NORTH EAST										
Hartlepool and Stockton-on-Tees	-2.2050	0.0275	1.1967	0.2314	0.0609	-1.3192	0.1871	0.7236	0.4693	0.0789
South Teesside	-1.4147	0.1572	1.5244	0.1274	0.0639	-1.1182	0.2635	0.0760	0.9394	0.1039
Darlington	-2.1685	0.0301	0.7376	0.4607	**0.0290**	-1.2807	0.2003	1.4552	0.1456	0.0859
Durham CC	-1.7030	0.0886	1.1797	0.2381	**0.0449**	-1.3421	0.1796	1.0281	0.3039	0.0539
Northumberland	-2.0517	0.0402	0.6433	0.5200	**0.0360**	-1.9894	0.0467	2.4001	0.0164	**0.0080**
Tyneside	-1.5845	0.1131	0.6014	0.5476	**0.0460**	-1.8116	0.0700	0.3691	0.7121	**0.0160**
Sunderland	-1.1273	0.2596	1.5323	0.1255	0.1329	-1.2459	0.2128	0.8357	0.4033	0.1129
NORTH WEST										
Cumbria	-2.2323	0.0256	0.9193	0.3579	0.0729	-1.2001	0.2301	0.7023	0.4825	0.0849
Halton and Warrington	-1.2596	0.2078	-0.0706	0.9438	0.0939	-0.6542	0.5130	0.0256	0.9796	0.2537
Cheshire CC	-2.1918	0.0284	-1.8517	0.0641	**0.0419**	-0.5368	0.5914	-0.3096	0.7569	0.2747
Greater Manchester South	-2.0169	0.0437	-0.3741	0.7083	**0.0200**	-1.3628	0.1729	-1.2033	0.2289	0.0799
Greater Manchester North	-2.3062	0.0211	0.5240	0.6003	**0.0060**	-1.9743	0.0483	0.9105	0.3626	0.0510
Blackburn with Darwen	-2.0660	0.0388	1.7788	0.0753	**0.0360**	-1.4018	0.1610	0.9960	0.3193	0.0609
Blackpool	-1.4617	0.1438	0.8896	0.3737	0.0639	-0.5744	0.5657	0.5472	0.5843	0.3157
Lancashire CC	-1.4482	0.1476	-0.5559	0.5783	0.0699	-0.6531	0.5137	0.0755	0.9398	0.2418
East Merseyside	-1.3242	0.1854	1.3640	0.1726	0.0879	-0.9834	0.3254	-0.8836	0.3769	0.1588
Liverpool	-1.3293	0.1837	-0.1250	0.9005	0.0979	-0.9394	0.3475	-0.0064	0.9949	0.1868
Sefton	-1.3544	0.1756	-0.2613	0.7938	0.0819	-0.7797	0.4356	0.4403	0.6598	0.1928
Wirral	-1.1874	0.2351	-0.2501	0.8025	0.1069	-0.8286	0.4073	0.2315	0.8170	0.2058
YORKSHIRE AND HUMBERSIDE										
Kingston upon Hull, City of	-0.9598	0.3372	1.8207	0.0686	0.1409	-1.0857	0.2776	1.0974	0.2725	0.1169
East Riding of Yorkshire	-1.4735	0.1406	-0.3910	0.6958	0.0629	-1.1595	0.2463	0.7989	0.4243	0.0959
North and N E Lincolnshire	-1.5953	0.1107	1.7237	0.0848	0.0529	-1.4224	0.1549	1.2034	0.2288	0.0669
York	-2.4084	0.0160	-0.3603	0.7187	**0.0060**	-1.5732	0.1157	-0.5024	0.6154	**0.0430**
North Yorkshire CC	-2.6185	0.0088	-1.6884	0.0913	**0.0020**	-1.6331	0.1025	0.7734	0.4393	0.0610
Barnsley, Doncaster, Rotherham	-1.8575	0.0632	2.9896	0.0028	**0.0250**	-0.4988	0.6179	0.5075	0.6118	0.3307
Sheffield	-1.5622	0.1182	-0.6498	0.5158	0.0539	-0.4567	0.6479	0.1405	0.8882	0.3297
Bradford	-1.9894	0.0467	1.3223	0.1861	**0.0180**	-1.5413	0.1233	0.6114	0.5409	**0.0460**
Leeds	-2.3860	0.0170	-0.0715	0.9430	**0.0060**	-1.8994	0.0575	-0.5788	0.5627	**0.0160**
Calderdale, Kirklees.Wakefield	-2.5650	0.0103	1.1666	0.2434	**0.0040**	-1.8286	0.0675	0.2492	0.8032	**0.0390**

TABLE A3: Local Spatial Correlation Statistics

Deleted: 6

	Index of occupational composition					*Index of productivity*				
	Getis Ord		*Local Moran's I*			*Getis Ord*		*Local Moran I*		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
EAST MIDLANDS										
Derby	-1.5419	0.1231	0.1699	0.8651	0.0689	-0.0455	0.9637	-0.0007	0.9995	0.4985
Derbyshire	-1.8263	0.0678	-0.3307	0.7409	0.0660	-0.2420	0.8088	0.1901	0.8492	0.4066
Nottingham	-1.3648	0.1723	0.8362	0.4030	0.0769	-0.3513	0.7254	0.1177	0.9063	0.3497
Nottinghamshire	-1.5455	0.1222	-0.6118	0.5407	0.1070	-0.6843	0.4938	0.3271	0.7436	0.2677
Leicester	1.1403	0.2542	-1.7494	0.0802	0.1369	-1.9537	0.0507	-1.3778	0.1683	**0.0170**
Leicestershire CC and Rutland	-1.2356	0.2166	-0.5337	0.5936	0.1009	0.1270	0.8989	0.0198	0.9842	0.4486
Northamptonshire	2.7308	0.0063	-0.6532	0.5136	**0.0050**	4.7808	0.0000	1.6028	0.1090	***0.0001***
Lincolnshire	-1.6743	0.0941	0.2180	0.2232	0.1270	-1.3893	0.1648	1.3479	0.1777	0.0539
WEST MIDLANDS										
Herefordshire, County of	-0.4022	0.6875	0.0065	0.9948	0.3556	-0.4626	0.6437	0.5102	0.6099	0.3546
Worcestershire	0.8643	0.3874	0.0488	0.9610	0.2048	1.5230	0.1277	-0.8357	0.4033	0.0779
Warwickshire	2.7001	0.0069	0.7147	0.4748	**0.0060**	3.2376	0.0012	2.9413	0.0033	**0.0020**
Telford and Wrekin	-1.1818	0.2373	1.0948	0.2736	0.1049	0.3438	0.7310	-0.2236	0.8231	0.3796
Shropshire CC	-1.0143	0.3104	-0.1366	0.8913	0.1379	-0.4406	0.6595	0.2665	0.7898	0.3337
Stoke-on-Trent	-1.0608	0.2888	2.1313	0.0331	0.1379	-0.4379	0.6615	0.4704	0.6381	0.3267
Staffordshire CC	-1.0218	0.3069	0.2748	0.7835	0.1518	0.3866	0.6990	-0.0645	0.9486	0.3397
Birmingham	1.2465	0.2126	0.4596	0.6458	0.1189	2.3753	0.0175	2.0807	0.0375	**0.0070**
Solihull	2.2865	0.0222	2.6683	0.0076	**0.0170**	3.1173	0.0018	3.6500	0.0003	**0.0020**
Coventry	2.0704	0.0384	-0.5800	0.5619	**0.0190**	3.1310	0.0017	2.7779	0.0055	***0.0001***
Dudley and Sandwell	0.6685	0.5038	-0.7847	0.4326	0.2817	1.1606	0.2458	-0.4305	0.6668	0.1309
Walsall and Wolverhampton	-0.4272	0.6692	0.4358	0.6630	0.3217	0.4907	0.6236	-0.1862	0.8523	0.3237
EAST										
Peterborough	0.8739	0.3822	-0.3225	0.7471	0.1858	2.1306	0.0331	1.0161	0.3096	**0.0340**
Cambridgeshire CC	2.9293	0.0034	4.5192	0.0000	***0.0001***	4.7525	0.0000	3.3811	0.0007	***0.0001***
Norfolk	-0.6295	0.5290	0.4782	0.6325	0.2837	-0.1002	0.9202	0.0276	0.9780	0.4905
Suffolk	1.5496	0.1212	-0.7244	0.4688	0.0679	1.7938	0.0728	-1.0708	0.2843	0.0569
Luton	4.4418	0.0000	-2.8115	0.0049	***0.0001***	6.3164	0.0000	6.5026	0.0000	***0.0001***
Bedfordshire CC	3.7853	0.0002	4.7698	0.0000	***0.0001***	5.9831	0.0000	1.9315	0.0534	***0.0001***
Hertfordshire	4.7446	0.0000	7.2415	0.0000	***0.0001***	6.6413	0.0000	11.3000	0.0000	***0.0001***
Southend-on-Sea	4.5971	0.0000	3.1662	0.0015	***0.0001***	6.1557	0.0000	-1.9331	0.0532	***0.0001***
Thurrock	5.7415	0.0000	-2.6552	0.0079	***0.0001***	6.1337	0.0000	6.9859	0.0000	***0.0001***
Essex CC	5.5477	0.0000	4.5184	0.0000	***0.0001***	6.2073	0.0000	3.4913	0.0005	***0.0001***

TABLE A3: Local Spatial Correlation Statistics

Deleted: 6

	Index of occupational composition					Index of productivity				
	Getis Ord		Local Moran's I			Getis Ord		Local Moran I		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
LONDON										
Inner London - West	5.2173	0.0000	18.7455	0.0000	***0.0001***	6.9654	0.0000	30.5316	0.0000	***0.0001***
Inner London - East	5.5848	0.0000	9.6419	0.0000	***0.0001***	6.4575	0.0000	24.6873	0.0000	***0.0001***
Outer London - East and NE	5.6047	0.0000	1.8524	0.0640	***0.0001***	6.1560	0.0000	5.8580	0.0000	***0.0001***
Outer London – South	5.4375	0.0000	9.8340	0.0000	***0.0001***	6.4803	0.0000	7.6446	0.0000	***0.0001***
Outer London - West and NW	4.9846	0.0000	9.8692	0.0000	***0.0001***	6.2960	0.0000	14.7916	0.0000	***0.0001***
SOUTH EAST										
Berkshire	5.3728	0.0000	10.0769	0.0000	***0.0001***	6.4668	0.0000	13.7705	0.0000	***0.0001***
Milton Keynes	3.2872	0.0010	2.2322	0.0256	***0.0001***	5.6296	0.0000	5.4760	0.0000	***0.0001***
Buckinghamshire CC	4.7971	0.0000	11.4010	0.0000	***0.0001***	6.8222	0.0000	9.0613	0.0000	***0.0001***
Oxfordshire	4.0513	0.0001	6.9510	0.0000	***0.0001***	6.1578	0.0000	5.5834	0.0000	***0.0001***
Brighton and Hove	4.2569	0.0000	8.3410	0.0000	***0.0001***	6.3417	0.0000	0.9971	0.3187	***0.0001***
East Sussex CC	4.3835	0.0000	2.5935	0.0095	***0.0001***	5.7659	0.0000	-5.1966	0.0000	***0.0001***
Surrey	5.2333	0.0000	12.7187	0.0000	***0.0001***	6.4801	0.0000	14.3239	0.0000	***0.0001***
West Sussex	4.2219	0.0000	2.9546	0.0031	***0.0001***	5.7521	0.0000	4.2989	0.0000	***0.0001***
Portsmouth	4.9755	0.0000	-1.4820	0.1383	***0.0001***	5.7789	0.0000	3.1835	0.0015	***0.0001***
Southampton	5.3196	0.0000	-2.0910	0.0365	***0.0001***	6.2985	0.0000	5.4350	0.0000	***0.0001***
Hampshire CC	5.1907	0.0000	4.1362	0.0000	***0.0001***	6.0403	0.0000	5.1297	0.0000	***0.0001***
Isle of Wight	-0.0061	0.9952	0.0088	0.9930	0.4765	1.2446	0.2133	-0.6489	0.5164	0.0809
Medway	6.3069	0.0000	-6.5344	0.0000	***0.0001***	6.3442	0.0000	0.7841	0.4330	***0.0001***
Kent CC	5.0127	0.0000	1.5889	0.1121	***0.0001***	5.9999	0.0000	2.6593	0.0078	***0.0001***
SOUTH WEST										
Bristol, City of	0.8920	0.3724	0.4408	0.6594	0.1818	0.4840	0.6284	0.2634	0.7922	0.3117
North and North East Somerset, South Gloucestershire	0.9656	0.3342	0.7078	0.4791	0.1778	0.0331	0.9736	0.0465	0.9629	0.4695
Gloucestershire	-0.5201	0.6030	-0.2964	0.7669	0.3007	0.1039	0.9172	0.0727	0.9420	0.4555
Swindon	3.8639	0.0001	0.5517	0.5812	***0.0001***	5.1886	0.0000	5.7927	0.0000	***0.0001***
Wiltshire CC	2.4285	0.0152	1.3972	0.1623	**0.0070**	2.0653	0.0389	-1.2493	0.2116	**0.0280**
Bournemouth and Poole	0.9734	0.3303	-0.0450	0.9641	0.1638	1.2294	0.2189	0.3954	0.6925	0.1249
Dorset CC	0.3497	0.7266	0.1040	0.9172	0.3377	0.5051	0.6135	-0.4295	0.6675	0.2777
Somerset	-0.2177	0.8277	0.1178	0.9063	0.3856	-0.9516	0.3413	0.8133	0.4161	0.1379
Cornwall and Isles of Scilly	-0.9689	0.3326	0.2759	0.7827	0.1319	-0.9991	0.3177	1.6395	0.1011	0.1249
Plymouth	-0.4726	0.6365	0.7178	0.4729	0.3167	-1.9537	0.0507	1.6171	0.1059	**0.0020**
Torbay	-0.6676	0.5044	0.2959	0.7673	0.2717	-0.9131	0.3612	1.4442	0.1487	0.1618
Devon CC	-0.9861	0.3241	-0.1755	0.8607	0.1698	-1.9190	0.0550	1.4487	0.1474	0.0600

Deleted: 6

TABLE A3: Local Spatial Correlation Statistics

	Index of occupational composition					*Index of productivity*				
	Getis Ord		*Local Moran's I*			*Getis Ord*		*Local Moran I*		
NUTS3 area	G_i	*p1*	$z(I_i)$	*p1*	*p2*	G_i	*p1*	$z(I_i)$	*p1*	*p2*
WALES										
Gwynedd+Anglesey	-0.3323	0.7397	0.2089	0.8345	0.3816	-0.8338	0.4044	0.9510	0.3416	0.1738
Conwy and Denbeighshire	-1.1130	0.2657	-0.4930	0.6220	0.1189	-0.5475	0.5840	0.5762	0.5645	0.2987
South West Wales	-0.5607	0.5750	0.1798	0.8573	0.3027	-1.4608	0.1441	1.5326	0.1254	**0.0450**
Central Valleys	0.0908	0.9276	-0.2234	0.8232	0.4785	-1.1243	0.2609	0.9010	0.3676	0.1259
Gwent Valleys	0.0328	0.9739	-0.0919	0.9268	0.4715	-1.0799	0.2802	1.0512	0.2932	0.1129
Bridgend and Neath Port Talbot	-0.1539	0.8777	0.1361	0.8917	0.4486	-1.1266	0.2599	0.8930	0.3719	0.1229
Swansea	-0.8777	0.3801	0.1054	0.9161	0.1938	-1.1567	0.2474	0.8506	0.3950	0.0949
Monmouthshire and Newport	0.6259	0.5314	0.2870	0.7741	0.2667	0.4177	0.6762	-0.0571	0.9545	0.3417
Cardiff and Vale of Glamaorgan	0.0600	0.9521	0.0556	0.9556	0.4875	-0.6472	0.5175	-0.1759	0.8604	0.2517
Flintshire and Wrexham	-1.4165	0.1566	1.3214	0.1864	0.0739	-1.2509	0.2110	0.2136	0.8309	0.0939
Powys	-0.0049	0.9961	0.0075	0.9940	0.4935	-1.5502	0.1211	2.0811	0.0374	0.1180
SCOTLAND										
Aberdeen City, Aberdeenshire and North East Moray	-0.9711	0.3315	-0.3303	0.7412	0.1349	-0.4449	0.6564	-0.3166	0.7515	0.4106
Angus and Dundee City	-0.6373	0.5239	0.6330	0.5267	0.2647	-0.2975	0.7661	0.1597	0.8731	0.4386
Clackmannanshire and Fife	-0.9162	0.3595	0.4618	0.6442	0.1608	-0.7804	0.4351	0.4156	0.6777	0.1838
East Lothian and Midlothian	-0.8036	0.4216	0.7472	0.4549	0.2228	-0.7250	0.4685	0.5413	0.5883	0.2418
Scottish Borders, The	-0.9893	0.3225	0.4459	0.6557	0.1588	-0.5318	0.5949	0.7265	0.4675	0.3077
Edinburgh, City of	-1.5135	0.1302	-2.7435	0.0061	*0.0480*	-1.1127	0.2658	-0.8953	0.3706	0.1139
Falkirk	-0.8882	0.3744	0.5435	0.5868	0.1818	-0.9639	0.3351	-0.1945	0.8458	0.1658
Perth and Kinross and Stirling	-0.7156	0.4743	-0.1424	0.8868	0.2358	-0.0821	0.9346	0.0776	0.9382	0.4845
West Lothian	-0.7571	0.4490	0.8786	0.3796	0.2438	-0.8418	0.3999	0.0172	0.9863	0.2228
East and West Dunbartonshire, Helensburgh and Lomond	-0.9169	0.3592	-0.7508	0.4528	0.1688	-0.5143	0.6071	-0.3111	0.7557	0.3067
Dumfries and Galloway	-0.8884	0.3743	0.8042	0.4213	0.2078	-0.3611	0.7180	0.1620	0.8713	0.3706
East Ayrshire and North Ayrshire Mainland	-0.7448	0.4564	0.8021	0.4225	0.2328	-0.0226	0.9820	0.0307	0.9755	0.4745
Glasgow City	-0.8108	0.4175	-0.0297	0.9763	0.1968	-0.8567	0.3916	-0.3957	0.6923	0.1878
Inverclyde, East Renfrewshire and Renfrewshire	-0.7855	0.4322	-0.2453	0.8062	0.2178	-0.3061	0.7595	0.0765	0.9390	0.4016
North Lanarkshire	-0.8214	0.4114	0.7547	0.4504	0.2168	-0.6957	0.4866	0.4344	0.6640	0.2418
South Ayrshire	-1.1230	0.2615	0.1789	0.8580	0.1349	-0.2676	0.7890	-0.0137	0.9891	0.4446
South Lanarkshire	-0.9424	0.3460	0.3187	0.7500	0.1828	-0.8280	0.4077	0.0122	0.9902	0.1958
Highlands	-0.9181	0.3585	0.2384	0.8116	0.1621	-0.8794	0.3792	0.0396	0.9684	0.3488

* p-values significant at the 0.05 level and using the Bonferroni correction (i.e. 0.00043) are marked in bold and bold-italics respectively.

Page 1: [1] Deleted Eleonora Patacchini 2/22/2006 10:00:00 AM

Appendix 1: Methods of Exploratory Spatial Data Analysis

Global spatial autocorrelation

When the variable under investigation is measured on a continuous scale, the measurement of global spatial autocorrelation is usually based on Moran's I and Geary's c statistics (Cliff and Ord, 1981).

Moran's I is defined as

$$I = \frac{n}{S_0} \frac{\sum_i \sum_j w_{ij} \left(x_i - \bar{x} \right)(x_j - \bar{x})}{\sum_i \left(x_i - \bar{x} \right)^2} \qquad (1)$$

where n is the number of observations, x_i denotes the observation at site i for the variable of interest X, and w_{ij} denotes the elements of the spatial weights matrix. S_0 is a scaling factor equal to the sum of all the elements in the weight matrix. The spatial weight matrix may be row-standardised such that the elements $\tilde{w}_{ij}$ in each row sum to 1 in order to normalise the size of the neighbourhood set for each site. In this case S_0 = n and the expression (1) simplifies to a ratio of the spatial cross-product to the variance. Moran's I is a cross product coefficient scaled to be less than one in value, with an expected value $E(I) = -1/(n-1) \approx 0$ for n sufficiently large. Values for Moran's I larger (smaller) than the expected value indicate positive (negative) spatial correlation.

An alternative measure of global spatial autocorrelation is given by Geary's c coefficient which is based on squared deviations. Geary's c is defined as

$$c = \frac{(n-1)}{2S_0} \frac{\sum_i \sum_j w_{ij} \left(x_i - x_j \right)^2}{\sum_i \left(x_i - \bar{x} \right)^2} \qquad (2)$$

The expected value for Geary's c is 1. Values of Geary's c less than one indicate positive spatial correlation, while values larger than one suggest negative spatial correlation.

Inference is typically based on a standardised z-value of the statistic computed by subtracting the expected value and dividing by the standard deviation in the usual way. Assuming that the variable of interest is normally distributed, the z-value follows a standard normal distribution, and the significance of the test statistic may be judged by comparing the computed z-value with its probability in the standard normal tables (for a theoretical discussion and detailed expressions for the moments of the (asymptotic) distributions of I and c under various assumptions, see Cliff and Ord (1981).). An alternative approach, referred to as the conditional randomization approach, is to assume that each value observed could equally likely have occurred at all locations. A reference distribution for the Moran's I or the Geary c statistic is then generated empirically by randomly reshuffling the observed values over all possible locations and by re-computing the statistic for each new sample. A pseudo significance level is obtained as $p = \frac{T+1}{M+1}$ where T is the number of the computed values of the test statistic that are equal to or larger than the observed value and M is the number of permutations. Note that the highest level of significance (i.e. that corresponding to $T=0$) is determined by the chosen number of replications, M. Given a certain number of permutations, a low value of this pseudo-significance level implies that the observed `value of the statistic (I or c) is extreme with respect to its reference distribution, and therefore the null hypothesis of spatial randomness should be rejected. Because of the theoretical limitations in using a normality approximation, the permutation approach is also usually considered (see Anselin 1988 and 1995 for more details).

Local spatial autocorrelation

Both Moran's I and Geary's c statistics are global statistics, in the sense that the overall pattern in the data is summarized in a single statistic, and as such they may be of limited interest. Such global statistics may summarise a number of possible disparate spatial relationships for a given set of data. A number of local indicators of spatial association that measure spatial dependence in a region of the study area have been developed (Anselin, 1995a; Getis and Ord, 1995). These statistics detect significant associations between a single x_i and its neighbours and are suited to the task of identifying clusters of high or low values of the variable or the existence of atypical localizations (spatial outliers) in the form of sites of high (low) values surrounded by areas of relatively low (high) values for the variable of interest.

The Moran Scatterplot and the Local Moran's I_i

A simple visual depiction of the relationship between the value of X at site i and the average of its neighbours is obtained by plotting $(x_i - \bar{x})$ on the horizontal axis against its spatial lag $\sum_{j=1}^{n} \tilde{w}_{ij}(x_j - \bar{x})$ on the vertical axis, where $\tilde{w}_{ij}$ denotes the elements of the (row-standardised) spatial weights matrix (i.e. the Moran Scatterplot).

The Moran's I measure of global spatial autocorrelation is formally equivalent to the estimate of the slope coefficient of the linear regression of Wz on z where W denotes the (row-standardized) weight matrix and z denotes the vector of observations in deviation form i.e. $(x_i - \bar{x})$. Using a row-standardized weight matrix (which implies S0=n), Moran's I in matrix form reduces to $I = z'Wz/z'z$ where z denotes the vector of observations in deviations form, i.e. $(x_i - \bar{x})$. Given this, standard regression diagnostics may be used to detect outliers and to identify individual areas that exert strong influence on the global Moran's I statistic. 'Studentized' or normed residuals (i.e. the absolute value of the ith residual divided by the square root of the residual sum of squares) and leverage measures based on the diagonal elements in the Hat or Projection Matrix may be used to detect outliers. The Hat or Projection matrix $H = X(X'X)^{-1}X'$ where X is the nxk matrix of observations on the k explanatory variables in the regression. The extent to which such observations are influential may be assessed using Cook's distance criterion di which measures the effect on the estimated slope coefficient of excluding the ith observation (see Anselin, 1996 and 1995b for more details).

The four different quadrants of the so-called Moran scatterplot correspond to the four types of local spatial association between a spatial unit and its neighbours: HH, contains areas with a high value surrounded by areas with high values, LH, contains areas with a low value surrounded by areas with high values; LL consists of low value areas surrounded by other areas with low values; HL consists of areas with high values surrounded by low value areas. Quadrants HH and LL correspond to positive spatial autocorrelation indicating spatial clustering of similar values; while the quadrants HL and LH correspond to negative spatial autocorrelation with groupings of dissimilar areas. While the Moran scatterplot shows the spatial regime (position across quadrants) of each location, it does not give an indication on the statistical significance of these local spatial association schemes. The existence of any significant local spatial pattern is detected by the use of local spatial correlation statistics (LISA).

.The 'local' version of Moran's *I* statistic for each spatial unit i is defined as

$$I_i = \frac{(x_i - \bar{x})\sum_j \tilde{w}_{ij}(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2 / n} \tag{3}$$

where n is the number of observations, x_i denotes the observation on unit i for the variable X and $\tilde{w}_{ij}$ denotes the elements of the (row-standardised) spatial weights matrix as before. It follows that the global Moran *I* is related to the local version as follows

$$I = \sum_i I_i \,,$$

i.e. the sum of all local Moran values is proportional to the global Moran's statistic. The local Moran focuses on the correlation between the value of the variable X at site i and its neighbouring values. In the presence of global spatial autocorrelation, inference should be based on the conditional permutation approach described above (Anselin, 1995). A significant and positive value for I_i indicates a local spatial clustering of similar values of the variable X, either high or low. If the local Moran's statistic, I_i , is significant and negative then the value of X at site i and those for its neighbours are dissimilar. To detect whether an area with a positive local statistic is in the HH or LL spatial regime and whether an area with a negative local statistic is in the HL or LH spatial regime, one has to look at its position in the Moran scatterplot, hence the complementarity between local Moran's statistics and Moran scatterplots.

Getis Ord Statistics

Alternative measures of local spatial correlation are the Getis Ord statistics (Getis and Ord, 1995), which are based on a comparison of the sum of values within the neighbourhood set with the corresponding global value.

There are two forms of the statistic

$$G^*_i = \frac{\sum_j w_{ij} x_j - \bar{x}\, S_0}{\sqrt{\left(\sum_j \left(x_j - \bar{x}\right)^2 / n\right)} \sqrt{\left(n \sum_j w_{ij}{}^2 - S_0^2\right) / (n-1)}}$$

and

$$G_i = \frac{\sum_{j \neq i} w_{ij} x_j - \bar{x} S_{-i}}{\sqrt{\left(\sum_{j \neq i} \left(x_j - \bar{x}_{-i}\right)^2 / n-1\right)} \sqrt{\left((n-1) \sum_{j \neq i} w_{ij}^2 - S_{-i}^2\right) / (n-2)}}$$

where $\bar{x}$ *and* $\bar{x}_{-i}$ denotes the sample mean of x with and without the ith observation included respectively. Similarly S_0 and S_{-i} denote the sum of the spatial weights with and without the ith weight included. Under the null hypothesis of spatial randomness, the sum of values of X among the neighbourhood set would not differ systematically from the sum of values for the sample as a whole. The G^*_i statistic includes site i in the summations that are the basis of the comparison, while it is excluded in the case of the G_i statistic. The spatial clustering of high values of X results in positive values for the G^*_i and G_i statistics, and negative values for G^*_i and G_i are indicative of a clustering of relatively low values for X. Inference can be based on the asymptotic approximation to the standard normal distribution as suggested by Getis and Ord (1995), although the presence of global spatial autocorrelation may affect the power of the test.

---Page Break---

Page 2: [2] Deleted Eleonora Patacchini 3/10/2006 5:53:00 PM

TABLE A1 : Moran Scatterplot – Tests for Outliers

Spatial weight matrix: estimated travel time ≤ 90 minutes		
NUTS3 area	*Leverage Measure**	*Cook's distance*

	H_i	criterion d_i
GVA per (employee) hour worked:		
Inner London -West	0.1057	0.0076
Berkshire	0.0778	0.0011
Outer London - West and North West	0.0652	0.0008
Buckinghamshire CC	0.0611	0.0042
Surrey	0.0496	0.0120
Stoke-on-Trent	0.0427	0.0011
Inner London – East	0.0414	0.0151
Average hourly earnings		
Inner London -West	0.3129	1.9274
Inner London - East	0.1211	0.0221
Surrey	0.0652	0.0016
Berkshire	0.0536	0.0000
Outer London - West and North West	0.0514	0.0000
Buckinghamshire CC	0.0425	0.0101

* Observations with a high degree of leverage are those for which the associated diagonal element in the projection matrix exceeds 2(k/n) = 0.0336. A value of the Cook's statstics in excess of 0.7 is regarded as denoting a significant observation.

Page 6: [3] Deleted **Eleonora Patacchini** **3/10/2006 6:02:00 PM**

TABLE A4 Robustness analysis for local indicators of spatial association

	GVA per (employee) hour worked						***Average hourly earnings***				
	From k=90 to k=120						*From k=90 to k=120*				
	Not significant	***HH***	***LL***	***HL***	***LH***		***Not significant***	***HH***	***LL***	***HL***	***LH***
Not significant	0.971	0.000	0.014	0.000	0.014	***Not significant***	0.917	0.072	0.000	0.000	0.010
HH	0.000	0.916	0.000	0.083	0.000	***HH***	0.000	1.000	0.000	0.000	0.000
LL	0.000	0.000	1.000	0.000	0.000	***LL***	0.000	0.000	1.000	0.000	0.000
HL	0.000	0.000	0.000	1.000	0.000	***HL***	0.000	0.000	0.000	1.000	0.000
LH	0.000	0.000	0.000	0.000	1.000	***LH***	0.000	0.000	0.000	0.000	1.000
	From k=120 to k=150						*From k=120 to k=150*				
	Not significant	***HH***	***LL***	***HL***	***LH***		***Not significant***	***HH***	***LL***	***HL***	***LH***
Not significant	0.825	0.05	0.075	0.025	0.025	***Not significant***	0.917	0.072	0.000	0.000	0.010
HH	0.000	1.000	0.000	0.000	0.000	***HH***	0.000	1.000	0.000	0.000	0.000
LL	0.000	0.000	1.000	0.000	0.000	***LL***	0.000	0.000	1.000	0.000	0.000
HL	0.000	0.000	0.000	1.000	0.000	***HL***	0.000	0.000	0.000	1.000	0.000
LH	0.000	0.000	0.000	0.000	1.000	***LH***	0.000	0.000	0.000	0.000	1.000

Figure 1: Spatial Distribution of Income in Great Britain

Quintiles by NUTS3 Administrative Areas

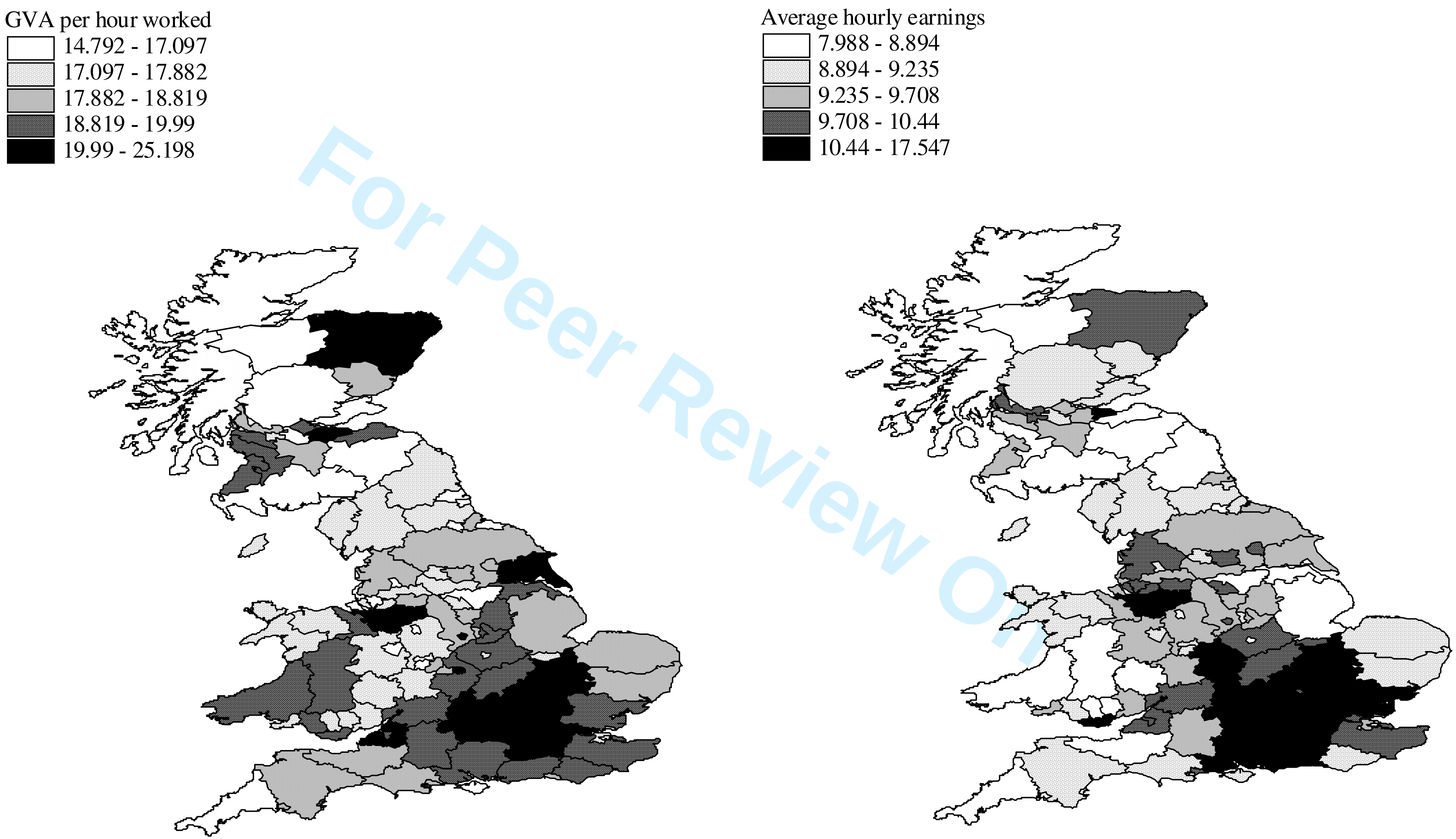

Figure 2: Moran Scatterplots for Income

Spatial proximity lag: travel time of < 90 mins

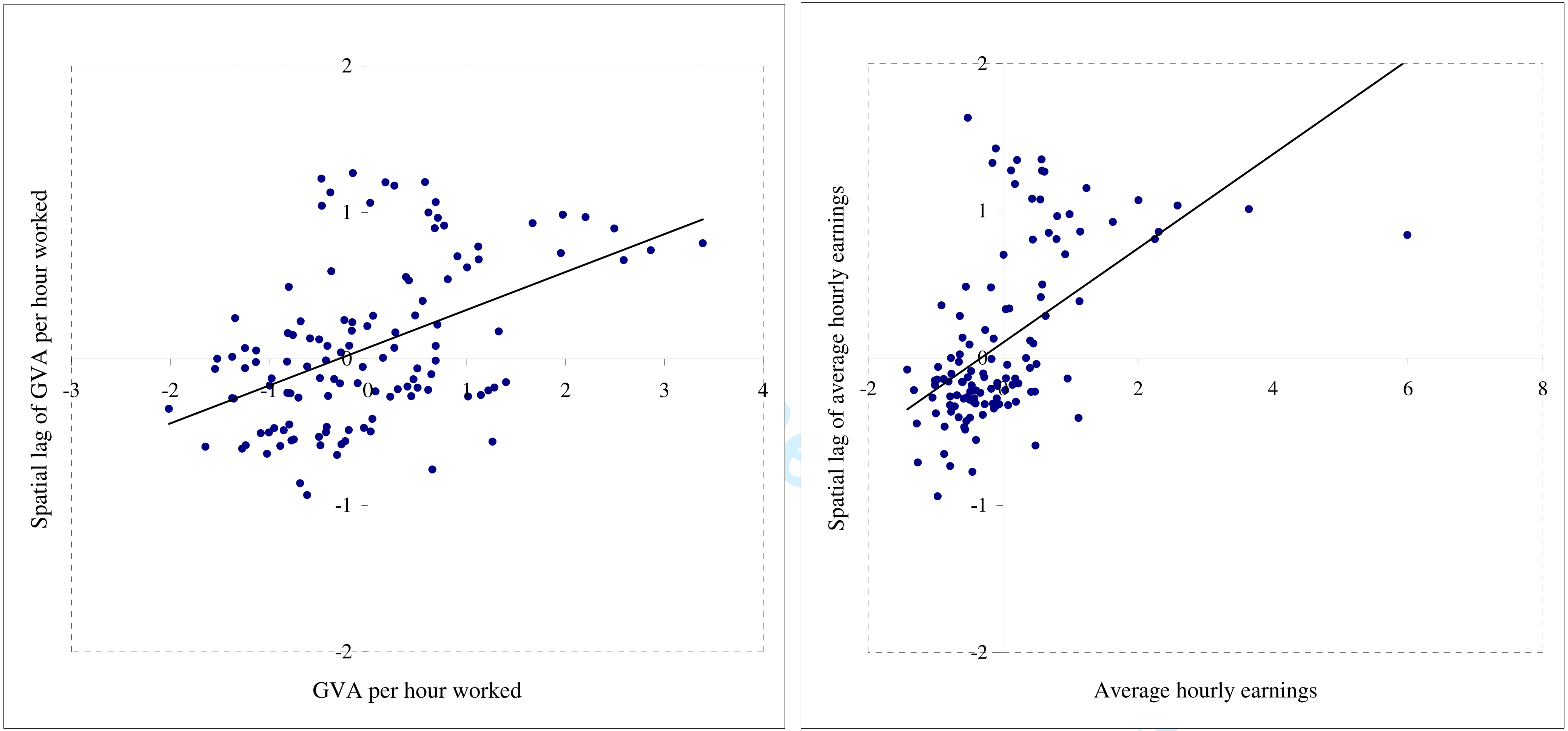

Figure 3: Local Indicators of Spatial Association for GVA per Hour Worked and Average Hourly Earnings (Significant at 0.05 level)

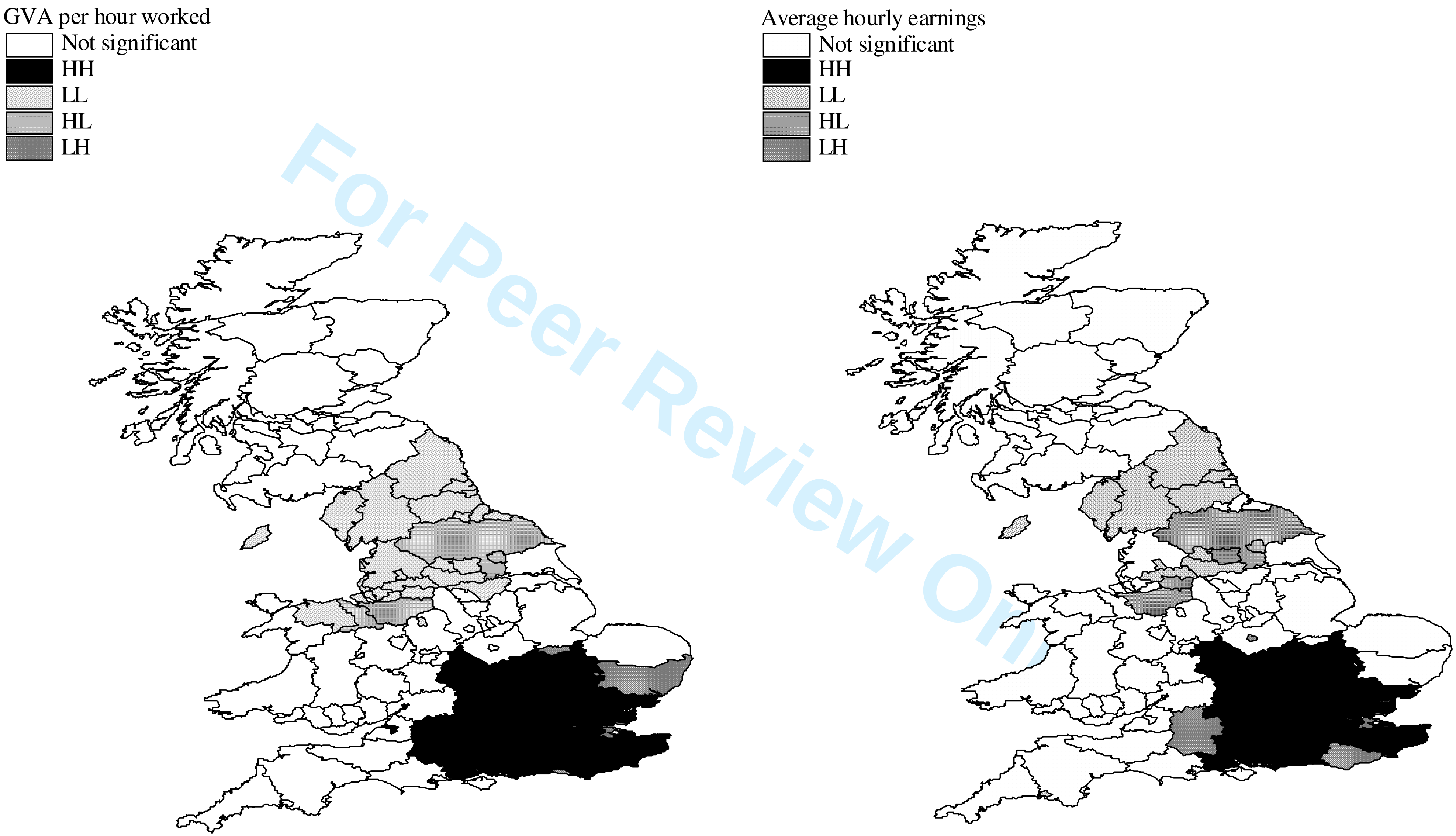

Figure 4: Local Indicators of Spatial Association for Occupational Composition and Productivity Indices (Significant at 0.05 level)