Article:
Ajebon, MO and Norman, PD (2015) Beyond the Census: A Spatial Analysis of Health and Deprivation in England. GeoJournal: an international journal on human geography and environmental sciences. ISSN 0343-2521

https://doi.org/10.1007/s10708-015-9624-8

Beyond the Census: A Spatial Analysis of Health and Deprivation in England

Mildred Oiza Ajebon [1] and Paul Norman [2]

1 Department of Geography, University of Durham, Durham, DH1 3LE, UK

Email: m.o.ajebon@durham.ac.uk

Tel: +44 (0) 740 58 90795

2 Centre for Spatial Analysis & Policy, School of Geography, University of Leeds, Leeds, LS2 9JT, UK

Corresponding author:

Paul Norman

email: p.d.norman@leeds.ac.uk

Tel:+44 (0)113 34 38199 Fax:+44 (0)113 34 33308

Beyond the Census: A Spatial Analysis of Health and Deprivation in England

Abstract

Whilst the UK is planned to have a census in 2021, it may well be the last and there is official acknowledgement that the country's statistical system should be enhanced by greater use of administrative data. Thus, there is a pressing need to understand whether alternative data sources are fit for social science purposes. This study assesses the potential of utilizing administrative statistics for investigating the relationships between health and socioeconomic distributions for small areas; a type of study regularly carried out using census data.

Pairs of administrative measures of health and deprivation indicator variables are compared with census equivalents for Lower Super Output Area geographies in England in 2001. The administrative datasets are then used to derive health measures and deprivation indexes for the time points 2001, 2006 and 2010. Inequalities in health are then analysed using administrative data derived area measures with health found to be poorer in areas with socioeconomic disadvantage. Overall, the administrative datasets used here reveal very similar patterns of health and deprivation to the decennial census suggesting they are viable alternatives and have great potential to enhance the country's statistical system given their availability outside census years.

Keywords: Deprivation and Health; Decennial Census; Administrative Statistics

Beyond the Census: A Spatial Analysis of Health and deprivation in England

Introduction

The decennial census in the UK is the most comprehensive source of socio-demographic information at varying geographical levels. The census has been essential for the allocation of public services based on the needs of subnational areas and for business targeting in the UK. However, the decennial census approach has been criticised in many ways. In May 2010, Sir Michael Scholar, Chair of the UK Statistics Authority (UKSA) wrote to the Minister for the Cabinet Office to say that,

> "As a Board we have been concerned about the increasing costs and difficulties of traditional Census-taking. We have therefore already instructed the ONS to work urgently on the alternatives, with the intention that the 2011 Census will be the last of its kind"(ONS, 2011a).

High population mobility and more complex ways in which people live make the procedure of the UK's decennial Census taking more challenging and the concept of a snapshot of the entire population at a point in time less relevant (Dugmore et al, 2011). Rapid population change, cost constraints, and the need for more timely and up-to-date statistics are increasingly driving the evolution of alternative methods for the collection and dissemination of detailed population and sociodemographic data across all geographical scales in the UK. Other factors driving the need for change include privacy, accuracy, completeness and duplication of existing data. Recent advances in information and communications technology which has improved the efficiency in the ways in which both public and private records (referred to as 'administrative statistics') are stored and organised suggests potential for alternative approaches (ONS, 2011a). A system that organises information already being collected about the population from the day-to-day administration of public services and storage of customer data may be a more cost effective way of meeting the high demand for detailed and timely statistics. Essentially, the collection and dissemination of administrative data has great potential to, "satisfy the growing thirst for population intelligence" (Harper and Mayhew, 2012: 184).

The 'Beyond 2011' programme began in April 2011 to assess the feasibility of using administrative statistics as an alternative to the UK's traditional census. Informed by the programme, a recent announcement that the UK government has welcomed the recommendation from the National Statistician that there will be a 2021 Census (predominantly online) but that the country's statistical system should be enhanced by greater use of administrative data. Whilst good progress has been made to demonstrate that administrative sources have utility to aid in counting the population (Harper and Mayhew, 2012), there is a pressing need to understand whether alternative data sources are fit for social science purposes in the analysis of population attributes. This paper contributes to the evidence base.

It is logical then to compare administrative data with the census to check for similarities in distributions. Since it is unrealistic to consider the full range of topics covered in the census, this study focuses on indicators normally included in health and deprivation measures and for which variables from administrative sources can be sourced. The work focuses on comparisons of health and deprivation measures from the 2001 Census using indicators for which alternative sources are available for the time points 2001, 2006 and 2010 in England. There are statistical descriptions to establish whether administrative measures of health and deprivation emulate the decennial census. The paper examines similarities between administrative measures of health and deprivation and census-based measures and explores whether spatial and temporal variations can be identified and monitored using these administrative sources. The small area units of analysis are the Lower Super Output Areas (LSOAs), a statistical geography for which administrative statistics are regularly published.

We base the work on the 2001 Census since subsequently there is almost a decade of administrative data to access. In 2011, whilst census data for that year could be used, there were revisions to the census geographies including the LSOAs but not to the geographies used for the release of administrative data which still used the 2001 boundary definitions. To have a 2001-2011 time-series including both censuses requires conversions between the geographies which are possible (Norman et al, 2003) but beyond the scope of this work.

Relationships between health and deprivation

Attempts at quantifying geographical variations of need have been prominent in studies of population health with ample evidence that socially disadvantaged groups are more likely to suffer poorer health outcomes compared with those who are well-off in society. The presence of regional variations in deprivation is well founded as studies have also shown socioeconomic structuring of the population to influence spatial patterns (Walsh et al, 2010). Health outcomes appear disproportionately poorer in some areas in the UK, notably Glasgow, Manchester and Liverpool, than can be explained by levels of deprivation (Whynes, 2009). There were attempts at linking socioeconomic conditions and health outcomes based on the 1971 Census data but an expansion of investigations into social inequalities in health followed the 1981 Census and the increasing use of residential postcodes in statistical records (Carstairs, 1995). As a result, the past three decades have seen increasing use of census-based approaches in the study of health and deprivation based on ward geographies. The work of Jarman (1983), Townsend et al (1988) and Carstairs and Morris (1990) represent landmark approaches in the use of deprivation indices developed from census variables to explain geographical variations in mortality in Britain. Indicators used include unemployment, household overcrowding, lacking cars, non-home ownership, lone pensioners, single parenthood and low social class. The Index of Multiple

Deprivation (IMD) represents a shift in the quantification of local area deprivation in England (and the other home nations) based on administrative statistics to guide public policy initiatives and for targeting resources to areas of need (IMD, 2010). Post 2001, IMDs have been developed using LSOAs in 2004, 2007 and 2010. The availability of census-based measures and the IMD underpin many studies of small area of deprivation during the last decade (Noble et al, 2006; Stafford et al, 2008; Tarozzi and Deaton, 2009; Norman, 2010).

With increases in the availability of administrative data relating to health and economic activity, various studies have explored whether there are similar patterns to those found using census data (e.g. Nordbotten, 2010). Norman and Bambra (2007) examined the utility of the professionally diagnosed Incapacity Benefit (IB) data as an updatable indicator of census measure of morbidity. The results showed that IB claims are reasonably consistent with the census health indicators and there are strong relationships between incapacity benefit claims and census measures of area type. Smith et al (2010) examined the utility of the General Household Survey (GHS) for predicting inter-censal measure of small area health expectancies. The results of the comparison of health expectancy as measured by the 2001 Census and 2001-05 GHS showed consistencies and suggest that the latter could be used as an updatable indicator of population health.

Thus, attempts have been made to determine links between health outcomes and deprivation using census-based approaches. However, comparisons of census measures of health and deprivation with equivalent administrative data in a geographical context have largely been limited to the analysis of mortality and of morbidity using limiting long-term illness (LLTI) and incapacity benefit (IB) data. Relationships between other pairs of deprivation indicators are yet to be fully explored (Bécares et al, 2011). No appraisals were uncovered on the potential of utilizing cross-sections of administrative statistics on different aspects of the population or on a timeline in England. This study fills this gap by investigating the relationship between health outcomes and deprivation using administrative data sources at different time points.

Data and Methods

Spatial data and sources

The geographical units of analysis for this study are the 32,482 Lower Super Output Areas (LSOAs) in England. The LSOAs are statistical geographical units for which administrative statistics for small areas are published on a regular basis making them suitable for comparison over time (Neighbourhood-Statistics, 2004). The GIS boundaries have been downloaded from the UK Outline Reference Database for Education and Research Study (UKBORDERS) available at EDINA. The sociodemographic datasets were acquired from national sources; ONS 2001 Census data (http://casweb.mimas.ac.uk/) and administrative data from Neighbourhood Statistics

(www.neighbourhood.statistics.gov.uk/) and NOMISWEB (www.nomisweb.co.uk/). The administrative datasets used in this study can be classified into:

- **Health:** Total event counts of Incapacity Benefit (IB) claims as an equivalent of the census Limiting Long Term Illness (LLTI) question;
- **Unemployment:** Job Seekers Allowance Claimant (JSAC) to emulate census unemployment;
- **Housing:** Council Tax Bands (CTBs) as an indirect measure of census tenure;
- **Lone Parent:** Lone Parent claimants records from Works and Pensions Longitudinal Survey (WPLS) as the equivalent of census low parent household;
- **Pensioners:** Pension Claimants as an equivalent of census pensioner households;
- **Resident Population:** Revised ONS Mid-year population estimates (MYEs) as the measure of resident populations of the LSOAs.

Exploration of spatial data

Initial explorations include a test for normality using skewness and Kolmogorov Smirnov. The percentages of all the variables are negatively skewed and have been log transformed to near normal distributions. Spatial and temporal patterns of pairs of variables are examined in correlations and maps.

Standardised Illness Ratios (SIRs)

To measure the spatial distribution of population health in England, indirectly Standardised Illness Ratios (SIRs) for LSOAs have been calculated for both LLTI and IB. The aim here is to compare the spatial and temporal patterns of ill-health for small areas using the 2001 Census LLTI SIRs as a benchmark and compare these with IB SIRs of 2001, 2006 and 2010. Although the relationship between LLTI and 'Not Good Health' (NGH) is strong (0.71, $p = 0.00$), LLTI has been more widely used so findings can be compared with previous studies (e.g. Norman and Bambra 2007). At national level, the age-groups (16-24, 25-49, 50-59, 60 and over) used are those for which the IB data are available and these were kept consistent between data sources for comparisons to be made.

Measures of socioeconomic disadvantage

The relationship between health and socioeconomic distributions for the time-series 2001, 2006 and 2010 are measured here by implementing a deprivation index as a measure of socioeconomic disadvantage of LSOAs in England. The patterns identified are compared with the Townsend deprivation index and its relationship with IB and LLTI. The Townsend index is chosen because it is a widely used index but with the limitations and the concerns relating to the choice of indicators and methods well known (Bradford et al, 1996; Deas et al, 2003). Directly comparable administrative statistics could not be obtained for three Townsend indicators: non-home ownership, non-car

ownership and household overcrowding. However, administrative indicators of less advantaged housing circumstances, the lower end of the council tax band (Band A and B), and other indicators of disadvantage such as Lone Parents were included in the alternative measure of deprivation. Using relevant denominators, percentages of all input variables were derived and log transformed to near normal distributions. The transformed variables were standardised by calculating z-scores [(Observation – Mean) / (Standard deviation)] to prevent any variable from dominating the model. The final index scores of the alternative deprivation indexes were derived by summing the z-scores of the variables equally weighted. This is consistent in method with the Townsend index (Townsend et al, 1988). The relationship between the alternative deprivation index and IMD 2010 is also examined.

Results

Census and administrative indicators of health and deprivation: Do they tell the same story?

Table 1 presents data for comparisons between variables. In line with previous studies (Bambra and Norman, 2006; Norman and Bambra, 2007), the LLTI variable shows a strong relationship with working-age IB claims (0.70, $p = 0.00$). The strongest correlations (0.98, $p = 0.00$ and 0.96, $p = 0.00$) exist between equivalent measures of total numbers of resident population and pensioners. This is to be expected since the mid-year population estimates are estimates derived from a census base population. The weakest relationship (0.55, $p = 0.00$) exists between census measures of households in rented accommodation and number of dwellings in the lower end of council tax bands (A and B). The low correspondence will be because the council tax band is not directly comparable with the census rented tenure measure. However, the low council tax band variable has strong relationships with LLTI, non-car ownership and unemployment (0.63, 0.64 and 0.61 all at $p = 0.00$) suggesting an association with socioeconomic disadvantage. Strong positive relationships are observed between unemployment and lone parent claims, non-car ownership and non-home ownership (0.76, $p = 0.00$ and 0.82, $p = 0.00$ and 0.75, $p = 0.00$).

[Table 1 about here]

Figure 1 shows temporal variations in the number of persons claiming health and economic deprivation benefits. The total number of IB claimants was 2,919,775; 2,919,770 and 1,728,785 in 2001; 2006 and 2010 respectively. Whether the observed decline in the number of persons claiming IB in England reflects changing health or changing eligibility for benefit claims is not certain. In contrast, unemployment, pension claims and lone parent claims rose between 2001 and 2010. The rise in the number of persons claiming economic benefits may be attributed to the recent economic downturn in the UK in the last decade (CEBR, 2012, ONS, 2011b). The counts of dwellings in the lower end of the council tax bands remain at a stable level of 43.7% for all years.

[Figure 1 about here]

Figure 2 shows differences with and between the census and administratively derived health measures for persons above the age of 16 for England. The highest counts for census-based variables are for LLTI and lowest for the permanently sick and disabled (PSD). Total IB claims exceed those of severe disablement allowance for the administrative based measures. IB is highest for 2006 and severe disablement fell steadily between 2001 and 2010.

[Figure 2 about here]

Figure 3 illustrates age-specific rates of self-reported health measures and IB claims for England. It can be seen that LLTI has the highest rates at all ages from 16 and above which is closely followed by NGH. PSD has the lowest rates for all ages and all categories and is most similar to IB. IB shows a similar pattern of illness rising steadily up to age 60. As expected, rates are lowest for ages below 25 for all sources. A temporal evaluation shows a decline of IB rates for all ages. The 2006 IB rates (though lower) are similar to the 2001 IB for ages 16-59. The 2010 IB rates are the lowest of three IB measures.

[Figure 3 about here]

The correlations in Table 2 show strong positive relationships between all measures of ill-health. The correlations are increasing more strongly between administrative-based IB and the 2001 Census measures of PSD (0.93), LLTI (0.90) and IB (0.82) for ages above 16. IB appears to be most similar to PSD compared with LLTI and NGH. The relationships between IB and census measures of health tend to wane over time. For instance, the relationship between IB and LLTI declined from 0.90 in 2001, 0.87 in 2006 to 0.84 in 2010 suggesting the geography of health may be changing a little during the decade.

[Table 2 about here]

Figure 4 shows the distributions of standardised illness ratios of LLTI and IB for the time-series 2001, 2006 and 2010. Compared with the national average of 100 depicted in white, the green areas have lower levels of IB and the red areas have higher levels of IB. Both the census LLTI and the alternative IB identify similar distributions of population health for England. The spatial patterns observed here are similar to the broad patterns identified elsewhere (Norman and Bambra, 2007, Smith et al, 2010, Stafford et al, 2008) with the southeast having the lowest illness ratios. The two measures identified similar patterns which coincide with the narrower range of health inequalities in rural areas compared with urban areas where health inequalities are markedly defined. The LSOAs with best health are in areas like Elmbridge, South Oxfordshire, Southwark, Hart and Maidenhead which are mainly in the 'Home Counties' or semi-rural communities. The LSOAs with worst health are largely in old industrial areas in Local Authorities (LAs) such as Easingston, Gateshead, Liverpool, Manchester, Rochdale and Blackpool.

[Figure 4 about here]

The 32,842 LSOAs are classified into quintiles to examine the consistency between the standardised ratios derived from census and administrative measures of health in quantitative terms. Quintile 1

represents the 6,496 LSOAs with the lowest illness ratios for all health measures and quintile 5 are the 6,496 LSOAs with the highest ratios. Table 3 shows a cross-tabulation of the time-series IB with census health measures. The highlighted cells are the leading diagonals of the matrices representing LSOAs that have been cross-classified into the same quintile for both IB and other health measures. Cells away from the leading diagonal identify where LSOAs have been classified into different quintiles. The higher the number of cells contained on the leading diagonal, the more consistent the health measure is with IB. The results for all cross-tabulations show that more than 55% of the LSOAs fall into the same quintile. As with the correlation result above, the most similar measure to the 2001 IB remains the PSD with over 73% of LSOAs on the leading diagonal. A Kappa value of 0.67 suggests a strong agreement between the quintiles. Very few LSOAs are more than two quintiles different. The most dispersed measure to the 2001 IB is the NGH census measure with 55% of LSOAs on leading diagonal with 45% of the remaining LSOAs 0.44 agreement between the quintiles.

[Table 3 about here]

Patterns of small area deprivation beyond the census

This section compares regularly used measures of deprivation (Townsend Index and IMD 2010) with a measure constructed from administrative data; an 'Alternative Deprivation Index' (ADI). Similarities in the spatial and temporal patterns identified are examined to determine the utility of using administrative statistics to measure composite socioeconomic distributions in England instead of the census. If administrative statistics are to be used as an alternative way of measuring socioeconomic distributions instead of the census, then we must determine how well a relatively simple administrative-based measure of deprivation emulates existing indexes. The maps in Figure 5 show the LSOA level Townsend deprivation index and the ADI time-series (2001, 2006 and 2010) for England. The LSOAs have been divided into quintiles, the dark red areas represent the most deprived 20% LSOAs in England and the dark blue shades depict the least deprived 20%. The white areas represent LSOAs around the national average deprivation scores in all models.

[Figure 5 about here]

Similar patterns of deprivation are identified by the ADI compared to other indexes though the geographical distribution of the Townsend deprivation appears to be smoother compared with the ADI. The Townsend Index shows London as having higher levels of deprivation compared to the ADI. It has been argued that the inclusion of the non-car ownership indicator in deriving the Townsend index is likely to exaggerate deprivation in London since most households in London are less likely to use private cars compared to other regions in England due to traffic congestion (Buck and Gordon, 2004).

Most urban areas are shown in all models as having higher levels of deprivation compared to rural and suburban locations. The ADI also found several urban conurbations contain many highly deprived LSOAs. These include London, Manchester, Leeds, Liverpool, Birmingham, Bradford, Tendering,

Easington and Middlesbrough, Newcastle and Blackpool. These areas have a common history of heavy industrialization like mining and manufacturing which have been experiencing consistent decline in the past decades (IMD, 2010, Walsh et al, 2010). Areas such as Hart, Greenwich, Oxfordshire, Forest Heath, and Cambridgeshire contain the least deprived LSOAs in England. The range of deprivation models analysed here show patterns of deprivation in England to be complex. The most deprived areas are shown to be surrounded by pockets of less deprived places distributed through all regions in England with some LAs containing the most deprived and least deprived LSOAs in England. A notable example is the North-East Government Office Region which shows wide sub-regional variation in the patterns of deprivation.

The correlation matrix in Table 4 shows that positive relationships exist between the Townsend deprivation index, the IMD 2010 and the time-series ADI. The strength of the relationships between these indexes is shown to wane with increase in years suggesting the effect of time as places change. The steepest gradient across deprivation measures are between IMD and the time-series ADI (2001, 2006, 2010) with significant positive correlations of 0.82, 0.80 and 0.74 compared to a flatter gradient between the Townsend Index and the ADI measures which demonstrates lower associations with correlations of 0.72, 0.65 and 0.58 respectively. The lowest correspondence exists between the Townsend Index and the ADI 2010. The stronger relationship between the IMD and the ADI is understandable since they are both constructed from administrative databases though the ADI is a much simpler approach.

[Table 4 about here]

Tables 5 and 6 reflect the patterns revealed in the correlations analysis above. The tables show the levels of similarity between the indexes by illustrating which LSOAs have changed their ranking from 2001 - 2010 and the magnitude of this change. A greater similarity exists between the quintiles of ADI and the IMD with greater dispersion in the classifications with increase in time. Some LSOAs have experienced change in their levels of deprivation with a small proportion of LSOAs moving up to three quintiles. About 301 of the most deprived LSOAs in 2001 experienced lower deprivation of up to two quintiles in 2010.

[Tables 5 & 6 about here]

Discussion

The study represents one of the first attempts at examining a cross-section of readily available administrative statistics (beyond IB health measures) as an alternative for the census on different aspects of the population and on a timeline in England. The alternative health and deprivation models are spatial datasets in their own right showing socioeconomic patterns which have a wide range of applications and potential for further research. The study has the merit of using LSOAs, a consistent statistical geography which allows changes to be monitored over space and time. This also enables comparison to be made with the IMD 2010. The methods used are fairly straightforward and can

easily be replicated or revised in other studies. The findings of the study confirm national administrative statistics on health and deprivation used in this research to be of high utility given the strong associations between the datasets and equivalent census measures. The results of the bivariate analyses of all self-assessed health measures (LLTI, NGH and PSD) of the 2001 Census demonstrates strong relationships with IB claims, suggesting that the latter can be a useful and objective alternative to census measures of health.

The time-series analyses show a decline in the number of persons claiming benefits. This trend is contrary to the findings of previous studies (Bambra and Norman, 2006; Norman and Bambra, 2007) which have found IB claims to be on the rise. Whether the observed changes in the rate of IB claims as well as other income related benefits are due to 'real change' or as a result of changing regulations and conditions for claiming benefits is not certain. The benefit data relating to lone parents and pensioners are also found to demonstrate strong relationships (0.91 and 0.97 at $p = 0.01$ level) with census equivalents. The lower end of the council tax bands (A and B), is found to have a relatively weaker but moderately positive association with equivalent census counts of households in rented accommodation. This is expected since the two measures are not directly comparable but does show some association with other indicators of socioeconomic disadvantage. Also, the data do not fully comply with the stringent national statistics standards so would not be expected to produce highly reliable standards. IB rates are seen to be strongly related with other measures of deprivation.

The series of health and deprivation models constructed using administrative data-based indicators identified similar spatial patterns compared with regularly used approaches constructed from census data. The study has shown that IB-based standardised illness ratios identify similar patterns with illness ratios derived from census measures of health of LLTI, NGH and especially for those reporting to be 'permanently sick or disabled' (PSD). Though some changes are observed in the spatial and temporal patterns of ill-health between 2001 and 2010, the broad patterns remain similar with higher illness rates observed in urban conurbations compared with more rural locations. The IB has the advantage over the decennial census of regular availability and can thus be used for monitoring short term patterns of heath distributions in England. In line with previous studies, illness rates for Northern England remain higher than the more southern parts with the exception of London. Furthermore, the broad patterns of administrative measures of deprivation are found to be consistent with census-based patterns identified elsewhere (Mclennan et al, 2011). The similarities relate to the concentration of higher levels of disadvantage in LSOAs in urban areas compared with more rural ones. Greater inequality in socioeconomic distributions is also confirmed in urban areas with most cities containing pockets of affluent LSOAs in deprived areas and people living in deprived LSOAs in otherwise affluent neighbourhoods (Stafford and Marmot, 2003, Townsend et al, 1988).

London is revealed to have more socioeconomic advantage than the Townsend and IMD deprivation indexes shown. Patterns of deprivation identified by various indices are determined by the degree of representation of the indicators used in their construction. The non-home ownership and the non-car ownership variables included in the Townsend Index may artificially raise deprivation in some places. Not owning a car or a home might not necessarily equate to socioeconomic deprivation especially in London where young families and professional migrants are more likely to rely on public transport systems and live in rented accommodation compared with more rural locations. Hence, the two indicators might be more reflective of households in particular life stages than of deprivation. It is recommended that future studies explore; in specific terms, the factors accounting for the difference observed in London's deprivation patterns in relation to health. The study confirms the strong relationship between IB health measures and the administrative-based models of socioeconomic conditions of the population. This is consistent with previous findings (e.g. Curtis and Rees Jones, 1998) which clearly show poor health to increase with social disadvantage. Despite data limitations and conceptual flaws associated with deprivation indexes, this study, in addition to previous findings, shows that administrative statistics have good potential for providing up-to-date statistics on a regular basis for health studies. Although data protection barriers and ethical issues surrounding the use of administrative statistics remain a major limitation to health research, the ambition of replacing the traditional census approach with a system that organises existing administrative databases in the UK is worth exploring.

Strengths of the study

The study validates the similar but not identical correspondence between incapacity benefit data and census measures of health. In addition, it examines the relationships between a range of census and administrative statistics on housing and other benefit claims questioned in literature. These include administrative statistics relating to council tax valuations, pensions, lone parents and job seekers allowance which are shown to be strongly related to census equivalents (households in rented accommodation, pensioners, lone parent with dependent children and count of unemployed persons). The study also shows the rate of income related benefit uptake to be a useful indicator of socioeconomic disadvantage given the strong positive relationships between benefit claims and popularly used deprivation indicators. Alternative deprivation models observed socioeconomic conditions in London to be better than indicated by the Townsend index of 2001 or IMD 2010. The study also uncovers temporal variations in health and socioeconomic distributions which suggest that the census becomes less adequate for monitoring short term changes. Greater dispersions are observed in the levels of similarity between pairs of administrative and census-based models with an increase in years. The study represents one of the first attempts at examining a cross-section of readily available administrative statistics (beyond IB health measures) as a complement to the census on different aspects of the population or on a solid timeline in England. The study has the merit of using LSOAs, a

consistent statistical geography which allows changes to be monitored over space and time. This also enables comparisons to be made with the IMD 2010. The methods used are fairly straightforward and can easily be replicated in other studies.

Limitations of the study

This research has various limitations. One relates to the lack of availability of administrative data that are directly comparable with the census. As Dugmore et al (2011) observe, some census topics (such as housing indicators) are not readily available from existing administrative databases. In addition, alternative data sources are not sufficiently developed to provide complete information on population attributes. Administrative statistics used in this research do not represent 100% of the population. The mid-year population estimates which were used as the denominator for calculating rates are directly obtained from the 2001 Census. This might have biased the strength of relationship observed between census variables and equivalent pairs of administrative statistics. Administrative data providers should explore ways of achieving a full coverage of the population in the future and producing outputs in formats compatible with the census. Most datasets are available at the local authority and parliamentary constituency levels with a limited number for small area geographies. Due to confidentiality constraints and ethical issues relating to the access of small area data, a wider range of relevant administrative data on health and deprivation could not be obtained or were rounded to an extent such that data were unusable.

Users of population statistics can lobby about definitional aspects and the inclusion of questions in census and official surveys. This is not the case with administrative data, the specification of which is beyond the influence of users. This study has been heavily reliant on benefit related data. Benefit systems are constantly undergoing welfare reforms and changing regulations which redefine the categories of people who can access these benefits. Therefore, the demographic information derived is subject to any future changes to administrative systems. The inconsistency in data over time reduces reliability and increases uncertainties in the temporal analyses undertaken in this study.

The range of administrative data available for small area geographies and for the time-series, 2001, 2006 and 2010 on the subject matter is also a limitation. Future studies should incorporate a wider range of variables where possible to increase the reliability and versatility of results. The appropriateness of the deprivation variables used and the ways of combining these into scores are subject to question with only a narrow range of deprivation aspects captured. The limitation of adequacy of variables for constructing socioeconomic models results more from lack of availability of relevant administrative statistics which could have been included. Nevertheless, this study, shows the explanatory power of the models in relation to health events to be sufficiently robust.

Acknowledgements

This work used Census data obtained via MIMAS' CASWEB and GIS boundary data obtained via EDINA's UKBORDERS; services supported by ESRC and JISC. These data are Crown copyright and are reproduced with permission of OPSI. Adapted data from the Office for National Statistics and obtained via the Neighbourhood Statistics and NOMIS websites licensed under the Open Government Licence v.2.0 are also used. All data are subject to Crown Copyright.

We are grateful for the feedback from the review process which has helped to improve this paper.

Conflicts of interest

None declared

References

BAMBRA, C. & NORMAN, P. (2006). What is the association between sickness absence, mortality and morbidity? Health & Place, 12, 728-733.

BÉCARES, L., NAZROO, J., ALBOR, C., CHANDOLA, T. & STAFFORD, M. (2011). Examining the differential association between self-rated health and area deprivation among white British and ethnic minority people in England. Social Science & Medicine.

BRADFORD, M. G., ROBSON, B. T. & TYE, R. (1995). Constructing an urban deprivation index: a way of meeting the need for flexibility. Environment & Planning A 27: 519-533.

BUCK, N. & GORDON, I. (2004). Does spatial concentration of disadvantage contribute to social exclusion? City Matters: competitiveness, cohesion and urban governance, 237-253.

CARSTAIRS, V. (1995). Deprivation indices: their interpretation and use in relation to health. Journal of Epidemiology and Community Health, 49, S3-S8.

CARSTAIRS, V. & MORRIS, R. (1990). Deprivation and health in Scotland. Health Bulletin, 48, 162.

CEBR. 2012. Unemployment to keep rising to 2016 across UK, except in the South East, East and London: Five more years of pain ahead as weak private sector job creation and public sector cutbacks push up unemployment. [Online]. London. [Accessed July 2014 via http://www.cebr.com/].

CURTIS, S. & REES JONES, I. (1998). Is there a place for geography in the analysis of health inequality? Sociology of Health & Illness, 20, 645-672.

DEAS, I., ROBSON, B., WONG, C. & BRADFORD, M. (2003). Measuring neighbourhood deprivation: a critique of the Index of Multiple Deprivation. Environment and Planning C, 21, 883-904.

DUGMORE, K., FURNESS, P., LEVENTHAL, B. & MOY, C. (2011). Beyond the 2011 Census in the United Kingdom. International Journal of Market Research, 53(5), 619-650

HARPER, G. & MAYHEW, L. (2012). Applications of population counts based on administrative data at local level. Applied Spatial Analysis and Policy, 5(3), 183-209

IMD (2010). The English Indices of Multiple Deprivation 2010. Neighbourhood Statistical Release, 1 - 21. [Accessed July, 2014]. Available from: http://www.communities.gov.uk/publications/corporate/statistics/indices2010.

JARMAN, B. (1983). Identification of underprivileged areas. British Medical Journal, (Clinical research ed.), 287(6385), 130

MCLENNAN, D., NOBLE, M., BANES, H., DAVIES, J., GARRATT, E. & DIBBEN, C. (2011). The English Indices of Multiple Deprivation 2010. In: GOVERNMENT, C. A. L. (ed.). University of Oxford: [Accessed July, 2014]. Available from: http://www.communities.gov.uk/publications/corporate/statistics/indices2010technicalreport

NOBLE, M., WRIGHT, G., SMITH, G. & DIBBEN, C. (2006). Measuring multiple deprivation at the small-area level. Environment and Planning A, 38, 169-185.

NORDBOTTEN, S. (2010). The Use of Administrative Data in Official Statistics–Past, Present, and Future–With Special Reference to the Nordic Countries. Chapter, 17, 205-223.

NORMAN, P. (2010). Identifying change over time in small area socioeconomic deprivation. Applied Spatial Analysis and Policy, 3, 107-138.

NORMAN, P. & BAMBRA, C. (2007). Incapacity or Unemployment? The Utility of an Administrative Data Source as an Uptadable Indicator of population Health Population , Space and Place, 13, 333-352.

NORMAN P, REES P & BOYLE P (2003) Achieving data compatibility over space and time: creating consistent geographical zones. International Journal of Population Geography. 9(5): 365-386

ONS (2011a). Beyond the 2011 Census Project. Office for National Statistics. [Accessed July, 2014]. Available from: http://www.ons.gov.uk/ons/about-ons/what-we-do/programmes---projects/beyond-2011/index.html.

ONS (2011b). Labour market statistics: September 2011.

SMITH, M. P., OLATUNDE, O. & WHITE, C. (2010). Monitoring inequalities in health expectancies in England–small area analyses from the Census 2001 and General Household Survey 2001–05. Health Statistics Quarterly, 46, 53-70.

STAFFORD, M., DUKE-WILLIAMS, O. & SHELTON, N. (2008). Small area inequalities in health: Are we underestimating them? Social science & medicine, 67, 891-899.

STAFFORD, M. & MARMOT, M. (2003). Neighbourhood deprivation and health: does it affect us all equally? International Journal of Epidemiology, 32, 357.

TAROZZI, A. & DEATON, A. (2009). Using census and survey data to estimate poverty and inequality for small areas. The review of economics and statistics, 91, 773-792.

TOWNSEND, P., PHILLIMORE, P. & BEATTIE, A. (1988). Health and deprivation: inequality and the North, Croom Helm London.

WALSH, D., BENDEL, N., JONES, R. & HANLON, P. (2010). It's not just deprivation': Why do equally deprived UK cities experience different health outcomes? Public health.

WHYNES, D. K. (2009). Deprivation and self-reported health: are there 'Scottish effects' in England and Wales? Journal of Public Health, 31, 147-153.

Table 1 Correlations of equivalent pairs of variables showing declining and weakening relationships for the time-series 2001, 2006 and 2010

VARIABLE	CORRELATIONS		
	2001	2006	2010
Health: LLTI / IB	0.70	0.70	0.69
Unemployment: Unemployment / Job Seekers Allowance Claimant	0.88	0.86	0.84
Housing: Rented Tenure / Council Tax Band A&B	0.55	0.50	0.50
Lone Parent: Lone Parent with Dependants / Lone Parents Claimant	0.91	0.88	0.81
Pensioner: All Pensioner Household / Pension Claimants	0.96	0.89	0.81
Population: Resident Population / Mid-Year Estimates of population	0.98	0.77	0.60

Note: All correlations are significant at p = 0.01. Limiting Long term Illness; Incapacity Benefit (IB)

Table 2 Correlations between incapacity benefit SIRs (2001, 2006, and 2010) and other measures of health for England

	VARIABLE	2001 CENSUS HEALTH MEASURES			ADMINISTRATIVE HEALTH MEASURES		
		LLTI	NGH	PSD	IB-2001	IB-2006	IB-2010
2001 CENSUS	LLTI	1					
	NGH	0.80	1				
	PSD	0.91	0.85	1			
ADMIN. DATA	IB-2001	0.90	0.82	0.93	1		
	IB-2006	0.87	0.82	0.90	0.92	1	
	IB-2010	0.84	0.80	0.87	0.89	0.93	1

Note: All correlations are significant at p = 0.01. Limiting Long Term Illness (LLTI); Incapacity Benefit (IB); Not Good Health (NGH); Permanently Sick or Disabled (PSD); Standardised Illness Ratios (SIRs)

Table 3 Crosstabulations of the quintiles of incapacity benefit standardised illness ratios with other census based health measure.

LLTI-SIR 2001	IB-SIR 2001				
	Q1	Q2	Q3	Q4	Q5
Q1	4836	1469	175	14	2
Q2	1489	3356	1501	140	10
Q3	147	1571	3439	1297	43
Q4	17	92	1358	4099	930
Q5	7	8	24	946	5512

65.40% on leading diagonal

Contingency Coefficient = 0.78, Kappa = 0.57

LLTI-SIR 2001	IB-SIR 2006				
	Q1	Q2	Q3	Q4	Q5
Q1	4625	1572	265	32	2
Q2	1622	3055	1580	230	9
Q3	226	1672	3055	1450	94
Q4	17	187	1540	3626	1126
Q5	6	10	57	1158	5266

60.42% on leading diagonal

Contingency Coefficient = 0.76, Kappa = 0.51

LLTI-SIR 2001	IB-SIR 2010				
	Q1	Q2	Q3	Q4	Q5
Q1	4393	1636	412	49	6
Q2	1743	2757	1636	348	12
Q3	321	1774	2713	1541	148
Q4	30	314	1614	3245	1293
Q5	9	15	122	1313	5038

58.86% on leading diagonal

Contingency Coefficient. = 0.73, Kappa = 0.45

Not Good Health-SIR 2001	IB-SIR 2001				
	Q1	Q2	Q3	Q4	Q5
Q1	4396	1594	406	89	11
Q2	1605	2755	1690	400	46
Q3	405	1663	2651	1590	188
Q4	83	446	1558	3136	1273
Q5	7	38	192	1281	4979

55.16% on leading diagonal

Contingency Coeff. = 0.72, Kappa = 0.44

Permanently Sick or Disabled-SIR 2001	IB-SIR 2001				
	Q1	Q2	Q3	Q4	Q5
Q1	5328	1090	69	6	3
Q2	1127	4122	1179	63	5
Q3	36	1254	4147	1022	38
Q4	4	30	1097	4547	818
Q5	1	0	5	858	5633

73.20% on leading diagonal

Contingency Coefficient = 0.81, Kappa = 0.67

Note. The cells represent the counts of LSOAs in a particular quintile. The highlighted areas represent LSOAs that have been classified into the same quintile by both health measures.

Limiting Long Term Illness (LLTI); Incapacity Benefit (IB); Not Good Health (NGH); Permanently Sick or Disabled (PSD)

Table 4 Correlations between ADI (2001 – 2010) and other deprivation indexes in England

Deprivation Models	Townsend Index 2001	IMD 2010	Alternative Deprivation Index 2001	Alternative Deprivation Index 2006	Alternative Deprivation Index 2010
Townsend Index 2001	1				
IMD 2010	0.88	1			
Alternative Deprivation Index 2001	0.72	0.82	1		
Alternative Deprivation Index 2006	0.65	0.80	0.93	1	
Alternative Deprivation Index 2010	0.58	0.74	0.89	0.94	1

Note: All correlations are significant at 0.01 level. Index of Multiple Deprivation (IMD); Alternative Deprivation Index (ADI)

Table 5 Crosstabulations of LSOAs by quintiles of the Alternative Deprivation Index (ADI) and the Townsend Index in 2001, 2006 and 2010

Townsend Index 2001	IMD 2010				
	Q1	Q2	Q3	Q4	Q5
Q1	4490	2085	416	22	0
Q2	1662	2846	1938	259	2
Q3	305	1300	2866	1755	57
Q4	39	241	1143	3331	1488
Q5	0	24	134	1129	4950

57% on leading diagonal

Townsend Index 2001	Alternative Deprivation Index 2001				
	Q1	Q2	Q3	Q4	Q5
Q1	3624	2388	901	98	2
Q2	1496	2240	2299	656	17
Q3	709	955	1942	2447	230
Q4	486	565	780	2226	2185
Q5	182	348	575	1069	4063

43% on leading diagonal

Townsend Index 2001	Alternative Deprivation Index 2006				
	Q1	Q2	Q3	Q4	Q5
Q1	3392	2357	1084	175	5
Q2	1419	2248	2146	848	46
Q3	751	1014	1866	2290	362
Q4	603	523	808	2094	2214
Q5	331	354	593	1089	3870

42% on leading diagonal

Townsend Index 2001	Alternative Deprivation Index 2010				
	Q1	Q2	Q3	Q4	Q5
Q1	3351	2313	1117	224	8
Q2	1375	2089	2218	952	73
Q3	759	1008	1647	2271	598
Q4	634	578	785	1819	2426
Q5	377	508	730	1230	3392

38% on leading diagonal

Table 6 Crosstabulations of LSOAs by quintiles of the Alternative Deprivation Index (ADI) and the IMD 2010 in 2001, 2006 and 2010

IMD 2010	Alternative Deprivation Index 2001				
	Q1	Q2	Q3	Q4	Q5
Q1	3901	1920	622	52	1
Q2	1677	2422	1903	482	12
Q3	711	1496	2364	1779	147
Q4	194	567	1316	3028	1391
Q5	13	91	292	1155	4946

51% on leading diagonal

IMD 2010	Alternative Deprivation Index 2006				
	Q1	Q2	Q3	Q4	Q5
Q1	3736	2027	675	58	0
Q2	1656	2391	1892	543	14
Q3	801	1410	2280	1814	192
Q4	286	556	1334	2831	1489
Q5	17	112	316	1250	4802

49% on leading diagonal

IMD 2010	Alternative Deprivation Index 2010				
	Q1	Q2	Q3	Q4	Q5
Q1	3684	1892	750	79	1
Q2	1635	2273	1952	614	22
Q3	815	1332	2146	1918	284
Q4	320	716	1177	2536	1747
Q5	42	191	472	1349	4443

46% on leading diagonal

Alternative Deprivation Index 2001	Alternative Deprivation Index 2010				
	Q1	Q2	Q3	Q4	Q5
Q1	4748	1559	183	6	0
Q2	1463	3048	1772	212	1
Q3	248	1518	2977	1666	88
Q4	27	326	1369	3370	1404
Q5	10	45	196	1242	5004

59% on leading diagonal

Figure 1 Temporal variation in the percentage mean of variables obtained from administrative sources: England in 2001, 2006 and 2010

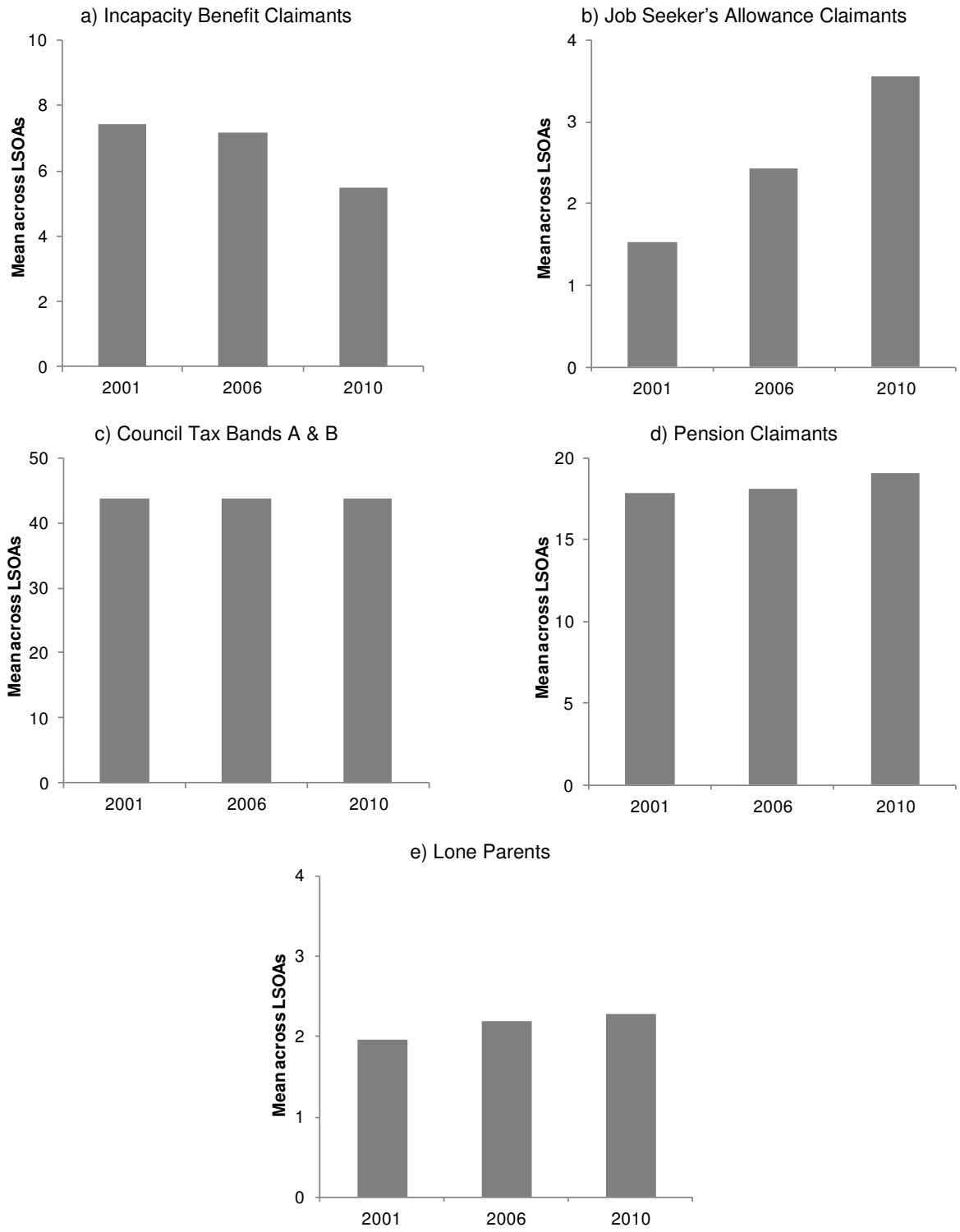
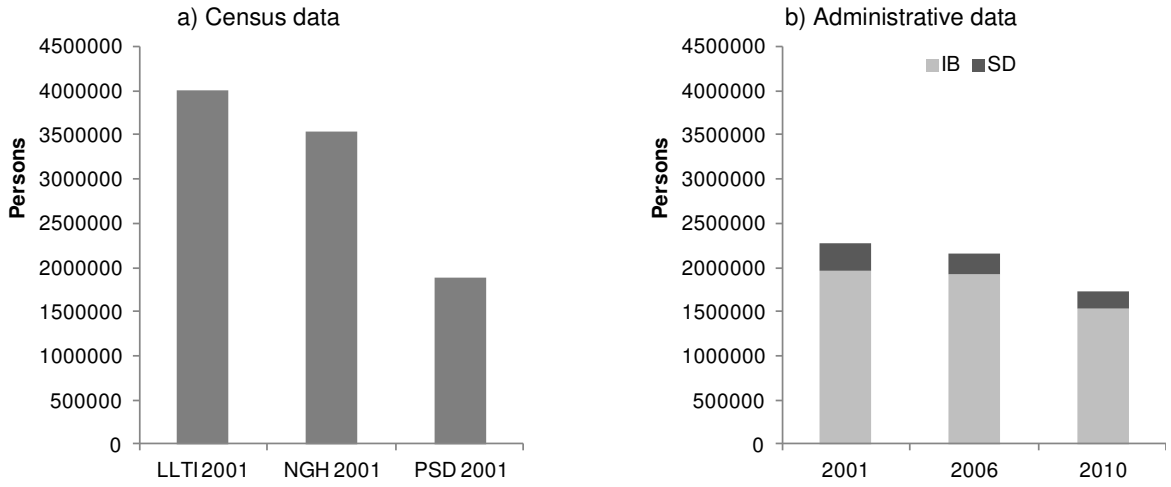
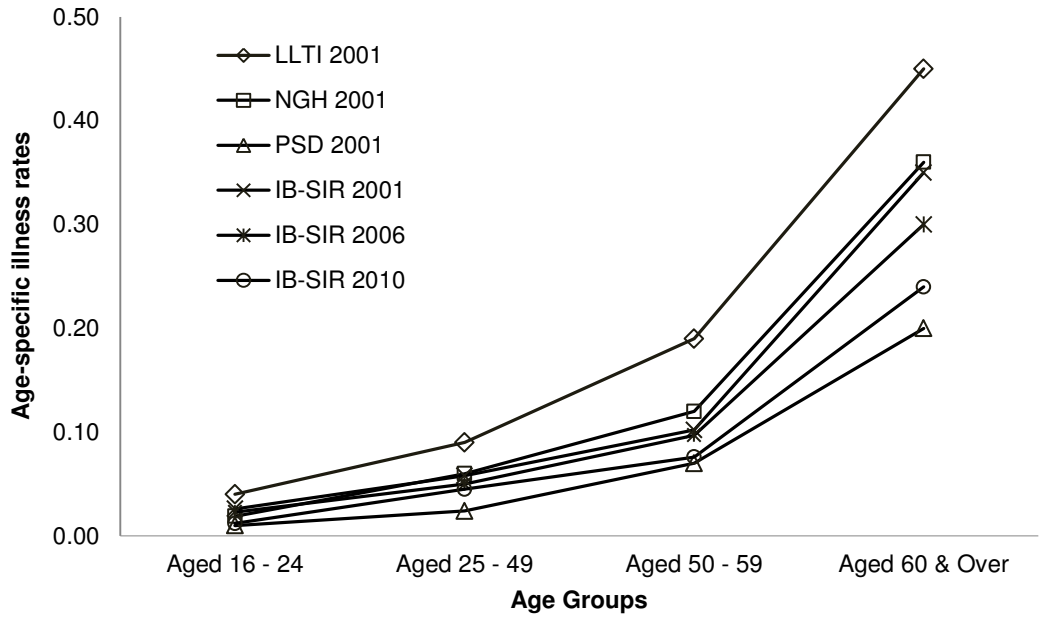


Figure 2 Event counts of census and administrative indicators of health conditions: England in 2001, 2006 and 2010



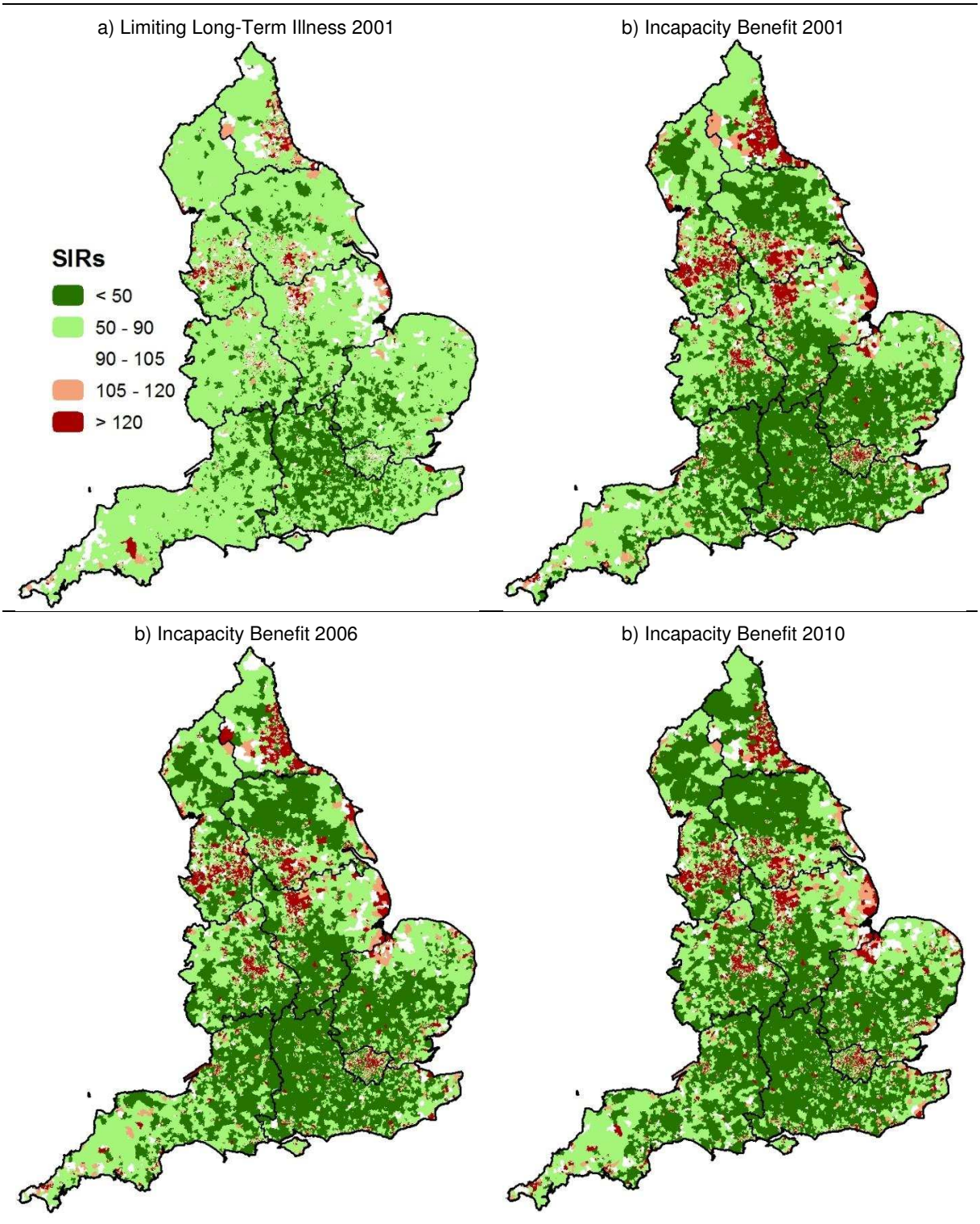
Note: Limiting Long Term Illness (LLTI); Not Good Health (NGH); Permanently Sick or Disabled (PSD); Incapacity Benefit (IB); Severe Disablement Allowance (SD)

Figure 3 Age Specific Illness Rates for census and incapacity benefit sources: England in 2001, 2006 and 2010



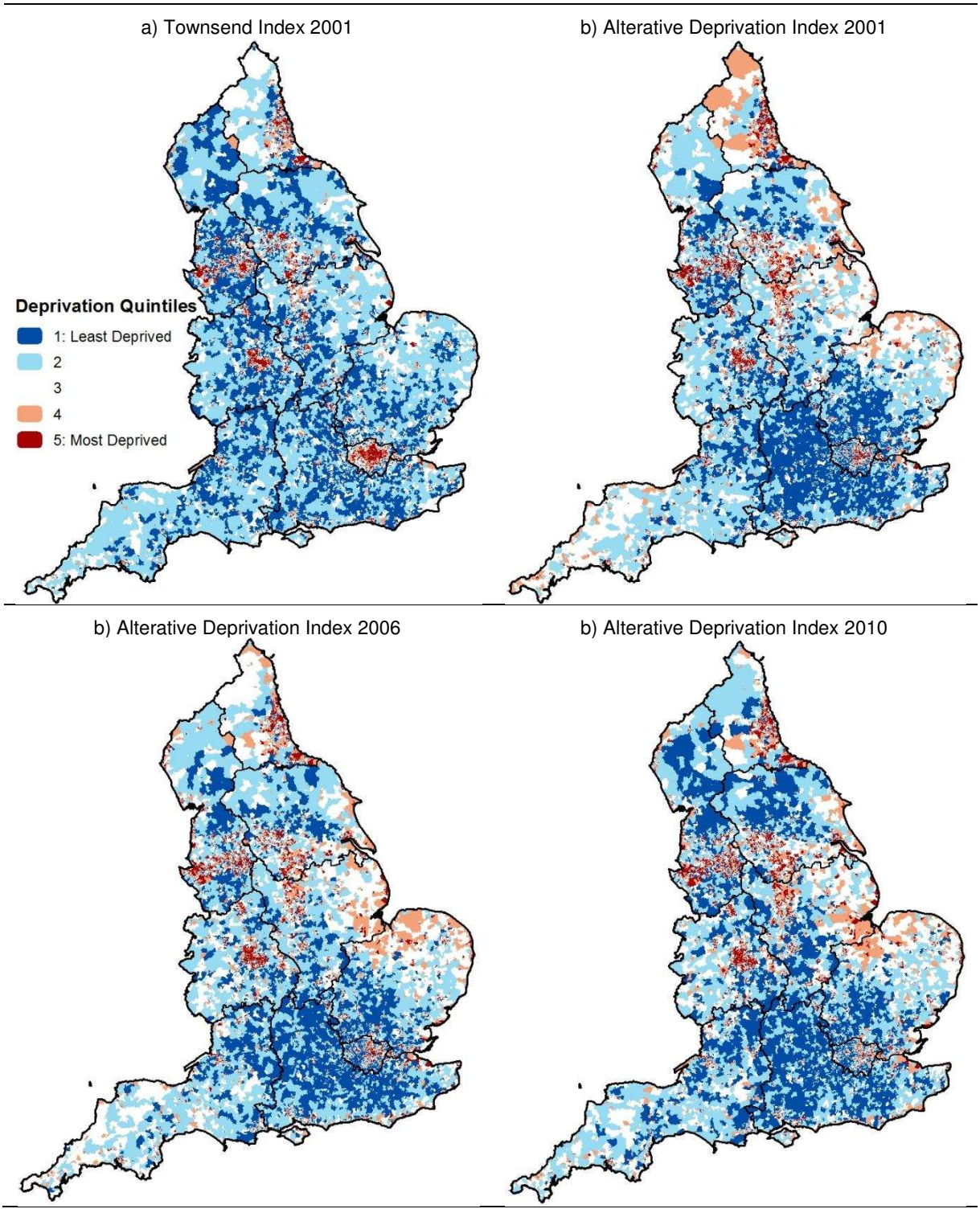
Note: Limiting Long Term Illness (LLTI); Not Good Health (NGH); Permanently Sick or Disabled (PSD); Incapacity Benefit (IB)

Figure 4 Distributions of Standardised illness ratios (SIRs) of LLTI and IB claims by LSOAs: England in 2001, 2006 and 2010



Note: Green areas represent lower than national average illness ratios and red areas indicate higher than national average ratios. White indicates the national average

Figure 5 Distributions of the Townsend Index and Alternative Deprivation Index (ADI) by LSOAs: England in 2001, 2006 and 2010



Note: Blue areas represent least deprived 20% LSOAs and red areas indicate the most deprived 20% LSOAs in England. The White shade indicates LSOAs with near national level deprivation levels