Census 2001 Review and Evaluation

September 2005

One Number Census: Evaluation Report

ONS is carrying out a review and evaluation of the 2001 Census in England and Wales which will culminate in a Data Quality report being published, followed by a General Report.

Plans for individual reports on specific aspects of the Census operation and a timetable for release have been published.

Each report is written in isolation and is subject to amendments as processing progresses and further information comes to light.

Reports will be released on the ONS website in the form of a high level Executive Summary and a more detailed Evaluation Report.

Content Page

Census Customer Services
ONS
Titchfield
Fareham
Hants PO15 5RR

Telephone: ++44 (0) 1329 813800
Fax: ++44 (0) 1329 813587
Minicom: ++44 (0) 1329 813669
E-mail: census.customerservices@ons.gov.uk
Website: www.statistics.gov.uk/census2001

Census 2001 Review and Evaluation

Project Objective

The One Number Census (ONC) project had the goal of providing a methodology and processes to identify and adjust for the number of people and households not counted in the 2001 Census. The extent of this underenumeration was identified using a large survey covering approximately 320,000 households, the Census Coverage Survey (CCS). Standard statistical estimation techniques were then used to produce an adjusted and quality assured database from which the final census results were produced. These results also formed the new 2001 base for the Office for National Statistics (ONS) population estimates.

The 2001 Census results were the focus of considerable debate. From their publication in September 2002 through to September 2004, ONS undertook further work on both the ONC results and the previous population estimates resulting in adjustments to both. Publication of this evaluation report follows the completion of that work.

This report evaluates the processes and operations based on the evaluation work carried out both from 2001 to 2002 and between 2002 and 2004. It examines the lessons learnt from implementing the methodology and managing the processes, and provides recommendations on issues for future research. The recommendations will be fed into development plans for the 2011 Census.

The ONC project was designed and implemented across the whole of the UK, however this report evaluates the success of the process in England and Wales only.

Background

Every effort is made to ensure everyone is counted in a census. However, no census is perfect and some people are missed. This underenumeration does not usually occur uniformly across all geographical areas or across other sub-groups of the population such as age and sex groups. Therefore, it is accepted best practice across census taking countries to assess the extent of any underenumeration, typically by means of a post-enumeration survey (PES). The levels and characteristics of non-response to censuses and surveys are changing in much of the developed world so assessing the scale and type of underenumeration is becoming increasingly important.

In the 1991 Census, 3.8 per cent of the population was missed overall, made up of 1.6 per cent imputed into absent households and a further 2.2 per cent estimated to have been missed. Although this compared favourably with other countries, what was notable was that underenumeration was significantly greater in certain types of areas and age-sex groups. In 1991, for the first time, records were imputed for households that were known from the Census operation to exist but had not returned a form. This resulted in the imputation of 806,000 persons, or 1.6 per cent of the population. This source of undercount was included in the census outputs and was therefore part of the raw census counts.

However, the 1991 Census post-enumeration survey, the Census Validation Survey (CVS), was not sufficient to measure the full extent and distribution of underenumeration. There were many reasons for this, but the primary issue was that its design was not fully independent of the census. The difficulties caused by the CVS' inability to identify the full extent of the undercount took time to resolve and four different resident population counts were available following the 1991 Census. Initially 'raw' census counts were published, including the 1.6 per cent imputed persons. Then a set of provisional estimates was published, up-rating the census count in line with the CVS. Subsequently a decision was made, in England and Wales, to base the national population estimates on demographic estimates, rather than up-rated census counts, and to produce consistent counts at Local Authority (LA) level using a mathematical model. This model used the rolled-forward 1981 Census-based population estimates to provide target sex ratios, against which the census counts were compared and thus revised. These third counts were then subject to a final revision following further analysis. The demographic analyses provided evidence that a further 2.2 per cent of the population had not been counted by the 1991 Census, in addition to the imputed persons.

At geographic levels below Local Authority level no official estimates of underenumeration were produced, although those produced by the Economic and Social Research Council funded project 'Estimating with Confidence' have been widely used (Office of Population Censuses and Surveys (OPCS) 1993, 1994; Simpson, Cossey and Diamond, 1997). The official census tabulations therefore remained inconsistent with the population estimates.

During the extensive consultation carried out prior to the 2001 Census, users of census data were clear that they wanted one definitive set of estimates to come out of the Census. Furthermore, given the unexpected problems with measuring underenumeration in the 1991 Census, users demanded a robust method for estimating this source of bias so that they could have confidence in the accuracy of census results. The ONC project provided a way to address these issues. It led to the strategy of a completely redesigned PES and full adjustment of the census database for the estimated underenumeration, so that all published census statistics would add to the national estimate of the population. It built on the advances made in previous censuses. In 1981, automatic editing and imputation of missing variables was introduced, followed by absent household imputation in the 1991 Census (imputing records for households known to exist from the Census operation) and in 2001 to the addition of imputed households and people estimated to have been missed by the ONC. Furthermore, the mid-year population estimates were to be based on the ONC population estimates and there needed to be a simple link between them.

The aim was also to have one final set of census estimates, rather than the four different estimates in 1991. That aim was not achieved. The reasons for this and the problems it caused are discussed in the 'Assessment and Lessons learnt' section.

Overall Methodology

The ONC methodology was devised by ONS in consultation with experts mainly from Southampton University. The project was set up in 1996 following the completion of the 1991 Census evaluation, and the emerging need to address the underenumeration issues that users had raised. The ONC Steering Committee, which included a number of external experts and representatives from Local Authorities, Academia and other National statistical Institutes, oversaw and guided the development of the methodology and management of the project. Full details of the methodology can be found at www.statistics.gov.uk/census2001/pdfs/oncguide.pdf. All other Steering Committee papers can also be found on the National Statistics website, detailing all the methodological research undertaken throughout the project at www.statistics.gov.uk/census2001/IntroOneNumber.asp#ONC Steering Committee. The day to day progress of the ONC project was overseen by the One Number Census Project Board, which included senior ONS management and representatives from other areas of ONS as well as some external experts.

The ONC strategy was relatively straightforward. A focused post-enumeration survey was employed as the tool for measuring underenumeration by providing an independent enumeration of a sample of areas. The results of the PES were then matched, at individual level, to the corresponding 2001 Census data. The combined census and PES information was used to produce an estimate of the numbers of people missed by the Census. The people estimated to have been missed were then added to the database by copying existing individuals. This type of strategy is commonplace amongst census taking countries, with the exception of the database modification. For instance, the US Census Bureau has a long history of using a PES to assess underenumeration; the first such survey had followed the 1950 Census. However, the output has always been to inform users of the accuracy of the count rather than to correct the count itself.

The implementation of such a strategy in a census context was not straightforward, and a number of fundamental questions, alternatives and assumptions were tested. Some alternative strategies were pursued during the development (such as using administrative records within a triple system approach, or modelling underenumeration down to Enumeration District level). The research into these alternatives are not described in full detail here - the series of ONC Steering Committee papers and papers presented at various national and international conferences reflect all of the work undertaken to develop the chosen approach. The key stages and features of the final ONC methodology are described below.

Underpinning the ONC methodology was the assumption of independence between the count of population given by the Census and CCS. For the ONC to work well, there must be no systematic relationship between the chance of a household or individual being enumerated in the Census and of being captured by the CCS. However, dependence was found to exist to a sufficient extent that adjustment was needed – see Assessment and Lessons Learnt section.

Key Stages of the One Number Census Methodology

The One Number Census methodology has 5 key stages:

- Census Coverage Survey;
- Matching the CCS and Census;
- Estimation of populations for each EA and LAD;
- Imputation of records for households and individuals; and
- EA/LAD Quality assurance.

Stage 1 - Census Coverage Survey

The 2001 PES focussed solely on coverage and was known as the Census Coverage Survey (CCS). It involved a complete re-enumeration of a sample of postcodes with the aim of interviewing all households and achieving a 100 per cent count of households and persons in the selected areas. The CCS did not include large Communal Establishments, and a different method was used to Quality Assure their populations as described in the Quality Assurance section. The practical implementation of the survey is not discussed here, since a separate project developed the field methodology and processes. An evaluation of the CCS project can be found at www.statistics.gov.uk/census2001/proj_ccs.asp

The CCS differed from the 1991 PES in a number of respects. Firstly the quality of census responses was measured in a separate survey, enabling the 2001 CCS design to focus purely on measuring underenumeration. Secondly, the CCS sample size was much larger. The 1991 PES was not able to provide sufficiently accurate population estimates due to its small sample size, which was related to the length of interview needed for the collection of information regarding quality of response. It was recognised that if underenumeration was to be measured properly, the CCS sample size had to be large. Lastly, the differential nature of underenumeration had to be captured by the CCS, and therefore estimates of reasonable precision must be achievable across geography and demography.

Estimation Areas

Ideally, the sample design would be based around Local Authority Districts (LADs) so estimates of underenumeration could be obtained directly for each. However, there are 376 LADs in England & Wales and the sample size required to support such direct estimates would be unfeasible. It was estimated that such a survey would require a sample size of over 1 million households, which was considered impractical and too costly. Therefore, England & Wales was divided into 101 areas, each with a population of about 500,000 persons. These areas were known as Estimation Areas (EAs) and were made up of whole Local Authority Districts (LADs) or groups of smaller LADs. A balance had to be struck here between the size and homogeneity of the EAs and the practical and policy aspects of grouping areas. This grouping allows direct estimates of reasonable precision to be obtained for each of these EAs, the cost being that some assumptions would have to be made about the patterns of LAD underenumeration to enable estimates for the LADs to be calculated from the sample. The sample would therefore have to be selected to ensure that each LAD had some sample areas. These issues are discussed further in the estimation methodology section.

Sample Design

A sample of postcodes was drawn from within each EA. Therefore, in effect, there were 101 independent samples. Postcodes were chosen as the sampling unit for both statistical and practical reasons - there was no perfect household sampling frame, and most households know the postcode to which they belong. The main disadvantage of sampling postcodes was that very little information was available about them with the exception of the approximate number of addresses. This meant that direct stratification of postcodes, with

the exception of size, was not possible. Therefore, the sampling strategy adopted within each of the EAs was a two-stage sample design that linked each postcode to a 1991 Census Enumeration District (ED), for which 1991 Census information was available.

First, a sample of 1991 EDs were chosen stratified according to a Hard to Count (HtC) index, as underenumeration was expected to be again disproportionately distributed across areas. This stratification is described in more detail below. The design attempted to minimise the sampling variance across the strata to enable the production of population estimates of an acceptable precision. The 1991 EDs were also stratified by key age-sex groups to further control the spread of the sample, using 1991 Census data.

Secondly, once the 1991 EDs had been selected, the second stage of the sampling was the random selection of between 3 and 5 postcodes within each of the selected 1991 EDs, depending on the size of the postcodes within the ED.

Hard to Count Index

The HtC index was constructed from 1991 Census variables found to be associated with underenumeration, and which had a strong correlation with non-response in the major census tests leading up to 2001. It was constructed nationally into 3 levels - 40 per cent of England and Wales were defined as 'easy', 40 per cent as 'medium' and the top 20 per cent as 'hard'. The objective of using such an index was to ensure that the sample was spread across all area types so that underenumeration could be measured robustly. Prior to 2001, there was relatively little knowledge of hard to count patterns, and this index represented the best stratification that could be constructed at the time. Therefore it was important to ensure that the CCS was carried out in both easy and hard to count areas. The hardest to count areas were over-sampled to reduce any differential in variances across the three strata. The index was designed to be fairly robust to change in the patterns of undercount between 1991 and 2001. However, large scale change in an area may have meant that the use of the 1991 information was not robust (see 'Assessment and Lessons Learnt' section).

Sample Sizes

The sample size issue was an important part of the research programme. This work concluded that to obtain estimates of reasonable precision (approximately 1 per cent Relative Standard Error, the estimated standard error of an estimate expressed as a percentage of the estimate itself) for each Estimation Area population total, a national sample size of approximately 4,000 Enumeration Districts, which are equivalent to 300,000 households, would be required. This chosen sample size was balanced against the costs of undertaking such a survey.

This total sample size was proportionally allocated to each hard to count stratum within every Estimation Area, based upon a national sample design which was used only to determine the sampling fraction for each of the hard to count strata. These fractions were then applied to the Estimation Area ED populations to determine sample sizes, although a minimum constraint of 7 EDs within a HtC stratum was applied. The application of the design and the sampling constraints resulted in sampling fractions of 3.4 per cent for the HtC 1 stratum (easy), 3.7 per cent for the HtC 2 stratum (medium) and 4.5 per cent for the HtC 3 stratum (hard), which reflect the different variability within each strata.

Between 3 and 5 postcodes were then selected at random within each of the EDs, resulting in a total sample of 16,400 postcodes. The sample sizes achieved can be found at www.statistics.gov.uk/census2001/onc_qa/ccs_sample_sizes.asp

Survey Methodology

CCS fieldwork began three and a half weeks after Census Day. It was a paper based interviewer survey, which asked only basic household and demographic questions to enable a short questionnaire, and therefore a short interview. The survey methodology was focused on achieving the highest response rates possible, in particular it included a calling strategy to maximise contact rates. Comprehensive training was provided for the CCS interviewers, again focused on achieving 100 per cent enumeration. The CCS field-force, which consisted largely of inexperienced interviewers, was

supplemented by 80 experienced ONS interviewers, allocated to teams in areas expected to pose particular problems in obtaining responses.

An important principle of the CCS was its independence from the Census. This was an essential requirement for the estimation of undercount within the ONC, an issue discussed more fully in the estimation section. Without independence there was a risk that the CCS did not identify those households and persons that the census missed. To achieve this in practical terms, CCS interviewers were not provided with the Census address lists for their areas. Maps of the CCS postcodes were supplied to interviewers and they attempted to interview every household at every address within the postcode, in effect independently re-enumerating the area. As they did not have an address list, they were also required to check the boundaries of their postcodes to ensure they correctly included every household.

In addition, efforts were made to ensure that the CCS did not have any 'contamination' effect on the Census and also that the CCS did not utilise any information gathered in the Census. Measures put in place were:

- the sample postcodes were kept confidential;
- census managers were not able to work on the CCS;
- CCS managers were not allowed to work on the Census; and
- census enumerators were allowed to act as CCS interviewers, but were prevented from interviewing in the same area they had enumerated and were not informed of their postcodes until their census fieldwork had finished.

An evaluation of the CCS project can be found at www.statistics.gov.uk/census2001/proj_ccs.asp

Stage 2 - Matching the CCS and Census

Once the Census and CCS operations had completed, and the data collected in each captured, the CCS household and person records were matched with those from the Census using a combination of automated and clerical matching. The outcomes from the matching would feed directly into the estimation of the numbers of households and persons missed by the Census. In order to ensure the estimates based upon this matching were unbiased, the aim of the matching process was to be as accurate as possible, with a target 'incorrectly unmatched' rate of 0.1 per cent. If a person was in both census and CCS, it was important that these records were matched - missed matches inflate the final population estimates. For instance, a 0.1 per cent missed match rate leads to a 0.1 per cent positive bias within the population estimates arising from the estimation strategy. Therefore a five-phase strategy that focused on ensuring accuracy through multiple levels of quality assurance was developed.

The five key phases of the matching process for each unique postcode were:

Phase 1 - Exact Matching

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

Phase 2 - Probability Matching

Probability matching was then applied to records that were not matched at Phase 1, using the same variables. A probability weight was assigned to each pair of 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 was identical with the exception of one detail, which may have been due to a recording error, then a high probability weight was assigned. Any household pairs with a high probability weight were linked and the individuals within them compared. Only households with very similar characteristics were matched at this stage.

The initial probability weights used were derived from the 1999 Census Rehearsal. These weights were then updated at key stages throughout the matching process, using 2001 matched data.

Phase 3 - Clerical Resolution

Pairs of households and individuals with a reasonable level of agreement, but not sufficiently similar for probability matching, were presented for clerical resolution. At this stage operators simply determined whether the pair of records shown constituted a matching pair or not. They were not expected to search for matching records.

Phase 4 - Clerical Matching

This involved a clerical search for any census records corresponding to CCS households and individuals that remained unmatched. Operators had access to images of Census and CCS forms to assist with decision making and were supported by a Computer Assisted Matching System (CAMS), a system developed specifically for the ONC matching process.

Phase 5 -Quality Assurance

All matches created by the matchers were checked by expert matchers until the matchers had correctly matched (or unmatched) 200 records. From this point only 10 per cent of their matches were quality assured, although if a match was found by others to be incorrect at this stage, the matcher's threshold was reset to 200. Matchers were unable to designate any record as being unmatchable. If they could not find a match for a record they would defer the record to an expert matcher for checking.

The Expert matcher's work was quality assured by the Quality Assurers on the same basis as matchers, 100 per cent of all matched and unmatched records were checked until they had correctly designated 200 records. Once this level had been reached only 10 per cent of the Expert's work was checked. The Quality Assurers would also check all unmatchable records across the whole Estimation Area to ensure that no potential matches had been missed. Further checks across multiple EAs were also carried out in cases where the CCS sample lay close to the border.

The Quality Assurers also checked types of match which were deemed to be difficult to make. This included one to many matches, which occurred when the Census recorded a property as containing one household, and the CCS recorded the same people, but living as two or more households within the property, or vice versa.

The final element of quality assurance was the double matching strategy. This involved independently re-matching an EA and comparing the outputs. This enabled any discrepancies between the two versions to be investigated before the output was sent for estimation. This ensured that any errors identified were corrected prior to Estimation, and enabled the identification of any training or system issues which required rectifying. Initially it was planned to only double match the first 10 EAs, however due to changes in the processing schedule it was possible to double match all but the last few EAs. These last EAs were subjected to an additional quality assurance run through, to make sure that the quality of output was consistent with those that were double matched.

Stage 3 - Estimation of populations for each EA and LAD

The ONC has three estimation phases.

- Within Phase one, Dual System Estimation (DSE) methodology was used to combine 2001 Census and CCS counts to estimate the true population in the sampled CCS postcodes.
- Phase two generalises these DSE counts from the sampled areas to the whole EA population using a modified standard survey estimator. This produced underenumeration estimates by age and sex for each Estimation Area in England and Wales, the level at which the CCS was designed to provide direct estimates of an acceptable precision.
- Phase three of the ONC estimation was the allocation of the estimated EA underenumeration to the component LADs, ensuring consistency with the EA estimates produced from Phase two. These LAD estimates were used as the constraining population totals for the ONC imputation process, and are the primary subject of the Quality Assurance process described later.

Phase 1 - Dual System Estimation (DSE)

Estimates of the total population in postcodes covered by the CCS were based on a methodology known as Dual System Estimation (DSE), a standard method for estimating underenumeration. This was the approach used by the US Census Bureau following both the 1980 and 1990 US Censuses. Dual system estimation combined the Census and independent re-enumeration (the CCS) counts to estimate the true population, allowing for people missed by both the census and the CCS, in the CCS sample areas. Although the method was theoretically straightforward, in practice it can have some problems:

a) The DSE assumes that the probabilities of being counted by either or both the CCS and the census were homogeneous across the target population. This was unlikely for most populations.

b) Unbiased estimation requires statistical independence between the Census and the CCS individual capture probabilities. This was not achieved - see below.

c) It was necessary to accurately match the two data sources to determine whether individuals on the lists were counted on one list or two. Errors in matching become biases in the DSE.

These conditions need to be met to ensure unbiased estimates. Assumption a) was approximated by splitting the population into groups by age and sex within the sampled postcodes of each HtC category and therefore, the DSE is applied at a very low level of aggregation. This, along with operational independence as described above, was intended to help assure that assumption b) was well approximated. Research into the impact of dependence (i.e. failure of assumption b)) showed that the overall approach was robust to some dependence between the Census and CCS, provided the response rates achieved by the Census and CCS were high. Details of this research can be found in Steering Committee paper SC0103 [www.statistics.gov.uk/census2001/pdfs/SC0103.pdf]. However, that assumption of high response rates was not achieved in practice in some areas, leading to the need to adjust for dependence. This problem and its solution are discussed later.

Some matching error (Assumption c)) was unavoidable. However, the matching process as described above was designed to ensure that matching errors, particularly missed matches, were minimised.

Phase 2 - Estimation Area Estimation

For the second stage of estimation, generalisation of the DSE counts from the sampled areas to the whole population can be carried out using standard survey estimators. The adjusted DSE count (or ratio) for each sampled postcode was used as the 'dependent' variable in a zero-intercept regression model, which links this count with the census count for that postcode. This ratio model was based on the assumption that the 2001 Census count and the dual system adjusted CCS count within each postcode are proportional to each other. Given that it was known from the 1991 Census that undercount varies by age and sex as well as by local characteristics, a separate ratio model within each age-sex group for each HtC category within each Estimation Area was used.

The standard ratio estimator described above was found to be sensitive to model failure in the presence of small census counts, and there were potential biases when estimating large postcodes. Further research resulted in modifications to the standard estimator to provide protection against these problems, as well as a strategy for dealing with outliers. Details of this research and the estimator can be found in Steering Committee paper SC0003A - www.statistics.gov.uk/census2001/pdfs/SC0003A.pdf

The modified ratio estimator was used to produce Estimation Area population estimates for each age-sex group in each HtC group. Diagnostic information was also produced to allow detailed assessment of the results for each individual estimate. If sample size issues or problems were identified, some HtC groups and/or age-sex groups could be combined to provide more robust (although synthetic) estimates.

Estimates of sampling error were also constructed to facilitate the production of confidence intervals associated with the population estimates. Since the design and estimation strategies were relatively complex, a jackknife type estimator based on successively dropping Primary Sampling Units was employed.

Phase 3 - Local Authority District Estimation

Many EAs consist of more than one LAD. With the exception of a few large LADs, LADs did not contain sufficient CCS postcodes to enable accurate direct estimates of underenumeration to be made and therefore small area estimation techniques were required.

The application of standard direct estimators, such as those used to estimate the Estimation Area totals, yield unbiased estimates. However, for smaller areas such as LADs, where the sample sizes are smaller, these estimators have very large standard errors. This led to the development of techniques that borrow strength from related areas to make indirect estimates that increase the effective sample size and thus decrease the sampling error level associated with the estimates. These techniques are based on the idea that a statistical model fitted to data from a large area (i.e. Estimation Area) can be applied to a much smaller area to produce a synthetic estimate for that area. The problem with this approach was that while the estimators based on the large area model have small variance they are usually biased for any particular small area. A compromise involves the introduction of small area specific effects into the large area model. These allow the estimates for each small area to vary around the synthetic estimates for those areas. This helps reduce the bias in the estimate for a small area at the cost of a slight increase in its variance.

As described in the previous section, direct estimation for the Estimation Area was based on a simple ratio model linking the 2001 Census count for each postcode with the DSE-adjusted CCS count for the postcode. This model was extended to allow for the multiple LADs within an Estimation Area by including a fixed LAD effect. The LAD adjusted synthetic model used was one that includes an overall age-sex effect and an LAD specific effect to distinguish between the LADs. These LAD effects are assumed to cancel out at Estimation Area level. This means that the model consisted of an overall Estimation Area age-sex slope parameter, and an adjustment to this slope to take account of the differences between the LADs. The approach was implemented separately for each HtC index stratum within an Estimation Area, and resulted in LAD population estimates by HtC and age-sex group. These were then passed to the Quality Assurance and imputation processes. See Steering Committee Paper SC0003B [www.statistics.gov.uk/census2001/pdfs/sc0003B.pdf] for more detailed methodology.

Stage 4 - Imputation of records for households and individuals

To adjust the Census database for estimated underenumeration, records for households and individuals estimated to have been missed by the Census were imputed. Imputation is already accepted as standard practice where questions are left unanswered or are found to be invalid; the ONC extended this to impute wholly missing households and individuals. The method was based on a calculation of coverage weights that reflect the propensity of individuals and households to be enumerated in the census. These weights were then applied to the counted population to drive a donor imputation system that copied existing households and individuals based upon the coverage weights.

The ONC population estimates defined the number of people to be imputed into each EA and LAD along with some basic information about underenumeration patterns for other characteristics.

However, it was important that the detailed characteristics of those households and individuals missed by the Census were identified so that a comprehensive analysis could be produced. Thus there was a distinction made between the two ways in which individuals were missed: either the household was missed or contact was made with the household but not all members were counted. Household and individual coverage weights were constructed using the matched Census and CCS data, and these determined the key characteristics of missed households and individuals. The remaining characteristics were then filled in using a donor imputation system, which replicated real records based upon the coverage weights. The final database provided a complete set of responses for each household and person, whether directly responding or imputed by the adjustment for underenumeration.

The ONC imputation process had four main phases:

- derivation of coverage weights; this made use of information collected in the CCS to estimate the extent of and characteristics associated with undercount in the Census. The matched Census and CCS data were used to model underenumeration and to calculate coverage weights. Two models were developed: one for household coverage and another for individual coverage within counted households;
- imputation of missed households (and the individuals within them) based on the household level coverage weights; imputed households were geographically placed into either a physical property identified by census enumerators where no census response had been received (known as a dummy form), or, if all of these dummy forms had been filled up, into areas where similar households already existed;
- imputation of missed individuals, i.e. individuals missed from households where the household had been counted by the Census. Donor individuals for each type were then selected and used to impute individuals into the types of households that were likely to have missed people from their census return. This process added people to real households; and
- calibration to estimates of the population. This ensured that the overall distribution of imputed individuals and households was the same as the ONC estimates of households and individuals missed by the 2001 Census.

The result was an individual level database that represented the best estimate of what would have been collected had the 2001 Census not been subject to underenumeration. Tabulations derived from this database automatically include compensation for underenumeration measured by the CCS, for all variables and all levels of geography.

Stage 5 - EA/LAD Quality assurance

An extensive quality assurance process was undertaken for the ONC estimates. The procedures described here formed a fundamental part of the ONC methodology. They were intended to deal with the possibility that the results of the ONC estimation may not be plausible, either in some areas of the country or in the nation itself.

The overall strategy was to use the highest quality data sources: administrative, demographic and qualitative, against which to compare the census results. This was done in a consistent manner across all Local Authorities and demographic groups, allowing comparisons and patterns to be analysed. Information and intelligence was gathered from all of these sources (including the Census and CCS) to provide a full picture of the area and data concerned. The administrative sources were not intended to replace census estimates at any time. The volatility of these sources had been established in the ONC development phase, when investigations found none of them reliable enough to use as a 'triple system estimator'. The research showed that inaccuracies within the administrative data at individual level would lead to significant biases within the estimated populations. With hindsight, ONS should have identified the need for further investigation in some areas. For example, use of council tax has been found to be helpful by the matching and LA studies

Quality Assurance Process

The process itself involved a series of quality checks for each EA, aided by data, grouped by age, sex and geography, drawn from the annual mid-year population estimates produced by ONS and aggregate level administrative data. The administrative data sources and those supplying them were: patient registers (National Health Service Information Authority); people of pensionable age receiving benefits and child benefit (Department for Work and Pensions); children at school (Department for Education and Skills and the National Assembly for Wales); students in further and higher education (Learning Skills Council and Higher Education Statistics Agency); Prisoners (Home Office); British armed forces personnel (Defence Analytical Services Agency) and American armed forces personnel and their dependants (United States Air Force).

The various data sources were 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, were compared with these 'diagnostic ranges'. These diagnostics provided the best indicators of population that were available prior to the census. It was not expected that the census would be adjusted to these diagnostics. A range of descriptive information was also gathered to give a fuller picture of the area under consideration, for example information about

the fieldwork for the Census and the CCS and feedback received from LAs commenting on past mid-year population estimates from the ONS. Demographic ratios were also calculated. All of this information was presented to Quality Assurance panel meetings consisting of specialists from ONS and Southampton University. The QA panel used consistent criteria to determine whether or not further action was required based upon the evidence for each EA and LAD.

Contingency

Where there was a difference between the ONC estimates and the diagnostic ranges, extensive checks of the ONC results and diagnostic ranges were to be undertaken with respect to, for example, sample sizes and outliers to explore whether contingency action was required if any issues were identified. The qualitative information gathered from the administrative sources, Census and CCS fieldwork, data capture processes and ONC matching and estimation processes, for example, were also to be examined.

There was also a predetermined contingency strategy in the event of the CCS not providing a robust estimate of undercount, such as in a situation where the fieldwork was poor. Information from similar LAs that had already passed the QA process was to be used to make adjustments. This process was called 'borrowing strength'. The similar LAs, or borrowing strength areas, for each LA were set out and agreed as part of the ONC consultation. The borrowing strength contingency strategy was invoked for two LAs. See the 'Assessment and Lessons Learnt' section for further details.

In the event that the national quality assurance of census results failed, there was also a predetermined national contingency strategy. This was developed to address the situation where evidence suggested that the degree of dependence between the Census and CCS was too high. A diagnostic would be, for example, that all estimates were a little low (or high). During the development of the methodology, it was assessed that the risk of this happening was extremely low, as it was believed to be extremely unlikely that the CCS would fail to detect underenumeration everywhere. The contingency strategy in this case adjusted the Dual System Estimator with plausible target sex-ratios, or other demographic information, in a similar way to the methods used to adjust the 1991 Census.

Consultation

A wide-ranging user consultation process was carried out by the ONS, as it was recognised early in the ONC project inception that user acceptance was vital. As a result the ONC methodology has had the most peer reviews and user consultations of all National Statistics - an example are the sixty publicly available research papers representing over 5 years of work on the National Statistics website. This process communicated the research programme underpinning the ONC to users of the UK censuses. They generally welcomed and supported this methodology as both improving the quality of the census counts and population estimates and ensuring that the published data are both consistent and transparent as far as users are concerned.

Census users were directly consulted at several stages in the methodological development process through census user group meetings and special workshops, including:

- the Steering Committee which included key representatives and experts from the academic and user community;
- the consultation process with census user groups;
- the paper Brown et al, 'A Methodological strategy for a one-number census in the UK' published in the Journal of the Royal Statistical Society Series A (JRSS-A) in 1999;
- the paper Steele et al, 'A controlled donor imputation system for a one-number census' published in the Journal of the Royal Statistical Society Series A (JRSS-A) in 2002;
- the paper Holt et al, 'Risk in official statistics - a case study of the one-number census project' published in the Journal of the Royal Statistical Society Series D (JRSS-D) in 2001;
- a variety of papers presented at international conferences;
- 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' and its 2001 update;

- the consultation with local government Statistical Liaison Officers in Winter 2000 on the areas that should be used for the contingency strategy;
- The Autumn 2001 Consultation paper 'A Quality Assurance and Contingency Strategy for the ONC';
- a workshop devoted to the One Number Census project held at Leeds University in May 1998; and
- special workshops held in conjunction with Census Output Consultation Roadshow meetings.

Results

This section presents some brief high level results from the ONC process. Further results have been published on the National Statistics website at www.statistics.gov.uk/census2001/IntroOneNumber.asp

At a national level, 6.1 per cent of the total population estimate were estimated to have been missed by the 2001 Census, implying a response rate of 93.9 per cent. A breakdown of the age and sex of these individuals is in Table 1, along with a measure of the relative width of the 95 per cent Confidence Interval around the population estimate. Further detail for smaller geographical areas down to LADs can be found at www.statistics.gov.uk/census2001/annexb.asp#byarea

Table 1 - Response rates and Relative 95% Confidence Intervals for Male and Female 5 year age bands.

		Males		Females	
Age group	All persons response rate %	Response rate %	Relative 95% Confidence Interval %	Response rate %	Relative 95% Confidence Interval %
All	93.9	93.3	0.2	94.4	0.2
0-4	90.4	90.4	0.4	90.4	0.4
5-9	92.4	92.4	0.4	92.4	0.5
10-14	93.6	93.4	0.5	93.7	0.4
15-19	92.6	92.5	0.4	92.7	0.4
20-24	87.8	87.0	0.7	88.6	0.6
25-29	88.8	87.6	0.6	89.9	0.5
30-34	91.7	90.5	0.5	92.8	0.4
35-39	93.3	92.2	0.4	94.3	0.3
40-44	94.6	93.9	0.4	95.4	0.3
45-49	95.7	95.3	0.3	96.2	0.3
50-54	96.6	96.2	0.3	97.0	0.3
55-59	97.1	96.7	0.3	97.4	0.2
60-64	97.3	97.1	0.3	97.5	0.3
65-69	97.6	97.5	0.3	97.7	0.2
70-74	97.8	97.8	0.3	97.8	0.2
75-79	97.8	97.8	0.3	97.8	0.3
80-84	97.2	97.0	0.4	97.3	0.3
85 and over	97.2	96.8	0.5	97.3	0.4

As a result of these estimates, around 1.3 million additional households were imputed (5.9 per cent of the total household estimate). These imputed households contained 2.5 million imputed individuals (4.9 per cent of the total population estimate). A further 0.6 million individuals were imputed into counted households (1.2 per cent of the total population estimate). Note that these results are those from the ONC process alone, and do not take account of the 193,000 and 107,000 subsequently estimated to have been 'missed', as detailed in the next section.

Assessment and Lessons Learnt

Overall strategies for estimating the population

When first results were published in September 2002, it was the best estimate that could be achieved by the ONC process in the time available. However, there was a large gap of 1.1 million between those ONC estimates and the mid-year population estimates (MYEs), which were based on rolling forward the 1991 Census counts by measuring annual population flows. Some LAs had large revisions to the population counts for 2001 and throughout the previous decade. This led to concerns being expressed by some LAs about their lack of confidence in the ONC results. In retrospect, ONS needed to have been better resourced to more immediately respond to those concerns.

Initially the gap between the ONC estimates and the MYEs was judged to be the result of both the demographic adjustments to the 1991 base after the Census and the known weakness in migration statistics. However, ONS was unable to explain empirically its conclusion that the difference was primarily the result of unmeasured migration outflows and gave a commitment to further work and analysis.

This further work resulted in revisions to the external migration series in June 2003 and revisions to the 2001 mid year population estimates based on analysis of the longitudinal study in September 2003. Further final revisions were made based on the Local Authority Studies in September 2004.

The explanation of the difference between the 2001 Census based mid-2001 population estimate and the rolled forward population estimate is summarised numerically in the table below, together with links to the reports that describe each of these components.

Intercensal Discrepancy	**1,140,000**	**Link to relevant report**
1991 Adjustment	351,000	www.statistics.gov.uk/downloads/theme_population/Meth_PopES_82_2000.pdf
Migration Adjustment	305,000	www.statistics.gov.uk/downloads/theme_population/Methodology%20for_Revised_International_Migration_Estimates.doc
Longitudinal Study Adjustment (and others in September 2003)	193,000	www.statistics.gov.uk/about/methodology_by_theme/Revisions_to_Population_Estimates/default.asp
Unexplained difference	**291,000**	
LA Population Studies	107,000	www.statistics.gov.uk/downloads/theme_population/2001Censuslapopulationstudies.pdf
Longitudinal Study Consequential Adjustment (and other adjustments in September 2004)	-25,000	www.statistics.gov.uk/downloads/theme_population/LAStudy_LS_ConsequentialAdjustment.pdf
Remaining unexplained difference	**209,000**	

After making the adjustments already published, the majority of the 1.1 million differences were explained. However, the table above shows that there remain about 209,000 unexplained differences. This is likely to be largely attributable to remaining difficulties in estimating migration accurately and issues associated with the usual residence definition.

The first two of the four adjustments revised the population estimates downwards. The other two adjustments, however, the Longitudinal Study adjustment and the LA Population Studies, revised the population estimates upwards because the studies found that the Census had missed people in ways that the ONC had not been able to compensate for.

The Longitudinal Study (LS) links census and vital event information for 1 per cent of the population of England and Wales from 1971. The longitudinal study adjustment examined linkage at each Census after 1971 (that is 1981, 1991 and 2001) and the profiles of those not found on each occasion. It also looked at how much of the cumulative losses in the study, as a result of unrecorded migration and mismatched census records, can be explained by the repetition, in subsequent decades, of the type of losses that occurred in the first decade of the study. It then looked at how much of what cannot be explained in this way could be accounted for by available figures on census undercounts in 1991 and 2001. This analysis concluded that the ONC had underestimated the population by about 193,000 probably due to within-household dependence. A full description of this study can be found at www.statistics.gov.uk/about/methodology_by_theme/Revisions_to_Population_Estimates/downloads/Methodology_for_revision_to_mid-2001.pdf

The LA Population Studies analysed all LAs in England and Wales to identify where the 2001 Census population estimates might have been at risk and, from this analysis, 32 LAs were selected for further study to determine the significance and scale of that risk. The results of the LA studies show that the One Number Census (ONC) worked well in most areas, but that there were a few cases where it was not able sufficiently to adjust for exceptional circumstances. ONS concluded that better estimates of the population could be made in 15 areas; Manchester and Westminster were the authorities with the largest changes with much smaller revisions to another 13 LAs. For a full description of these studies see www.statistics.gov.uk/downloads/theme_population/LAStudy_FullReport.pdf

The analysis showed that there was a need for revisions to the 2001 Census based population figures of around 107,000 for England and Wales as a whole (64,000 of which were further revisions in addition to the results provisionally announced previously). The revisions are just outside the 95 per cent confidence interval for the population as a whole, as estimated from the ONC sample in 2002, of +/- 0.2 percent (or +/- 104,100).

There will always be issues about the ability of a census estimate to accurately reflect the true picture particularly for small areas with high underenumeration. More could have been done in advance to focus on the estimated confidence intervals, the assumptions underpinning them, particularly with regard to response, and the likely impact if those assumptions were not met or were pushed to their limits. It may also be that the term 'one number census' raised unrealistic expectations that everything, both the Census and the population estimates would be perfect first time.

Both the Statistics Commission and the Local Government Association have published reviews which conclude that the methodology used in 2001 was the best available and no alternative approach would have produced more reliable results overall. Following an interim report in 2003 on the 2001 Census in Westminster, the Statistics Commission published a final report in January 2005 in which they said 'The 2001 Census was, in most respects and on the evidence available, a success. It produced robust local estimates across most of the UK. But for a relatively small number of areas, particularly in some inner city ones, there is now evidence that the methods used were not equal to the challenges they faced'. In the report, the Statistics Commission sought to draw out lessons to be learned for the future, some of which are also identified in this report and some of which are broader issues around address lists and cross- public sector working. ADD LINK

It is possible that the twin objectives of the ONC, to form the base for both the Census results and the mid-year population estimates, led to one detracting from the other. At the time this was felt to be very important, to avoid the difficulties caused in 1991 with four different population counts. The Census is a micro level dataset, measuring the population at one point in time, and consistency within the dataset is important. The population estimates, however, are a macro level dataset which may need to be revised in the light of further knowledge. One option for the future would be to produce ONC-style Census results plus a first set of population estimates and then have a second phase where time can be taken to assess the results thoroughly, with the view to (if necessary) producing a final set of population estimates perhaps later.

Further discussion of options for improving population estimates for the future can be found in the study 'A demographic statistics service for the 21st century'. www.statistics.gov.uk/about/Methodology_by_theme/downloads/Demographic_Statistics_Service.pdf. As a result of this, ONS has begun a programme of work looking to improve population statistics. This includes reducing the risk that, when population estimates based on the 2011 Census become available, users are not again faced with large differences between these and the mid-year population estimates based on the 2001 Census. The ongoing programme of work includes: identifying the implications for the 2011 Census; continuing to work with local authorities on research to improve population estimates; implementing the recommendations of the National Statistics Quality Review on International Migration. A central theme of the work is improving our understanding of sources such as administrative data and investigating the scope for utilising those sources to help improve population estimates at the national and local authority level, and also increase knowledge of their reliability.

Statistical

Stage 1 - Census Coverage Survey

Sample Design

The CCS sample design was a success in that it provided the information for estimating the level of underenumeration, given the assumptions on which it was based. The design strategy successfully allowed the measurement of undercount across all area types and for every Local Authority District where the assumptions underpinning the design were not violated. There are some issues with the design which mainly have arisen from its implementation, which are discussed here.

Practical Sampling problems

There were some issues discovered in the field that required the sample of postcodes to be altered. The different reasons for these are described below:

- Two census managers were shown the CCS sample for their areas, jeopardising the independence of the Census and CCS. An alternative sample was drawn within these areas to overcome this risk.
- Some postcodes consisted entirely of non-residential addresses. This was expected, as there are known problems with the PAF in identifying non-residential addresses.
- Some postcodes were in areas where the Team managers (TM) felt it was too dangerous to send interviewers. Obviously these cases needed to be carefully handled, and replacing the postcodes was only carried out if the TM was very sure that the area could not be enumerated safely.
- When carrying out the fieldwork checks, there was some confusion over how to classify halls of residence. This was mainly because of the definition of Communal Establishments (CEs) - the concept of managed accommodation was difficult to judge in these cases because there was no manager. Therefore some TMs classified them as large CEs (and thus out of scope) and others defined them as multiple households.
- A few postcodes were entirely demolished and others were unfinished new housing. This was again expected as the PAF was not always up to date.
- There was one sampled postcode in a tower block that had been evacuated because it was due for demolition.

These problems highlight the need for close links between the sampling and fieldwork teams, and the requirement for a back-up set of sampling units which can quickly replace sampled areas if problems such as these arise.

The HtC index and sample balance

The Hard to Count (HtC) index was in general an appropriate first level of stratification even though, in the absence of any suitable alternative, the data used to construct the index was from the 1991 Census. There were some issues in areas that had undergone substantial redevelopment since the 1991 Census, as shown by the evaluation work carried out in the Local Authority Studies. In such areas, the stratification may have been less efficient, which led to a greater risk of estimation bias because of poor sample balance. The solution to this problem in the LA studies was to post-stratify, which reallocated the sample and population into strata that are formed using the more up-to-date census data. During the ONC, a crude strategy for achieving this was trialled in areas where it was believed the HtC index may be out of date, but the results showed little difference. The feasibility of post-stratification should be further explored for the future. In addition, the availability of more up-to-date small area data sources is increasing and it is important to utilise this information within any future design to reduce the risk of estimation bias.

Stage 2 - Matching the CCS and Census

Implementation

The Matching process defined by the original specification was very successful. The automatic probability matching and search facilities included within the system worked extremely well. The original methodology was augmented as knowledge was gained though the process. The probability weights used in the probability matching stage were updated as EAs were completed. This enabled a higher proportion of records to be matched automatically, without clerical intervention. 73 per cent of households were automatically matched (59 per cent of households required no clerical checking and 14 per cent requiring clerical approval).

The Computer Assisted Matching System (CAMS) performed extremely well throughout the matching process. CAMS worked in Microsoft Windows and was easy to navigate and use. A comprehensive training program and manual were written and well received by the matching staff. The system was robust, and rarely caused any problems despite its complexity.

Developing CAMS in-house proved to be a great success. It provided the flexibility to solve problems and make changes quickly and easily where necessary. Several changes were made to CAMS in response to requests from matchers. These changes included the location of buttons on the matching screens and shortcuts to carry out searches. Another significant change made to the system was the introduction of Management Information files. These recorded all actions carried out in CAMS, who carried them out and at what stage. This output enabled detailed analysis of the performance of staff and the effect of adjusting the probability matching weights.

The Matching team operated at its peak with 19 staff. All matchers and expert matchers were employed through an agency contract. Using an agency gave the flexibility to increase or change the staffing profiles at short notice, which became important as the timetable became compressed due to processing order changes. The contract with the agency worked extremely well and most of the agency staff stayed for the full term of the project.

A single EA took between one day and one week to match, depending on the complexity of the area. Inner city areas tended to be more complicated to match than rural areas. This depended largely on the proportion of students, ethnic minorities, and transient populations that required matching. ONC matching was completed on schedule, despite changes to the processing order.

Methodology

The most significant difficulty experienced in the Matching process was achieving the required level of accuracy. The target for matching accuracy was an error level of below 0.1 per cent. The error rate was measured by comparing the two matching outputs from the double matching strategy, and identifying inconsistent matching outcomes. This assumes that matching error was purely due to clerical error rather than systematic methodological errors, which are almost impossible to measure without a third source of information that provides the true match. In order to meet the accuracy target for each EA, matching protocols were introduced based on best practice identified from the most successful matching staff. These protocols outlined the searches that each level of staff should have completed in order to classify a

record as unmatchable. Although it proved challenging, the accuracy targets were met for all EAs, based on the measurement provided by the double matching. This was a significant achievement.

Stage 3 - Estimation of populations for each EA and LAD

Implementation

The estimation process itself worked well. Systems were in place in time and were easy to understand and run. The outputs enabled effective analysis of the estimation process, and provided summary analyses for the QA panel. Expertise was built up quickly and the close knit team was able to share experiences and analysis techniques.

One run of the estimation system took approximately 30 minutes. Analysis and checking of the estimation outputs took around half a day, although this was longer if additional analysis had been requested by the QA panel (see the Quality Assurance section). The impact of the compressed processing timescales was mitigated by the secondment of additional staff, which enabled the Estimation process to finish on schedule. The seconded staff were able to pick up the procedures quickly, due to the well designed system and the knowledge of the Estimation team. It would have benefited the ONC process if some of the analysis requirements of the QA panel had been anticipated in advance, for example early indications of the national/regional pattern of undercount as the EAs were completed, to reduce the level of ad hoc analysis required. However, all requirements could not be foreseen and therefore it was important to employ a flexible, expert team in this area as was the case.

Methodology

The estimation methodology worked well to provide a consistent analysis of the underenumeration identified by the CCS. The outcomes met prior expectations, and provided a rich source of information on underenumeration patterns across the country and within each Estimation Area. One area of the estimation methodology identified during the LA studies as being a potential source of bias was the outlier strategy. This aspect of any future methodology should be fully explored and cater for all types of outlier.

Dependence

The ONC relied almost completely on the CCS as the source of information on underenumeration and subsequent work has shown this to be a shortcoming, especially when the undercount is much larger and more complex than expected. Investigations during the development of the ONC found that none of the other sources which might have been used, for example in a triple system estimator (Census, CCS and a third national population source), was reliable across the country as a whole.

For this reason, Dual System Estimation was adopted and underpinning this methodology was the assumption of independence between the count of population given by the Census and CCS. For the ONC to work well, there must be no systematic relationship between the chance of a household or individual being enumerated in the Census and of being captured by the CCS. However, intuitively it does seem likely that the chance of being missed by the CCS was higher if the person in question was missed by the Census. For example, if certain types of people are deliberately avoiding being counted by any official exercise (illegal immigrants, men living with single mothers claiming benefits, fathers avoiding paying child support), they are more likely to be missed by both the Census and CCS, despite interviewers best efforts.

The measurement of such 'dependence' was extremely difficult; internationally there has been some theoretical work but no research, to our knowledge, has attempted the practical application of measuring dependence. The research programme paid a great deal of attention to the impact that dependence might have on the population estimates arising from the ONC. This showed that for reasonably high coverage levels in both Census and CCS the impact of some dependence was not significant. As a result, the risk of needing to invoke a national contingency adjustment due to dependence was assessed to be low. However, despite the operational measures to ensure independence and the high response rates in the CCS, dependence was found. Therefore, a strategy was developed to measure dependence to assess this source of bias.

This initial strategy for measuring dependence was based on adjusting the Dual System Estimator with plausible target sex-ratios, or other demographic

information. However, the quality assurance process had shown changes in the observed sex ratios over time in the censuses of 1971, 1981, 1991 and 2001, so that confidence in an adjustment based on an assumed sex ratio was diminished. Thus in 2002, when dependence was found and the initial sex-ratio contingency strategy was rejected, alternative strategies were explored.

A second strategy involved the matching of dummy forms, or households where the Census or CCS found no residents. However, the quality of dummy form information collected during the 2001 Census fieldwork was such that highly accurate matching was not possible. Therefore a third strategy was implemented. The Postcode Address File (PAF) and Census address file were combined to produce an estimate of the expected number of households. The number of households estimated by the ONC process was compared with the expected number of households to test whether the number of households estimated by the ONC was less than the number expected. Where this was the case, the dual system estimates were re-calculated using the observed levels of dependence. Because of the instability of the PAF, the measurement of dependence was carried out at Government Office Region (GOR) level and by Hard to Count group. A fuller description of the methodology can be found here - www.statistics.gov.uk/census2001/pdfs/dependency_paper.pdf. This work indicated that there were significant levels of dependence between the Census and CCS, and therefore a decision was taken to adjust the ONC estimates for dependence.

The estimate for dependence in England and Wales added an additional 230,000 to the ONC population estimates. This addition was not a top up - it required complete re-running of the ONC estimation and imputation with modified dual system estimation formulae. There are many lessons to be learnt from this. Firstly, there must be recognition that a national adjustment such as this is likely to be needed for future exercises and any contingency resource should be protected. Secondly, critical assumptions must be subject to rigorous testing and sensitivity analyses to provide a realistic assessment of their impact if violated. Thirdly, an assessment of dependence should form an integral part of any future dual system estimation methodology, as the ONC has indicated that the assumption of independence is not realistic. Lastly, it will be important to find ways of making better use of administrative data as a third source, as this will reduce the potential for this source of estimation bias although care must be taken not to introduce new biases. Greater input of local knowledge will be helpful with regard to use of administrative data (see Quality Assurance section).

Overcount

The ONC primarily focused on measuring undercount. Overcount has not traditionally been a problem within the censuses of England and Wales, and therefore measurement of it was given a low priority. The CCS collected information about potentially overcounted individuals by asking whether there was anywhere else they might have been counted in the census. A matching study was undertaken based on the responses collected, resulting in an estimate of less than 0.1 per cent overcount. Further studies have indicated that this might have been an under-estimate of the level of overcounting. Whilst the overall level was not significant, it is likely to be more of an issue in the future, especially if multiple methods of responding to the census are offered.

Stage 4 - Imputation of records for households and individuals

Implementation

The process as a whole worked generally well, although the development of the system was problematic.

The methodological development was hampered by the late delivery of rehearsal data, which resulted in a compressed timetable for developing the 2001 imputation system ready for live running. With new and innovative methodologies it was always possible that unforeseen issues would arise. When the first 2001 Census data were delivered, the imputation system required wholesale changes, outlined here - www.statistics.gov.uk/census2001/pdfs/sc0201.pdf. These changes were necessary to enable the system to calibrate exactly to the Local Authority age-sex population estimates. The main issue was the link between household underenumeration and individual underenumeration causing a serious over-imputation of persons within the households estimated to have been missed. The solution to this issue evolved gradually as at each step it was necessary to understand and analyse the impact the changes made before implementing further changes to the system.

The resulting system performed well for the majority of processing, although there were some cases that required further system changes to allow convergence. Each area took approximately 4 or 5 hours to run, although the areas requiring most imputations needed 24 hours. The analysis arising from the system was good, and provided sufficient information on the patterns of underenumeration for all variables. This high level Quality Assurance for key variables was extremely useful for providing assurances that the imputed records were, at a high level, plausible.

Methodology

The basic imputation methodology worked well to provide a database that was fully adjusted for underenumeration.

However, a problem with the imputed data was identified when detailed census outputs were being produced. The imputation system had used the same donor household and people for successive imputations and placed them within a small area, the worst case being where around 200 copies of the same household had been created. This skewed the distribution of all variables (e.g. single year of age). In extreme cases the skewness was likely to be noticeable at ward level for single year of age. This presented an unacceptable risk of disclosing personal information and therefore had to be corrected.

The problem arose because of the nature of the system - being a constrained process, designed to copy and place households and persons to meet prescribed age-sex and household totals. It uses dummy forms as the first choice for placing households, and if there are no suitable dummy forms it places the copy of a household in the same 2001 Enumeration District as the original household. Therefore, if a household has been copied repeatedly and no dummy form was available for placement, the household was placed in the same area as the original. The system did not attempt to minimise the number of repeat donors, and since it was a constrained system it was sometimes forced to copy a particular household more than once. This was not examined until the problem was identified, as the simulation work and testing had not highlighted it as an issue. The need to constrain the number of times any individual record can be used as a donor is one of the lessons ONS are carrying forward to future Census work.

With such a new and innovative methodology, it was always possible that unforeseen issues would arise. However, thanks to the dedicated and knowledgeable teams across the census areas, the issue was investigated and resolved quickly and professionally. The lesson to be learnt here is that close links between the methodological development and system development are essential for innovative methodologies. For 2001, the team that developed the methodology also developed the system. This allowed problems to be identified and resolved quickly.

Stage 5 - EA and LAD Quality Assurance

Operationally, the quality assurance process was successful: nothing on its scale had been done before in any other country and international experts have commented on its thoroughness and professionalism. There are issues of the ever-present time/quality balance, however; the inevitable pressure to produce census results can conflict with the need for in-depth demographic analysis.

The subsequent work by ONS has demonstrated that more analysis was required to identify the need for further adjustments in a small number of areas. The difficulty was that these issues were not resolved until 2004. For the 2011 Census it will be important to build in more in-depth analysis and other improvements that resolve these issues more speedily, including taking account of other sources and local knowledge as appropriate.

There are some lessons arising from the QA process and these are described below.

Population subgroup adjustments

The quality assurance process included analysis for each local authority of a number of specific population subgroups known from 1991 to be prone to underenumeration. These were full-time students, home armed forces, foreign armed forces and their dependants and prisoners. The estimates for these subgroups were compared with data from other official sources to determine whether the results were plausible. This was especially important for students, armed forces and prisoners as the CCS did not cover large communal establishments. Adjustments were made

where significant under-enumeration was identified for these groups, using data and qualitative information sometimes from the institutions themselves (e.g. MoD data on the numbers of people paying mess fees at each of the large communal accommodation blocks). In total this resulted in 27,000 persons being added to the estimates for England and Wales as a whole.

Additional data and analysis

More information about the Census data collection would have helped the QA process. The lack of reliable quantitative information from the Field Management Information System (FMIS) hampered some judgements (discussed more fully in the Data Collection Development evaluation report).

Other concerns were that local data and knowledge were not sufficiently used in the QA process. Much of this goes back to the fact that there was no adequate national third source, and the requirement to be consistent nationally. For the future it will be important to find ways of making use of the administrative data that are now available as a third source at the data collection and estimation stages, building on the way these data were used within the 2001 Census QA process, and to have a greater pool of local knowledge throughout.

The Quality Assurance process also placed demands on the work areas that supplied the data. Often, the panel would ask for additional analyses relating to a particular source or for some exploratory work to confirm or reject a hypothesis. The ONC team was able to cope with this work through careful management, the establishment of a dedicated analysis team, and the commitment and expertise within ONC, the wider Census areas, and other areas of ONS that provided support. Such resource needs to be built into the plans for the 2011 Census. Also, more senior management resource should be planned into the QA process, since it was within this framework that the headline census results were subjected to rigorous evaluation. The use of external expertise within the QA should also be considered.

Contingency measures

The sub-national contingency strategy was required for two LAs; Sheffield because of problems in the field with the CCS and Shepway because of doubts about the balance of the CCS sample in that area. In the former, strength was borrowed from the agreed borrowing strength LAs, and in the latter it was possible to borrow strength from the other LAs within the EA. These measures were broadly successful, although their acceptability has been questioned by at least one of the LAs concerned. For the future, this may be an area where local knowledge may be able to support a sub-national contingency strategy.

Publication of QA information

Although a substantial amount of information on the QA process was provided, ONS had not expected or planned for the huge demand for information from LAs and others. With hindsight we might have predicted this better and therefore been able to respond more systematically. For the future, provision of QA information and the ability to deal with a large demand for information should be built into the plans.

Project operation and organisation

Operationally the ONC was a major success. A new and ground-breaking methodology was successfully implemented and for the first time (both here or anywhere in the world) all census results were adjusted for measured underenumeration. Three key statistical elements (matching, estimation and imputation) were developed and operated to a high degree of professionalism. However, there were significant hurdles that required careful management and commitment from the ONC, Census and Demography teams. Operational difficulties with other aspects of the Census, such as delays in receipt of postback forms are described in other relevant evaluation reports.

The Census enumeration itself did not go as planned. The final postback rate of 88 per cent was remarkably high, but it led to significant disruption in the postal flow, with forms being received quite sporadically in many areas. The poorest response rates were way below

planning calculations, and in some cases violated the assumptions on which the ONC sampling methodology was based. No post-enumeration survey could be expected to estimate and adjust, with a high level of precision, an undercount of 25 per cent or more. Without the CCS to estimate undercount, this would have put at risk the value of the census in the majority of areas.

Organisationally, the project management of the whole ONC was successful and was well integrated with the rest of the Census operations. Relations with the academic team at Southampton were extremely good and created an excellent partnership that helped to develop skills within ONS.

The methodological development was ground breaking not only in the field of statistics but also the organisation, management and collaboration with experts both across government and the user community. The implementation was also a success due to the excellent teamwork within ONS between the ONC project, Census division and Population and Demography division. It is important that these aspects of the project are not overlooked for any such future ventures, as the human resource aspects of an innovative statistical exercise are just as vital as the formulae and systems that make up such a process.

Project Timetable

There is a perception amongst some users that the ONC led to a much longer timetable and to the results being delayed. The ONC process added around three months elapsed time to the original output timetable, mainly to carry out national level quality assurance and as contingency for methodological problems. In the event the publication of first results was delayed from the planned August 2002 release by a further month, mainly due to the impact of the delayed processed data deliveries. The three months that had been included within the ONC project plan as a contingency was used to absorb these census processing delays, leaving very little contingency for methodological problems. Even so, through careful management of the issues outlined above within ONS and the adaptability of ODPM, the project ultimately was able to provide the results that allowed the population estimates to be delivered in time for the Standard Spending Assessment (SSA).

In implementing any major innovations, there is always a greater risk of unexpected problems and there were two such occurrences in the ONC. One was when the imputation method caused a risk of disclosing personal information, and the other was the need to adjust for dependence. These are discussed in more detail later in this report. Both of these took significant time and resource to successfully resolve, although they did not cause additional delays because of staff's commitment, expertise and flexibility. Two unexpected major hitches is probably less than might be expected for the level of innovation involved but it does illustrate the need to build significant contingency into the timetables for innovative processes, to a greater extent than in this case. With hindsight, it may be that the additional three months in the timetable was not long enough to allow sufficient contingency time for the unexpected.

A related lesson is that the ONC system development was carried out too late because the methodology was not finalised until a late stage. There were unexpected delays in delivery of the rehearsal data, essential for finalising the methodology and also testing and preparing the ONC systems for live data. This resulted in the adaptation of timetables, escalation of risk levels and the final methodology being signed off later than originally planned. Furthermore, timetables were accelerated to prepare and test the systems once the rehearsal data were available to ensure they were ready for the live data.

Following the delays to the basic processing timetable, it was recognised early on by ONC management that a compressed timetable, as a result of delays in data capture processes, would require additional resource if critical publication dates were to be met. Staff were seconded from other areas to address this. The seconded staff had to be trained but were able to pick up the procedures quickly, due to the clarity of the process and the expert knowledge of the core team. The Matching team (matchers and expert matching quality assurers) were all agency staff, which gave the flexibility to increase or change the staffing profiles at short notice. This allowed the throughput to be increased to cope with the compressed timetable, eliminating further delays.

Conclusion

ONS recognises the lessons from this project, which were, in summary:

- contingency planning is critical, particularly in the event of delayed data delivery or the requirement for a national adjustment;
- innovative projects inevitably carry risks in their implementation, and it is essential to ensure that appropriate resources are in place throughout development, implementation and live running;
- critical assumptions that underpin methodologies, such as independence, must be carefully evaluated and robust measures put in place;
- the potential for administrative data to be used in the future requires a full programme of research;
- the demand for information arising from such a project must not be underestimated and should in future be built into planning; and
- the quality assurance of estimates arising from the census requires the collection of robust management and local information and greater senior management resource.

The key messages arising from this evaluation that feed into future developments are summarised as:

- the ONC project was a good model for conducting an innovative large scale statistical exercise;
- the human resource aspects of a large statistical undertaking are vital for success;
- the information generated by the ONC process provide a rich source for developing future censuses, and this needs to be harnessed effectively;
- a Post Enumeration Survey based approach is not likely to be able to measure all types of undercount in future censuses - further research is needed to address this; and
- greater use of administrative data and local knowledge will reduce the risk of implausible or biased population estimates.

It was a huge achievement to carry out the ONC, particularly what was accomplished in the timescales. The further studies undertaken have provided lessons to be learnt but have also demonstrated that the ONC approach provides a platform on which to build for the future.

The success should be judged in the context of the Census operation that actually took place, with levels of underenumeration quite variable and sometimes extremely low, as the postal flow and the later census payroll problems created difficulties that the One Number Census could not always answer. However, in the context of a census, with large migrant flows in the last half of the decade, and considerable differences in residence status, it is clear that the Census proper on which the population statistics as a whole are founded, needs to be broadened in its basis as a platform for adjustment and analysis.

Census 2001 Review and Evaluation

Census Topics	Target Dates for Release
Legislation	Published
Non-Compliance (Executive Summary Only)	Published
Data Needs	Published
Geography	Published (Executive Summary)
Publicity	Published
Data Collection Development	Published
Data Collection Support	Published
Census Coverage Survey	Published
Processing	Published
Annex: Quality of Data Capture and Coding	Published
Downstream Processing	Published (Executive Summary)
Data Quality	
- Question non-response rates	Published
- Disclosure Control (Executive Summary only)	Published
- Data Validation (Executive Summary only)	Published
Edit & Imputation	Published
One Number Census	
- Quality Assurance	Published
- Lessons learnt (Executive Summary only)	Published
Output Policy	Published (Executive Summary)
Output Production	
- Part 1:Review of Output Released to date	Published (Executive Summary)
- Part 2:including Sample of Anonymised Records (SARs)/Origin Destination Matrices	Published
Census Access	Published
Programme Management	Published (Executive Summary)
Quality Report	Published
General Report	Published

Please note that the dates for release of individual evaluation reports noted above are target dates, and therefore subject to change. For the latest information please visit www.statistics.gov.uk/census2001/reviewevaluation.asp