



ONS(ONC(SG))97/01

## **ONE NUMBER CENSUS STEERING COMMITTEE**

### **A One Number Census - a strategic overview**

1. The attached paper provides an overview of the current strategy for achieving a One Number Census in 2001. It sets out how the estimates from the various stages of the proposed methodology will be integrated, the role that consultation will play, and as an Annex, the way the 1997 Census Test will be used to help with the ONC research programme.
2. **The steering committee are asked to:**
  - a) **note the paper; and**
  - b) **provide any comments on the currently proposed strategy.**

**Andy Teague and David Thorogood  
Census Division  
ONS**

**Room 4200W  
Segensworth Road  
Titchfield  
Fareham  
HANTS  
PO15 5RR**

**May 1997**

# ONE NUMBER CENSUS STEERING COMMITTEE

## A One Number Census - a strategic overview

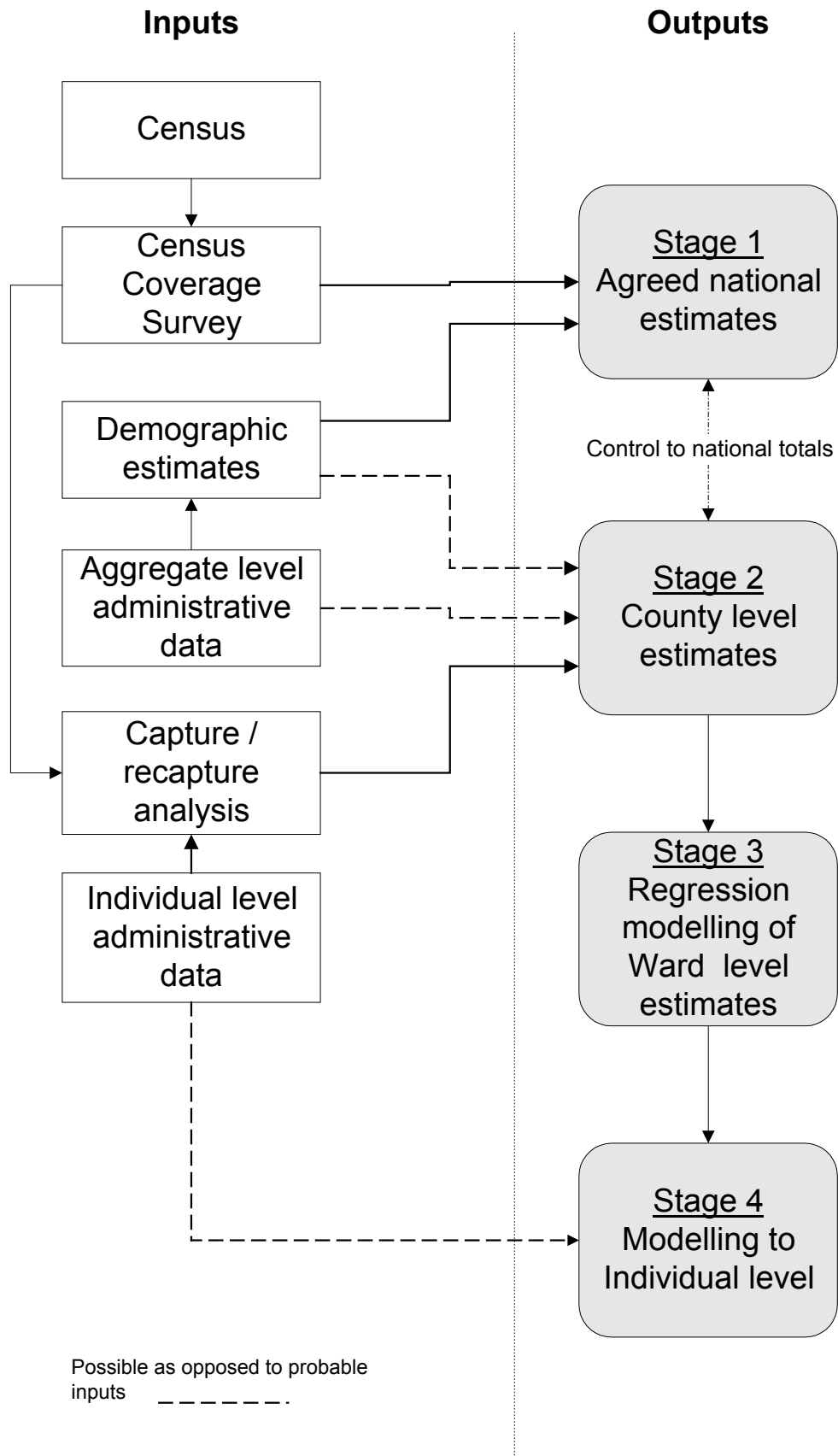
### Introduction

1. The One Number Census should be viewed first and foremost as a contingency that the 2001 Census will fail to enumerate the whole population. While that is inevitable - no Census will ever achieve complete coverage (partly because nobody knows what the truth is!) - this does not mean that the Census Offices are not attempting to count everyone. On the contrary, significant efforts and new techniques are being employed to minimise the level of underenumeration, and in particular to reduce its differential nature.
2. The establishment of the ONC project was born out of the problems with the 1991 Census and the role of ONS and the other Census Offices to produce the best statistics on the population of the UK. The Census in 2001 plays a major part in this - it will provide a significant amount of detail about households and people at the small geographical level.
3. The first stage in the ONC will be to produce the best estimate of the national population by age and sex. Subsequent stages will produce estimates for lower levels of geography and for other characteristics of people and households, ultimately to impute households that are estimated to have been missed (and not identified by enumerators) and people estimated to have been missed from households. This last stage would allow all statistics based on the 2001 Census to aggregate to "One Number". The four stages are:

- |          |  |
|----------|--|
| Stage 1. | Producing national level estimates                                     |
| Stage 2. | Producing County or Local Authority District level estimates           |
| Stage 3. | Modelling of local area estimates (for example, local authority wards) |
| Stage 4. | Modelling to individual level.   |

4. These are illustrated in Figure 1 (overleaf). One of the major challenges of the project will be to develop a strategy for producing the best estimate possible at each stage, and to ensure that the estimates at each stage are constrained to all other stages.

5. **Figure 1 - The Stages of a One Number Census**



## Scope

6. The ONC research project has so far concentrated on developing methodologies for measuring the coverage of the 2001 Census In England and Wales. Once a full methodology for England and Wales is proposed (later in 1997), the project will be extended to cover Scotland and Northern Ireland, in consultation with the other UK Census Offices - GRO(S) and NISRA.

## Agreeing a national level estimate

7. ONS produces an estimate of the population in England and Wales each year. These are traditionally based on the preceding Census, and updated by the number of births since the last Census<sup>1</sup> less deaths plus an allowance for net migration.
8. To estimate the population by age and sex at the national level in 2001, counts from the 2001 Census adjusted for estimated net undercoverage from the Census Coverage Survey (CCS) and administrative records using capture/recapture analysis techniques (see below), will be compared with the demographic estimates of the population. The proposed strategy is that if the error bounds for the adjusted Census count and the rolled forward estimates overlap, then the adjusted Census count should be accepted. It may be that for some age groups, the adjusted Census count is believed, while for others the demographic estimate is believed. Other techniques such as sex-ratio analyses will be used to back up the decisions taken.
9. The demographic estimates will essentially be the rolled forward estimate of the population, produced in the same way as in any intercensal year. As part of the process of producing the best population estimates, alternative sources will be reviewed such as administrative records on the elderly population from DSS Retirement Pension records and child benefit records. In addition, research is under way into the possible use of cohort analysis techniques (which are independent of any Census) to estimate the population.
10. The main issue that needs to be addressed is how the errors attached to the demographic estimates could be calculated. While it is uniformly accepted that there is little in the way of error in the birth and death records, the estimates of net international migration are (in the main) produced from the International Passenger Survey and as such are subject to sampling and non-sampling errors. Further, if administrative records are used, then any error in these will need to be estimated.

---

<sup>1</sup> Because of the well-documented problems with coverage of the 1991 Census (*see 1991 Census Validation Survey Report* for further details), the national level estimates are currently based on the 1981 Census.

## **Integrating the national and sub-national (county) estimates**

11. The estimation of net underenumeration at county level using the CCS is discussed in paper ONS(ONC(SG))97/02. This paper recognises that it is unrealistic to assume that the CCS could provide 100% coverage. Therefore data from administrative sources (such as the Family Health Service Authority (FHSA) registers and Council Tax Registers) will be linked to the Census and CCS records at the individual level and three or four way capture/recapture analysis used to reduce the error attached to the CCS-adjusted Census county estimates.
12. If the demographic estimates at the national level are decided to be the most accurate (see above), then the county level (CCS-based) estimates would need to be scaled to ensure that everything adds up. This would be most logically done on a pro-rata basis. But if the required adjustment is too large (what is meant by “large” needs to be decided), then demographic analysis techniques or administrative records could in theory be used at the sub-national level to help distribute the extra net undercount.
13. The relatively poor quality of migration data compared with births and deaths data is acknowledged to be the main weakness with demographic estimates. This relates particularly to sub-national geographies. The quality of different sources of migration data is being examined in ONS to enable more reliable sub-national demographic estimates. Aggregate counts from administrative sources may also be suitable for use at county or regional level as an adjunct (see ONS(ONC(SG))97/05 for further details).

## **Small Area Estimation and imputation**

14. The production of adjusted census counts for small areas (and ultimately to adjust the Census database) represents the final goals of the ONC process. Some initial regression-based ideas for this are discussed in papers ONS(ONC(SG))97/02 and 97/03. The modelling processes proposed will, by their nature, control to the higher level total of the preceding stage. They would need to be ultimately extended down to the lowest geographical building block - probably the postcode unit. A further stage will then be necessary to impute the characteristics of households that have not been estimated by the modelling process and to impute people into existing households. The editing and imputation routines being developed for the processing of enumerated households can, in theory, be used for this process but further research is necessary.
15. It is recognised that the estimates at this level will be subject to a fair degree of uncertainty, and therefore it may be decided that the ONC process should not be extended below some intermediate level (such as local authority wards - 6,000 population on average). If this is the case, then population estimates by some key variables - age, sex, ethnicity and employment status for example - at the intermediate geographical level might be the limit of the ONC. The

Census counts themselves would be unadjusted, but detailed guidance will be available to indicate where underenumeration is likely to have a significant effect. A strategy for adopting this fall-back position will need to be devised, probably built around the estimated error in the estimates.

## **Consultation**

16. Consultation will play a key part in the acceptance of the methodology. Initial feedback has been positive from public sector customers, less so from the private sector, which was expected as the latter are more concerned with timely data than complete coverage.
17. A consultation strategy is being drawn up but initial plans are for three main phases:
  - a) in the Autumn of 1997 to consider the methodological proposal
  - b) in Summer/Autumn 1998 to consider the proposal in the context of timing of the output, the marketing strategy for the Census and associated costs. This will be the key phase. It is likely to be carried out over a two to three month period and focused through the members of the Census Advisory Groups, and will inform the decision to be taken at the end of 1998 on the extent to which the ONC will be adopted.
  - c) a final pre-2001 phase in 2000, taking on board findings from the trial linked to the 1999 Census Dress Rehearsal.
18. In addition to this, presentations will be made at BSPS, RSS and international meetings.

**Andy Teague and David Thorogood**  
**Census Division**  
**ONS**

**May 1997**

### The 1997 Test and the One Number Census

1. The 1997 Test is primarily concerned with assessing various methodological options for improving the coverage of the Census, and for testing the acceptability of possible new questions for the 2001 Census:
  - new methods to deliver and collect census forms;
  - new form designs;
  - acceptability of questions on income and ethnic group.
2. While these are aims to maximise the coverage of the census enumeration, the 1997 Test will provide some evidence and opportunities for research into the One Number Census.
3. Evidence will be collected from the 1997 Test about the likely impact the new methods and questions have on the coverage of people and households. We will obtain some evidence on the types of household and people where coverage is likely to be worst. The 1997 Test will also provide us with the opportunity to test a couple of methodological elements of a ONC, all be it only on a small scale basis at this stage.
4. Firstly, it will be possible to link the 1997 Test database to some external records - Council Tax Dwelling Lists, and data from Family Health Service Authority registers and the National Health Service Central Register. Experiments have been done using 1991 Census and FHSA data, matched by sex, date of birth and postcode, for capture/recapture analyses. These have not been wholly successful due to the poor quality of the FHSA registers in 1991. The main problem with the registers has been that they are often out of date, particularly in inner-city areas and for young males. In other words the people are in the wrong place. This has improved somewhat since 1991 so it is worth repeating the exercise of linking on sex, date of birth and postcode.
5. To try and make further improvements, names are being captured in the 1997 Census Test. The prime aim will be to assess the practicalities of linkage using both names and date of birth, and date of birth and postcode, and in turn the effect this has on the estimates produced by capture/recapture techniques. The degree to which names helps in improving the estimates will be assessed to establish whether names should be captured on the census database in 2001.
6. Secondly, we be carrying out a small scale trial of an independent coverage survey. The proposal for the coverage survey in 2001 is that it should be postcode unit-based, that is a re-enumeration of selected postcode units. A sample of some 20 postcode units will be surveyed in Brent, one of the 1997 Census Test Areas, to assess the practicalities of carrying out an independent postcode-based survey. The survey records will be linked to the Census Test database.