



Area (2004) 36.2, 101–110

# Editorial: the 2001 UK census: remarkable resource or bygone legacy of the 'pencil and paper era'?

Paul Boyle\* and Danny Dorling\*\*

\*School of Geography & Geosciences, University of St Andrews, St Andrews KY16 9AL

Email: P.Boyle@st-andrews.ac.uk

\*\*Department of Geography, University of Sheffield, Sheffield S10 2TN

Revised manuscript received 15 January 2004

*National censuses are expensive. They are conducted infrequently. They collect information that some feel infringes their human rights, and people are required by law to complete them. The outputs are not perfect, and in some situations may be misleading. Some suggest that censuses hark back to a period when regularly collected administrative data were not available. These are some of the views held about national censuses. Why, then, would others argue that they are an essential resource? In this paper, we consider some of the pros and cons of conducting national censuses, before introducing a series of papers that draw on early data available from the 2001 UK census. We argue that these papers, and the wealth of research that will be conducted in the future with 2001 census data, make a strong case for supporting the compulsory collection of personal information about the 'entire' population every ten years.*

**Key words:** 2001 UK census, one number census, human rights, administrative data

## Introduction

Censuses date back to ancient times. The word 'census' is Latin and the Romans carried out a number of censuses (remember Joseph, Mary, a birth event, and a return migration!) The earliest 'modern' censuses were conducted at the beginning of the eighteenth century in Scandinavian countries and the first American census was taken in 1790. The first British census was held in 1801 and they have been conducted every ten years since that time, except for 1941 when the Second World War intervened, and 1966 when a quinquennial census was conducted. The snapshot of the 'entire' national population that a census provides is a unique opportunity to gather information about demographic, social, economic and housing characteristics and how they vary across small geographical areas.

The value of such a comprehensive data collection exercise is often taken for granted by those academics and applied practitioners who take advantage, frequently on a daily basis, of the detailed outputs that result. For example, in the absence of detailed socio-economic information on patients, epidemiological and public health researchers often rely on categorizing patient outcomes by the deprivation score of the small area in which they reside (Boyle *et al.*, this issue); these scores are most commonly derived from decennial census data. Retailers regularly conduct research on brand popularity, or attempt to target customers with particular characteristics, using geodemographic classification systems that, again, are mainly derived from census data (the data that are the most robust of all the information resources they use). Planners rely on small area population estimates, which are validated every ten



years from census information, to make decisions about the distribution of finite resources throughout their catchment areas. And academics from a range of disciplines regularly make use of the complex array of information provided in sources such as the Sample of Anonymised Records (SAR), or the Office for National Statistics England and Wales Longitudinal Study (LS), which provide a range of variables at the individual level for a large random sample of the population. The list of uses made of census information is virtually endless and for those who make use of this resource its value cannot be underestimated.

However, there are others who would argue that national censuses are part of a bygone age, providing information that can be collected in more efficient and cheaper ways. Many countries, including some in Western Europe and Scandinavia, have rejected compulsory national censuses in favour of utilizing routinely collected, linked administrative data and the UK census offices are now actively examining whether future censuses should be abandoned in favour of such an approach. If data on individuals already exists, surely it is more cost-efficient to invest in linking this information together? And, if some information is not available from these standard resources, could rolling surveys be used to replace the national census?

Given this debate, it is an apt time for this special issue. This series of papers provides some of the first results from the 2001 UK census, a small sample of what a conventional census can be used for. These studies take advantage of the range of demographic, economic, social, cultural and health related questions that, we argue, could not be examined reliably using other administrative data. The papers were first presented at the Royal Geographical Society/Institute for British Geographers annual conference held in London in September 2003. Because some census data still remain to be released at the time of writing this editorial introduction (January 2004), the studies were inevitably constrained to differing degrees by the lack of available data. In particular, no interaction data (migration and commuting flows) have been released to date. Even so, the breadth of material is impressive and, we argue, the studies demonstrate that the decennial census is a uniquely valuable resource that we should be wary of abandoning.

### **The 2001 UK census: a vital resource?**

A census is essential. It is the only time when data are collected nationally at a very local level. This means

that they can be used to allocate resources to a wide variety of geographies. The census data also provide the base from which population numbers can be estimated for the 10 intercensal years. The census is certainly the biggest and probably the most important data collection exercise carried out in the United Kingdom, as it is in most countries around the world. (Diamond 1999, 9)

While expensive, and criticized for their imperfections, many social scientists agree that national censuses are a unique resource, the importance of which is easy to underestimate. Diamond (1999) identifies a number of reasons for conducting censuses. First, they provide detailed information about the entire population (except for a small percentage who fail to complete a form), which is released for small geographical areas. No surprise, then, that geographers are probably the largest single group of social scientists in the UK who specialize in census data handling, provision and analysis. Second, and consequently, they provide a valuable tool for resource allocation – the Department of Environment, Transport and the Regions allocate resources to local authority districts partly on the basis of population estimates derived from the census. In fact, the census is the most influential data component in the allocation procedure. Third, censuses provide denominator information from which birth, death and disease rates can be calculated (Hennell, this issue) and which are vital for our understanding of local area variations in health inequalities (Boyle *et al.* 2004). Fourth, no other data source provides reliable information about migration and commuting flows between small geographical areas, an important element in town and country planning. Fifth, census data are increasingly being used by commercial organizations for the creation of geodemographic profiles and customer targeting despite the private sectors' access to vast sources of individual consumer data. Sixth, the fact that censuses have been conducted for 200 years, and because so many of the questions remain unchanged for long periods of time, means that the changing circumstances of people in small geographical areas can be charted. Finally, although their consistency is a particular strength, they do provide an opportunity to collect nationally comprehensive data about socio-economic questions that become particularly pertinent. Hence in 2001 a small number of new variables were added to the form, although these did not include the question academics most wanted to be added: income (Dorling 1999).

The 2001 UK census, conducted on 29 April, was, we believe, amongst the most advanced national censuses ever attempted worldwide. As with previous censuses, the overall aim was to provide reliable, consistent estimates of the number of people, and the characteristics of those people, in each and every locality in the country. Such information cannot currently be obtained from other sources for such small areas, which is a crucial point for geographers especially. Indeed, the 2001 data were disseminated for 175 434 Output Areas spread across the UK and this was by far the greatest level of geographical detail provided in any census; in 1991 data were disseminated for only 113 465 Output Areas (called Enumeration Districts in England and Wales). However, the most advanced aspect of this census was the capture-recapture imputation procedure used to estimate the characteristics of households and people who did not complete the initial census survey (Brown *et al.* 1999). This procedure negated the need in 2001 for academics to carry out the work to correct the census (Martin *et al.* 2002; Mitchell *et al.* 2002). The census authorities also had access to both methods and the information necessary to conduct a much better job of imputation than independent researchers could have done.

Leaving the task of correcting data aside, the task of gathering this information was, of course, enormous and began many years before the census was conducted; not surprisingly, deliberations over the next UK census in 2011, should there be one, are well underway. Thus, as Cook (this issue) explains in detail, the 2001 census involved: a lengthy process of consultation prior to the development of the new census form; the recruitment of 70 000 enumerators working on the ground to help coordinate the form collection; the introduction of a new post-back system of form return which achieved an impressive 88 per cent success rate prior to further follow up; a final response rate of 94 per cent; a massive post-census coverage survey, compared to those conducted in previous censuses; the implementation of the innovative new 'one number census' approach for dealing with missing information promptly; the complete coding of 100 per cent of questions for the entire population, rather than for a sample as in previous years; a revolutionary 'census access programme' designed to make the data freely available to all, removing the need for the bureaucratic systems of charging that existed in previous censuses; the release of data in record time, with most of the area-based statistics

being provided within 30 months of the completion of the data collection exercise.

In addition, new questions were added to the census on general health, carers and religion (Dixie and Dorling 2002), although there were some variations between the constituent countries of the UK in how these (and some of the original questions) were handled. For example, the new question(s) on religion was asked differently in Scotland as compared to England and Wales and differently again in Northern Ireland (see Figure 1). Unlike the other questions on the form, this question was not compulsory. Also, information was gathered in Scotland on the journey to school by children, which was not collected in England and Wales.

The variety of uses that census data have been put to, and could be put to, are remarkable. Most research that uses census data does not explicitly record that source (for instance, in medical research) because its availability is taken for granted. However, a recent search of academic journal articles recorded in the ISI web of science, where the source was explicitly acknowledged, using the key words '2001 UK census' found papers on a number of topics including, for example: defining areas for neighbourhood regeneration programmes (Harris and Johnston 2003); methodological implications of the 2001 census for marketing purposes (Leventhal 2003); how the Samples of Anonymised Records (SARs) and Customised Output from the 2001 UK population census will be of use to central and local government planning, especially in housing related work (Gardiner 2001); and the unique methodology used for the One Number census procedure (Brown *et al.* 1999; Steele *et al.* 2002). It would appear that there is powerful evidence for the salience of census data in academic and applied research but, despite the obvious benefits that censuses provide, there are some who question the need for such an expensive data collection exercise, particularly when the quality of the output can be questioned as it failed to capture 100 per cent of the population in 2001.

### The 2001 UK census: an expensive imposition?

Britain's census could be scrapped after the government agreed with an influential committee of MPs which concluded that the exercise belongs in the 'pencil and paper era'. A review of the census, which was first held in the year of the 1801 Act of Union,

England and Wales	Scotland	Northern Ireland
<p><b>10</b> What is your religion?</p> <p>◇ This question is voluntary ◇ ✓ <i>one box only.</i></p> <ul style="list-style-type: none"> <li><input type="checkbox"/> None</li> <li><input type="checkbox"/> Christian (including Church of England, Catholic, Protestant and all other Christian denominations)</li> <li><input type="checkbox"/> Buddhist</li> <li><input type="checkbox"/> Hindu</li> <li><input type="checkbox"/> Jewish</li> <li><input type="checkbox"/> Muslim</li> <li><input type="checkbox"/> Sikh</li> <li><input type="checkbox"/> Any other religion, <i>please write in</i></li> </ul> <div style="border: 1px solid black; width: 100px; height: 15px; margin-left: 20px;"></div>	<p><b>13</b> What religion, religious denomination or body do you belong to?</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> None</li> <li><input type="checkbox"/> Church of Scotland</li> <li><input type="checkbox"/> Roman Catholic</li> <li><input type="checkbox"/> Other Christian, <i>please write in</i></li> </ul> <div style="border: 1px solid black; width: 100px; height: 15px; margin-left: 20px;"></div> <ul style="list-style-type: none"> <li><input type="checkbox"/> Buddhist</li> <li><input type="checkbox"/> Hindu</li> <li><input type="checkbox"/> Muslim</li> <li><input type="checkbox"/> Jewish</li> <li><input type="checkbox"/> Sikh</li> <li><input type="checkbox"/> Another Religion, <i>please write in</i></li> </ul> <div style="border: 1px solid black; width: 100px; height: 15px; margin-left: 20px;"></div> <p><b>14</b> What religion, religious denomination or body were you brought up in?</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> None</li> <li><input type="checkbox"/> Church of Scotland</li> <li><input type="checkbox"/> Roman Catholic</li> <li><input type="checkbox"/> Other Christian, <i>please write in</i></li> </ul> <div style="border: 1px solid black; width: 100px; height: 15px; margin-left: 20px;"></div> <ul style="list-style-type: none"> <li><input type="checkbox"/> Buddhist</li> <li><input type="checkbox"/> Hindu</li> <li><input type="checkbox"/> Muslim</li> <li><input type="checkbox"/> Jewish</li> <li><input type="checkbox"/> Sikh</li> <li><input type="checkbox"/> Another Religion, <i>please write in</i></li> </ul> <div style="border: 1px solid black; width: 100px; height: 15px; margin-left: 20px;"></div>	<p><b>8</b> Do you regard yourself as belonging to any particular religion?</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Yes ▶ Go to <b>8a</b></li> <li><input type="checkbox"/> No ▶ Go to <b>8b</b></li> </ul> <p><b>8a</b> What religion, religious denomination or body do you belong to?</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Roman Catholic</li> <li><input type="checkbox"/> Presbyterian Church in Ireland</li> <li><input type="checkbox"/> Church of Ireland</li> <li><input type="checkbox"/> Methodist Church in Ireland</li> <li><input type="checkbox"/> Other, <i>please write in</i></li> </ul> <div style="border: 1px solid black; width: 100px; height: 15px; margin-left: 20px;"></div> <p style="text-align: right;">▶ Go to <b>9</b></p> <p><b>8b</b> What religion, religious denomination or body were you brought up in?</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Roman Catholic</li> <li><input type="checkbox"/> Presbyterian Church in Ireland</li> <li><input type="checkbox"/> Church of Ireland</li> <li><input type="checkbox"/> Methodist Church in Ireland</li> <li><input type="checkbox"/> Other, <i>please write in</i></li> </ul> <div style="border: 1px solid black; width: 100px; height: 15px; margin-left: 20px;"></div> <ul style="list-style-type: none"> <li><input type="checkbox"/> None</li> </ul>

**Figure 1** The 2001 UK census question on religion in England and Wales, Scotland and Northern Ireland

will examine whether the £254m exercise could be replaced by greater use of existing computer records. (Watt 2002)

There is no doubt that national censuses are expensive. The 1991 UK census cost approximately £140 million and the most recent estimates suggest that the 2001 UK census cost £259 million (Cook, this issue). It is therefore quite correct that the value of this exercise should be questioned at regular intervals. As Simpson (2003) points out, the quality of a research tool such as the census can be assessed in many ways. It should have a defensible, documented design; it should be conducted rigorously; it should provide credible and verified results; and

it should contribute new knowledge about the UK that cannot be derived from other data sources. The 2001 UK census has been attacked on all these fronts, and others too. The need for a census also has to be questioned given the uncosted burden of the time taken by the population to complete it, the invasion of their privacy and even more basic criteria such as how it is possible to conduct an exercise every decade that relies on many tens of thousands of people being available to work for short periods on very low wages to carry out the survey.

Of course, the nature of such a large exercise means that inefficiencies become press-worthy events. The decision in 2001 to reduce the size of the field force used to collect census forms by one third compared

to 1991, and rely instead on a post-back system, was a new departure. Once the forms had been returned by post the enumerators were expected to focus on those households that had not returned their forms, thus concentrating their resources on areas that were traditionally hard to count. However, despite the fact that the new post-back method achieved a surprisingly high response rate of 88 per cent, the postal service failed to return the forms quickly enough for the census managers to know who had sent them back before the next stage of the process was initiated. This inevitably attracted considerable negative publicity and was a major practical problem in the census process (Cook, this issue). The fact that the envelope provided to post the form back was a little wider than many post boxes could easily accept is also worth remembering in the future!

The type of information that is collected is also often criticized. Most obviously, the lack of a question on income has frustrated many academic and applied practitioners, especially as the question has been regularly included in other national censuses such as in the US. Of course, a careful assessment of the implications of including such a question was conducted, and it was felt that the reliability of the returns may have been influenced if an income question was added, as the UK population is apparently less willing to share information on their income than those in other nations. Even so, it seems inevitable that the demand for an income question will be even more vocal in 2011, assuming there is another census, and academics will continue to question the assumption that the population is unwilling to provide this information (Dorling 1999).

Some of the variables that were included in the census have also been questioned as being of less use than others or of being unnecessarily omitted. Brown and Boyle (2000) note that while neither the 1990 US census nor the 1991 UK census included questions on sexuality, the 1990 US census provided information about same-sex couples derived from information collected on relationships within households. However, similar information was not released from the 1991 British census and, in fact, if a household appeared to be a 'genuine' same-sex couple, the policy was to recode the non-head of household as an 'unrelated' member of that household. Of course, this decision was taken, it was said, because of a scientific concern about reliability rather than as a direct exercise of heterosexism or homophobia (it may even have been taken in the belief that

suppressing information on their geographical concentration was in the interest of same-sex couples), but it is interesting that same-sex couples were retained in the US census and they have also been retained in the 2001 UK census output.

Other variables have also attracted some critical attention. For example, in the 1991 British census, the addition of a question on ethnicity was broadly welcomed, but it has since been criticized for failing to distinguish between race and culture and ignoring the category of 'Black British', which many second-generation black people used to describe themselves. Although the question was improved for the 2001 UK census, capturing such a flexible and situational construct with a single question is difficult, as the concepts invoked by the term include culture, heritage, language, community, religion, nationality, lineage, geographical origin and shared symbolic characteristics such as foods and clothing (Ahmad 1999). The concerns raise the question of whether collecting information on ethnicity is at all worthwhile (Ahmad and Sheldon 1993).

Various other concerns have been raised about the census, including the reliability of the data capture. While the census captured the vast majority of people in the UK on census night, it inevitably missed some people who were not a random sub-set of the national population. The initial results, based on comparisons with post-census coverage survey, suggest that the average undercount was higher in 2001 than in 1991; that areas with a poor response rate in 1991 had an even poorer response rate in 2001; that the geographical spread of non-response was wider in 2001 and that while those that were missed in 2001 had generally similar characteristics to those who were missed in 1991 (young male adults, especially those who were unemployed, private renters or recent migrants), a broader cross-section of the population was missed in 2001 (Simpson 2003).

The One Number census was, of course, designed to impute missing people, and the missing characteristics of those who were captured, but who failed to complete the whole form. However, the results have been questioned, with some Local Government Authorities being particularly vocal in their dismay at the final estimates (Figure 2). Inevitably the results are questioned mainly by authorities who are set to lose most resources as a result of the new census and not by those surprised to find their share of national resources will increase. The census authorities have succumbed to some of this criticism,

## Anger over census 'bungle'

**Patrick Wintour, chief political correspondent**  
**Wednesday January 8, 2003**  
**The Guardian**

Westminster council yesterday served notice that it would seek an order from Gordon Brown directing the office of national statistics to rerun the census in the borough, after it claimed the survey had been bungled.

**Figure 2** Anger over census 'bungle'

awarding more people to some of those authorities who complained most bitterly, but they did not increase the estimate by many and certainly the number missing was tiny as compared to 1991 (when the 'missing million' was a significant problem).

In Westminster the estimated population from the census of 181 000 was 65 000 less than that estimated in 2000 and it was claimed that this would cause a reduction in the Council's annual government grant of about £6 million. Other urban boroughs were also concerned by the census count, and there has been a recent decision by the Office for National Statistics (ONS) to revise the count upwards in Manchester by 24 500 as it was pointed out that the address list used by the ONS to identify those that should have been included in the census exercise was two years out of date in this district and that approximately 14 000 addresses were missed.

Of more concern to research than the slight adjustments which have been made in recent months to the population figures is that, despite the fact that there was only a 94 per cent response rate and there was a subsequent (albeit well planned) imputation process, the ONS decided that releasing tables with small numbers could pose a risk to people's confidentiality. For the first time, therefore, a decision was taken to adjust small numbers in the tabular output for areas, such that all 1s and 2s were altered to 0s or 3s – a procedure that has been used in some countries such as New Zealand (data were 'blurred' in 1991, but in a much more benign way). As the 'small cell adjustment' procedure has not been released, it is impossible to undo this adjustment. And, as the table totals are the sum of their rounded cells, variable totals will vary across different tables. This will have greatest impact in the Special Migration and Workplace Statistics that record flows of people between small areas. As the majority of flows of one or more in the 1991 British census

were either one or two, their random replacement with 0s and 3s results in a drastic reduction in fitness for purpose (Duke-Williams 2003). Fortunately, for those interested in using the census in Scotland, this adjustment procedure has not been implemented, although the result for interaction data is rather unsatisfying, as flows from England and Wales to Scotland will not be rounded, while flows in the opposite direction will be. The current situation is unsatisfactory and clearly cannot be supported for long given the inconsistencies between different countries in the UK.

However, while academics complain about the suppression of small numbers, the very nature of collecting personal information from the entire population has been questioned from a human rights perspective. Perhaps surprisingly, though, while some have been concerned about the need for compulsory data collection from the population, the media were generally supportive of the exercise in the UK, in stark contrast to the stance among the US press who were much less convinced about the efficacy of their 2000 census. There was a heated media debate prior to the 2000 US census that focused on the 'big brother' nature of census data collection and questioned whether people should be 'forced' to share information about themselves. It is interesting to speculate whether the UK press would have given similar attention to the issues had we not been gripped by the foot and mouth outbreak at the time.

There was also a surprisingly heated debate in the US about census imputation. While the One Number census was adopted in Britain with relatively little concern among the press (although a few articles drew attention to it after the first census data were released), similar issues gripped the US media because of the political connotations. In a nutshell, the democrats were in favour of imputing missing people and the republicans were strongly against such a strategy (which they claim would be against the constitution; Dorling 2002). Note, of course, that most of the people who were missed were almost certainly those from poorer backgrounds who may have been more inclined to be democrat supporters. In the US, census data are used in the construction of political boundaries, while in the UK we continue to rely on the electoral roll (which thus favours the Conservative party when seats are redrawn; Pattie *et al.* 1996). The result in the US is that imputation has not been used to improve estimates, unlike the complex procedures adopted in the UK.

Given the various concerns raised above, including the cost of conducting the decennial census, National Statistics are now seriously considering whether future censuses are required. Similar questions are also being asked in a variety of other nations (Cook, this issue). Other possibilities include the utilization of regularly collected administrative data; rolling surveys of a sample of the population; or continuous surveys that are conducted annually and gradually work their way around the country. We would not dispute that there are exciting possibilities ahead that do not rely on decennial census information. However, as the papers that are summarized in the section below demonstrate, there is much that can be achieved with census data that may not be possible using other sources of information. Our personal view is that it will be impossible to collect enough reliable, detailed information about small geographical areas for censuses to be abandoned and, although it remains early days in the discussion, it looks increasingly likely that reliance upon administrative data will not be achieved before 2011. Most importantly, administrative data sources did not reveal that there were a million fewer people living in the UK in 2001 than we thought there were; only a very well taken census could do that. So there is probably at least one more census ahead of us, although it is interesting to speculate on the form that it will take.

### Papers in this issue

As described above, the papers in this special issue cover a wide range of the research that is possible using census data, given the timing of the data output. These studies reflect some of the diverse interests of those academics who take advantage of census information, but there will be many hundreds of other studies conducted over the next few years that will add to the body of census-related work.

Len Cook, the National Statistician with overall responsibility for the England and Wales census, provides a general overview of the 2001 UK census and suggests some of the possibilities for the future. He argues that we are experiencing a period of significant social and demographic change and that as households become more mobile, and frequently own more than one home, the business of collecting census data has become increasingly difficult. Obviously, the problem of accuracy is a perpetual one for census offices and we are glad that the bold step

was taken in 2001 to introduce the 'One Number census', designed to include a more accurate estimate of the population immediately. However, as described above, many have voiced dismay at the differences between the 1991-based population estimates provided for their areas and the 2001 census count. This is perhaps one of the most important factors encouraging the census offices to look seriously at different strategies for collecting information about the 'entire' population and, at the very least, serious consideration is being given to the most efficient way of producing a reliable address register (to prevent problems such as that which occurred in Manchester), and whether a national population register might also be a realistic future option (the Home Office's deliberations about identity cards may well have a significant impact on this). It seems crucial to realize that the business of censuses has changed and, not least because we now have more reliable information available for validating census output, aiming to provide the 'perfect number' may not be a realistic ambition.

Hennell (this issue) continues on the theme of the reliability of different population counts, which are essential for the distribution of a variety of resources within the health service. He compares the 2000- and 1996-based population projections at the national scale and for individual local authorities and demonstrates that there may be two distinct forms of systematic error in official UK migration statistics between 1991 and 2001. While the national patterns of variation between the 2001 census count and the projections appear to be mainly explained by mis-estimation of international out-migration, local differences appear to be more specifically related to mis-estimation of international in-migration. In particular, it does appear that there may have been a mis-estimation of the volume of out-migration of younger adult males from the United Kingdom throughout the period 1982–2000, and that these persons represent a major challenge for future population counts. The question, of course, is how we best handle the errors that we believe to have occurred. Hennell is particularly concerned about the enumeration of the large number of adults who do not have a permanent address, are 'staying with friends', 'sleeping on other peoples' floors' and so on. Census authorities will have to come to terms with the day-to-day mobility of a more transient population.

Rees and Butt (this issue) focus on the changing number of people in different ethnic minority groups

through the period 1981–2001. These groups continue to grow rapidly, and reasonably consistently, at an average of 40 per cent per decade. While some groups such as Black Africans and Bangladeshis have grown rapidly, others have expanded more moderately, but the White population has barely grown at all and the White British population has probably declined during the two decades. A considerable literature has examined the geographical distribution of minority groups through time, with a general conclusion that decentralization of non-whites from metropolitan and urban centres has been slow but is occurring (e.g. Rees and Phillips 1996). Interestingly, this paper shows that while Black and ethnic minority groups were concentrating into metropolitan areas between 1981 and 1991, between 1991 and 2001 deconcentration began for most groups. Indeed, five out of the ten ethnic minority groups they considered became less segregated from the white population in 2001 than they were in 1991. In sum, ethnic diversity continues to increase in most places at a steady pace and this is a pattern that we can expect to see continue in the future (see Dorling and Rees 2003 2004; Mohan 2004). The paper is also a good example of research that would be difficult to examine using any other regularly collected administrative data.

Higgs *et al.* (this issue) use 1991 and 2001 census data for Wales to examine temporal and spatial changes in Welsh speaking during the decade. Various commentators have welcomed the apparent 2 per cent rise in Welsh speaking that occurred during this period, but such an interpretation would ignore the fact that the question was altered in 2001 and also that the language continues to decline in use where it is most likely to be used (where the highest proportions speak Welsh). In this case, the census becomes less useful for examining temporal changes in the answer to a particular question, but is still useful for allowing small area geographical comparisons to be made. The paper also addresses the usefulness of the questions themselves, and points out that for sensible studies of Welsh speaking, language use surveys which show how the Welsh language is actually being used in the home, workplace and community may provide a better indication of whether people really use the language they claim that they speak. Higgs *et al.* call for the integrated use of census and survey data.

Ballas (this issue) uses the 2001 census data in a simulation exercise designed to examine population change between 1991 and 2001 in two northern

cities (Leeds and Sheffield) and integrates the census data with a survey. The British Household Panel Study (BHPS) includes a variety of detailed information about households that is not captured in the census. However, the expense of conducting the BHPS means that it is based on a relatively small sample of approximately 5000 households in the whole of Britain. It is therefore impossible to use the BHPS to tell us about the characteristics of people in small areas within cities. Using the 2001 census small area data for the two cities as a constraining factor allows the characteristics of households to be simulated from the BHPS for such small areas. The exercise is used to examine the changing trends in income inequalities and child poverty between and within the two cities. This paper therefore demonstrates that one, often ignored, use of the census is the ability to combine information with data from other sources. The result is that geography can be added to other socio-economic survey data sets such as the BHPS.

Martin (this issue) examines how neighbourhood-level data are being used increasingly for policy making. The census is likely to become the basis for much of this work but, as Martin points out, given that census areas do not necessarily reflect local sentiments about neighbourhood boundaries, it can be difficult to reconcile needs to provide information about 'real' local places, and the intricacies of collecting and releasing census data. Thus, an obvious challenge for the future is to combine quantitative and qualitative approaches to neighbourhood definitions. Indeed, while the new 2001 UK census geography will provide a much improved basis for the consolidation of data from official administrative sources, additional challenges need to be faced when considering future census data collection.

Coombes and Raybould (this issue) consider urban/rural variations in employment rates. Following the assertions of the Chancellor that there are many job vacancies in communities neighbouring places with large numbers of unemployed, they examine local variations in the characteristics of people looking for work and local trends in job availability. The 2001 UK census data allowed a comparison of the differences in employment rates between working age people categorized by gender and age and level of qualifications and the local job opportunities that exist. They make the obvious point that those unemployed as a result of the decline of traditional manufacturing industries will not necessarily have the skills required to work in

new industries that may grow simultaneously. Focusing on the *change* in job availability between 1991 and 2001 shows that the metropolitan conurbations continue to have the lowest employment rates and faltering local economies. However, London has performed much better than the other conurbations, with a growth in employment opportunities and lower unemployment rates, and this growth shows no sign of diminishing. Crucially, they show that there is simply an overall shortfall of labour demand in the provincial conurbations and many other northern urban areas, and there is also little evidence of substantial surplus demand for low-skill workers in other parts of the country. It seems that Chancellor's claims cannot be supported from the 2001 UK census data.

Boyle *et al.* (this issue) examine the changing patterns of mortality in Scotland between 1981 and 2001. While it has been established for some time that there is a deprivation gradient in health and mortality, there is growing recognition that the gap between the most and least deprived is widening. However, it has also been shown that health and mortality outcomes are worse in places experiencing population decline. This study examines whether there is a deprivation gradient in Scottish mortality and also whether this gap is widening through time. In addition, and importantly, they examine whether any apparent widening of the mortality gap may be explained by patterns of population change. The results confirm that a deprivation gradient exists; that it has widened since 1981; and that, while the gap is widening most in those places that are experiencing population decline, it is also widening elsewhere. This is another example of research that relies on having small area data for the creation of reliable population denominators and the creation of deprivation scores. Without census data most epidemiological work on the UK could either not be carried out or not be scaled up to assess its overall importance.

## Conclusion

On balance, it is our belief that the decennial census is a vital resource for a remarkable amount of demographic, economic, health and social research, which would not be possible using alternative sources of information. The census is particularly useful for understanding small area variations, and this is becoming even more exciting as the ability to compare change through time is

a realistic possibility (Martin *et al.* 2002; Norman *et al.* 2003). The unique data provided every ten years have formed the heart of numerous academic, commercial and government projects. Census data help us distribute resources, if not votes, fairly by allowing us to identify neighbourhoods that require support; they help retailers target consumers more accurately; they provide accurate population denominator information for use in epidemiological, political, social, planning, economic and geographical studies; and so on. This is not to say that other sources of information should be ignored, or that other methods of data collection should not be considered. A reliably updated national address list would be a wonderful resource that would help improve the accuracy of future censuses and surveys and would also be useful in other contexts. Indeed, the creation of a national population register would be valuable for many reasons, not least because we could learn much about population mobility through time.

Undoubtedly advances will need to be made in the way we capture information and handle people who are increasingly difficult to attribute to a single address. There have also been significant changes in the nature of international migration in recent times, with a growing number of young people coming from the EU, and the old Commonwealth and more people choosing to retire abroad for at least part of the year. Many people now have second homes, which are increasingly being used on a regular basis, rather than as holiday destinations. Increasing numbers of people have a flat 'in the city' only for use during part of the week. Others move between different household's residences, such as children who spend time with separated parents. And increasing numbers of young adults and students live away from home for most of the year. Collecting useful information about the population at a particular time point is difficult in these circumstances, but it is a challenge that must be continued.

## Acknowledgement

Useful comments were made on a draft version of this paper by Paul Norman.

## References

- Ahmad W 1999 Ethnic statistics: better than nothing or worse than nothing? in **Dorling D and Simpson L** eds *Statistics in society: the arithmetic of politics* Arnold, London 124–31

- Ahmad W and Sheldon T** 1993 Race and statistics in **Hammersley M** ed *Social research: philosophy, politics and practice* Sage, London
- Boyle P J, Curtis S, Graham E and Moore E** 2004 *The geography of health inequalities in the developed world: views from Britain and North America* Ashgate, London
- Brown J J, Diamond I D, Chambers R L, Buckner L J and Teague A D** 1999 A methodological strategy for a one-number census in the UK *Journal of the Royal Statistical Society Series A – Statistics in Society* 162 247–67
- Brown M and Boyle P J** 2000 Governmentality, sexuality and the census in **Brown M** ed *Closet space* Routledge, London 88–115
- Diamond I** 1999 The census in **Dorling D and Simpson L** eds *Statistics in society: the arithmetic of politics* Arnold, London 9–18
- Dixie J and Dorling D** 2002 New questions for the 2001 census in **Rees P H, Martin D and Williamson P** eds *The census data system, 2002, part VI, planning for 2001 census outputs* John Wiley & Sons, Chichester chap 20 283–93
- Dorling D** 1999 Who's afraid of income inequality? *Environment and Planning A* 31 571–4
- Dorling D** 2002 Bushmanders and bullwinkles. How politicians manipulate electronic maps and census data to win elections by Monmonier M (2001) *International Journal of GIS* 16 401–6
- Dorling D and Rees P H** 2003 A nation still dividing: the British census and social polarisation 1971–2001 *Environment and Planning A* 35 1287–313
- Dorling D and Rees P H** 2004 A nation dividing? Some interpretation of the question *Environment and Planning A* forthcoming
- Duke-Williams O** 2003 The 2001 census origin-destination statistics Unpublished paper to the British Society for Population Studies annual conference Bristol, 10–12 September
- Gardiner C** 2001 Informing policy making: new approaches to analysing the 2001 census *Local Government Studies* 27 71–88
- Harris R and Johnston R** 2003 Spatial scale and neighbourhood regeneration in England: a case study of Avon *Environment and Planning C – Government and Policy* 21 651–62
- Leventhal B** 2003 Developments in outputs from the 2001 census *International Journal of Market Research* 45 3–19
- Martin D, Dorling D and Mitchell R** 2002 Linking censuses through time: problems and solutions *Area* 34 82–91
- Mitchell R, Martin D, Dorling D and Simpson S** 2002 Bringing the missing million home: correcting the 1991 SAS for undercount *Environment and Planning A* 34 1021–35
- Mohan J** 2004 A nation dividing? Some questions of interpretation *Environment and Planning A* forthcoming
- Norman P, Rees P and Boyle P J** 2003 Achieving data compatibility over space and time: creating consistent geographical zones *International Journal of Population Geography* 9 365–86
- Pattie C, Dorling D, Johnston R 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
- Rees P and Phillips D** 1996 Geographical spread: the national picture in **Ratcliffe P** ed *Ethnicity in the 1991 census: volume three. Social geography and ethnicity in Britain: geographical spread, spatial concentration and internal migration* HMSO, London 23–109
- Simpson L** 2003 Are the census outputs fit for purpose? Working paper Cathie Marsh Centre for Census and Survey Research, University of Manchester
- Steele F, Brown J and Chambers R** 2002 A controlled donor imputation system for a one-number census *Journal of the Royal Statistical Society Series A – Statistics in Society* 165 495–522
- Watt N** 2002 Antiquated census could be scrapped *The Guardian* 28 May