UK Data Archive Study Group Number 33335 - Census 2001 SARs

Census2001

A Guide to the One Number Census

Office for National Statistics
General Register Office Scotland
Northern Ireland Statistical and Research Agency

September 2001

©CROWN COPYRIGHT

A Guide to the One Number Census

CONTENTS

About this Booklet 1
Section 1: Introduction 2
What is a One Number Census? 2
Why is a One Number Census necessary? 2
How will a One Number Census be achieved? 3
Section 2: Census Coverage Survey 5
Section 3: Matching the Census and CCS Data 8
Section 4: Design Group and Local Authority District Estimates 11
Section 5: Imputation in the One Number Census 12
Section 6: Quality assurance 14
Annex A: Questions Frequently Asked about ONC Methodology 16
Annex B: The Research Team and Steering Group 18
References 19

ABOUT THIS BOOKLET

This booklet has been written to inform users of 2001 Census data and the mid-year population estimates about the One Number Census. The Guide is an update to the Spring 1999 consultation paper following the finalisation of the One Number Census methodology. It is designed to answer the following questions:

- What is a One Number Census?
- Why is a One Number Census necessary?
- How will a One Number Census be achieved?

Annex A includes questions frequently asked on the One Number Census.

SECTION 1: INTRODUCTION

What is a One Number Census?

1.1 The One Number Census (ONC) project aims to integrate the 2001 Census counts with the estimated level of underenumeration in the Census – that is the number of households and people not counted. First it will provide a new base for the mid-year population estimates at the Local Authority District (LAD) level, and second it will adjust the Census database itself for the estimated undercount so that all statistics add to 'One Number' - the national estimate of the population.

1.2 The primary source in estimating the underenumeration will be the Census Coverage Survey (CCS). This has a notably different design to the Census Validation Survey undertaken following the 1991 Census.

1.3 The project derives its name from the intention to eliminate different population counts so that all census outputs add to one number - the national population estimate which has been adjusted for underenumeration. Furthermore the mid-year population estimates will be based on these ONC population estimates and there will be transparency between them.

1.4 The ONC Project considered the methodology separately for England and Wales, Scotland and Northern Ireland. The high-level methodology presented here is consistent throughout the UK. Although there are some differences, these are mainly associated with the design and practical aspects of the Census Coverage Survey and the use of different geographies during the quality assurance process in Scotland.

Why is a One Number Census necessary?

The Census

1.5 The primary aim of the Census is to provide a snapshot of the entire population and thus provide high quality counts of the population for small areas across the country. These statistics are used for many important purposes. In particular, they provide a new base for the mid-year population estimates which are used to help distribute central government support to local authorities (in 2001 £36 billion was distributed to local authorities by DTLR). They also allow central and local government as well as health authorities to plan health, education, transport and housing needs for the years ahead.

1.6 Every effort was made to ensure everyone was counted in 2001 and a number of initiatives were introduced to maximise coverage. However, no census is perfect and some people will have been missed. Therefore, it is accepted practice throughout almost all similar census-taking countries to assess the extent of any underenumeration, usually by means of a post-enumeration survey.

What happened after the 1991 Census?

1.7 In 1991 two problems affected the Census in Great Britain.

- Over two per cent of the population was missed overall, although this compared favourably with other countries.
- Underenumeration did not occur uniformly across all areas or age-sex groups. For example, in some city areas over 20 per cent of young males were estimated to have been missed (OPCS, 1994).

1.8 The Census Validation Survey (CVS), as the 1991 Census follow-up survey was called, did not identify the extent and distribution of the underenumeration.

1.9 These problems took time to resolve and several different population counts were available following the 1991 Census, including the published census count and the census count uprated in line with the CVS. In particular, after publishing the census counts and a set of provisional estimates, a decision was made, in England and Wales, to base the national population estimates on demographic estimates and to produce consistent counts at local authority level using a mathematical model. At geographic levels below local authority level no official adjustments were produced, although those produced by the Economic and Social Research Council funded project 'Estimating with Confidence' have been widely used (OPCS 1993, 1994; Simpson, Cossey and Diamond, 1997). In Northern Ireland, where the extent of the underenumeration was not as great, the subsequent mid-year population estimates were uprated by a factor of 0.66%. As a result of these changes throughout the UK, the official census tabulations remained inconsistent with the population estimates. The aim of the One Number Census project is to develop procedures that will avoid such problems after the 2001 Census.

How will a One Number Census be achieved?

1.10 The process used to achieve the above aims is described in the following sections. The key stages, illustrated in Figure 1 overleaf, can be summarised as follows:

a) A Census Coverage Survey (CCS), undertaken independently of the Census, was designed to establish the coverage of the 2001 Census. For the CCS, the UK was divided into one hundred and twelve areas, each with a population of about 500,000. These areas are known as 'design groups' and are made up of whole LADs or groups of smaller LADs. The CCS took place in all of these design groups. The CCS is covered in more detail in section 2.

b) The CCS records are matched with those from the Census using a combination of automated and clerical matching. The matching is covered in more detail in section 3.

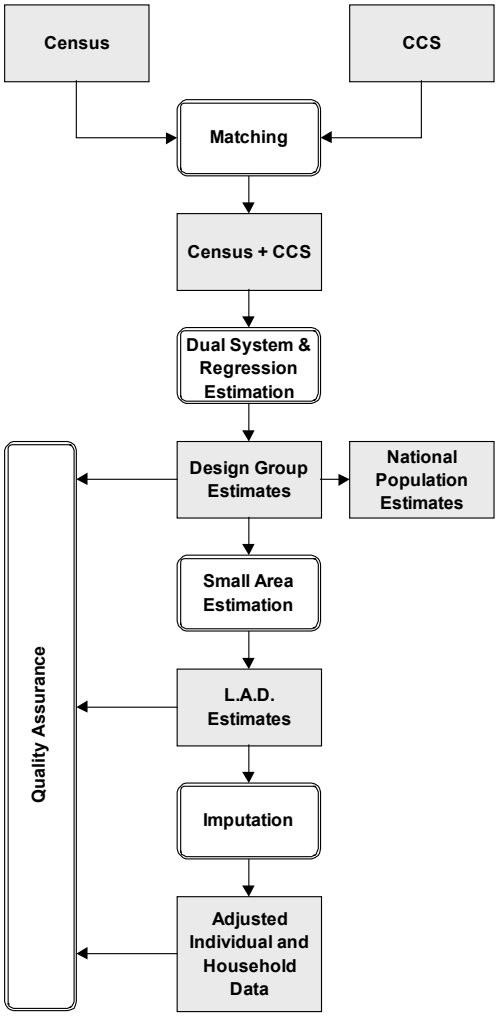
c) Populations for each design group, by age and sex, are estimated using a combination of standard estimation techniques. See section 4 for more detail.

d) Small area estimation techniques are used to estimate Local Authority District populations by age and sex. This estimation is also covered in section 4.

e) Households and individuals estimated to be missed by the Census are imputed to produce a fully adjusted Census database. The imputation process is covered in more detail in section 5.

f) All ONC population estimates are quality assured using demographic analysis and aggregate level administrative data. The quality assurance process is covered in section 6.

Figure 1 – The One Number Census process



Consultation

1.11 It is important that users of census data have confidence in the estimates of the population produced by the One Number Census. Acceptance was sought in a number of ways, including;

- the composition of the research team and the Steering Group (Annex B);
- the consultation process with census user groups;
- a Series A Royal Statistical Society paper (Brown *et al*, 1999);
- several RSS seminars;
- the Spring 1998 Census Consultation paper '2001 – A One Number Census';
- the Spring 1999 Consultation paper 'A Guide to the One Number Census'.
- a workshop devoted to the *One Number Census* project in May 1998; and
- special workshops held in conjunction with Census Output Consultation Roadshow meetings during April/May 1999 and September 2001.

1.12 Members of the research team can be contacted with questions or queries on the ONC process via the addresses at the back of this booklet.

SECTION 2: CENSUS COVERAGE SURVEY

Introduction

2.1 The key element in the ONC methodology is the Census Coverage Survey (CCS). Important features of the CCS include:

- It was designed to enable census population counts to be adjusted for underenumeration at the national, local and small area level.
- It comprised an intensive enumeration of a representative sample of 19,500 postcode units across the UK. The sample of postcodes was drawn from all Local Authority Districts (LADs).
- It consisted of a short, paper-based interviewer-completed questionnaire (as opposed to the Census self-completion questionnaire) designed to minimise the burden on the public. This was vital since the CCS, unlike the Census, is a voluntary survey.
- It was operationally independent of the Census enumeration exercise.
- It was undertaken during a four week period starting 3½ weeks after Census Day.

2.2 This chapter details the sampling methodology used and explains why this particular approach was adopted. The sample size of the survey is also discussed. Some key aspects of the survey methodology are highlighted and response rates reported.

Design

2.3 The primary aim of the ONC is to estimate the population on Census Day, by age and sex, for all LADs. If the CCS design was applied to individual LADs, the required sample would be prohibitively large. Therefore, in order to select an efficient sample to produce LAD population estimates, LADs were grouped to form 'design groups', each with a population of approximately half a million people. There are 101 design groups in England and Wales, eight in Scotland and three in Northern Ireland. The CCS design strategy was applied within each of the design groups. The sample was designed to enable accurate direct estimates of underenumeration for 37 age-sex groups[1] at the design group level.

2.4 The CCS is a postcode based survey. However, reliable postcode level information beyond number of addresses is not known. Therefore, information from the 1991 Census was used as a proxy to select 1991 Enumeration Districts (EDs) according to likely difficulty of enumeration in a first stage selection.

2.5 As underenumeration is disproportionately distributed across areas, the EDs are stratified according to a Hard to Count (HtC) index. In England & Wales and Scotland, this index is constructed from the following 1991 Census variables, which are associated with underenumeration:

- Multi-occupancy - defined as addresses containing more than one household, e.g. a set of bedsits (England &Wales, Scotland)
- Unemployment (England &Wales)
- Language difficulty - derived from country of birth (England &Wales)
- Private rented accommodation (England &Wales, Scotland)
- Number of households imputed in 1991 (England &Wales, Scotland)

[1] The age-sex groups to be estimated are: all 0 year olds and then 1-4 year olds and 5 year age groups up to 85+ for males and females.

- Ethnic group (Scotland)

2.6 The index in England, Wales and Scotland categorises all EDs into three groups representing 40%, 40% and 20% of EDs nationally - the hardest to count areas being the 20% group. The sample is then chosen from each of these strata.

2.7 Northern Ireland use an 8-level stratification index based upon the following factors which were found to be most related to undercount in the 1997 Census Test:

- Religion
- Urbanity / Rurality
- Deprivation

2.8 Within each of the HtC categories the 1991 EDs were clustered on the basis of their age-sex distribution. Evidence from the 1991 Census showed that young males aged 20-29, children aged 0-4 and women aged 85 and over were relatively more likely to be underenumerated. The EDs were clustered into groups with similar populations in these key age-sex groups and the sample selected from these clusters. This provided an efficient design and spread the sample with respect to the size of these key age-sex groups. Finally, a number of postcodes from each ED were chosen at random. These selected postcodes formed the CCS sample.

Sample Size

2.9 The sample size of the CCS must be sufficiently large that the accuracy of ONC population estimates is acceptable. The larger the sample size, the more accurate the population estimates, however this must be balanced against the cost and practicalities of carrying out a larger CCS.

2.10 The ONC research programme examined the optimal sample size representing the best value for money in terms of accuracy. This research indicated an overall sample size of around 300,000 households for England and Wales would provide an acceptable level of accuracy (relative errors of around 1%) for the total population of the design groups (around 0.1% for the national population). For example, a design group population of 500,000 would have, on average, a standard error of 5,000. The sample selected contained 320,000 households within 16,400 postcodes in England and Wales, 40,000 households in 2,400 postcodes in Scotland and 10,000 households in 700 postcodes in Northern Ireland.

Practicalities

2.11 CCS fieldwork began on 24 May 2001 - three and a half weeks after Census Day. The timing of the fieldwork period was dictated by the need to wait until census fieldwork finished, balanced by the advantages of conducting the survey as soon as possible after Census Day.

2.12 The CCS postcodes were aggregated to form interviewer workloads, typically between 90 and 200 households.

2.13 Interviewing was carried out in two stages: first, interviewers identified every address within the postcode; second, they sought to obtain an interview with a member of each household within the identified addresses.

2.14 Unlike the Census, identification of addresses within the interviewers' areas was not guided by any list. Instead, maps of the CCS postcodes were supplied to interviewers

and they confirmed the physical extent of the postcodes on the ground by calling on addresses. To ensure interviewers visited every household in their allocated postcodes they were instructed to contact households adjacent but outside the postcode boundary to ensure that all households in the selected postcodes were included in the CCS. This process defined the border of the postcode on the ground, and avoided the identification of households in the CCS being dependent either on administrative lists, which may have been out of date, or the postcode maps, which may have been inaccurate.

2.15 The CCS interview asked basic demographic characteristics of everyone living in a household, questions about the accommodation and simple relationship information. As a result, the questionnaire was short and simple and interviews could be carried out on the doorstep within 5 to 10 minutes, depending on household size.

2.16 To ensure census staff did not make a special effort to obtain response in areas to be covered by the CCS, the CCS sample postcodes were kept confidential. Census managers were not able to work on the CCS. Similarly, CCS managers, who were informed of their sample postcodes before the start of the census fieldwork, were not allowed to work on the census. Census enumerators were allowed to act as CCS interviewers, but were not informed of their postcodes until their census fieldwork had finished; they were also prevented from interviewing in the same area they had enumerated.

2.17 The quality of the CCS depended on the ability of the field force to conduct their work correctly, and great efforts were made by CCS HQ to ensure that the training provided to the field staff was adequate.

2.18 The CCS field force in England and Wales was supplemented by 80 experienced interviewers from the ONS's Social Survey Division. These interviewers were allocated to teams in areas expected to pose particular problems in obtaining responses. In Northern Ireland the entire field force was made up of experienced interviewers.

2.19 Interviewers were instructed to make as many calls as necessary to obtain an interview, and to call at different times and on different days to maximise the probability of making contact. If, on the final day of fieldwork, they were still unable to make contact with a household they delivered a self-completion form for the householder to complete and return to CCS HQ by post.

2.20 The CCS in England and Wales obtained interviews from an estimated 89% of the residential households identified. The inclusion of posted CCS forms increased this response rate to 91%. In Scotland, the CCS achieved an overall response rate of 95% and in Northern Ireland this figure was 92%.

SECTION 3: MATCHING THE CENSUS AND CCS DATA

3.1 Estimates of the total population will be based on a methodology known as dual system estimation (section 4). It is inevitable that some households and people will be missed by both the Census and CCS but dual system estimation can be used to estimate this number by considering the relative numbers of the people observed by;

- both the Census and CCS;
- the Census but not the CCS; and
- the CCS but not the Census.

3.2 In order to identify the numbers in each of these groups it is necessary to match the records from the CCS with those from the Census. It is essential that this matching process is accurate as the number of mismatches has a direct impact on the final population estimates.

3.3 The 2001 matching exercise will involve a combination of automated and clerical matching. The matching process for a single CCS postcode is outlined below and illustrated in Figure 2. There are four key stages:

Stage 1 - Exact Matching

3.4 Automatically link CCS and Census households and individuals where key details match exactly. The key details used to exact match households are postcode, address name/number, type of accommodation, number of people and surname of household representative. The details used to exact match individuals are forename, surname, day and month of birth, marital status and relationship to the household representative. Households will only be considered matched at this stage when all individuals within the household pair have been linked.

Stage 2 - Probability Matching

3.5 CCS and Census records that were not matched at Stage 1 of the process are then run through a probability matching process. Probability matching examines the same variables as used in exact matching. A probability weight is assigned to each pair of CCS and Census records based on the level of agreement between them. The higher the probability weight, the closer the agreement between the two records. For example, if a pair of records is identical with the exception of one detail, which may be due to recording error, then a high probability weight will be assigned. Any household pairs with a high probability weight are linked and the individuals within them compared. Only very similar households and individuals will be considered as matched at this stage.

3.6 The initial probability weights used have been derived from the 1999 Census Rehearsal. Once some 2001 Census data has been linked, it will be possible to update the weights to be more appropriate to the data being matched. 'Training' the weights in this manner will also help to allow for any regional differences in the data.

Stage 3 - Clerical Resolution

3.7 Pairs of households and individuals with a reasonable level of agreement are presented for clerical resolution. At this stage operators will simply be asked to

determine whether the pair of records shown constitute a matching pair or not. They will not be expected to search for matching records.

3.8 The file of exact, high and medium (with clerical verification) quality matches can be used to update the probability weights once again.

Stage 4 - Clerical Matching

3.9 The final stage of the ONC matching process involves a clerical search for any census records corresponding to unmatched CCS households and individuals. Operators have access to images of Census and CCS forms to assist in decision making.

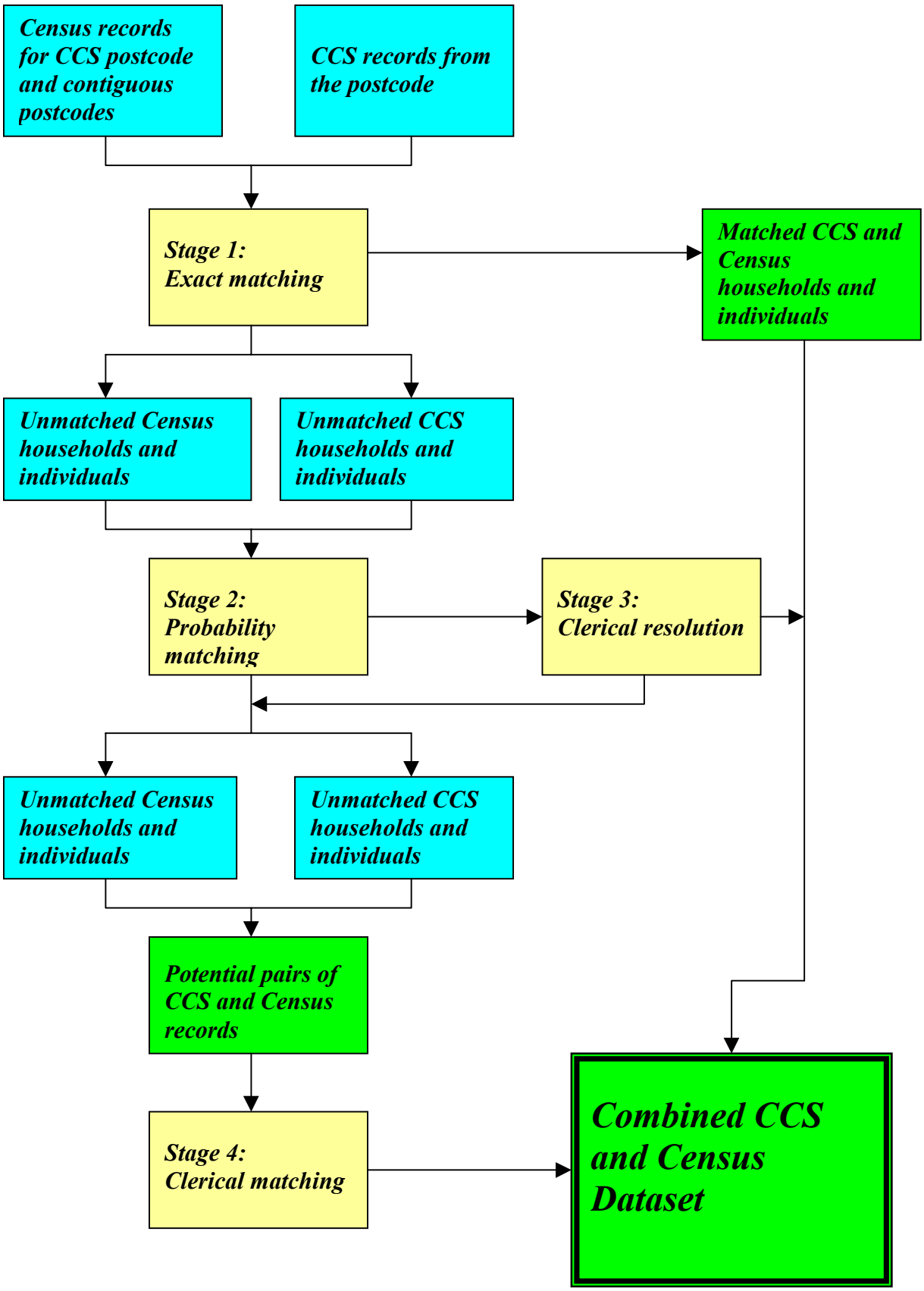
Quality Assurance

3.10 As previously mentioned, the accuracy of the matching process is critical to the accuracy of the ONC population estimates. Quality Assurance procedures, similar to those used in the US, have been built into the matching process to ensure that the necessary high levels of accuracy are met.

3.11 The output of the clerical matchers will be checked by expert matchers to ensure that all matched pairs of records are legitimate matches. These experts will also check that all unmatched records do not have a possible match using extensive database searches.

3.12 A small number of supervisors will check the work of the expert matchers. These supervisors will also assist in marginal matching decisions. These processes should ensure accuracy and a consistent approach.

Figure 2. An illustration of the ONC matching process



SECTION 4: DESIGN GROUP AND LOCAL AUTHORITY DISTRICT ESTIMATES

4.1 The next stage in the ONC process is to derive estimates of the population for all Local Authority Districts (LADs) using the combined Census and CCS data generated by the matching.

Stage 1 – Design Group Estimation

4.2 The output from the matching process will be used to estimate underenumeration by age and sex for each CCS postcode. This will be achieved using Dual System Estimation (DSE).

4.3 The use of DSE requires a number of conditions to be met to ensure the minimisation of error in the estimates.

- Firstly, independence between the Census and CCS is required for an unbiased estimate. As a result the Census and CCS were operationally independent. Simulation work was undertaken to examine the impact of dependence between the Census and CCS on the ONC population estimates. These simulations demonstrated that, even for quite extreme levels of dependence, the impact on the population estimates is small provided both the Census and CCS have high response rates.
- Secondly, for an age-sex group within a postcode, the chance of a person being in the Census or CCS is assumed to be the same across all people. This is a reasonable assumption since the majority of postcodes are small and contain similar types of people.

4.4 Given these assumptions, DSE combines those people counted in the Census and/or CCS and estimates those people missed by both.

4.5 Once postcode level DSEs have been obtained, a modified regression equation is calculated for each age-sex group within each Hard to Count group to estimate the relationship between the census counts and the dual system estimates. This equation is then used to estimate design group underenumeration for each age-sex group in each HtC group.

4.6 The output from this process will be estimates of the population for each design group by age and sex, together with an indication of their accuracy. All of the subsequent processes described below will be consistent with these population counts.

Stage 2 – Local Authority District Estimation

4.7 Since many design groups consist of more than one LAD, estimates of the age-sex population for each LAD will need to be made. This forms the second stage of the estimation process.

4.8 Many LADs do not contain sufficient CCS postcodes to enable accurate direct estimates of underenumeration to be made. Small area estimation techniques can be applied to produce accurate LAD level population estimates.

4.9 The small area estimation technique used by the ONC uses information from the whole design group to model the undercount within LADs. The resulting population estimates are calibrated to the design group estimates, and their accuracy can also be calculated.

SECTION 5: IMPUTATION IN THE ONE NUMBER CENSUS

5.1 The aim of the One Number Census is the creation of a single, person level database. This database will be used to generate all statistical output from the Census. In the past this database only reflected those individuals enumerated by the Census (although in 1991 provision was made for imputing households identified by enumerators as not returning a census form). However, the information on the characteristics of missed persons obtained in the CCS will allow the creation of a database which represents our best estimate of the entire population, whether counted by the Census or not. This will be accomplished by a process of imputation, where additional individuals will be added to the Census database to account for those missed by the Census.

5.2 Imputation is already accepted as standard practice where questions are left unanswered or are found to be invalid. Improved procedures for this purpose in the 2001 Census have been described elsewhere (Vickers and Yar, 1998). The ONC will extend imputation to whole households and people, based on the data collected in the CCS and the ONC population estimates described in section 4.

5.3 The ONC population estimates define the number of people to be imputed along with some basic information about coverage patterns for other characteristics. However, it is important that we identify the detailed characteristics of those households and individuals missed by the Census. For instance, it is anticipated that the characteristics of people within entirely missed households will differ from those missed from within otherwise counted households. Accurate matching of Census and CCS data will allow the identification of these key characteristics. Once these features have been identified, prediction of both numbers and characteristics of missed individuals in the population not covered by the CCS will be possible.

5.4 The ONC imputation process can be summarised in three stages.

Stage 1 – Imputation of missed households

5.5 The first stage of the process imputes missed households and the individuals within them. A weight is allocated to each Census household corresponding to the likelihood of households of that type being missed by the Census. These weights are derived from an analysis of households missed by the Census, but captured by the CCS. The Census households are ordered by these weights and cumulative actual and weighted counts calculated. The cumulative counts are compared and, if the weighted count exceeds the unweighted count by more than 0.5, an imputed household is created with the characteristics of the current household.

5.6 Imputed households are placed into either a physical location identified by census enumerators where no response was received (e.g. absent household, refusal, non-contact), or into areas where similar households already exist.

Stage 2 – Imputation of missed individuals

5.7 The second stage of the imputation process focuses on individuals missed from households counted by the Census. A weight is created for each individual based on information obtained from analysis of the matched CCS/Census data, which reflects the likelihood of people with their characteristics being omitted from the census return for their household. These weights are used to impute individuals into the types of

households that are likely to have missed people from their Census return. It is this process that adds people to real households.

Stage 3 – Calibration to estimates of the population

5.8 A crucial requirement of the imputation process is that the overall distribution of imputed individuals and households should equal the ONC estimates of households and individuals missed by the 2001 Census. This calibration is accomplished by adjusting household and individual weights and by a final stage in the process, which either removes excess imputed individuals and households or tops up the Census database where necessary.

5.9 The result is an individual level database that represents the best estimate of what would have been collected had the 2001 Census not been subject to underenumeration. Tabulations derived from this database will automatically include compensation for underenumeration for all variables and all levels of geography, and will be consistent with the ONC population estimates.

SECTION 6: QUALITY ASSURANCE

6.1 A quality assurance process will be undertaken to ensure that the ONC population estimates are sensible and of the right overall magnitude. This will involve a series of quality checks, aided by data, grouped by age, sex and geography, drawn from the sources listed below.

Demographic Estimates

- Annual mid-year population estimates produced by the Population & Demography Division of the Office for National Statistics (ONS), the Population Statistics Branch of the General Register Office for Scotland (GROS) and Demography and Methodology Branch of the Northern Ireland Statistics and Research Agency (NISRA).

Administrative Records

- Numbers of people listed on patient registers supplied by the Health Authority in England and Wales, the Northern Ireland Central Health Index and the Community Health Index in Scotland.
- Numbers of British armed forces personnel supplied by the Defence Analytical Services Agency.
- Numbers of American armed forces personnel and their dependents supplied by the United States Air Force.
- Numbers of children for whom child benefit is being paid supplied by the Department for Work and Pensions and the Northern Ireland Department of Health and Social Services and Public Safety.
- Numbers of people drawing the state pension (or an alternative benefit) supplied by the Department for Work and Pensions and the Northern Ireland Department of Health and Social Services and Public Safety.
- Numbers of children at school supplied by the Department for Education and Skills, the Welsh Assembly, the Scottish Executive and the Northern Ireland Department of Education.
- Numbers of students in higher education supplied by the Higher Education Statistics Agency and the Northern Ireland Department for Employment and Learning.

6.2 Administrative records will be used only at an aggregate level. Some were identified as having the potential to provide information for particular subgroups of the population for comparison with those derived from the Census, such as school age children and armed forces personnel. Other identified sources, such as patient registers, cover the whole population.

6.3 The various data sources will be used to calculate a range of plausible values for the number of people of each sex within five-year age groups in each geographical area. The ONC population estimates, with confidence intervals, will be compared with these diagnostic ranges.

6.4 In addition, a range of descriptive information will be gathered to give a fuller picture of the area under consideration. This will include: information about the fieldwork for both the Census and the Census Coverage Survey; information on the ratio estimators from the regression equations within the ONC estimation process; details of the

adjustments made to the 1991 Census estimates; and feedback received from local authorities commenting on past mid-year population estimates from the ONS.

6.5 Demographic ratios such as the number of males compared to the number of females and the number of children compared to the number of adults will also be checked. This will be done on the same geographical basis as the checking of population estimates and again will compare the estimates with figures drawn from demographic estimates and administrative sources.

6.6 The processes described so far take place after estimates have been made for each design group and Local Authority District (LAD). There are further checks to be made after the imputation of households and individuals, as these processes do not explicitly control all characteristics of the records created.

6.7 After the imputation process, the distribution of single year ages within each of the five-year age ranges will be compared with distributions calculated from the demographic estimates and administrative sources. Estimates for specific population subgroups, such as students and members of the armed forces, will also be studied, comparing the estimates with the relevant administrative data.

6.8 All of this information will be presented to a panel consisting of specialists from each of the Census Offices and from Southampton and St Andrews Universities. They will consider the evidence for each design group and LAD before either accepting or rejecting the estimates.

6.9 In the event of any estimates being rejected at any stage of the process, a predetermined contingency strategy will be invoked. Details of the contingency strategy, plus further details of the quality assurance process, are included in the paper 'A Quality Assurance and Contingency Strategy for the One Number Census' (ONS, 2001).

ANNEX A: QUESTIONS FREQUENTLY ASKED ABOUT ONC METHODOLOGY

The following gives a set of questions frequently asked about ONC methodology.

Will the ONC also measure and adjust for overenumeration?
The ONC has been designed specifically to address underenumeration. Historically, UK Censuses have not suffered from overenumeration. However, work will be undertaken to assess the level of overenumeration in the 2001 Census and adjustments to the population estimates will be made if necessary.

What are the differences between the ONC in E&W, Scotland and NI?
The ONC strategy is broadly similar across the UK. However, GROS in Scotland and NISRA in Northern Ireland are both responsible for their own Census and CCS. Therefore, there are some differences based upon the differing local conditions. For instance, the Hard to Count index is constructed differently in both Scotland and Northern Ireland but its application within the CCS Design is identical. Since Religion is of particular interest within Northern Ireland, the estimation and imputation processes have been adapted to include this variable within the undercount models.

How does the ONC treat Communal Establishments and other special groups?
Large Communal Establishments (CEs), such as prisons, are outside the scope of the ONC. Small CEs, such as guest houses, were covered by the CCS and high level adjustments will be made to account for residents missed from counted CEs if necessary. In addition to this, the population residing within CEs and other special groups such as the armed forces, students and nurses and doctors living within hospitals will all be closely examined as part of the ONC quality assurance process.

How will the CCS address the problem of finding those people who do not want to be counted?
CCS training focused on achieving responses from people who were reluctant to take part in the survey. This is reflected in the high response rate achieved. However, if someone is determined not to be counted, it is highly unlikely that they will be captured by the CCS. Dual System Estimation includes an estimate for the number of people and households remaining uncounted by both the Census and the CCS.

How will the CCS deal with people who migrate between Census day and CCS fieldwork?
The questions on the CCS related to whoever was present on Census night. If a household or individual has moved, proxy information is collected where possible.

What is the strategy for resolving major discrepancies between ONC estimates and Mid Year estimates?
This strategy is covered in detail in the paper 'A Quality Assurance and Contingency Strategy for the One Number Census' (ONS, 2001).

Will all questions not asked in the CCS be imputed?
Yes. This is a consequence of the ONC imputation process which copies complete Census households and people.

Will the ONC result in a delay to outputs?
The main national and local results will be released, to a pre-announced timetable, as speedily as possible once processing is complete and the total population of the country has been determined. It is anticipated that the ONC will add six weeks to the overall processing timetable.

Will unadjusted Census counts be released? Will a measure of the extent of imputation be given in the outputs?
An indication will be given as to how much the Census counts have been adjusted for underenumeration. However, unadjusted Census counts will not be published.

ANNEX B: THE RESEARCH TEAM AND STEERING GROUP

Development of the ONC methodology was undertaken by a joint team of government and academic researchers, under the direction of **Professor Ian Diamond** and **Professor Ray Chambers** at the University of Southampton and including **Dr Ludi Simpson** of Bradford City Council. The Project Board, chaired by **Mr Tim Jones**, then Director of Methods and Quality Division at ONS, reported to a Steering Committee which oversaw the methodological development. The Steering Committee included representatives of the academic and local authority communities and a senior representative of the Australian Bureau of Statistics as well as officials from the UK Statistical Service. The members of the Steering Committee were:

Chair:
Dr John Fox (Chair), Group Director, Census, Population and Surveys, ONS (succeeded by **John Pullinger** from February 2000).

External members:
Dr Jim Cuthbert, Consultant, formerly Government Statistical Service
Professor Denise Lievesley, Director of the UNESCO Institute for Statistics
Professor Mike Murphy, London School of Economics
Mr Tim Skinner, Australian Bureau of Statistics
Professor Mike Titterington, Glasgow University
Mr Steve Turner, Tees Valley Joint Strategy Unit

Official members:
Dr Norman Caven, Registrar General for Northern Ireland
Mr James Meldrum, Registrar General for Scotland (succeeded by **Mr John Randall** from February 1999)
Mr Julian Calder, Group Director, Survey and Statistical Services
Mr Graham Jones, Director of Census
Mr Tim Jones, Director of Methods and Quality and Chair of the ONC Project Board
Ms Judith Walton, Director of Population and Vital Statistics
Dr Marie Cruddas (Secretary), Statistician, Census Division

REFERENCES

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. *J. R. Statist. Soc* ***A, 162****, 247-267.*

ONS (1998) Evaluation of the main objectives of the 1997 GB Census Test. AG(98)01

ONS (1999) A Guide to the One Number Census, Census Consultation Paper.

ONS (2001) A Quality Assurance and Contingency Strategy for the One Number Census.

OPCS (1993) Rebasing the annual population estimates. *Population Trends*, **73**, 27-31.

OPCS (1994) Undercoverage in Great Britain. *1991 Census User Guide* **58**. HMSO, London.

Simpson, S., Cossey, R. and Diamond, I. (1997) 1991 population estimates for areas smaller than districts. *Population Trends*, **90**, 31-39.

Vickers, P. and Yar, M. (1998) The development and evaluation of the donor imputation system (DIS) for the 2001 UK Census of population and housing. *Proceedings of the Joint IASS/IAOS Conference,* Mexico, September 1998.