

APPENDIX E : PROBLEMS OF USING CENSUS DATA

There have always been problems with using Census data, but these tend to have been ignored, because it was the only data source available. Changes in the way in which the Census was produced in 2001 mean that it is possible to be more specific about issues which should be borne in mind when using Census data.

These issues have three different aspects:

- issues arising from the 'One Number Census' imputation process;
- issues arising from modifications made to the data in order to maintain confidentiality;
- issues arising from the date of the Census compared to the other data sources.

Because the Social Need Study relies upon the calculation of rates, rather than comparing absolute numbers between zones, the effect of these factors upon its reliability will depend upon the extent to which they affect the characteristics of the population. For example, even if the total number of households and the number without cars are not reliable, does this affect the **percentage** of households without a car?

The 'One Number Census' Imputation Process

In previous Censuses the numbers published were the people actually counted. In 2001, a 'One Number Census' was produced. This attempted to impute the characteristics of households and people who were missed by the Census, in order to give details for everyone who was resident in the area.

A number of local authorities, including Nottingham City Council, have doubts about how the imputation process has worked, considering that they have other data sources which suggest that the Census population figures are too low. ONS have now accepted that there have been problems in certain areas and amongst certain age-groups. At the time of writing, this amounts to revisions to the numbers of younger adult men (the most difficult group to enumerate) in the 2001 Mid-Year Estimates of Population (which are based upon the 2001 Census), adding a further 20,000 to the population of Manchester (the first area to be examined in detail) and initiating a three year review of the figures for other areas.

Of the authorities in Nottinghamshire, this mainly affects Nottingham City. The revision to the numbers of adult men has resulted in 2,000 men aged 25 to 34 and 400 aged 35 to 49 being added to the 2001 Mid-Year Estimates for the City. The maximum change in any other Nottinghamshire District was 100.

No corresponding revisions have been made to the Census data, nor is it possible to say how these revisions should be spread around the wards or social need zones, as the Mid-Year Estimates are only published at local authority district level.

The City Council considers that some other age-groups have also been underestimated in the Census. For instance, evidence from school rolls and Child Benefit data suggests that the number of 5 to 9 year olds has been underestimated by about 5%.

Unfortunately, there is no way of knowing what the characteristics are of any households and people missed by the imputation process, apart from the fact that a disproportionate number of them are younger adult men. These may have largely been in households which have already been included or they may constitute one person households which have been missed.

Although it is impossible to know definitely, it seems unlikely that any of the indicators used have been significantly affected by errors in the imputation process, but as the undercount is an issue mainly in Nottingham City, it may be that figures for the City area are rather less reliable.

Data Modifications and Confidentiality

To ensure it is not possible to deduce information on individuals, small numbers in the Census have been modified. ONS have not given a threshold under which a number may be subject to modification.

This modification means that any small figures or aggregations of small figures, e.g. the aggregation of some output area data to social need zones, cannot be regarded as the actual figures. Checks have, however, been carried out to ensure that data for zones in each District sum reasonably closely to District figures.

The main effects are likely to be on indicators where the actual numbers involved are lowest, as only small numbers have been modified. Because of this, the two indicators most likely to be affected are 'young people unemployed' and '16 and 17 year olds not in full-time education'. The rates for these indicators should, therefore, be treated with particular caution.

Date of the Data Sources

The Census was carried out on 29th April 2001. Most other information used in the Study is more recent, meaning there are potential inconsistencies, particularly where denominators from the Census are used to calculate rates for non-census indicators (e.g. the rate of domestic burglaries per 1,000 dwellings).

This would only affect the analysis if conditions in different parts of Nottinghamshire had changed in a markedly different way from the area as a whole between Census Day and the dates of different data sources. There is no evidence that this is the case (e.g. unemployment has continued to fall in nearly all areas). The only significant exception is that the number of university students has risen considerably, meaning that the social composition of those parts of Nottingham, and to a lesser extent Broxtowe and Rushcliffe, inhabited by students, may have continued to change. However, this is unlikely to have significantly affected the Study's results, even in those areas.