

Advisory Groups Paper (97)04

ADVISORY GROUP PAPER

METHODOLOGICAL RESEARCH TOWARDS A 'ONE NUMBER CENSUS'

1. This paper describes the methodological research being undertaken by ONS to:
 - a) estimate the level of net undercoverage in the 2001 Censuses in the United Kingdom; and,
 - b) adjust the Census enumerated database, and counts from it, to take account of the estimated net under a *one number census* (ONC).
2. The minimum objective of the project must be to agree a methodology to measure the level of undercoverage necessary to rebase the population estimates in 2001. Having ensured that this minimum objective can be achieved then the overall aim of the project is to develop an agreed methodology to fully adjust the Census for net undercoverage.
3. **Advisory Group members are asked to:**
 - **note this information paper, and**
 - **note the plans for consultation via the Advisory Groups of which the main phases will be in the Autumn of 1997, 1998 and 2000.**

Roma Chappell
Census Division, ONS
01329 813529

Background to the decision to research ONC methodologies

1. ONS are carrying out methodological research to:
 - a) estimate the level of net undercoverage in the 2001 Censuses in the United Kingdom; and,
 - b) adjust the Census enumerated database, and counts from it, to take account of the estimated net undercoverage - *a one number census* (ONC).
2. The minimum objective of the project must be to agree a methodology to measure the level of undercoverage necessary to rebase the population estimates in 2001. Having ensured that this minimum objective can be achieved then the overall aim of the project is to develop an agreed methodology to fully adjust the Census for net undercoverage.
3. In recent censuses, coverage adjustment has only been made to allow for production of the rebased population estimates at local authority district level.¹ Until recently this has been considered to be sufficient, but the 1991 Census suffered from a higher than expected level of undercount. While the estimated total undercount of 2.2% was not excessive, problems were caused by the wide variation in the estimated undercount between population subgroups. Young males, the elderly, and the population in urban areas were estimated to have been the most poorly enumerated, with over 20% of the population missing for some population sub-groups. This led to confusion in interpreting the census counts, as users were concerned about their validity.
4. The Office for National Statistics has therefore undertaken to explore the feasibility of adjusting all the census counts as well as those that feed into the population estimates so that all counts agree. This is far from easy to do and requires a substantial amount of research and consultation with key customers over the next two years, before a decision can be taken on whether or not to proceed with a One Number Census. Much of the research needs to be carried out in 1997, so that a full methodological proposal can be put together. Work in 1998 will concentrate on evaluating the quality of the likely results, assessing the impact on the production timetable and consulting with users.

Methodological Research

Introduction

5. The full **aim of the One Number Census** approach is to integrate coverage estimation with both Census results and Population Estimates to achieve one population and one household count. Under a full ONC all Census output would be adjusted for undercoverage. This would be best achieved by supplementing the database of enumerated households and people by imputing, via the One Number Census process, those estimated to have been missed.

¹ Having said this, there was an improvement in 1991 in the Census resident population base. Households that were wholly absent on Census night and who failed to make a voluntary return, were imputed.

Other sources - Census coverage survey, demographic analyses and administrative records

6. The evidence for the size of the national undercount in the 2001 Census and the distribution of this at sub-national level will come from a number of sources. A key source will be an independent post enumeration survey. The **objectives** of the survey will be to assess the coverage of households and people within households. The sample size must be large enough to give precise estimates at the required geographical level and the requisite sample size and associated costs are being researched.
7. It is intended that the survey should take the form of a short interview to check on the coverage of households and on people within households and to collect basic demographic characteristics (primarily age, sex, ethnic group, economic activity and marital status). This simplified **Census Coverage Survey** (CCS) questionnaire can then be a pre-coded, machine readable form which will be quick and cheap to process. Another important advantage of a simplified questionnaire is that it will allow for the CCS to be ready to go out into the field soon after Census Day, thereby reducing the chance of errors due to lapses of memory by the public
8. **Demographic analyses** should also be carried out to provide an independent check of the population counts at national level. The rolled forward population estimates are an obvious example but it is important to recognize that the base for these will need to be the 1981 Census (or earlier). It is essential, therefore, to back up the rolled forward estimates with other evidence.
9. It is planned that **administrative records** will be used as alternative sources of information on the population as a whole and for certain population subgroups. The quality (e.g. completeness and frequency of updates) and availability (at the appropriate time) of the registers must be established so that their use can be assessed. Assessing the quality of administrative sources in order to use them to validate the Census is not easy because the Census has often been used as a 'Gold Standard' against which to assess register based figures. It will be essential, therefore, in assessing the likely accuracy of Census counts, to use several sources of information. It will also be necessary to ascertain the availability and suitability of registers for the different countries of the UK. For example certain registers may only be available for England and Wales, and not for Scotland or Northern Ireland. Legal aspects of the use of administrative records will also need to be addressed.

National Level Estimation

10. The estimates of population at the national level from all sources must be considered so as to arrive at an agreed national population count. This is a key aspect of the project and users' views will have a major input. There will be no difficulty if all the main sources are broadly consistent. In the event of conflicting evidence, then either one source would have to be believed over the other sources or there would need to be a way of combining the estimates.

Sub-National Level Estimation

11. The next step is to estimate undercoverage in the Census at sub-national level:
 - Administrative records, the Census Coverage Survey and the Census can in theory be linked and **capture/recapture** analysis techniques used to estimate those missing from all the linked sources. The basis of the capture/recapture technique is that if the chance of being enumerated in the Census is independent of the chance of being enumerated in another source, such as the Census Coverage Survey and one or two administrative registers, then these sources can be combined to estimate the number of people missed by all of them.
 - After the Census in 1991, **sex ratio analyses** were carried out at a high area level corresponding to groups of Counties. These were useful input in determining the level of undercoverage as they clearly highlighted the shortfall in the Census counts. Although the analyses were unstable at lower geographical levels in 1991, further research into ways of using sex ratio analyses will be

carried out.

- Additionally, the possibility of using **sub-national rolled forward estimates** needs to be reassessed. In 1991 these were of insufficient quality to validate the Census. The improvements to FHSA (Health Authorities) registers may provide better estimates of migration within the country which in turn will allow better rolled forward population estimates.

Small Area Estimation

12. To achieve the full One Number Census, coverage adjustment would have to be made below sub-national geographical level, such as to ward level or lower and ultimately at the individual level. People who are estimated to have been omitted from the Census will be imputed either as completely missed households or as individuals from enumerated households. The methodological approach to this stage is being researched and will be the subject of further consultations later this year and next year. Ian Diamond (University of Southampton) is a member of the One Number Census project and his previous work on small area estimation, undertaken by Estimating with Confidence, will be taken into account in this research.

Co-ordination

13. Most of the research tasks are timed so that the different aspects can be brought together to provide a clearer proposal in the Autumn of 1997. In particular, the recommendations on estimation to produce a national level count adjusted for undercoverage and sub-national estimation need to be carefully coordinated and the methodology for controlling the sub-national estimates to the national count agreed
14. A full ONC approach would undoubtedly impact on the timetable for producing census results. Having produced the initial database of enumerated records, the ONC requires an extra step to supplement the database with others estimated to have been missed. This process would take some time (the current estimate is nine months with census results being available from end 2002). The trade-off is that higher quality estimates would be produced. Early indications are that the major customers would find a delay of this length acceptable, providing the ONC methodology works.
15. The risks of a full ONC, of being able to deliver to customer expectations and on time are recognised by ONS. An assessment of all the risks involved will be a key element in the final decision. If it is decided to proceed with full implementation, then contingencies must be planned and the conditions for implementing these made clear to customers at the outset. If the full aim is considered to be too ambitious then the minimum objective must be to provide coverage adjusted population estimates (by age and sex) at the local authority district level.

Consultation

16. **Customer consultation**, particularly with the major users of Census data and population estimates and projections, is essential to ensure that customers accept ONC adjustment methodologies, the resulting coverage adjustments to Census counts, and the impact a ONC will have on the timing of the availability of Census results. A Project Steering Committee has been set up to oversee the work of the ONC Project Board. This committee will give strategic direction and technical advice as necessary and will involve representation from outside government (specifically local authorities, academia and another national statistics organisation) in order to provide a sound framework. Consultation will in the main be carried out using existing Census (and other ONS) Advisory Groups. Three main phases of this consultation are planned:
- ONC principles (Autumn 1997)
 - ONC methodologies (Autumn 1998)
 - The final approach. (Autumn 2000).

International Aspects

17. We are keeping abreast of international developments. As far as we know at the moment only ourselves and the US are pursuing research into the possibility of a one-number census, although there is a fundamental difference between the US and the UK approach.
18. The US approach is to enumerate the population, implementing measures to improve coverage so that the enumeration level achieved is above a high minimum acceptable level. Then, once this level is reached the regular enumeration procedures cease. The residual undercount would be estimated from a sample follow up and by applying statistical estimation and modelling techniques.
19. The UK approach is different in that there is no target below 100% for a minimum acceptable level of coverage for the regular enumeration exercise. We are aiming for as high a coverage level as possible, and to improve coverage in the areas which are difficult to enumerate, but that this will need to be backed up by coverage estimation methodologies to allow Census counts to be adjusted for the likely differential undercount.

Roma Chappell
Census Division, ONS
01329 813529