

# A century of inequality in England and Wales using standardized geographical units

Ian N Gregory\*, Daniel Dorling\*\* and Humphrey R Southall\*

\*Department of Geography, University of Portsmouth, Buckingham Building, Lion Terrace, Portsmouth, Hants, PO1 3HE and \*\*School of Geography, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT

Emails: Ian.Gregory@port.ac.uk, D.Dorling@geog.leeds.ac.uk

Revised manuscript received 19 February 2001

*This paper analyses geographical trends in relative poverty in England and Wales over the last century by comparing key quantitative indicators from key dates. The comparison is made possible by interpolating all the datasets onto a single standardized geography. Results suggest rising inequality in spite of the decline in absolute poverty.*

**Key words:** *England and Wales, inequality, GIS, standardized administrative units, long-term change*

## Introduction

This paper is concerned with very long-term trends in poverty and inequality in England and Wales over the last century. It seeks, in particular, to provide a tentative quantitative answer to the question of whether relative poverty has become more or less extreme geographically (Dorling and Woodward 1996). Such a study is only possible through the application of novel methods for data analysis based around the Great Britain Historical Geographical Information System (GBHGIS) (Gregory and Southall 1998), which we use to make data for different dates as closely comparable as possible. The range of variables available for use is inevitably limited because we require statistics that were measured for detailed geographical areas over the course of the century, and which remained both significant and broadly comparable over the whole period. Even so, we believe that both the methodology and the results are sufficiently interesting to justify this paper and its findings. If the three very different indicators of inequality used here all show the same trend towards increased geographical inequality, then it is likely that poverty as a whole is becoming more geographically unequal.

Our analysis is strictly quantitative, using data for the whole of England and Wales interpolated onto a set of standardized spatial areas. This permits us not only to clearly define the geography of relative inequality in different periods, but also to make some broad generalizations concerning long-term trends in relative inequality. In other words, we are concerned with establishing which geographical areas experienced the worst hardship, how much worse off they were relative to the best-off areas, and how this pattern changed over time. Owing to the limitations of early data the areas used are relatively large and crude, being approximately equivalent in scope to modern Travel-to-Work areas. This puts us at the top of a multi-scale approach to the study of poverty and inequality (Shepherd 1998) and should complement more detailed and localized studies.

The debate about how best to define poverty is as old as the study of poverty itself and many different methods have been put forward (see for example Gordon and Pantazis 1997; Lee 1998). Our time scale restricts us to three key indicators, each of which is discussed in more detail in the subsequent sections. Each can be associated with one of William Beveridge's five great evils, identified by him at the mid-point of our study:

- infant mortality with disease;
- overcrowded housing with squalor; and
- numbers in the lowest social class with want (Beveridge 1942, 6; Dorling and Tomaney 1995).

### *Infant mortality*

This is the death rate for children aged under one year, an essentially demographic variable and therefore one calculable for different dates with few problems of consistency. Infant mortality rates are far lower today than they were 100 years ago, but the deaths of young children are still a deep trauma for the families concerned. The rate remains a useful indicator of broader poverty, being particularly influenced by the mother's health and nutrition, and the environment into which the child was born (Staines 1998).

### *Overcrowded housing*

Our data concern the number of persons in each household relative to the number of rooms, information gathered by every census this century. The definition of a household and the method of counting rooms are both problematic, and today we might give greater emphasis to other aspects of poor housing, such as available amenities and environmental issues including condensation (Dorling 1995). However, patterns prove remarkably enduring, and space and privacy are still greatly valued in society.

### *Unskilled workers*

An obvious measure of economic hardship is what proportion of the population is either engaged in an unskilled occupation or, if unemployed, was last engaged in an unskilled occupation. Here we use the Registrar General's Social Classification, and specifically the proportion of people in Social Class V. This is routinely used as an indicator of poverty due to the low wages, economic insecurity and the lack of control over their work that these people endure (Bartley *et al.* 1998). Again, there are difficulties in producing consistent definitions, from the lack of detail provided by the 1911 census to the large numbers of people who had never worked and who were allotted no social class in 1991, but comparable statistics can be constructed.

Comparative indicators of Beveridge's other two 'great evils', ignorance and idleness are harder to construct. Ignorance had to be omitted both because of a lack of geographically disaggregated data on education and because of shifting standards.

Idleness clearly meant unemployment, but insuperable problems led to its exclusion here (see Gregory *et al.* in press).

Ideally, this paper would consider continuous time series but, quite apart from the sheer volume of data that would be required, much of our information comes from the census, carried out only once every ten years. There was no 1941 census, and the collection of many other statistical series was interrupted by the two World Wars. This has meant that our analysis is restricted to four key dates:

- 1 *The turn of the century*: a period of prosperity relative to other nations, but also perceived as having great disparities between rich and poor. We use the 1901 and 1911 censuses and the 1901 Registrar General's Decennial Supplement.
- 2 *The inter-war recession years*: a time which saw unemployment, especially in the North, first entering the political agenda but also the foundations of modern consumer society being laid. We focus on 1928, and the 1931 census.
- 3 *The 'never had it so good' 1950s*: In 1957, Macmillan continued 'Go around the country, go to the industrial towns, go to the farms and you'll see a state of prosperity such as we have never had in my lifetime'. We use 1958 and 1961.
- 4 *The present — booming Britain?*: We mainly use the 1991 census, giving us regular 30-year intervals, with some more recent data.

## **Using GIS to create standardized geographical units**

Our most recent data are available for very small geographical areas: wards, enumeration districts (EDs) or even, in the case of the mortality data, the precise postcode of the individual fatality's home (Dorling 1997). However, for most of the century we are limited to the data that appeared in published reports such as the census, or the Registrar General's Supplements. The types of area used depends largely on date and are as follows:

- 1 Until the First World War *Registration Districts* (RDs), of which there were around 630, were the principal units used by the Registrar General, and hence the census. They were originally defined in terms of towns and their spheres of influence (Lipman 1949).
- 2 From 1911, RDs were replaced by *Local Government Districts* (LGDs) which consisted of county,

municipal and metropolitan boroughs, and urban and rural districts. There were around 1500 LGDs although as noted below this fluctuated considerably. When originally formed LGDs were typically subdivisions of RDs: the urban areas were designated as boroughs or urban districts while the rest of the RD became a rural district. This means that although LGDs provide a much clearer urban/rural dichotomy than the RDs, their original pattern was nested within the RD structure.

- 3 Post-1970 data are frequently published for *Wards* or *Enumeration Districts*. This gives vastly more spatial detail than the other two sets of units: in 1991 there were around 8000 wards and 110 000 EDs in England alone (Coombes 1995). Unit postcodes provide a similar spatial resolution (Raper *et al.* 1992).

Even comparing data for the same system of units at different dates is problematic as a constant trickle of boundary changes took place. For example, in 1901 there were 1834 LGDs but by 1961 there were only 1466, and even the areas which remained in existence throughout often experienced major alterations to their boundaries.

We are able to compare datasets which use very different systems of reporting units by drawing on the GBHGIS. When completed this will contain a full record of the changing boundaries of the statistical reporting units of Britain from parish-level upwards (Gregory and Southall 1998 2000). Work to date has been limited to England and Wales, which is the only reason for the limited geographical focus of the study. When the system is complete, it will be possible to use the simple population counts which are available for the individual parishes to estimate the size and characteristics of the populations transferred by boundary changes (Gregory 2000a 2000b), but for now we assume that the population transferred is proportional to the area transferred (Goodchild and Lam 1980).

The data sources and the areas used to map each data set are shown in Table I. Standardization has two benefits. Most obviously, it allows direct comparison between different dates for a given administrative unit. Perhaps even more importantly, it allows the data to be compared consistently: the value given for any unit is the average value within that unit and therefore, as reporting units become geographically larger, local extremes are increasingly averaged away and the range of values is reduced. This means that data collected for the three different types

of units listed above are not comparable unless standardized onto a single standardized geography.

In this paper we use a simple technique called areal-weighting to interpolate data onto standardized units. This is based on the assumption that the density of a variable across an administrative unit is homogeneous across its entire area. Under this assumption the basic problem can be summarized as the need to redistribute a variable  $Y$  from a set of source zones  $S$  onto a set of target zones  $T$ . Values of  $Y$  for each target zone can then be estimated as:

$$\hat{y}_t = \sum_s \frac{A_{st} y_s}{A_s} \quad (1)$$

where

$A_s$  is the area of the source zone,

$A_{st}$  is the area of the zone of intersection between the source and target zones and  $Y$  is the variable being modelled (Flowerdew and Green 1994).

This can be implemented straightforwardly by using a GIS overlay operation to calculate the areas of the zones of intersection. The process inevitably introduces some error to the results so in an attempt to minimize this 1898 Registration Districts have been used as the target districts as these are the largest units used to publish data and the later units nest into their structure to a reasonable degree. Both of these have been shown to reduce the error introduced by areal interpolation (Fisher and Langford 1995; Sadahiro 2000).

The modern data required a slightly different methodology based on using either a corrected ED centroid or unit postcode centroid to aggregate up to 1981 wards (Dorling 1993). The resulting ward-level data were then re-allocated to the target districts based on the entire value being allocated to the RD which contained the largest part of the ward. This method is crude and can be improved upon when an enhanced GIS becomes available. However, for the time being it is felt that this is acceptable given the large degree of aggregation at each stage, from over 100 000 EDs or unit postcodes to around 9000 wards to 630 target districts.

Once the data are in this standardized form, we can identify the target districts containing the worst-off ten per cent of the population and the best-off ten per cent: our measure of relative inequality is simply the ratio of the average rate of the relevant indicator in the worst-off districts to the average for the best-off. Note that the worst-off targets will by no means contain all the individuals or

**Table 1** Data sources and base maps used

Data set	Source	Collected by	Mapped onto
Infant mortality 1891–1900	The RG's <i>Decennial Supplement</i> 1901	RD	RDs
Infant mortality 1928	The RG's <i>Statistical Review</i> 1928	LGD	1928 LGDs for the North, 1910 LGDs for remainder
Infant mortality 1958	The RG's <i>Statistical Review</i> , 1958	LGD	Approximated 1951 LGDs
Infant mortality 1990–92	Individual death records	Postcodes	1981 wards
Overcrowding 1901	1901 Census county reports	LGD	1901 LGDs for the North, 1910 LGDs for remainder
Overcrowding 1931 (both counts)	1931 Census county reports	LGD	1931 LGDs for the North, 1910 LGDs for remainder
Overcrowding 1961	The 1961 census	LGD	Approximated 1951 LGDs
Overcrowding 1991	1991 Census small area statistics	ED	1981 wards
Social class V 1911	1911 census employment Tables	Boroughs, urban districts over 5000 people, county totals for rural districts	1911 RDs
Social class V 1931	1931 Census employment report	12 regions split into county boroughs, other urban district and rural districts	1931 LGDs for the North, 1910 LGDs for remainder
Social class V 1951	1951 Census county reports	LGD	Approximated 1951 LGDs
Social Class V 1991	1991 Census small area statistics, 10% sample	ED	1981 wards

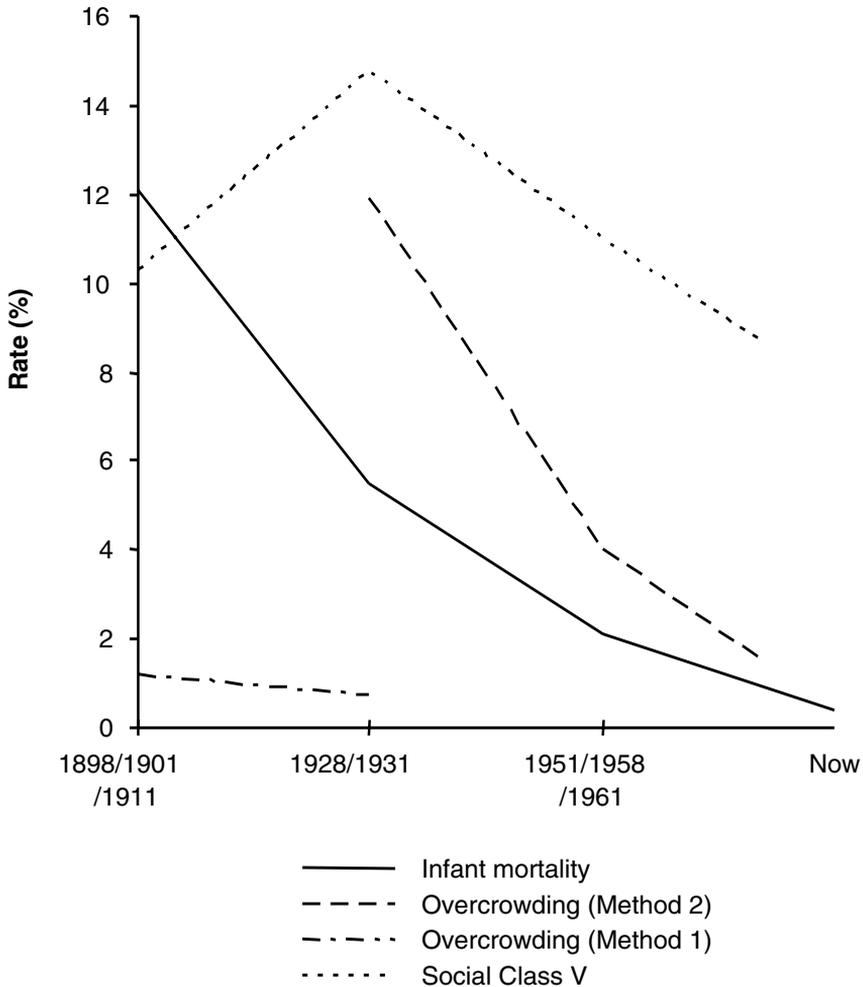
Notes: RG's stands for Registrar General's

'North' in the 'Mapped onto' column refers to the four northern counties of England plus Lancashire and Cheshire, and the East and North Ridings of Yorkshire

households in the worst-off decile of the population as a whole: many, and probably the majority of the people in these districts will not be that badly off, while some people in better-off districts will be among the decline of worst-off individuals. However, so long as the most detailed data available concerns geographical aggregates, not individuals, our focus is necessarily blurred and there are advantages to such an ecological approach in that it allows us to observe ecological processes such as de-industrialization, migration to areas with a better climate or, more abstractly, Thatcherism.

Mapping is a tool that has a long history in research on poverty (Dorling 1998). Booth for example used mapping both for presentational purposes, and to a more limited extent, for interpretative analysis (Booth 1889; Shepherd 1998). The advent of GIS has freed cartography from many of its traditional constraints and it provides us with a range

of visualization techniques (Hearnshaw and Unwin 1994; Dorling 1992). The most familiar is the choropleth map, where geographical areas are shaded according to the intensity of a value or rate recorded across them. However, conventional choropleth maps emphasize often sparsely-populated rural areas while cities can all but disappear, so in the results which follow we also include area cartograms. Cartograms are derived from the conventional map by a process in which the area of each district is made proportional to its population, or some other measure of its significance, while trying as far as possible to keep adjacent units together (Dorling 1994 1996). The result may be unfamiliar, but the shape of the country is broadly retained. One key feature of our descriptive population-based cartograms is that where the units are divided into five bands, or quintiles, indicated by different shades of grey, each quintile covers



**Figure 1** Median rates of infant mortality, overcrowded housing, and social class V calculated at 1898 RD level, turn of the century to the present.

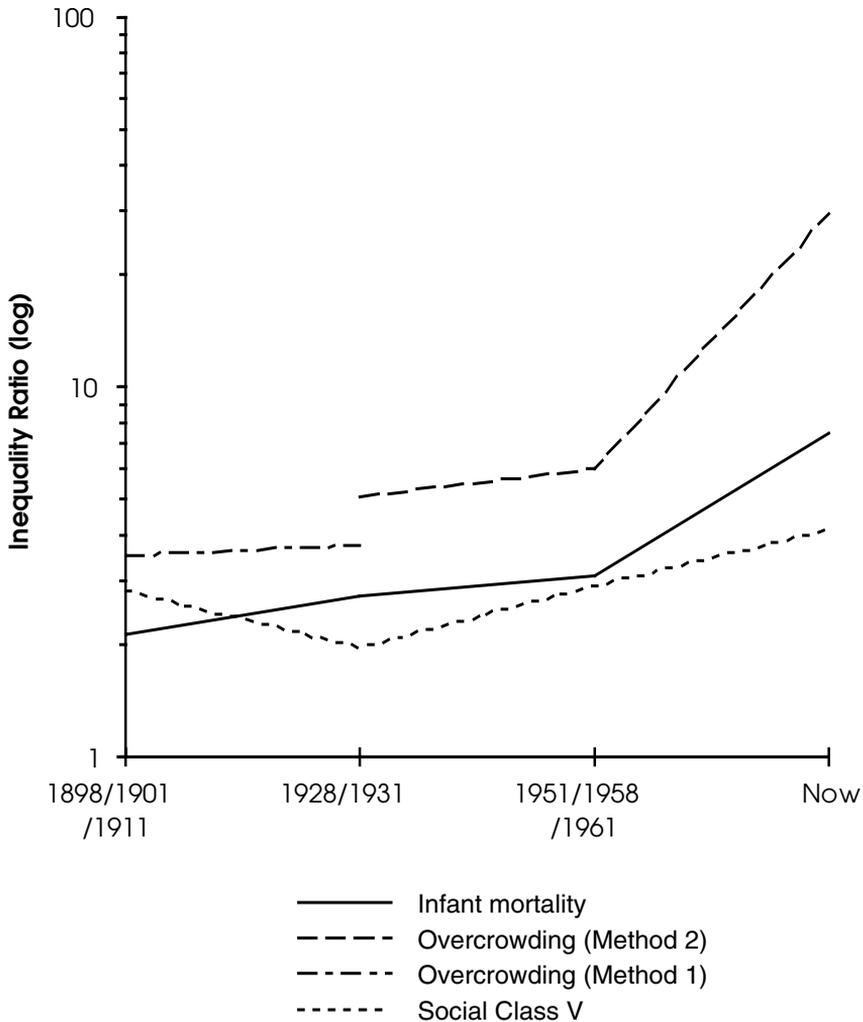
Note: Overcrowding is calculated by two different methods that are not directly comparable, method 1 for 1901–31 is based on households living in four rooms or less, method 2 for 1931–91 is based on the entire population

one-fifth of the total population and therefore there will be an equal area of each shade in each map. Conversely, the conventional maps often contain quite small areas of the darkest shade because the worst-off were concentrated in major cities.

For the remainder of this paper we discuss both the data sources and results for each variable separately. The conclusion then draws the themes together. Figure 1 shows the change in median rates for the three variables and Figure 2 gives the changing inequality ratios. Table 2 summarizes the key quantitative results.

### Infant mortality

Infant mortality data permit the most straightforward comparisons over time as definitions have changed little. Our earliest data come from the 1901 *Registrar General's Decennial Supplement*, listing deaths per 1000 births for ten years from 1891 to 1900, collected at RD-level. For 1928 and 1958 the data were taken from the *Registrar General's Statistical Reviews* for the individual year in question, again expressed as deaths per 1000 births. These were collected for LGDs. The 1990s' data are slightly different in that



**Figure 2** Inequality ratios for infant mortality, overcrowded housing, and social class V calculated at 1898 RD level, turn of the century to the present

Note: Overcrowding is calculated by two different methods that are not directly comparable, method 1 for 1901-31 is based on households living in four rooms or less, method 2 for 1931-91 is based on the entire population

they are based on actual death records for the period 1990 to 1992, using the fatality's postcode as a spatial reference (Shaw *et al.* 1998). Three years were used because modern rates are so low that small numbers can become a major statistical problem (Davey Smith and Dorling 1996 1997). The rate is expressed as deaths per 1000 babies, based on the 1991 census and corrected both for normal under-reporting (Pattie *et al.* 1996) and by using the national count of birth certificates to allow for the extra under-reporting typical of census counts of

babies. The reason the number of babies rather than births is used in the 1990s is to remove the impact of people migrating in the year that their baby is born, thus skewing the rates. This has become important as migration around the time of birth has become more common in the 1990s and families are now much smaller.

The century saw a massive fall in infant mortality, the rate for the median RD falling from 121 per thousand in the 1890s to four per thousand in the 1990s. The geographical pattern in the 1890s is very

**Table 2** Relative poverty in England and Wales, 1898 to the present

Variable	Descriptive measures				Best 10%		Worst 10%		Ratio
	Mean	Median	St Dev.	Skew	Cut-off	Mean	Cut-off	Mean	
Infant mortality 1890s	127.0	121.0	28.6	0.58	109.0	95.6	186.0	204.0	2.14
Infant mortality 1928	56.0	54.9	16.9	0.20	45.3	34.3	86.5	93.3	2.72
Infant mortality 1958	21.1	21.1	7.40	0.02	16.3	10.7	28.0	33.1	3.10
Infant mortality 1990s	3.86	3.73	2.65	0.14	2.39	1.04	5.59	7.86	7.57
Overcrowding 1901 <sup>1</sup>	1.31	1.19	0.50	0.73	1.01	0.82	2.49	2.93	3.53
Overcrowding 1931 <sup>1</sup>	0.82	0.73	0.34	0.82	0.58	0.48	1.68	1.79	3.73
Overcrowding 1931 <sup>2</sup>	14.5	11.9	7.97	0.97	9.3	8.0	31.2	40.6	5.07
Overcrowding 1961 <sup>2</sup>	4.63	4.01	2.57	0.71	2.81	2.16	9.69	13.01	6.01
Overcrowding 1991 <sup>2</sup>	0.85	0.43	2.08	0.43	0.28	0.20	2.00	5.78	29.4
Class V 1911	10.6	10.3	2.96	0.36	7.78	6.84	17.0	19.2	2.80
Class V 1931	14.7	14.8	2.52	-0.12	12.2	11.3	20.0	22.3	1.97
Class V 1951	11.8	11.0	3.90	0.62	8.72	7.48	17.8	21.6	2.88
Class V 1991	8.43	7.60	3.93	0.65	5.08	4.30	13.8	17.7	4.12

Note: The 'cut-off' is the rate above or below which the top or bottom districts containing 10% of the population are found. The two 'means' to the right of the table refer to areas falling respectively above and below these cut-offs. The 'ratio' is the mean for the worst 10% divided by the mean for the best 10%. The measure of skewness is calculated by subtracting the median, multiplying this by three and dividing by the standard deviation. Overcrowding is calculated by two different methods that are not directly comparable. Method 1 for 1901-31<sup>1</sup> is based on households living in four rooms or less, method 2<sup>2</sup> for 1931-91 is based on the entire population.

distinct: high rates in the industrial districts of South Wales, the West Midlands and the North, especially Lancashire and the West Riding. Tyneside was relatively healthy, and in London the worst conditions were in a small number of central districts. Rural areas were almost uniformly much healthier. By 1928, Lancashire and the West Riding had experienced significant improvement and some rural districts had fallen back in the ranking, particularly in parts of Wales and East Anglia. It seems that conditions in urban areas were improved by large-scale public health measures, while rural areas were beginning to be penalized by poorer medical care. The two more recent dates are more difficult to summarize except on a regional scale. The north-south, or rather core-and-periphery divide is particularly clear in 1958, with almost uniformly low relative rates in the South East, and, as Figure 3 shows, was still strongly present in c.1990. The urban-rural inequalities continue to increase to the detriment of the rural areas and indeed many rural areas now have among the highest death rates.

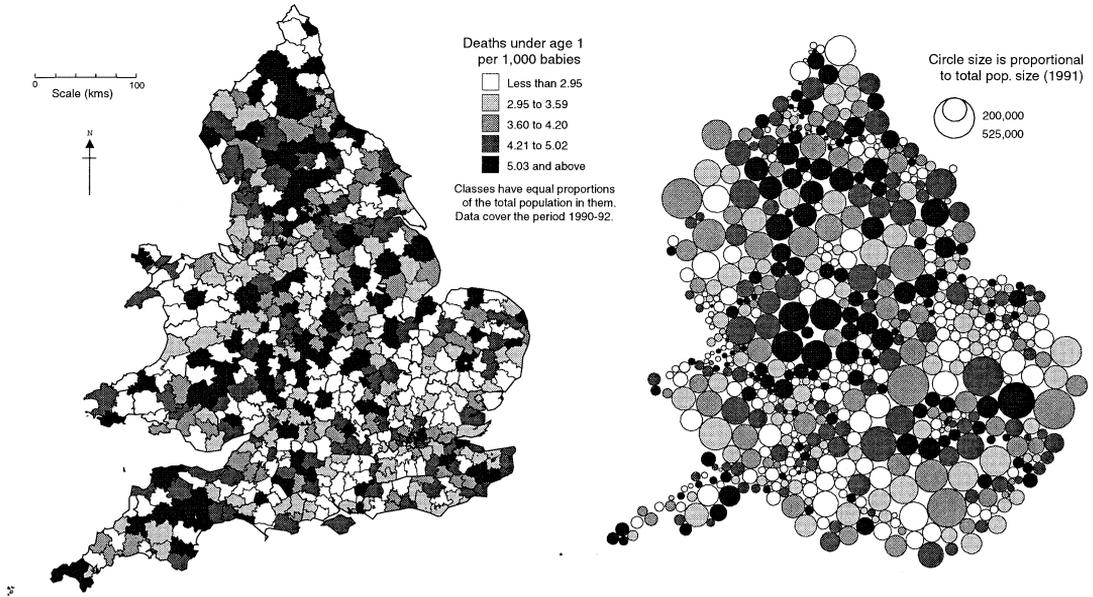
Figure 4 summarizes this by mapping the change in rank of areas between the districts over the 100-year period. The areas that have improved their

ranking significantly, by which we mean a gain of two or more quintiles, are predominantly in the traditional industrial areas and inner London. The areas that have got significantly worse are a broad spread of rural areas without any clear geographical pattern. One exception to the pattern is the city of Birmingham, whose high rates in the 1890s have persisted throughout.

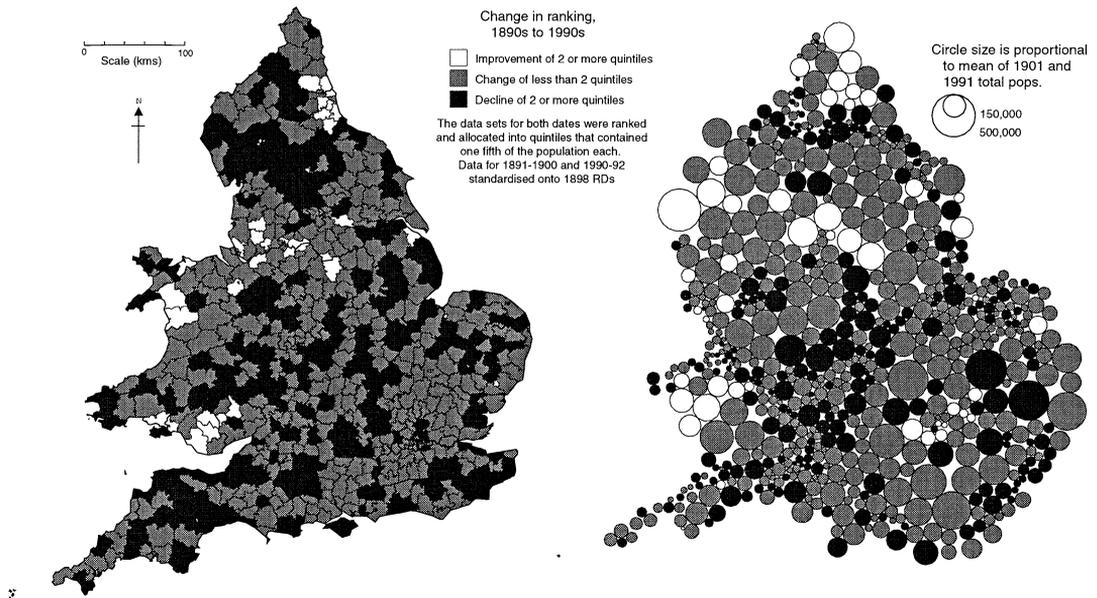
Perhaps the most interesting finding, however, is the trend in the inequality ratio. In the 1890s, the worst-off target districts had only twice as high a mortality rate as the best-off ones, but by the 1990s they had 7.5 times the rate of death. The increase has been sustained over all three of our sub-periods but has been particularly marked in the period since 1958, when the ratio was still only just over three.

### Overcrowded housing

This section examines the changing geography of housing via census statistics of overcrowding. Overcrowding is usually summarized in terms of persons per room: in the nineteenth century, overcrowding meant rates of over two persons per room, but now over one person per room is seen as unacceptable.



**Figure 3** Infant mortality, present on 1898 RDs



**Figure 4** Relative change in infant mortality, turn of the century to the present

In all fields of social statistics such changes in levels of acceptability can be seen (Dorling and Simpson 1998). We have, therefore, compromised and defined overcrowding as over 1.5 persons per room, the same standard as used by the 1931 census. For

the three later dates we were able to calculate the percentage of each district's population living at over 1.5 persons per room. For 1901 however the available data did not permit us to calculate this rate, the report tabulating numbers of persons only for

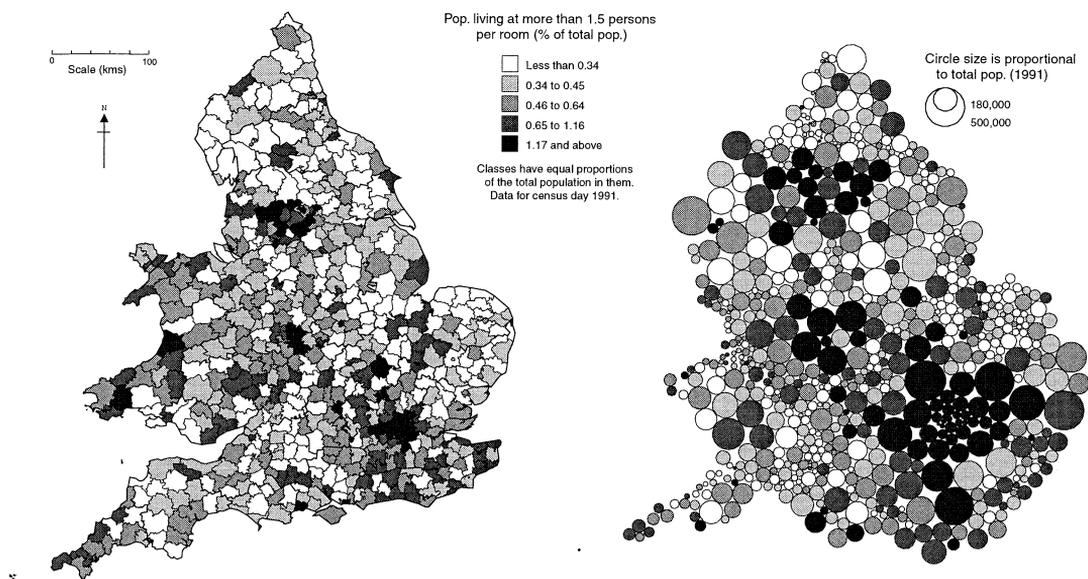


Figure 5 Overcrowding, present on 1898 RDs

'tenements' of four rooms or less; the *General Report* of the 1901 census [Cd. 2174, 39] describes tenements as 'separate occupations', or in other words a dwelling with a separate front door. From these data we could calculate the number of persons living at over 1.5 persons per room in these small dwellings, and as we had no population figures available this total had to be used to calculate a rate per tenement. The 1931 census allowed us to calculate a similar measure, based on taking the number of 'private families' living in four rooms or less and working out the density of occupation of these in terms of persons per room. This is slightly different to the 1901 definition because of the possibility of more than one family living in a single tenement: in 1931, where more than one family was enumerated within a 'structurally separate dwelling', each was classified in terms of the rooms the family occupied and not the total number of rooms in the dwelling. From this we could calculate the number of families occupying less than five rooms and this is given as a ratio of the number of families. We therefore include two sets of results for 1931, the first permitting the closest possible comparison with 1901, the second with 1961 and 1991.

The overall rates of overcrowding appear to have fallen even more rapidly than those for infant mortality; the median rate of people living at over 1.5

persons per room being nearly 12 per cent in 1931 and falling to about 0.5 per cent by 1991. The 1901 data suggests that rates were falling in the first 30 years of the period as well. The most striking feature of the geographical pattern is the concentration of the worst housing in both urban and rural areas of North East England. This remains little changed from 1901 to 1961, but disappears completely by 1991. It has been suggested that these poor conditions were an extension of the very poor conditions that prevailed over the border in Scotland, and partly reflected the nature of the housing stock, with the distinctive and small 'Tyneside flat'; certainly, in 1901 a much higher proportion of all 'tenements' were of four rooms or less than in any other region. The transformation since 1961 indicated the impact of specific policies, and in particular the urban renewal and slum clearance programmes in Tyneside associated with T. Dan Smith. Another interesting feature is the surprisingly low rates of over-crowding in the North West and Yorkshire, with the exception of Liverpool, while the West Midlands conurbation and London had large areas of overcrowding at their cores.

Compared to the notable geographical changes with infant mortality, however, the spatial distribution of overcrowded housing remains fairly constant over the period apart from the notable exception of the

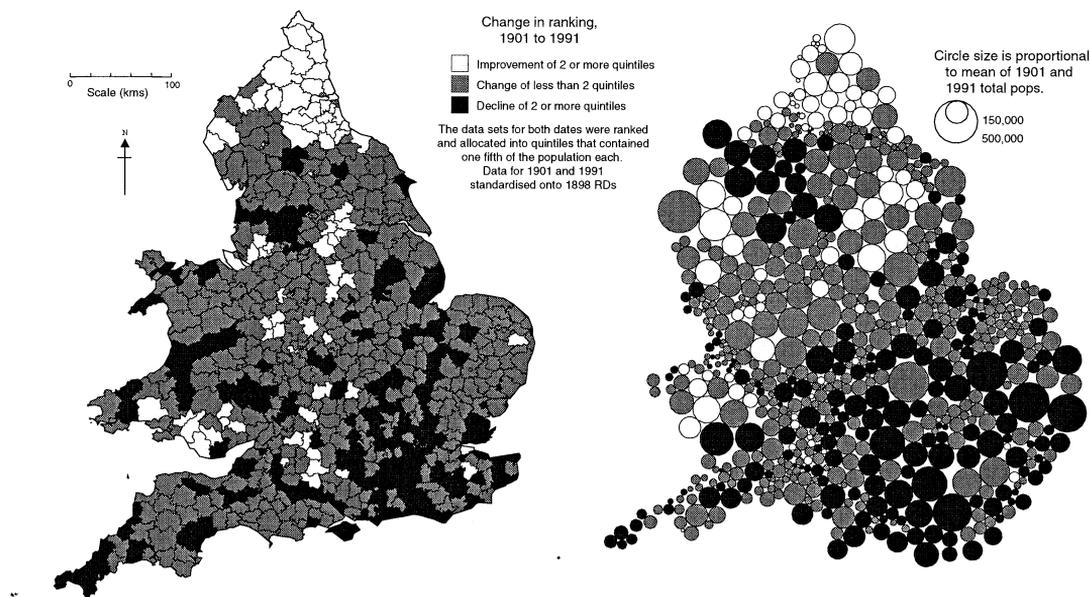


Figure 6 Change in overcrowding, turn of the century to the present

improvements in the North East. Figure 5 shows the pattern that existed in 1991 and Figure 6 shows how the patterns have changed over time: as well as the North East getting significantly better the South East, and particularly the London commuter areas have got significantly worse. Like infant mortality, this map also suggests that urban areas have improved relative to the rural areas. The inequality ratio again rises so that by 1991 the worst-off districts contained nearly 30 times the proportion of people living in overcrowded conditions as did the best-off areas. This trend is very similar to infant mortality in that the trend has been consistently towards greater inequality and the majority of the increase has been in the last 30 years.

### Unskilled employment

Social class V, unskilled workers, is perhaps the most difficult of our variables to define consistently. In our attempts to do this we have had to make several compromises. For other variables our census or decennial supplement data have been for 1901, 1931, 1961 and 1991, but owing to limitations with the data for 1901 and 1961 we have had to use 1911 and 1951, respectively, instead. The 1991 data were taken from the ten per cent sample of the census *Small Area Statistics*. The basic census data

are for economically active males in social class V, and to these we have added males who had not worked in the last ten years and those on a government training scheme as this makes the data more comparable with earlier years (Dorling and Woodward 1996). The 1951 data were taken from the County Reports and include the number of occupied and retired males in social class V.

Earlier censuses of course included a question on people's occupations but did not classify the responses in terms of social class and, in general, we have to estimate class structure from tabulations primarily concerned with industrial structure. However, 1931 census social class statistics are tabulated in the *1931 Registrar General's Decennial Supplement Part IIa, Occupational Mortality*: Table 12 divided the country into 12 regions: Greater London, the remainder of the South East, four regions for the North, two for the Midlands, the East, the South West, and two for Wales. Each of these regions was subdivided into county boroughs, other urban districts, and rural districts with the number of males aged 16 and up in social class V given for each of these classes. We used 1931 LGD populations from the census reports to reallocate the data from the 33 areas covered in the *Decennial Supplement* (not all regions included all three types of areas) to LGDs, and thence to RDs. Unfortunately, this technique

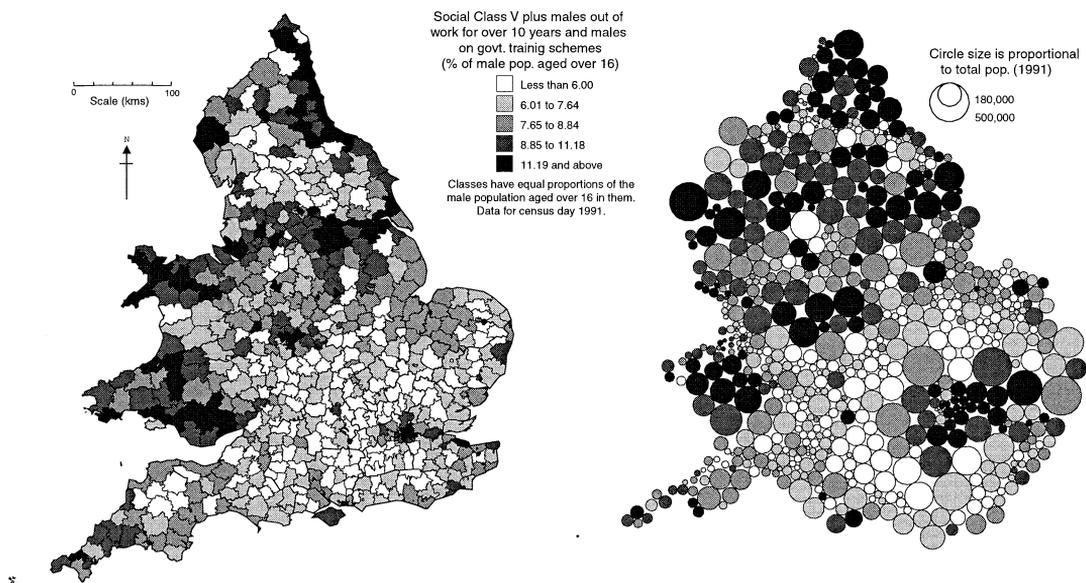


Figure 7 Social class V, present on 1898 RDs

inevitably reduces the range of values, and hence lowers the inequality ratio.

The only way to estimate class structure pre-First World War is from industrial statistics. The 1901 census tabulated different industries for each county, but for 1911 we do have a reasonably detailed set of occupational statistics which we could assign to social classes (Armstrong 1972). These are available for all urban LGDs with a population of over 5000, and for the aggregates of all other urban districts and, separately, of all rural districts in each county. The Tables classify the male population aged ten and above into 42 occupational categories. A few, such as 'Conveyance of men, goods, and messages: dock labourers, wharf labourers, coal heavers, coal-porters, and labourers', can be assigned as a whole to Social Class V. For the remainder, we estimated the proportion in Class V from the more detailed national statistics, which list some 400 categories. For example, the national Tables break 'Metals, machines, implements, and conveyances: general engineering and machine making' down into 14 sub-categories, of which only 'undefined labourers in engineering works' belongs to Class V, representing only c. 5 per cent of the general engineering category.

One final point to note is the variation in the minimum age covered by these tabulations: in 1911

they include boys as young as ten, in 1931 it is 14 and by 1991 16. However, this reflects changing ages of starting work and we have made no attempts to adjust for this variation.

The overall proportion of occupied males in Class V rose slightly from 1911 to 1931, which may reflect the limitations of our data, but then halved between 1931 and 1991: a significant drop, but smaller than for our other two variables. The geographical patterns for social class V in 1951 and 1991, as shown in Figure 7, are broadly similar with higher rates being concentrated in London and the old industrial areas, and lower values in rural areas particularly in the South East. In 1931, allowing for the particularly poor quality of our data, broadly the same patterns apply. The data for 1911 present an intriguing pattern: the highest values were in London and particularly the East End; almost all of southern England had higher rates than the Midlands or the North. This is reflected in Figure 8 which shows areas in the rural south in particular as having improved significantly since before the First World War, while Wales, the West Midlands, western parts of Norfolk, Nottinghamshire, Derbyshire, and southern Yorkshire, and what are now County Durham and West Cumbria have got worse. This arguably reflects major changes in the industrial bases of different areas, the northern areas losing the staple industries which employed

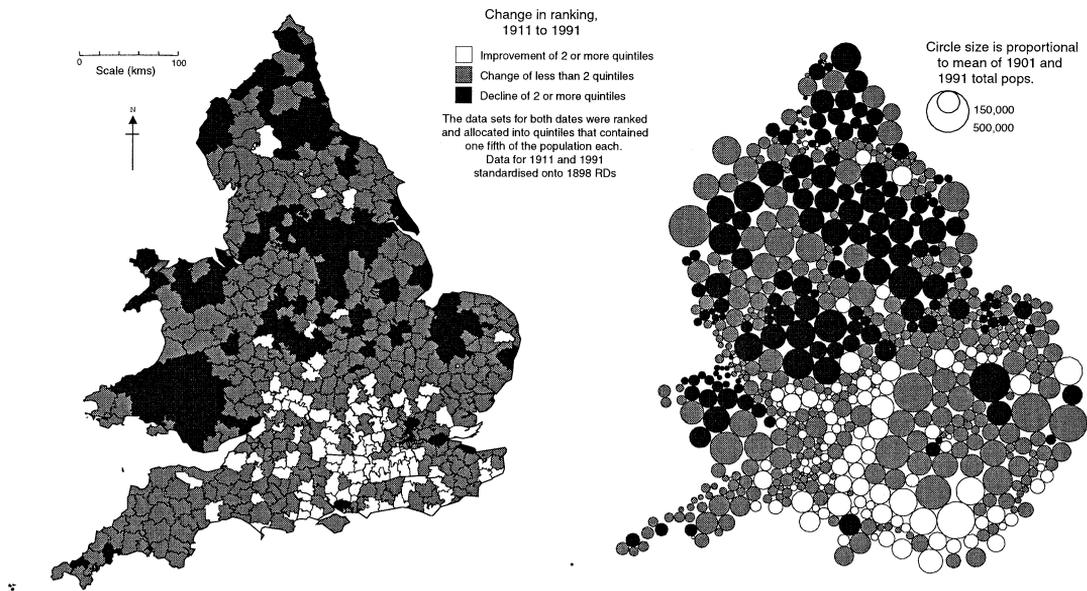


Figure 8 Change in social class V, turn of the century to the present

large numbers of skilled and semi-skilled workers (coal miners, for example, are in Class IV), while rural southern areas were colonized by white-collar commuters.

The inequality ratio for Social Class V tells a broadly similar story to our other measures: up to 1951, the worst-off areas had about three times as many workers in class V as the best-off. Here again, the 1990s' figure show increased inequality, in this case to over four.

## Conclusion

Our analysis clearly shows that the three variables all follow broadly similar patterns in terms of median rates, inequality ratios, and their geographical pattern. With the exception of the early years for social class V, where the data must be treated with particular caution, the median values of all three variables have dropped considerably over the century. Infant mortality has fallen to only three per cent of its 1890s rate, while overcrowding has dropped to six per cent of the 1931 rate.

The geographical patterns, again excluding the problematic early social class data, show a marked persistence of higher poverty in the periphery: the North, but also South Wales and often the remoter parts of the South West. In general, the

major conurbations are worse off than adjacent rural areas, although in the mid-twentieth century significant areas of rural deprivation are apparent. London generally contained substantial deprived areas relative to the affluent South East, but even so the worst conditions were to be found further north. There is, however, evidence to suggest that, while the urban areas may contain the highest rates, the relative unhealthiness of rural areas has increased substantially. This will in part be due to improvements in environmental conditions in the major urban areas but may also reflect the greater problems of improving access to health care facilities, especially maternity units, in rural areas. Figures 9 and 10 compare the population quintiles of infant mortality and overcrowding at the turn of the century and the present respectively. At the turn of the century the two patterns are broadly similar and where there were differences these tended to be found in broadly identifiable areas such as east Lancashire, where infant mortality is significantly worse than overcrowding and the rural North East where overcrowding is worse than infant mortality. For the present, Figure 10 shows that apart for the clear excess of overcrowding in and around London, the pattern is far more complex with differences being both more common and more scattered.

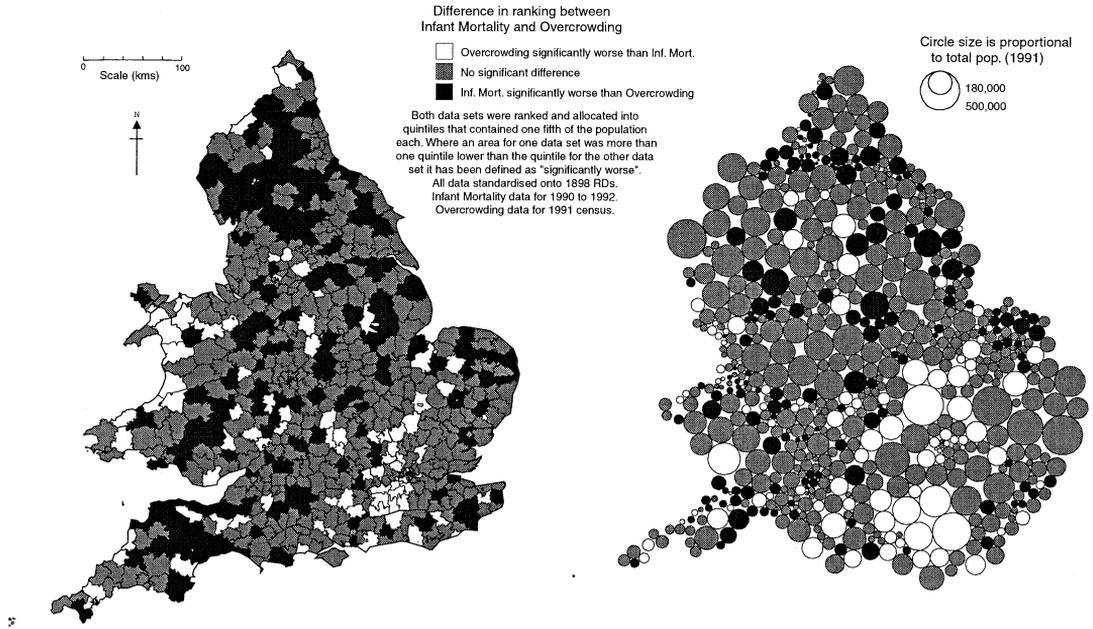


Figure 9 Difference between infant mortality and overcrowding, present

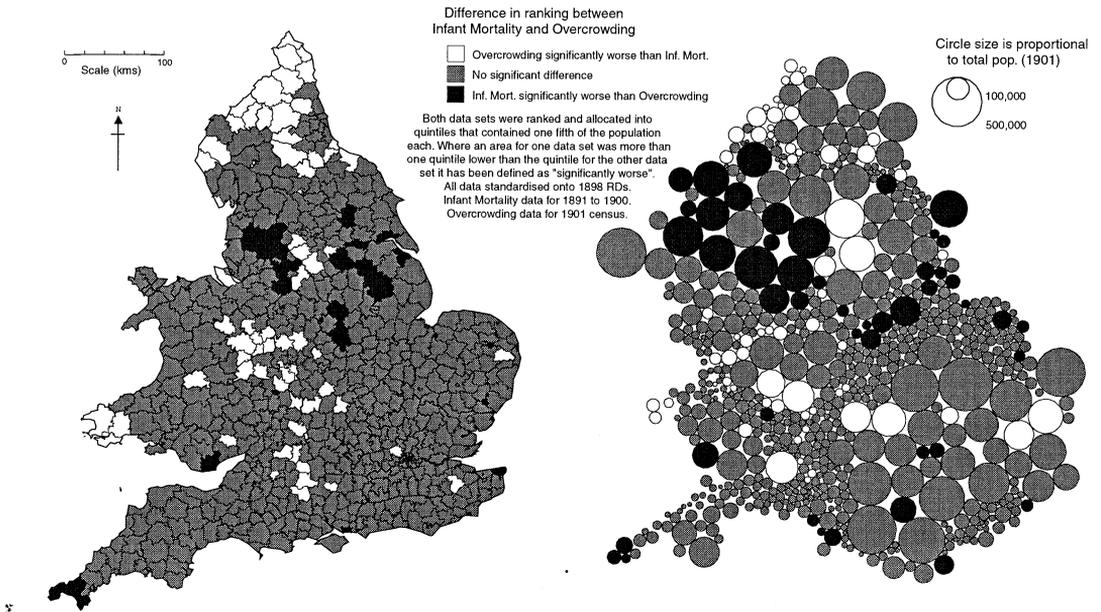


Figure 10 Difference between infant mortality and overcrowding, turn of the century

This paper is in many ways a provisional example of the ways in which an historical GIS can be used to analyse long-term change. Its key advance is the use of standardized geographical units to allow long-term comparison at a spatially detailed level. More work is required to improve the accuracy of this but

this is currently underway (Gregory 2000b). The potential for comparing over 100 years worth of data in this way is, however, large and should open up new understandings of the way the present has developed from the past.

Substantively, our index of relative inequality is inevitably crude but we believe that it is reasonably robust. In particular, it must be emphasized that there is no statistical reason why lower absolute levels of our measures of poverty should inevitably lead to higher levels of relative inequality; that this seems to have happened says something about our society. However, our results may in part reflect the particular scale of analysis we have used. While our analysis is at sub-county level, over much of the country we cannot distinguish country from town, and if at the start of our period social segregation was primarily at the micro-scale, one street or neighbourhood versus another, our calculations will largely ignore it. It can be argued that over the century longer distance commuting and spatial divisions of labour have steadily enlarged the scale at which social segregation occurs, making it more visible to this study. However, as our cartograms show, much of the population was and is concentrated into the large urban centres where Registration Districts cover relatively small and homogenous areas, so this problem should not be exaggerated.

Over the century absolute measures of poverty in England and Wales have obviously declined dramatically. At the same period, however, inequality — relative poverty — has grown and has been particularly marked over the last 30 years. One issue for further research is how much of this increase has happened since the coming of Thatcherism in 1979. The spatial pattern of inequality across England and Wales has become more complicated as hardship is no longer concentrated in a few key, predominantly urban areas. Instead, the problems of some rural areas in particular seem to have increased in their relative importance over the century. We believe, therefore, that despite its current limitations this paper provides significant evidence that over the last century an overall improvement in living standards has been accompanied by a steady rise in relative deprivation.

### Acknowledgements

The construction of the historical GIS which made this research possible was funded by the ESRC (grants R000221314, R000221703, R000237506 and

R000237757), the Leverhulme Trust and many other bodies. Data entry was funded by the Nuffield Foundation and the JISC Knowledge Gallery programme. We are deeply grateful to Eileen Longland for entering the data.

### References

- Armstrong W A** 1972 The use of information about occupation in **Wrigley E A** ed *Nineteenth century society: essays in the use of quantitative methods for the study of social data* Cambridge University Press, Cambridge 191–310
- Bartley M, Davey Smith G and Blane D** 1998 Making sense of health inequality statistics in **Dorling D and Simpson S** eds *Statistics in society* Arnold, London 234–43
- Beveridge W H** 1942 *Social insurance and allied services* HMSO, London
- Booth C** 1889 *Labour and life of the London people, volume I: East London with a coloured map* William and Norgate, London
- Coombes M** 1995 Dealing with census geography: principles, practices and possibilities in **Openshaw S** ed *Census users' handbook* GeoInformation International, Cambridge 111–32
- Davey Smith G and Dorling D** 1996 'I'm all right John': voting patterns and mortality in England and Wales 1981–92 *British Medical Journal* 313 1573–77
- 1997 Association between voting and mortality remains, refereed letter *British Medical Journal* 315 430–1
- Dorling D** 1992 Visualising people in space and time *Environment and Planning B* 19 613–37
- 1993 Map design for census mapping *The Cartographic Journal* 30 167–83
- 1994 Cartograms for visualising human geography in **Hearnshaw H M and Unwin D J** eds *Visualisation in GIS* John Wiley, Chichester 85–102
- 1995 *A new social atlas of Britain* John Wiley, London
- 1996 *Area cartograms: their use and creation. Concepts and techniques in modern geography series 59* University of East Anglia Environmental Publications, Norwich
- 1997 *Death in Britain: how local mortality rates have changed: 1950s–1990s* Joseph Rowntree Foundation, York
- 1998 Human geography—when it is good to map *Environment and Planning A* 30 277–88
- Dorling D and Simpson S** eds 1998 *Statistics in society, the arithmetic of politics* Arnold, London
- Dorling D and Tomaney J** 1995 Poverty in the old industrial regions: a comparative view in **Philo C** ed *Off the map: the social geography of poverty in the UK* Child Poverty Action Group, London 103–22
- Dorling D and Woodward R** 1996 Social polarisation 1971 to 1991: a micro-geographical analysis of Britain monograph in the *Progress in Planning* series 45 1–67
- Fisher P F and Langford M** 1995 Modeling the errors in areal interpolation between zonal systems by Monte Carlo simulation *Environment and Planning A* 27 211–24

- Flowerdew R and Green M** 1994 Areal interpolation and types of data in **Fotheringham A S and Rogerson P A** eds *Spatial analysis and GIS* Taylor and Francis, London 121–46
- Goodchild M F and Lam N S-N** 1980 Areal interpolation: A variant of the traditional spatial problem *Geo-Processing* 1 297–312
- Gordon D and Pantazis C** eds 1997 *Breadline Britain in the 1990s* Ashgate, Aldershot
- Gregory I N** 2000a in press Longitudinal analysis of age and gender specific migration patterns in England and Wales: a GIS-based approach *Social Science History*
- 2000b The use of areal interpolation to explore long-term demographic change *GeoComputation 2000: proceedings of the fifth international conference on GeoComputation* GeoComputation CD-ROM, Manchester
- Gregory I N and Southall H R** 1998 Putting the past in its place: the Great Britain historical GIS in **Carver S** ed *Innovations in GIS 5* Taylor and Francis, London 210–21
- 2000 Spatial frameworks for historical censuses—the Great Britain historical GIS in **Hall P K, McCaa R and Thorvaldsen G** eds *Handbook of historical microdata for population research* Minnesota Population Center, Minneapolis 319–33
- Gregory I N, Southall H R and Dorling D** in press A century of poverty in England and Wales, 1898–1998: a geographical analysis in **Bradshaw J and Sainsbury R D** eds *Getting the measure of poverty. The early legacy of Seebohm Rowntree* Policy Press, Bath
- Hearnshaw H M and Unwin D J** eds 1994 *Visualisation in GIS* John Wiley, Chichester
- Lee P** 1998 Where are the deprived? Measuring deprivation in cities and regions in **Dorling D and Simpson S** eds *Statistics in society* Arnold, London 172–80
- Lipman V D** 1949 *Local government areas 1834–1945* Basil Blackwell, Oxford
- Pattie C, Dorling D, Johnston R J and Rossiter D** 1996 Electoral registration, population mobility and the democratic franchise: the geography of postal voters, overseas voters and missing voters in Great Britain *International Journal of Population Geography* 2 239–59
- Raper J, Rhind D W and Sheperd J W** 1992 *Postcodes: the new geography* Longman, Harlow
- Sadahiro Y** 2000 Accuracy of count data transferred through the areal weighting interpolation method *International Journal of Geographical Information Science* 14 25–50
- Shaw M, Dorling D and Brimblecombe N** 1998 Changing the map: Health in Britain 1951–1991 *Sociology of Health and Illness* 20 694–709
- Shepherd I D H** 1998 Mapping the poor: The micro-geography of poverty in late-Victorian London Paper presented to the Rowntree Poverty Conference at the University of York, 18–20 March 1998
- Staines A** 1998 Poverty and health in **Dorling D and Simpson S** eds *Statistics in society* Arnold, London 244–51