

THE CENSUS COVERAGE SURVEY – THE KEY ELEMENT OF A ONE NUMBER CENSUS

Richard Pereira¹

ABSTRACT

The key measure of Census quality is the level of response achieved. In recent censuses worldwide this level has been in the high nineties percent. This was also true of Censuses in Britain in 1991 (98%). However, what was particularly noticeable about the Census in Britain in 1991 was the differential response rate and the difficulty in effectively measuring this. The One Number Census programme was set up in the UK to research and develop a more effective methodology to measure and account for underenumeration in the 2001 Census. The key element in this process is the Census Coverage Survey - a significantly larger and redesigned post-enumeration survey.

This paper describes the planning and design of the survey with particular emphasis on the implementation of the proposed field methodology in practice. It also provides a high-level overview of the success of the survey.

KEY WORDS: Census; enumeration; coverage; survey

1. INTRODUCTION TO THE ONE NUMBER CENSUS

It was estimated that the 1991 Census in Great Britain covered 97.8% of the population. Whilst the level of underenumeration (2.2%) compared well with other countries, it was not uniform 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).

Arriving at these estimates took some considerable time mainly because it became clear that the post-enumeration survey had failed to find many of those missed by the Census. During this period, several different population counts were made available including the published Census count, a count adjusted for underenumeration based on the findings of the post-enumeration survey (known as the Census Validation Survey (CVS)) and a count adjusted for underenumeration based on demographic analysis.

Annual population estimates in Britain are based on the Census allowing for underenumeration in the Census. In 1991 it was decided that the demographic analysis estimate should form the basis of the national population estimates (effectively meaning that the national population estimate was still based on the 1981 Census). At local authority level (c.120,000 population) population estimates were based on the 1991 Census counts uprated for underenumeration using a mathematical model and constraining the local estimates to the national estimate. The Census counts remained unadjusted and the main Census output tabulations inconsistent with the population estimates.

The differential level of underenumeration raised some questions as to the validity of the Census results themselves. While for the majority of applications, relative comparisons could be confidently made between areas and population subgroups, it was difficult to be definitive about this.

The UK Census Offices have, through the One Number Census (ONC) project, developed procedures that will avoid such problems after the 2001 Census by integrating the Census counts with the estimated levels of underenumeration. The result is a fully adjusted Census database at the individual record level (including imputed households and persons to reflect the estimates of underenumeration) that will provide a new base for the mid-year population estimates at the Local Authority District (LAD) level and for all Census output so that all statistics add to "One Number", the national estimate of the population on Census Day.

¹ Richard Pereira, ONS, Segensworth Road, Titchfield, Hampshire, UK, PO15 5RR

2. RESEARCH INTO METHODS FOR MEASURING UNDERENUMERATION

Extensive research was carried out into methods for measuring underenumeration in the 2001 Census including demographic analysis, administrative records and post-enumeration surveys. It quickly became clear (ONS, 1995) that a post-enumeration survey was the most effective approach but that the survey would need to be redesigned and larger than used in 1991.

The 2001 Census post-enumeration survey was known as the Census Coverage Survey (CCS)). Although coverage surveys were carried out after the 2001 England and Wales, Scotland and Northern Ireland Censuses, there are some minor differences. For the sake of brevity, the rest of this paper will describe the approach and results in England and Wales.

3. DESIGN OF THE CCS

When the CCS was designed to feed into the ONC the following key aspects needed to be covered:

- it had to be entirely independent from the Census itself - essential for the ONC methodology (a dual system estimation process) to work;
- there had to be full geographical coverage of underenumeration;
- it had to allow census population counts to be adjusted for underenumeration at national, local and small area level;
- it had to comprise a nationally representative sample of postcodes drawn from all LADs to enable population estimates to be made for all districts;
- the burden on the public had to be as low as possible;
- it had to be done efficiently within the Census budget

3.1 Design and sample selection

One aim of the ONC is to estimate the population by age and sex for all LADs. However, the sheer number of these meant that to apply the CCS design to each one would result in a prohibitively large sample size. Therefore LADs were grouped together to form design groups, each of which had a population of approximately 500,000 people. The CCS design was applied separately to each of these, and they were constructed to enable estimates of LAD underenumeration to be made indirectly.

The CCS is a postcode based survey but since unit postcode level information (beyond the number of addresses) is generally not known information from the 1991 Census was used as a proxy to select 1991 Enumeration Districts in a first stage selection.

Underenumeration is disproportionately distributed across areas and so the EDs were stratified according to a Hard to Count (HtC) index. This index ranges from 1 to 3 and is constructed from census variables known to be associated with underenumeration (multi-occupancy, private rented accommodation, young migrants from the 1991 Census data).

Within each of the HtC categories the 1991 EDs are clustered on the basis of their age-sex distribution. The EDs are clustered into relatively homogenous groups and the sample selected from these clusters. Five postcodes from each ED were chosen at random to participate in the CCS and the aim was to fully enumerate each of these. Simulations were carried out under this design and proved that it was sound.

Postcodes were clustered to form interviewer workloads for practical purposes. Each workload, to be worked by a pair of interviewers, was forecast to consist of between 100 and 200 households - large communal establishments such as hospitals were not covered principally because of the burden on survey respondents so shortly after the Census.

3.2 Sample size

The overall sample size was determined as a balance between the accuracy required for the ONC estimates and the cost of the exercise. Much simulation work was carried out (ONS, 1998) to determine the optimal size in this balancing act. The final findings indicated that a sample of around 20,000 postcodes (300,000 households) would provide an acceptable level of accuracy, giving relative errors of around 1 per cent, for the population of the design groups, or around 0.1 per cent for the national population.

3.3 Timing

The CCS was planned to take place immediately after the Census fieldwork had finished, some three and a half weeks after Census day. This timescale was deliberate so that the number of people moving house between the Census and the CCS would be as small as possible - it is estimated that just under one per cent of households move each month. It also meant that the Census was still fresh in the minds of the members of the public, reducing the likelihood of them forgetting exactly who was in the household on Census day, whilst ensuring that Census activity was finished.

For the same reason, the fieldwork period was planned to be very short - some three weeks, encompassing four weekends, the times considered most likely to find householders at home and available for interview.

4. TESTING THE APPROACH

Initial tests for the survey began in 1997 when a pilot exercise was conducted in Brent, part of London. This exercise surveyed a small number of households, only 450, but it successfully established the general feasibility of the survey. In particular it highlighted the importance of timeliness in conducting the CCS and provided a first insight that interviewers working together in pairs worked well.

A larger scale test took place in 1998 in Southampton. This test surveyed a larger number of households (around 2,000) and as well as reconfirming the findings from the 1997 pilot it resulted in a redesigned and easier to use interview form.

In 1999, as part of the Census rehearsal, the CCS was conducted in five representative areas around the country. In this exercise 18,000 households were interviewed using the proposed design to find out whether it was feasible. This was a successful operation with area response rates varying between 80 and 95 per cent. It proved that the design and basic methods that had been researched were practical and appropriate for this type of survey. At the same time it collected some valuable information for use in refining the field methodologies for the real exercise in 2001.

In particular:

- response rate data could be used to help determine an effective calling strategy;
- debriefing feedback from field staff could be used to ensure we had the right staffing structure and that training was targeted correctly;
- the survey data obtained could be used to ensure the questionnaire was reasonable.

A sequence of further small-scale local tests were conducted to check that refinements made to the field methodology as a result of the findings of the rehearsal were workable.

5. FIELD METHODOLOGY

The field methodology was designed such that we could obtain a high response to the CCS in terms of households interviewed and people found within these households, and ensure we would find people missed by the Census. This was challenging in view of:

- the voluntary nature of the CCS;
- the independence required from the Census
- the time limited nature of the fieldwork period.

This section describes some of the field methodology procedures that were developed to overcome these issues.

5.1 Finding people missed by the Census

The CCS had a number of features built in to ensure it would collect information on people who had been missed by the Census.

The interviewer-driven nature meant that dialogue between interviewer and interviewee occurred during which there was a greater chance of either party realising that somebody was being incorrectly excluded than would be the case for a self-completion questionnaire.

The survey form included a number of probe questions that were proven in the rehearsal to reduce the likelihood that somebody will be accidentally excluded. The householder was specifically asked whether there was anybody working away, in the armed forces, in hospital, on holiday or at university on Census day - those from experience believed to be most likely to have been missed.

It was important to have a solid base of survey experience in the staff so that they could readily recognise the importance of capturing information on everybody they needed to and use techniques to do so. This experience base was accomplished by seeking recruits with previous proven survey experience. Additionally, approximately 100 ONS interviewers were employed to help out in an advisory and assisting role in areas considered to be more difficult to enumerate. These trained social survey interviewers were able to use their valuable survey experience and training to good use in probing for missing people.

5.2 Team and Pair working

This was found, in the rehearsal, to be a key part of a successful CCS. There was a real team spirit present with all members pulling in the same direction and creating a real synergy to complete the survey work.

The team-working ethos was deliberately promoted for the 2001 operation. The training events centred around this with syndicates being set up from the start with members working a team as they would in the field. The training courses were residential events for Team and Field Managers thereby giving them a chance to communicate on a social as well as a working level.

Team meetings were scheduled in throughout the course of the survey period. At these, the team of interviewers working to a Team Manager were able to get together and discuss openly the way things were going, constructively solving any problems encountered.

Interviewers working in teams of 20 were paired up for conducting the survey work, each pair covering a workload of between 100 and 200 households. This pairing provided each interviewer with moral support throughout.

The great advantages of the team approach are

- there is greater resilience to resignations and loss of field staff
- there is a real momentum that is maintained in the work with team enthusiasm pushing it forward throughout the fieldwork period
- there is a health and safety benefit with team members being able to support others in the event of difficulties
- it promotes better communications and openness for discussions of problems encountered.

5.3 Data collection

The CCS was an interviewer conducted survey so that the public interface could be foremost. This was felt to be important for two reasons:

- To fully explain to members of the public exactly what the survey was all about, how important it was and how everyone who participated would be helping, thus increasing the willing response rate;
- To ensure the questions were all comprehended correctly by members of the public and therefore standardised across all interviews.

The questionnaire used was kept deliberately short. It asked a subset of Census questions - only basic demographic characteristics of everyone living in the household, questions about the accommodation and simple relationship information. As a result of this short and simple questionnaire the interviews were able to be carried out on the doorstep, with no need to enter the property, within five to ten minutes per house depending on the number of residents. The shortness of the interview kept the burden on the public down, and again helped to ensure a good response.

From the outset it was clear that a paper based collection system would be more appropriate than a computer-based collection system for the interviewers. The interview time was short, only five or ten minutes, and so the use of laptops and associated set-up time would be inappropriate and involve entering the interviewees house - something that might be seen as an intrusion. The associated logistic and financial implications of providing 4,000 interviewers with IT equipment for a short piece of work added weight to this decision.

5.4 Maps and property listing

To maintain the independence of the CCS from the Census no forms of address lists were used by interviewers. Instead, the first two days of the fieldwork period were used to list all the properties within the sampled postcodes. Interviewers were provided with maps outlining the rough extent of the included postcodes and were trained to contact households adjacent to but outside the marked boundaries to confirm their postcode. Postcodes in the UK act as delivery point only and do not necessarily have well defined geographical boundaries.

This on-the-ground property searching helped the interviewers to ensure that all households within the given postcodes were counted and included on the list of properties to interview. As a result of this method of property listing many properties within postcodes were found that otherwise would have been missed.

5.5 Calling strategy

Interviewers available time over the three week fieldwork period was naturally limited and so to make their time most effective a well-defined calling strategy was required. The calling strategy derived had to meet the following criteria:

- it had to lead to an effective use of interviewer time as measured by the number of interviews achieved over the survey period;
- it had to allow substantial efforts to obtain interviews from difficult cases and not introduce any bias into the nature of response;
- it should allow timely, and standardised, assessment of progress in obtaining responses;
- it must be robust to variation in workload size and response rates;
- it should not require provision of detailed information on workload size or characteristics from HQ to field-staff.
- it should be easy to explain to field-staff.

Response data from the 1999 rehearsal exercise were used to model the expected response at each survey day under different calling strategies. The key element was to work out an appropriate time and day for an interviewer to make a repeat call on a household if nobody was found for interview on last calling. For example, if an interviewer called on a house at 10am on a Tuesday and found nobody home then it would not be very effective to call again at 10am on Wednesday, the next day since it would be very likely that the householder would be out once more. It is sensible to vary the day and the time of day of calling for repeat calls following non-contact.

To work out an effective calling strategy the fieldwork period was considered as consisting of three important interview shifts: weekday daytime, weekday evenings and weekends. These fitted in with typical working and leisure time patterns and it is these shifts that were to be varied with repeat calls in the case of non-contact.

The rehearsal response data also demonstrated that there was a significant number of households (particularly in city areas) where nobody was available for interview until at least 9pm at night. Therefore, unlike many other surveys, interviews were going to have to take place at these "unsociable" hours where necessary. This had an impact on the health and safety aspect of the survey (see earlier paragraph on pair/team working) but did mean that planning for this was possible.

Interviewers were therefore instructed to ensure they had visited each non-contacted household during weekday daytime, weekday evening and weekend periods by the end of 7 June - two weeks after the start of the survey - and to have visited each non-contacted household during a daytime and evening of each day of the week (including Sunday), and at least once before 9am on a weekday, and after 9pm on a weekday by the end of the survey period.

Note that the strategy was designed as an additional tool for interviewers to use rather than being too proscriptive since the clustered postcode nature of the CCS sample and distributed workloads meant that flexibility within this framework was required. However, interviewers were trained to call back on a property as many times as possible over the fieldwork period, sometimes over ten times, to make contact.

Local knowledge also fed into this process. For example if the interviewers, who were ideally recruited from areas close to their workloads, knew about a regular weekly market in the vicinity then they could anticipate that interviewing on that day would not yield as many successful responses as other days of the week perhaps might.

5.6 Persuasion

A small amount of resistance was anticipated in co-operating with the survey. Typical reasons for non-compliance included:

- the questions might be considered intrusive
- it might be seen as duplication of Census work
- reluctance to comply for political reasons
- reluctance to comply on the basis of confidentiality

To counter these the interviewer training covered in depth the reasons for why the survey was required. It provided the interviewers with the knowledge necessary to explain what the potential benefits for co-operating are, and to give solid reassurance on the confidentiality of the collected data under the Census Act. There was also a public information helpline number that was given out to concerned members of the public that they could call requesting further information from headquarters.

Despite the information the interviewers were trained to be able to provide there were still some households that refused to participate. Many of these refusals were not adamant refusals. It was possible to convert some of them by:

- offering to the householder to call back at a more convenient time, making a fixed appointment;
- backing off early at the first sign of a refusal before it could become adamant with the view to calling back at a later time;
- using a different member of the interview team to call back on a subsequent visit to avoid the "first impressions" affect;
- using a skilled and persuasive ONS social survey interviewer to convert the refusal.

Even after using these techniques there remained some "hard-core" refusals for which interview was impossible. The numbers of these were monitored carefully and a summary of the rate is shown as part of the overall results in the final section.

5.7 Nonresponse strategy

Despite using the recommended calling strategy at the end of the survey period there remained a small number of households in the sample where contact had still not been made.

A self-completion survey form was developed in an attempt to capture information from these households. This included all the questions asked on the interviewer driven form presented in a simple way with instructions included such that the householder could complete them without the need for an interviewer being present, and post it back in a supplied freepost envelope.

On the final day of the fieldwork period interviewers were required to call back on any households in their workloads still not contacted, to complete the administrative information required on the front of the form (address, postcode, form ID, interviewer number), and to leave the form at the household for completion by the householder when found. Note that the interviewers were not informed of the exact nature of the non-response work and were not supplied with the postback forms until the last moment so that early distribution could not occur.

This was designed very much as a "last chance" strategy - none of the benefits of having an interviewer present would apply to form fillers - but numbers using this method were expected to be small by this stage in the survey.

6. IMPLEMENTATION OF PRACTICAL ISSUES

This section describes some of the more practical issues that were implemented in the overall CCS plan.

6.1 Staff structure

In order to carry out an interview with each and every household a solid management structure was required to keep things on track and ensure the work was carried out in an organised way. The following structure was used.

34 Field Managers (about 6 per RM)
|
273 Team Managers (about 8 per FM)
|
3,900 Interviewers (about 20 per TM)

The Field Managers were primarily recruited as overall survey managers and had to manage the CCS within a relatively large area of the country. Each could be considered to be responsible to a large degree for running a CCS in their own region. This would involve recruitment of Team Managers, aiding in training, ensuring all aspects of the survey were delivered on time and communicating with headquarters. Each Field Manager had a team of approximately 8 Team managers.

The Team Managers working to the Field Managers were responsible for managing a team of interviewers rather than managing a defined geographical area, though they still required some survey management as well as team management experience. The key tasks for the Team Managers included recruitment (on a one-to-one basis) and training of a team of up to 20 interviewers, holding regular team meetings to assess performance and get feedback from the field, and regular progress reporting on the number of interviews achieved. Importantly they also had a key role in the quality assurance of survey, checking interview forms on receipt and carrying out spot checks to ensure procedures had been properly followed.

The Interviewers were the people who actually carried out the doorstep interviews. Their key tasks included listing all properties found within their designated workloads, asking the survey questions, recording the answers given accurately, explaining to the public what the survey was all about and providing Team managers with progress updates.

Additionally, approximately 100 ONS interviewers were employed to help out in an advisory and assisting role in areas considered to be more difficult to enumerate as previously outlined.

Breaking down the survey into these discrete regionalised management units meant that the difficult task of overseeing the national level operation was made possible.

6.2 Recruitment

It was necessary to recruit over 4,000 field staff in a short period of time to carry out the interviews. The recruitment took a cascading approach.

- Field Managers were interviewed and recruited directly by HQ staff travelling around the country.
- Team Managers were recruited by HQ staff and Field Managers paired up.
- Interviewers were recruited by Team Managers on a one-to-one basis due to resource constraints.

This meant that HQ had very much an on-the-ground approach and direct say in who was recruited at the management levels. It also meant that FMs and TMs had a direct involvement in the recruitment of their specific teams of TMs and interviewers respectively.

6.3 Training

A cascading approach was taken for the training along lines similar to those used for the recruitment. From the outset the ideal was to keep the message as "undiluted" as possible by involving Headquarters staff directly in the training presentations at as low level as practically possible. This would ensure that all field staff received a common message in terms of their training.

A very detailed training plan for Field Managers was put together covering all aspects of the survey from how to recruit under open and fair rules to how to complete administrative claim forms. This was delivered directly by HQ staff both to keep the message common and to demonstrate the enthusiasm behind the project from the HQ team. The FM training took place over three two-day sessions, each one residential, and really helped build up the CCS teamworking spirit.

Training for Team Managers was delivered by Field Managers and HQ staff. It followed a similar pattern to the Field Manager training albeit slightly condensed. The same subjects were covered, particularly on recruitment since TMs would be carrying out one to one interviews for the recruitment of interviewers and therefore it was vital that the proper procedures were followed in this. The presence of HQ staff at these events provided the quality assurance that was required.

Each Team Manager was solely responsible for delivering training to his or her team of interviewers in a three day non-residential event, including a one day home study section. All material, including manuals, slides and modules, was prepared by HQ for these training sessions to ensure the same message was put across at each one. Quality was assured by HQ staff and Field Managers attending a sample of the events to act as observers.

Also as part of the interviewer training a video was produced that illustrated very clearly the way in which the interviews should take place, together with examples of how to deal with some of the less ordinary cases that might be encountered. The video was shown at each interviewer training event and received very favourable feedback.

6.4 Communications and progress reporting

The CCS was a large scale survey with only a limited time available in which to carry out the fieldwork. It was therefore essential that any problems encountered by field staff or methodology issues arising were reported quickly and clearly to headquarters so that solutions and/or amendments to instructions could be arranged. Also, it was important for the managers in the field to know at all times how well the

interviewing work was progressing so that resources could be targeted to underperforming areas in an effort to bring them up to speed.

At the same time, we needed a system for disseminating information from headquarters to the fieldwork team rapidly. There was no time to write and send letters by post and the sheer number of people involved meant that telephone calls to them all were out of the question.

We therefore developed an IT based solution, the CCS Team Reporting and Communication System (TRACS) to which managers in the field could connect using PCs. It was a relatively simple web-based reporting system that allowed information to be passed up the management chain from the field (for progress reporting), down the management chain (information dissemination and additional supplies) and sideways to field staff within the management layers (sharing of experiences, requesting advice).

Progress reports made via the TRACS system were used to assess the performance of each workload continuously as the fieldwork ran. This enabled us to quickly to identify difficult areas and to plan contingency accordingly. The response rate on any given survey day could be compared with predetermined expected response rates derived from data collected in the 1999 rehearsal exercise. Cumulative expected response curves were modelled for the different difficulty areas covered in the CCS sample and these provided the target performance measure for each workload. Where actual responses fell below the curve management action was required to bring it back up if we were to meet the final response required by the end of the survey period. This worked particularly well in London where the response was generally slower than elsewhere.

7. OTHER ISSUES

7.1 Foot and mouth disease outbreak

The foot and mouth outbreak in the UK rendered some rural parts of the country inaccessible to interviewers at the time of the survey.

There were considered to be two key points that would impinge upon interviewing:

- Farm houses included in the CCS sample,
- Rural areas and roads closed to the public.

To assess the impact of the first of these issues information on the location of all 180,000 farms in England and Wales was obtained from then Ministry of Agriculture, Fisheries and Food (MAFF). Comparing this with the CCS postcode list indicated that there were between 1,000 and 2,000 farms in the sample.

The only way to assess the impact of the second of these was to use the Team managers as on-the-ground intelligence. In the run-up to the interviewing period, TMs performed regular observation trips around their areas and reported any restrictions encountered back to HQ.

A new set of directions was disseminated detailing how all inaccessible properties should be dealt with. This centred on making contact with the householder by dropping off or posting a letter. Following this initial written contact a telephone interview was attempted, either being initiated by the householder after having received the contact letter, or by the Team Manager where a telephone number for the inaccessible property was known.

Despite being untested in previous trials this method proved relatively successful. A good number of telephone interviews were achieved and fewer properties lost from the sample than would have otherwise been the case.

8. RESULTS

It is difficult at this stage to detail exactly how successful the CCS has been in assessing how well the Census has done in counting people and households. The final analysis can only be done following the application of the ONC methodology to the data collected in the two exercises. However, we can say that the CCS did extremely well in achieving a good successful interview rate, in keeping the number of refusals to participate low, and in keeping the burden on the public also low. These factors hint that the CCS has performed well as an input to the ONC process.

Of particular note:

- A national response rate (number of successful interviews achieved as percentage of properties found) of 91 per cent was finally achieved, outstanding for a voluntary survey;
- In London and the South East, known to contain the most difficult area to enumerate, a response of 84 per cent was achieved;
- Of the 101 design groups covered, 58 had a response over 90 per cent. Only 9 had a response of less than 80 per cent, caused by isolated pockets of low response within the design group.
- The postback exercise for non-response was successful, increasing the number of households covered by nearly two per cent;
- The national refusal rate was only five per cent, varying between four and seven per cent with area.
- The overall response is certainly considered a success since the size of the CCS was chosen such that it would provide built-in contingency for lower performing areas by borrowing strength from others if need be (ONS 2001)

9. CONCLUSION

The following points are considered to have been particularly noteworthy:

- The team working concept was extremely effective with a good momentum maintained throughout the duration of survey,
- The calling strategy varying times and days across workload shifts helped the interviewers work efficiently and methodically,
- The flexible field force approach worked well, particularly near the end of the survey period in London when response rates were lower than elsewhere and required raising within a short time.

The Census Coverage Survey was the largest one-off household survey carried out in the United Kingdom. It achieved its primary objective in terms of overall response rates, whilst keeping the variation in the response between areas as low as possible.

REFERENCES

ONS (1998), "Census Coverage Survey: The Precision of Population Estimates for Different Sample Sizes and Design Areas", Steering Committee paper ONS(ONC(SC))98/12, ONS Titchfield

ONS (1995), "Validation options for the 2001 Census", Policy Board paper 033, ONS, Titchfield

ONS (2001), "A Quality Assurance and Contingency Strategy for the One Number Census", Advisory group paper, ONS Titchfield