Catalogue no. 11-522-XIE

Statistics Canada International Symposium Series - Proceedings

Symposium 2004: Innovative Methods for Surveying Difficult-to-reach Populations

2004

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2004
Innovative Methods for Surveying Difficult-to-reach Populations

HARD TO COUNT POPULATIONS IN THE 2001 AND 2011 UK CENSUSES

Owen Abbott and Neil Jackson[1]

ABSTRACT

The 1991 and 2001 UK Censuses suffered from falling response and, in particular, differential response which led some areas and population subgroups to have response rates significantly lower than the national average. This paper describes the efforts made in the 2001 Census to both maximise and measure the response in the hardest to count sectors of the population. It evaluates the patterns that emerged from the 2001 Census and then describes the research that will be undertaken to develop a strategy for both reaching and measuring hard to count populations in the 2011 UK Census.

KEYWORDS: Census; Hard to Count; One Number Census.

1. INTRODUCTION

The central objective of the 2011 UK Census, like its predecessor in 2001, is to provide high quality population statistics as required by key users such as policy makers and service providers, on a consistent and comparable basis for small areas and small population groups. Currently, these may be expected to include counts of people, dwellings, households and families, with a breakdown of key characteristics). To achieve this objective, the census operation will strive to maximise response rates whilst minimising differential response. Where there is a conflict of priority or direction, the impact on response rates will be a key determining factor in making a decision. This includes any post-enumeration survey or other coverage assessment process (ONS, 2004a).

The 2001 Census showed, as expected, that response rates vary by geographical area and by population subgroups. In order to achieve a high response, intelligence is required in advance to allow efficient targeting of resources. This enables types of areas, households and people that are likely to pose enumeration problems to be identified and thus appropriate strategies for enumeration to be developed. A project has been established with the objective of providing this intelligence to allow development of the data collection strategy for the 2011 Census. Key components of this project are the experiences and data from the 2001 Census.

This paper firstly reviews the 2001 Census experience, examining the initiatives put in place to attempt to maximise response and reach the hardest to count populations. It then reviews the outcomes arising from the innovative 'One Number Census' methodology used to estimate and adjust for undercount in 2001. The paper then describes how this experience and data will be harnessed within the 2011 Census enumeration.

2. MAXIMISING RESPONSE IN THE 2001 CENSUS

A number of initiatives, described here, were introduced in the 2001 Census to maximise response rates.

2.1 Post-back

The 2001 Census was the first in the UK to use a post-back methodology. Although, as in previous years, forms were hand-delivered to households by enumerators, for the first time householders were given a post-paid envelope and asked to return the completed form by post, 88% of households returned their form by this method. Enumerators

[1] Owen Abbott and Neil Jackson, Office for National Statistics, United Kingdom

were still used during the follow-up phase, but since they only needed to visit households that had not returned a form by post, fewer field staff were required and they were more effectively targeted.

One consequence of using post-back was that face-to-face contact between enumerator and householder was no longer essential in order to obtain a response. Forms could be returned even if the householder was never at home when the enumerator called. With an increasing number of people living alone, many of whom are young people who are often out in the evening, this was considered an important step forward in enabling response.

2.2 Publicity

A variety of media, both paid and free, were employed to publicise the Census, to reach as many people as possible. ONS also developed partnerships with a number of outlets, including national and local television and newspapers. For instance, the office worked with the makers of the UK's most popular television programme, Coronation Street, to develop a storyline in which one of the characters became a Census enumerator. At a local level, BBC Local Radio stations featured regular Census updates from local area managers.

There were special campaigns targeted at particular subgroups that are known to be difficult to enumerate. For instance, special material was designed for students and distributed in halls of residence and student unions. The parents of new-born babies, who often forget to include them on Census forms, were targeted in a series of events at hospitals across the country, and the mothers of all babies born on Census day were offered a free Census sleepsuit.

2.3 Other Initiatives

In order to increase the opportunity for non-English speakers and the visually impaired to respond to the Census, an information leaflet was produced in 26 languages as well as braille and large print. Although the Census form itself was only printed in English and (in Wales) Welsh, the information leaflet contained details of the questions and instructions for completing the form.

A community liaison initiative was put in place to make contact with community and minority groups. These groups included local authorities, the Commission for Racial Equality, organisations representing the disabled and elderly, and organisations such as the Citizens Advice Bureau who were in a position to assist those having difficulty completing their form. The purposes of this liaison included raising awareness of the Census among minority groups and providing guidance and support to field staff in reaching and dealing with these groups.

3. MEASURING RESPONSE IN THE 2001 CENSUS: THE ONE NUMBER CENSUS

Despite the initiatives designed to maximise response, the 2001 UK Census was expected to suffer from a degree of undercount. The strategy put in place to measure and adjust for this undercount was known as the One Number Census (ONC). The ONC methodology is discussed in detail elsewhere (see ONS, 2001 and Brown *et al* (1999)). A summary of its main features is given here.

The key element of the ONC was a very large post enumeration survey known as the Census Coverage Survey (CCS), which repeated a subset of Census questions. The CCS in England and Wales covered a sample of 320,000 households, comprising approximately 1.5 per cent of the population. The sample was spread across all 376 local authority districts (LADs) in England and Wales. The CCS collection methodology was designed to ensure operational independence between the CCS and Census. It used a doorstep interview shortly after the Census fieldwork and did not use address lists.

A combination of automated and clerical matching was used to link the CCS to the Census. The output from the matching exercise was used within a Dual System Estimation (DSE) framework to estimate the number of individuals missed by both the Census and CCS. A modified ratio estimator was used to estimate the population of areas not included in the CCS. This produced population estimates for each age-sex group at Design Group (DG) level. DGs were areas with a population of approximately 500,000, consisting of one or more LADs. Although the CCS was the largest survey ever carried out in the UK, it was not large enough to provide direct population

estimates for each age-sex group at LAD level, which was the output required. A small area model was used to allocate the population of a DG between its constituent LADs.

The ONC was a success, providing robust population estimates for the majority of Local Authorities. Further analysis of census results and other sources has led to revisions of the population estimates in areas where there were enumeration problems in the Census. The CCS achieved a national response rate of 91 per cent, outstanding for a voluntary survey. Based upon this, the ONC was able to produce robust estimates of the undercount in the 2001 Census for key demographic characteristics and geographical areas.

4. ESTIMATED RESPONSE IN THE 2001 CENSUS

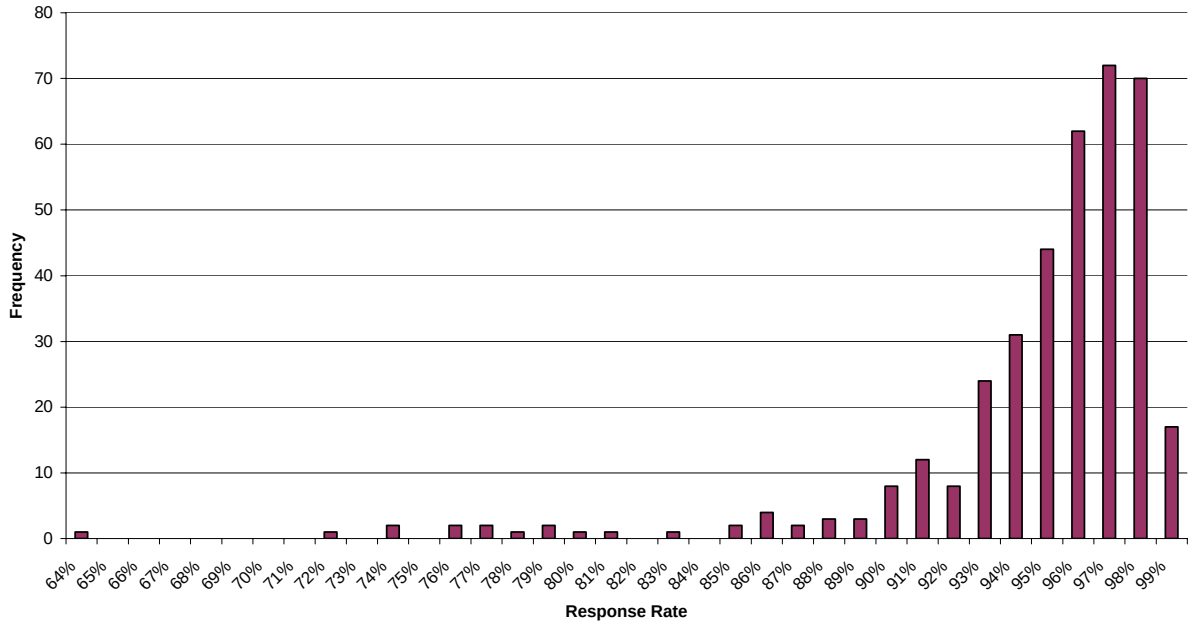
4.1 National Response Rates

The ONC estimated that 94 per cent of the residents of England and Wales were enumerated by the 2001 Census. However, this figure hides a wide variation in response rates between different geographic areas and socio-demographic groups. For instance, LAD response rates ranged from 64 per cent to 99 per cent, while response rates by age and sex varied between 87 per cent and 98 per cent (ONS, 2002). More detail on these figures is given in the rest of this section.

4.2 Sub-national Response Rates

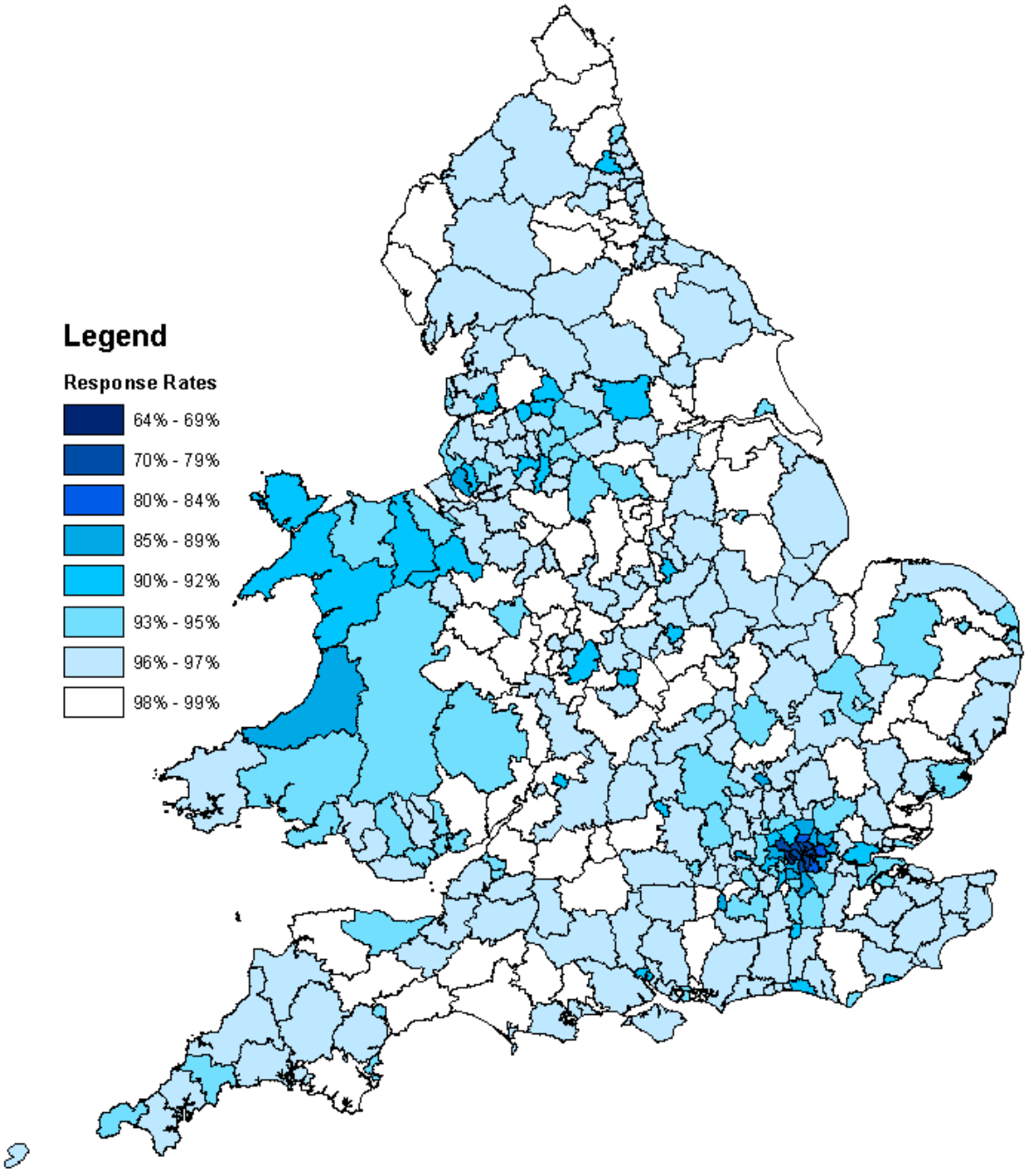
Within England and Wales, the lowest geographical level at which response rates were published was local authority district. The distribution of estimated response rates across all 376 LADs is shown in Figure 1 below. It can be seen that the distribution is highly skewed - most LADs have a response rate in the high nineties, but there is a very long tail. Almost all of the LADs in the tail are in London, and the areas in the low nineties are generally large metropolitan cities.

Figure 1: Distribution of LAD level response



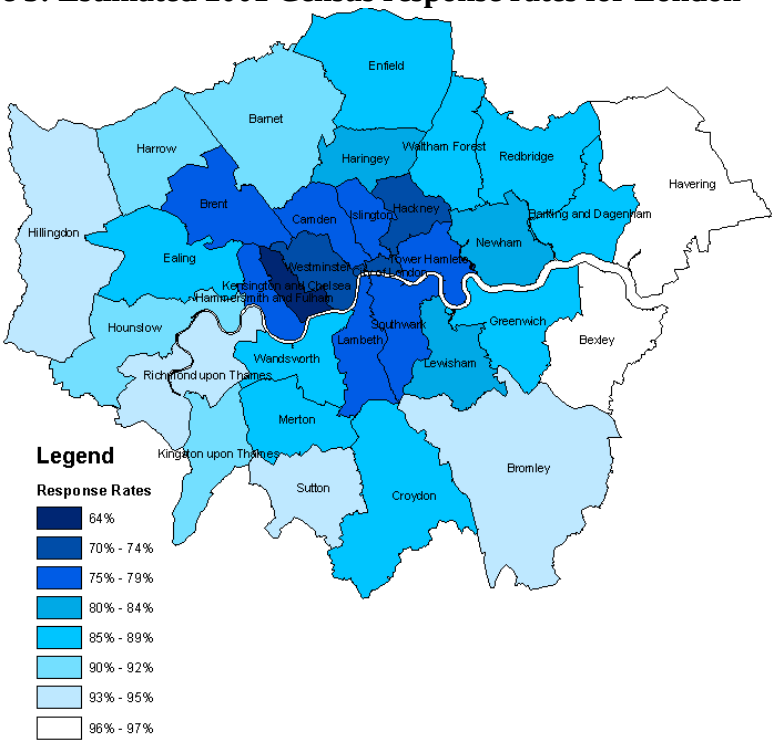
The division between London and the rest of England and Wales is illustrated by Figure 2, which maps estimated response rates by LAD. It can easily be seen that the lowest response rates are concentrated in London, no LAD outside London has a response rate below 85 per cent.

Figure 2: Estimated 2001 Census response rates for England and Wales



London consists of 33 LADs. These are divided into Inner London, which had a response rate of 78 per cent, and Outer London, whose response rate was 90 per cent. As can be seen from Figure 3, the response rate generally improves further from the centre. However, only two LADs had a response rate above the national average, and neither of these were in the top 50 per cent of response rates for England and Wales.

Figure 3: Estimated 2001 Census response rates for London



4.3 Socio-Demographic Response Rates

The national response rate for males was 93 per cent, compared to 94 per cent for females. The variability by age is far more marked. Figure 4 shows the national undercount rates by age and sex. It can be seen that males and females in their twenties were harder to count than any other age group. Babies, which are a known group prone to underenumeration, are also relatively poorly counted. Elderly people, particularly those in their seventies, are the easiest to count.

Figure 4: Undercount rates by age and sex

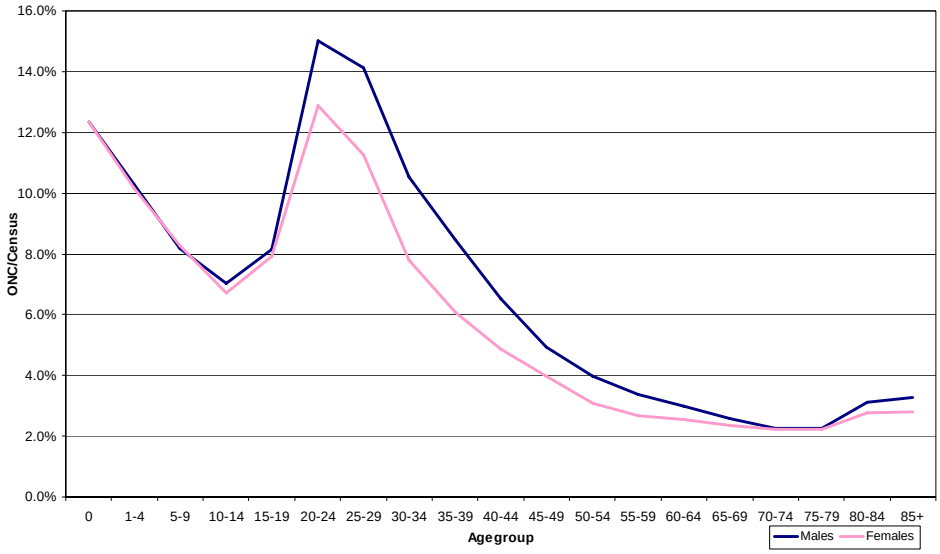


Table 1 shows some of the lowest response rates for selected other variables measured by the CCS at national level. Many of these, such as flats and rented accommodation, are strongly correlated with each other. Others, like single people and students, are correlated with the age groups with the lowest response rates. Almost all are characteristic of inner city areas, which agrees with the geographical response pattern.

Table 1: Selected low response rates for selected variables at national level

Variable	Value	Response rate
Household Ethnicity	All non-white	83%
	All Black African	72%
Activity last week	Unemployed	89%
	Student	87%
Marital status	Single	91%
	Separated	90%
Accommodation type	Purpose-built flat	86%
	Converted house	81%
Tenure	Rented from Council	90%
	Rented privately	81%

5. THE 2011 CENSUS ENUMERATION STRATEGY

5.1 Context

The context in which Censuses are conducted is becoming increasingly different. The world is changing, not only in terms of technological advancement, but also the attitudes of society and how people live and work. The population of England and Wales is becoming increasingly mobile and living patterns more complex. Increasing numbers of people are becoming associated with more than one address, sometimes commuting across international borders. As a result of this and lessons learned from the 2001 Census the design of the 2011 Census is significantly different from that of its predecessors.

The 1991 and 2001 UK Censuses suffered from falling response and, in particular, differential response which led some areas and population subgroups to have response rates significantly lower than the national average. This respondent behaviour is in keeping with that experienced by those conducting social surveys and with electoral turn-outs over this period. In all likelihood, contacting and eliciting a response from some sectors of society will continue to become increasingly difficult. This is the key issue when designing the 2011 Census.

5.2 Changes to the 2011 Census Design

In order to meet the challenge of maximising response in the context within which the 2011 Census will be conducted, changes to the design of the 2011 Census have been proposed. These changes are outlined in the information paper *The 2011 Census: A Design for England & Wales* (ONS, 2004b).

The key change that is designed to maximise response and minimise differential response is the aim for a mixed enumeration strategy. It is anticipated that a post-out post-back methodology, used for the first time in England and Wales, will be the main approach. In order for this to be feasible, the 2011 Census will be underpinned by a comprehensive national address list. The opportunity to respond via the internet will be offered for the first time. Furthermore, a comprehensive form tracking system will be used to manage the field operation, allowing real time information on response. The primary benefit of these innovations is that they provide the scope for optimal use of expensive, difficult to recruit, census enumerators in areas where response is poorest.

However, these changes present both opportunities and challenges, not only in supporting the testing and proving of the principles, but also in that they change fundamentally the nature of the operation. The post-out method relies on the accuracy of the underlying address list, and significantly reduces face-to-face contact with respondents. It is

expected that there will be areas of the country where the address list is not suitable for post-out methods, and there are population subgroups for whom a post-out method would reduce response levels. For these areas and subgroups, traditional enumeration methods may be used, or even tailored approaches may be developed. The key to success using the mixed enumeration approach will be to identify in advance where these areas are, who the hard to count are and, most importantly, why are they hard to count?

5.3 Enumeration Intelligence

In order to support the 2011 Census enumeration, the census programme has committed to research and identify hard-to-count groups and areas and develop robust initiatives and procedures to improve their enumeration. This project will provide intelligence for the enumeration process, including field methodology, questionnaire design, publicity and community liaison.

There are three key areas of research. First is the identification of characteristics associated with poor response. This research will build substantially on the experiences from the 2001 Census outlined in this paper. It is intended to extend our understanding of the characteristics associated with non-response by using additional data sources from the census operation and from external data sources, such as crime rates. This work is closely related to research being conducted both within the UK and internationally to tackle non-response to social surveys and censuses. The 2011 Census work will be co-ordinated with this wider research.

The second key area of research is to improve our understanding of why particular groups do not respond well to the census, such as language and literacy problems or the attitudes of some respondents. This is vital for the development of effective collection methods, questionnaire designs, accompanying translation questionnaires, leaflets and publicity to maximise response.

Once the barriers to response have been identified, the third key area of research is to identify what information is required, to what level of geography, in order to support the enumeration operation. The idea is to construct a national categorisation that identifies areas (or population subgroups) to support the mixed enumeration strategy. This categorisation will thus be the key determinant of how fieldwork resources are allocated. Other countries have used this approach. The United States constructed a Hard to Count score in 2000 based on 12 characteristics measured in the 1990 Census. This was then used to plan the enumeration and locate questionnaire assistance centres. Canada also used response rates from the previous census to decide where to implement an early follow-up procedure. The challenge for this research is to identify robust data sources at appropriate geographical levels, and decide upon the optimal number of levels the categorisation should have. An example of how this categorisation might look in England and Wales together with possible enumeration approaches is given in Table 2 below.

Table 2: Example enumeration categorisation

	If the Address List is more accurate in area	**If the Address List is less accurate in area**
Inner City - Young Men	Post-out and Collection/Interview	Pre-list Post-out and Collection/Interview
Inner City - Ethnic Mix	Hand Delivery and Post-back	Hand Delivery and Post-back
Elderly Area	Hand Delivery and Post-back	Hand Delivery and Post-back
Urban Mix	Post-out and Post-back with extra resources	Pre-list Post-out and Post-back with extra resources
Suburbs	Post-out and Post-back	Pre-list Post-out and Post-back
Rural area	Post-out and Post-back	Pre-list Post-out and Post-back
Disability	No area categorisation possible	
Anti-government	No area categorisation possible	

This project will provide this intelligence to feed into development work and major testing activities. The 2007 Census Test will provide the first opportunity to assess the robustness of the developmental categorisation and the effectiveness of targeted enumeration. The evaluation of the test will be key to implementing the enumeration intelligence research in the 2011 Census operation.

6. THE 2011 CENSUS COVERAGE ASSESSMENT STRATEGY

It is traditional that census undercount is measured and the outcome disseminated to users. Most census taking countries undertake this, usually using some form of post-enumeration survey (PES). Undercount levels have on the whole been increasing over the past few decades, mainly due to changes in society. More importantly, the differential nature of the undercount has been exacerbated, leading to increasing priority and focus on the methods for measuring this bias.

Despite our best efforts to minimise undercount in the 2011 Census, it is inevitable that a full count of the population will not be achieved. Therefore, in order to achieve our central objective of providing high quality statistics on a consistent and comparable basis, it is essential that we have undercount assessment strategies sophisticated enough to detect and estimate differential response down to small areas and population sub-groups.

The 'One Number Census' methodology used to estimate and adjust for undercount in 2001 was on the whole successful and there may be elements of the approach that can be developed for 2011. However, changes to the census design, such as the use of a national address register, a mixed enumeration strategy and developments in administrative data mean that applying the previous methodology directly may not be appropriate. The undercount assessment approach for 2011 needs to take account of more than the lessons learned from 2001 - the nature of the census has changed as has the context within which it is conducted.

Furthermore, changes in society and the census design may lead us to experience increased levels of overcount as experienced in other countries - particularly the USA. Therefore, the 2011 research will also consider how to estimate and adjust for overcount.

7. SUMMARY

Achieving a high level of response in any census is vital if it is to meet the needs of government and the user community. The innovations put in place to help maximise response in the 2001 UK Census went some way to achieve this. However, 6 per cent of the population did not respond, as estimated by the One Number Census. The differential nature of this non-response across geography and demography was a cause for concern, as it stretched the ability of the ONC to make robust estimates.

For the 2011 Census, it is intended to further develop initiatives to maximise response but also to minimise differential non-response. A key component is the research and development of enumeration intelligence, which will support the optimal allocation of resources. This project will provide intelligence for the enumeration process, including field methodology, questionnaire design, publicity and community liaison. Nevertheless, there will be an element of undercount in the 2011 Census and a robust measurement strategy is once again required.

REFERENCES

Brown, J. J., Diamond, I. D., Chambers, R. L., Buckner, L. J., and Teague, A. D. (1999), “A methodological strategy for a one-number census in the UK”, Journal of the Royal Statistical Society A, 162, pp. 247-267.

ONS (2001), “A Guide to the One Number Census”, http://www.statistics.gov.uk/census2001/pdfs/oncguide.pdf

ONS (2002), “National and Local Response rates (Annex B)”, http://www.statistics.gov.uk/census2001/annexb.asp

ONS (2004a), “2011 Census: Strategic aims and key research in England and Wales”, http://www.statistics.gov.uk/downloads/theme_population/Strategic_aims.pdf

ONS (2004b), “The 2011 Census: A Design for England & Wales”, http://www.statistics.gov.uk/downloads/theme_population/2011_design_information.pdf