

2001 CENSUS DISCLOSURE CONTROL**Introduction**

- 1 Advisory Group paper AG(01) 08 explained that ONS will round all Census output to either a zero or multiple of 3, and that thresholds would be increased to 100 people and 40 households for the Census Area Statistics. This was the earliest opportunity to inform users of the decision. This paper aims to give further background to that decision.
- 2 **Advisory Group Members are asked to note the background to additional disclosure control measures to be applied to 2001 Census output in England and Wales and the reasons why the National Statistician has felt it necessary to introduce these at this stage.**

Background

- 3 The Registrar General and National Statistician has an obligation not to reveal information collected in confidence in the Census about individual people and households, and has given public assurances that he will not do so. In presenting very detailed results from the Census, protecting this data is of key importance. Traditionally the confidentiality of Census output is prevented by a combination of a variety of disclosure control methods.
- 4 The general strategy for ensuring the statistical confidentiality of 2001 Census output was stated in the Government's March 1999 White Paper *The 2001 Census of Population* (Cm 4253):

“Precautions will be taken so that published tabulations and abstracts of statistical data do not reveal any information about identifiable individuals or households. Special precautions may apply particularly to statistical output for small areas. Measures to ensure disclosure control will include some, or all, of the following procedures:

- *restricting the number of output categories into which a variable may be classified, such as aggregated age groups;*
- *where the number of people or households in an area falls below a minimum threshold, the statistical output – except for basic headcounts – will be amalgamated with that for a sufficiently large enough neighbouring area; and/or*
- *modifying the data before the statistics are released.”*

Initial proposals for disclosure control in England and Wales

- 5 In the August 2000 Census Discussion Paper *Census Area Statistics*, ONS reported that measures would be taken to minimise the risk of the disclosure of information about identifiable individuals through the Census Area Statistics (CAS) whilst maintaining the statistical integrity and value of the statistics as far as was possible.
- 6 The primary method of disclosure control proposed at that time was the swapping of records on the output database at a pre-tabular stage so that it would never be certain that information related to any identifiable individual apparently shown in a cell of a table or in cells which have been summed, differenced or manipulated in any other way. Counts in all tables, although subject to this modification, would sum consistently to sub-total and total counts, as would counts for Output Areas (OAs) within a ward to the figures for the ward, and so on.
- 7 The Discussion Paper went on to propose that the primary method would be supplemented by two further statistical measures.
- 8 First, tables would be considered fit for release for Output Areas if:
 - cross-classifications are broad banded, particularly when three or more variables are cross-classified;
 - counts within a table would not generally fall into only a limited number of cells in the rows and/or columns; and
 - the number of cells in a table would not usually exceed the population to be counted in it; and
- 9 Second, the CAS tables would not be supplied for any area that failed certain population thresholds. Those proposed at the time were 50 usually resident persons *or* 20 resident households (compared with the thresholds of 50 residents *and* 16 households for the Small Area Statistics in 1991).
- 10 It was proposed that OAs would be created so as to have populations above the thresholds. In England and Wales, any ward or civil parish/community which fell below the thresholds would be amalgamated with a contiguous area to form an OA with a population which exceeded the thresholds, with advance consultation, where possible, with the local authority concerned. Counts of the number of residents (male and females) and resident households would be supplied for any areas which were 'suppressed' in this way.
- 11 The paper noted that in Scotland, OAs would be created so as to be above both thresholds. Each OA in Scotland would be assigned wholly to a ward, civil parish or other area type according to the assignment of postcodes within it.
- 12 The Discussion Paper went on to note that a further measure to control disclosure would be covered by a clause in end user licences making it a breach of the agreement to attempt to derive disclosive information or to purport to have done so. This condition currently applies to the 1991 Small Area Statistics.

13 However, the more recent Discussion Paper on Standard Area Statistics (September 2001) noted that the 2001 Census disclosure control measures were under review as part of an ONS-wide programme of methods and quality. At the round of Census Output Roadshow meetings this Autumn, ONS reported that the outcome of the review would be announced in due course.

Outcome of the ONS review of disclosure control

14 Maintaining the confidentiality of individual data underpins the trust that exists between data suppliers and any agency that acts as custodian of information about them. At ONS we are fortunate that businesses and the public have confidence that their information is securely held and that we do not release any data that could identify an individual. It is essential that this trust be maintained. The review was conducted to explore whether or not the measures that were in place would maintain this trust.

15 The review considered whether or not the measures then in place protected against both *real* and *perceived* disclosure. It considered:

- (a) the level of statistical detail in the proposed tables;
- (b) the level of geography for which the tables would be produced;
- (c) the increased availability of the data; and
- (d) the proliferation of other datasets against which Census data could be matched.

The review concluded that protection of the confidentiality of personal census information to meet the assurances given when it was collected remains paramount and will not be compromised in any way.

16 On the basis of this review the National Statistician decided that an additional measure of rounding all cells to either zero or a multiple of 3, should be introduced for all tables to be produced for England and Wales.

17 There will also continue to be minimum thresholds of numbers of persons and households for the release of output. The review concluded that these should be raised to 100 resident persons *and* 40 resident households for the release of Census Area Statistics.

18 These measures, together with the record swapping, will protect against both real disclosure risk and perceptions of disclosure that may arise in cells containing counts of '1'. Once formed, such perceptions may be difficult to dislodge.

19 Although it was not the primary consideration in the review, the introduction of rounding will allow the ONS to offer greater flexibility in output. It will, in particular, enable the Census Office to meet users' stated requirements for census information for overlapping geographical areas, for example, where these result from changes in administrative boundaries.

20 Work is continuing in order to ensure that confidentiality is protected in all cases of overlapping geographies, but the assumption is that Standard Tables will be issued for new wards and health areas introduced at any time after Census Day as well as for the wards and health areas existing at Census day, and for Parliamentary Constituencies,

National Parks and postal sectors as at Census day that may overlap other areas for which results are issued.

- 21 These arrangements for greater flexibility and for disclosure control apply to England and Wales. The full scope for flexible output, including the possibility of an interactive service for customised output will be covered in a further paper. There will also be more information on the implementation of the disclosure control measures.
- 22 A separate assessment of disclosure risks and output requirements has been made for Scotland by GROS, and, although the protection of confidentiality remains equally paramount in Scotland, somewhat different arrangements appropriate to circumstances in Scotland will apply.

Response to users' comments

- 23 In response to concerns expressed by Advisory Group members in the light of the announced additional measures the ONS would wish to assure users that it has listened carefully to their views expressed throughout extensive consultation on output requirements. However, confidentiality of census information is paramount. ONS has a responsibility to protect individual data, and the decision on what measures are required to do so is one for the National Statistician.
- 24 The National Statistician has assessed the risks and has decided that rounding is required in addition to record swapping and thresholds. All output will, however, be produced from one database and will give an accurate count of the population in 2001. Additionally, all tables that are produced will be consistent with this and in that respect will be compatible with the One Number Census strategy.
- 25 The rounding procedure will allow the Census Office to provide users with more flexible output and meet their stated requirements for overlapping geographies without further measures having to be implemented. Rounding will also allow ONS to continue with its Neighbourhood Statistics Strategy, which is essentially concerned about providing consistent data for small areas drawn from the Census and other sources.
- 26 The National Statistician has considered these additional measures necessary since the target size of output areas for which small area statistics are to be made available is almost half the average of the corresponding areas in 1991. The 2001 Census data will also be far more publicly accessible. Without these measures the risk of perceived disclosure would have been substantially increased.

Action

- 27 **Advisory Group Members are asked to note the background to additional disclosure control measures to be applied to 2001 Census output in England and Wales and the reasons why the National Statistician has felt it necessary to introduce these at this stage.**

**ONS
27 November 2001**