

**UN STATISTICAL COMMISSION and
UN ECONOMIC COMMISSION FOR EUROPE**

**STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES (EUROSTAT)**

CONFERENCE OF EUROPEAN STATISTICIANS

Joint ECE-EUROSTAT Work Session on Population and Housing Censuses
(Ohrid, The former Yugoslav Republic of Macedonia, 21-23 May 2003)

Session I – Invited paper

EVALUATING THE 2001 UK CENSUS

Submitted by the United Kingdom

BACKGROUND

- 1 The 2001 UK Census – the twentieth in a series going back to 1801 - was a traditional census in many ways.
 - It was organised centrally;
 - It employed a temporary field force across the whole country that attempted to deliver forms to each and every household;
 - Enumeration areas were designed such that they contained an appropriate number of households that could be covered by one person; such areas varied in size to reflect the assessed level of difficulty-to-enumerate.
 - Information was collected through a self-completion questionnaire, completion of which was statutory;
 - The form and content of the census questions was decided after: (a) a lengthy period of consultation with users on their requirements for information to meet a wide variety of needs; and (b) an extensive programme of public cognitive research and small- and large-scale question testing; and
 - Data was processed centrally.

- 2 Planning of the 2001 Census built upon the strengths of the successes from previous censuses, as well as the lessons learned from the 1991 Census. There were, in particular, a number of issues arising from the 1991 UK Census and subsequently, which needed addressing:
 - Some sub-groups of the population were more difficult to enumerate, in particular:
 - the elderly,
 - babies,
 - students,

- young male adults,
 - ethnic minorities, and
 - inner-city populations.
- Nationally, undercoverage was just 2 per cent, but more problematic was the differential undercoverage across parts of the country that caused difficulties in interpreting results, particularly when comparing one area with another;
 - Societal changes since 1991 created a more difficult environment in which to conduct a traditional census. Particular difficulties resulted from:
 - the increase in the number of one-person households;
 - changing work patterns;
 - prevalence of entry-phone security systems;
 - greater public concerns about data protection and human rights issues; and
 - a prevailing climate of decreasing public participation in, and greater hostility towards, civic activities.
- 3 Several aspects of the planning of the 2001 Census attempted to address one or more of these problems. This, and the need to reflect response to a greater range of users' requirements for information, increased public concerns, the need to reduce the burden on the public, and to incorporate up-to-date technology to enable the Census to provide better value for money, gave rise to a number of major innovations in 2001 making it far less 'traditional' than had previously been the case. These included:
- the decision to conduct the Census on a *de jure* basis and exclude visitors;
 - planning of Enumeration Districts using automated geographic systems;
 - separation of Enumeration Districts, designed solely for the purposes of data collection, and of Output Areas, specifically designed as building bricks for the release of statistical results;
 - outsourcing a number of large-scale operations, particularly:
 - the primary data processing operation to commercial organisations with established experience in large automatic data processing systems, including optical mark reading, automatic character recognition and computer-assisted coding technologies;
 - the payment of field staff; and
 - operating a telephone helpline service.
 - initiating a Community Liaison Programme in which local communities were informed about the purpose of the census and were encouraged to support the census, and from whom possible enumerators could be recruited.
 - re-designing the census form to make it easier to complete and to provide a greater degree of confidentiality;
 - providing for forms to be mailed back rather than be collected on the doorstep, thus freeing valuable field force resources to enable a more cost-effective effort to be put into collecting forms in those areas which were expected to be – and indeed proved to be – harder to enumerate and from which the initial response would be poor.
 - conducting the largest ever household survey in the UK as a follow-up coverage check, to enable not only a measurement to be made of the extent of non-response at the local

area level so that estimates of the missing population could be made, but also to provide details, through imputation, of the characteristics of those persons missed in the Census in a process known popularly as the One Number Census;

- devising an Edit and Imputation strategy with the aim of estimating for all missing data and resolving inconsistencies in responses for the people and households affected; and
 - the dissemination of data freely available through electronic media, principally via the ONS website, under a new Government initiative designed to facilitate greater public access to official statistics.
- 4 This paper looks at the results of the evaluation of several of these innovations and of some other aspects of the census operation, and considers how the results of these will help towards designing the overall format and content of any future UK Census. The evaluations and lessons learned are reported primarily from an England and Wales perspective, but are, in many instances, also relevant to the situation experienced in Scotland and Northern Ireland, where the census is a devolved matter.

AIMS AND SCOPE OF THE 2001 CENSUS EVALUATION PROGRAMME

Aims

- 5 Systematic and continuous review of procedures and operations has formed a fundamental part of the UK 2001 Census programme. Each stage of the process of planning and implementation has been subject to rigorous analysis, and the evaluation and findings will form an essential part of the metadata/support material compiled.
- 6 The main purposes identified for 2001 Census Evaluation are to:
- inform users of issues surrounding data collection and processing to assist them in interpreting data (through the provision of metadata via the Census Offices' website, availability of Project Reports and the production of a Data Quality Report);
 - address the need for public accountability (through the production of a General Report); and
 - signpost the potential for change in planning towards 2011 (through the production of an Evaluation Review).
- 7 In managing the wider aspects of evaluation, the Office for National Statistics (ONS) in England and Wales are ensuring that the user community is consulted as necessary, and that there is close liaison with the other two UK Census Offices in Scotland and Northern Ireland in order to co-ordinate activities and share information so as to avoid unnecessary duplication of effort and ensure consistency of approach.

Scope

- 8 The whole project attempts to ensure that individual Project-based Evaluation Reports are produced to a common approved standard within an agreed timetable. The following topic areas are scheduled to be covered by individual Evaluation Reports:
- Data Needs*
 - Publicity*
 - Legislation and Non Compliance*
 - Data Collection Development and Support*
 - Geographic Support

- Processing *
 - Edit and Imputation*
 - Census Coverage Survey and the One Number Census*
 - Output Policy and Dissemination (including Census Access)
 - Data Quality
 - Output Production
- 9 Some of the Evaluation Reports (*) are complete and are already available on the ONS Census website (www.statistics.gov.uk/census2001/reviewevaluation.asp). The others are scheduled for completion by the end of July 2003.
- 10 In addition to these free-standing evaluations, a summary of data quality of the 2001 Census will be published in a *Quality Report*, currently planned for June 2003, and a wider overview of the whole census operation will be set out in a *General Report* nearer the end of 2003.
- 11 To avoid the risk of omissions, individual evaluations are examined in combination with others, to ensure coverage of a range of cross-project themes such as:
- Security and Confidentiality;
 - User Consultation; and
 - Programme Management including Risk Management and Outsourcing.

LESSONS LEARNED FROM THE EVALUATIONS TO DATE

- 12 A number of lessons have been learned from the experience of the UK 2001 Census. Some of these have been noted within the individual Evaluation Reports completed to date, and others will be incorporated in the remaining Evaluations. All will feed into the process of assessing the strengths and weaknesses of the 2001 Census so that future models may be improved and designed to meet the continuously changing needs of users and circumstances of the population.
- 13 This paper focuses on a number of the more significant issues that arose during the 2001 Census, and, in particular, those that arose from the innovations introduced, and which may have implications for planning any future census.

Date of the Census

- 14 The date of the UK Census is set in legislation and needs to be determined well in advance so that all aspects of the census can be planned accordingly.
- 15 The date of the census directly affects the quality of the data collected. It has to be set so as to:
- (a) maximise the numbers of households present;
 - (b) ensure minimal interruptions and distractions to the delivery and collection of forms; and
 - (c) ensure sufficient availability of field staff.

It also affects the location of people such as students, rough sleepers and those on holiday.

- 16 As the main criterion for 2001 was that the UK Census must maximise coverage, the main holiday periods had to be avoided. Although the Census does not have to take place on a specific day of the week, a Sunday has traditionally been chosen as the most likely time that people will be at home.
- 17 Data for the base year of the new round of annual mid-year population estimates was required to be provided by August 2002. The later in the year that Census day fell, the shorter the processing period would be to ensure meeting this target.
- 18 Consultation with users had indicated a strong preference for a term-time census to ensure better coverage of students and more information on their term-time accommodation. Local Authorities with large term-time student populations wanted the census to count students as being resident in their areas. Thus, Sundays between 12 March and 22 April 2001 were ruled out if Census Day were to fall in term-time for most schools, colleges and universities.
- 19 Daylight hours need to be long enough for evening enumeration to be carried out in reasonable light and good weather conditions. The Census could therefore not be held during British Standard Time, in the period from end-October until the last weekend in March.
- 20 A further consideration was that Census Day should, if at all possible, avoid the Sunday either side of any local government elections. In 2001, such elections in England were planned for Thursday 3 May, and local council elections in Northern Ireland on Wednesday 16 May. No local elections were planned in Wales or Scotland.
- 21 The two possible months that would be suitable for holding the 2001 Census were, therefore, April and May. Within this period no single date met all the conditions, but 29 April and 13 May stood above the others as being the most suitable. In January 1996, the Census Offices proposed that 29 April was the date that would have the least overall disruptive effect on the coverage of the 2001 Census.
- 22 Consultations were held on deferring the county council elections planned for 2001, but it was decided at the time to make no change. But following a major outbreak of foot and mouth disease, the Government announced early in April 2001 that the local elections planned for May, together with an early General Election, would be held on 7 June. As a result the potential overlap of census and election campaign activities was largely avoided.
- 23 **However, in setting a date so early in the Census planning programme there is a risk that subsequent events independent of the Census may affect particular aspects of the operation, even the Census itself.**
- 24 In moving from a 'population present' base (*de facto* population) solely to a 'resident population' base (*de jure* population) the need to fix a single date for the Census that maximises the number of people likely to be at home is perhaps less important. In 2001 for example households which were wholly absent on Census night were required to complete a form on return to their address if that was within six months of the date of the census. This spreading of the date by which a form had to be returned somewhat blurred the focus of 29 April as Census Day (despite the publicity campaign - see below), and indeed many more forms were returned even before Census day than had been the case in any previous census.

- 25 **Consideration will be given to the benefits of returning to a *de facto* basis for enumerating the population in any future census, particularly as the number of ‘visitors’ in many areas is a significant component of the overall demographic profile.**

Question content and form design

Question content

- 26 Decisions on the selection of questions and, the design and layout of forms for the 2001 Census were based on a careful assessment of information requirements, detailed discussions with users and a comprehensive programme of question testing. All aspects of the work were governed by a combination of ethical and practical considerations including equity, public acceptability, the limitations imposed by a self-completion questionnaire, processing requirements and cost constraints. Furthermore, recommendations took account of Eurostat requirements, the quality of the expected data, the burden on the public, and statutory restrictions.
- 27 Despite the best efforts it was not possible to meet all users’ requirements for information and as a result a number of potential topics/questions were not included in the 2001 Census. In some cases this was because there was insufficient need for the information, in others because of concerns about public acceptability, problems with the accuracy or reliability of responses or difficulties in devising suitable questions. In particular, the case for including a question on income was rejected by the Census Offices, despite strong and widespread support from users, because of concerns about public acceptability and the implications for overall response and coverage.
- 28 Among the other topics which were considered but which were not included in the 2001 UK Census were questions on:
- Housing/households*
- age of dwelling;
 - value of the home
 - access to a garden or yard;
 - size of car;
- Individuals*
- nationality;
 - duration of residence in UK;
 - disability;
 - Scots language (a question on Gaelic language in Scotland *was* included);
 - main language used in the home;
 - number of miles travelled per year;
 - address five years before the Census (a question on address one year before the census *was* included);
 - duration of residence at current address;
 - private education, health insurance and pensions schemes;
 - health lifestyle; and
 - number of current jobs.
- 29 The importance of formal consultation, research and question testing were recognised from the outset of the programme for the planning and development of the 2001 Census. Staff from the Census Offices worked co-operatively with users to meet their information needs,

develop effective questions and design suitable questionnaires. Every effort was made to maintain regular contact with users, report progress and provide information on successive stages of the work. (Particular details on the consultation on the form of the revision to the question on ethnic group are given below.)

- 30 Work on the 2001 Census has highlighted the importance for the assessment of future data needs of the following:
- **There is a need for continuous and co-operative working arrangements and good communication including the use of timely and well-informed publicity;**
 - **The Census Offices need to build on the strengths of the formal consultative groups, and from the invaluable advice provided by technical experts;**
 - **Better co-ordination of the necessary legislative and census timetables was identified, along with the need for improved awareness of the possible implications of political developments; and**
 - **There should be contingency arrangements to deal with unexpected problems resulting from late revisions to questions to reflect changes in Government policy.**

Form Design

- 31 Detailed decisions about the design and layout of the Census forms were based on successive test results, detailed discussion with users, design consultants and print experts as well as thorough quality assurance and proofing procedures. Special attention was given to a combination of factors including:
- public acceptability;
 - technological considerations;
 - the burden on the public; and
 - financial constraints.
- 32 Early work drew on published material about how respondents to self-completion questionnaires were affected by layout and presentation as well as research in other census organisations. Initially, both the *matrix* format, used in previous UK Censuses, and a new *pages-per-person* layout were tested. External design consultants were then appointed to re-design both documents. The results of this exercise provided excellent working models for subsequent assessment, most notably in the 1997 Census Test, the statistical design of which allowed comparisons between reactions and responses to both the matrix and pages-per person layouts. Completion rates for both household and person questions, overall discrepancy rates, levels of objection and refusal as well as concerns about privacy showed that the *pages-per-person* format outperformed the *matrix* style form.
- 33 These findings were substantiated by an evaluation of technical and operational issues including data capture, scanning and recognition, which highlighted a number of problems with the matrix form. As a result it was concluded that the pages-per-person format would be more suitable for the 2001 processing systems.
- 34 A series of recommendations took account of results from the 1997 Census Test, research on estimated costs of printing and postage, acceptability, burden on the public as well as the preferences of Census users and field managers. Most proposals were adopted, including the preferred pages-per-person design and overall size of 20 pages.

- 35 However, coverage was reduced from six persons per form to five following the decision to use a new larger matrix-style question to collect relationship information. (The performance of this question is the subject of a separate paper by Ian Maté (GROS).)
- 36 The primary purpose of the 1999 Census Rehearsal was to check methodological and operational procedures rather than the design of the form. However, results from data processing and preliminary analysis of returns highlighted a number of problems with form completion, colour and design. Supplementary information from the Census Quality Survey demonstrated that the public's reactions to the form were favourable with over three quarters of respondents (77%) finding it simple and easy to understand.
- 37 Although the basic design and layout of the form were not changed a number of minor revisions were implemented to overcome a number of shortcomings identified in the 1999 Rehearsal and the subsequent Census Quality Survey, and to resolve specific printing and technical issues. These changes were accomplished without compromising public acceptability or burden.
- 38 The use of scanning technology for data processing had a significant effect on the way the census form was designed. Whilst the design needed to reflect the requirements of respondents and processing requirements it also needed to take account of advances in printing technology. New form design software speeded up the design process but also introduced new challenges of balancing the very important aspect of public acceptability of the form and the exacting design requirements needed for scanning technology.
- 39 **Evidence suggests that the new format of the 2001 Census form was successful in that:**
- **it lessened the burden on the public in completing it;**
 - **it facilitated higher levels of question completion;**
 - **it provided a greater degree of confidentiality in that each person was able to complete their own details on pages which were separate from other household members; and**
 - **it better assisted the automatic processing systems.**
- 40 Processing difficulties resulting from the variations in the content and layout of the forms to accommodate a number different questions across the constituent countries with the UK was not regarded as an intrinsic weakness of the form itself, but would have been the case with the adoption of the previous matrix-style format.
- 41 Particular concerns were expressed by groups representing the interests of the disabled that the census form did not readily accommodate the blind and other persons with visual impairment. Although, the census questions were available in large print and Braille format, there was a call to provide for the returns to be made in these formats.
- 42 Although design of the processing system did not enable for returns made in alternative formats and media to be processed automatically, the Census Offices did accept returns in these other formats and made arrangements for the responses to be coded manually.
- 43 **In order to conform to possible future statutory requirement on disability discrimination, such alternative means of making a return in future censuses will be investigated and tested. Furthermore, to comply with the UK Government's commitment to conducting its business and its delivery of services increasingly through electronic media, consideration will be given to offering the option of making**

returns in any future census via e-mail or through interactive TV systems – or by whatever alternative methods communication technology offers at the time.

Ethnic Group Question

- 44 Unlike the case in many other European countries the UK Census does not, traditionally, include a question on nationality, nor was it the purpose of the 2001 Census to enquire into national identity. The 2001 Census did, however, include a question on ethnic group as did the 1991 Census, the main aim of which was of identify areas with high levels of ethnic minority populations. This will help provide valuable information on which to plan and provide local services and to monitor racial disadvantage and social exclusion – a major aim of the 2001 Census.
- 45 Generally the subject and wording of the questions proposed for the 2001 UK Census created little controversy, once the initial concerns about the Government’s proposals with regard to the questions on income and religion had been resolved. However, prior to the 2001 Census, some significant concerns were expressed about the form and content of the ethnic group question. Although there was widespread consultation on, and support for, the question among the ethnic community, there was evidence that some minority groups were unaware of the significant consultation that had been carried out on this topic.
- 46 A Government White Paper had, in March 1999, set out the formal proposals to include a revised question on ethnic group (the version of the question included in the 2001 Census in England and Wales is shown below). The form and wording of the question had been decided after extensive consultation with census users on their needs for information and of public testing to assess acceptability. Further comments on the White Paper proposals had been invited, and the form and wording of the question were approved by the UK Parliament in the Census Order and the subsequent Regulations.
- 47 After publication of the White Paper, however, a case was made by representatives of the Sikh and Kashmiri communities that the question should separately identify these particular categories. Extensive correspondence was conducted during 1999 with concerned members of these communities to explain the background to, and rationale for, the proposed question and the purposes to which the resulting information would be put.

2001 Census question on ethnic group in England and Wales

What is your ethnic group?

- ◆ Choose one section from (a) to (e) then tick the appropriate box to indicate your cultural background

(a) White

British

Irish

Any other White background

please write in below

.....

(b) Mixed

White and Black Caribbean
 White and Black African
 White and Asian
 Any other mixed background
Please write in below

(c) Asian or Asian British

Indian
 Pakistani
 Bangladeshi
 Any other Asian background
Please write in below

(d) Black or Black British

Caribbean
 African
 Any other Black background
Please write in below

(e) Chinese or Other ethnic group

Chinese
 Any other
Please write in below

- 48 The matter was partially resolved with the commitment to include a question on religion in the Census that would include a ‘Sikh’ tick box. Assurances given by the National Statistician that written-in responses of ‘Kashmiri’ would be coded and separately identified in some outputs also eased the situation and avoided a threatened boycott of the census by these groups. The issue did not re-emerge to any significant extent at the time of the Census itself, **but the ‘Sikh’ issue has more recently re-emerged and needs to be addressed in any future census question.**
- 49 Also, shortly after the legislation establishing the 2001 Census questions came into force in June 2000, representations were made to the ONS expressing concern that the wording of the ethnic group question did not provide tick box response categories to enable people in Wales specifically to indicate their own identity. In response the ONS noted that although the question provided tick boxes within the White group only for ‘British’ and ‘Irish’, it also included a write-in category within each of the main ethnic groupings so that people could describe themselves in whatever way they chose.
- 50 **However, in the light of the strength of the views expressed on this issue, arrangements have been put in hand for the devolved National Assembly for Wales to take a more formal role in deciding the format and content of future census forms in**

Wales. This may result in a need to consider separately the legal and logistical arrangements for conducting a Census in Wales.

- 51 **In the meantime a new form of question on ethnicity, in which the components of national identity within the UK and more wider ethnic origins are separated into a two-tier format, is being introduced into national surveys with a view to adopting it as the *de facto* basis for collecting information on ethnicity in future censuses and from a wide range of other statistical and administrative sources.**

Changes in Data Collection (Field) activities

- 52 A major influence in shaping the Data Collection methodology for the 2001 UK Census was the evidence, as research and testing progressed during the preceding decade, that it was becoming increasingly difficult to make contact with people, especially - but not exclusively - in inner cities. This loss of contact threatened to reduce the high level of coverage needed in the Census. The reasons for lower levels of contact included:

- growing numbers of single-person households;
- changing working patterns;
- society becoming less compliant; and
- sub groups of the population who feel increasingly disenfranchised.

- 53 Thus the Census at the turn of the 21st Century presented special challenges, and it was recognised that more radical approaches were required in determining the methodology to be used for the 2001 Census.

Postback of forms

- 54 The most significant change in methodology was in the method of collecting completed forms that for almost two centuries had been the exclusive domain of the Enumerators. An existing national postal system and the increasing willingness and ability of the public to return documents without having to wait for a return visit from an enumerator, led to a postback methodology being introduced in the 2001 UK Census. A finely balanced decision was taken to use the local census management teams to receive and prepare completed forms for transfer to the central processing office, rather than to arrange for these forms to be posted direct.

- 55 The introduction of the postback method realised some major benefits, which included:
- a smaller field force than was required in the 1991 Census (in England and Wales for example, 73,000 temporary field staff were employed, compared with 115,000 in the 1991 Census) - thus easing the pressure on the recruitment programme and cost;
 - freeing up resources to focus on the areas which were difficult to enumerate, such as major inner city areas; and
 - targeting limited financial resources where they were most needed.

- 56 Postback as a collection methodology was judged to be a major success. In all about 22 million forms were mailed back by the public. This post-back response, some 88 per cent, far exceeded expectations of around 70 per cent from international experience. Field staff in the follow-up exercise, which was extended in some areas, collected a further 7 per cent

of forms. An estimated 3 per cent of forms were delivered to addresses that were either vacant or second residences.

- 57 The effect of this high return should have meant that far less effort would be needed to follow up on non-returns – this was the main objective. There was, however, a counter effect that caused considerable disruption to the fieldwork and which led to the public contacting the Census Helpline in far greater numbers than had been anticipated. That effect was due to backlogs in the postal flow to local field managers due, in part, to the disruption to the postal services in many areas resulting from an industrial dispute at the time of the Census. The follow-up by field staff was hampered by this and led to unnecessary call-backs in instances where householders had already posted back their form. Nevertheless, the teams worked flexibly to overcome these difficulties with considerable success.
- 58 Though postback was an effective method of data collection, a key lesson is to consider whether a ‘central’ postback alternative would have been better. The ‘local’ approach worked largely because of the strenuous efforts of the field force, but they were at the mercy of a postal flow that was beyond their control.
- 59 **Directing mail to a centralised facility could be more attractive as it would enable a focused follow-up of non-returns, but above all would remove the huge burden on local field teams. New technology may enable faster processing of mail where risk could be better managed with postal service providers. Quality and form control will, however, remain key issues.**
- 60 **Modelling for postback in the future must include less reliance on a single national postback rate, although the 70 per cent estimate was considered a reasonable rate for planning, having been based on international experience and testing. Variable rates were planned across very broad geographical areas only, but more needs to be done in modelling for a range of national and more local rates that can be tested and used to inform the service delivery process. More attention must also be paid to variability of service delivery from area to area, and this should be carefully considered with future postal service providers.**
- 61 **The ONS is already investigating alternative post-back arrangements and the possibility of ‘real time’ form tracking systems by which the status of individual forms could be centrally monitored from initial posting, through receipt, to scanning and processing.**
- 62 **Consideration should also be given to using a more flexible start date for the follow-up phase, in order to enable the greatest possible postal response before visits commence, as well as allowing more time for the effects of postback to work in the field.**

Form Delivery

- 63 Despite the decision to use postal services as the prime collection methodology, the traditional method of doorstep delivery by Enumerators was used rather than a postout method because of the lack of a totally reliable national list of households, and because personal contact with householders was judged to achieve a higher level of response.
- 64 The bulk of the delivery of forms was completed on schedule. Of the 103 Census Areas, 94 reported that delivery had already been completed or would be completed by the end of Census Day (29 April 2001). In the remaining 9 areas small pockets of households were reported where delivery had not been completed by Census Day but would be completed

by the end of the next day. Most delays appeared to be due to recruitment problems or late resignations of Enumerators. Difficulties gaining access to properties and delays in receiving additional supplies of forms were also cited as reasons. All 9 of the Census Areas that reported problems covered inner cities and 5 of them were in London.

- 65 The additional procedures put in place to handle these difficulties as well as the major outbreak of foot and mouth disease were successful, and delivery in affected areas was achieved without any detriment to the overall timetable.
- 66 **The success of postback suggests that a postout of future census forms may be a goal that is becoming more achievable.** However, the potential for this is a subject for later consideration, and does not fall within the remit of this paper, nor indeed does the consideration of the use of alternative media for capturing census information from the public.

Field Management

- 67 The importance of field management information cannot be underestimated, and, with the right tools, provides the Census headquarters with useful data to monitor progress and identify problem areas. Not only does information received from so large and widely dispersed a field force help to compile status reports on targeted activities during field operations, both for Census HQ and for senior field managers, but it also assists in the control and management of the field force and its tasks, and its financial provisioning.
- 68 Field Management Information System (FMIS) used in 2001 underpinned the need for and importance of timely, reliable and robust information. **This experience must be built on, extending the use of technology and identifying further uses of such systems in a future census, whilst paying attention to its fitness for purpose and usage in the field.**
- 69 The 2001 Census system was telephone based and, in the event, proved difficult to use and compounded the problems experienced through the delays to postback. **A key objective must, in future, be to ensure information systems are robust, have reliable functionality, and are appropriate in their application in the field as well as centrally. Such systems are essential for data collection purposes and serious consideration must be given to providing IT to district manager level to improve communications .**

Community Liaison

- 70 The 1991 Census and the subsequent testing programme for the 2001 Census had highlighted the fact that some sections of the community were becoming increasingly isolated and disenfranchised and were at risk of being significantly under-enumerated.
- 71 Changes in methodology for the 2001 Census meant that field staff would have less contact with the public than previously. There was, therefore, an increased need to target minority groups in order to provide them with reassurance about the aims of the Census, to encourage them to participate and to offer help with form completion.
- 72 The Community Liaison initiative was set up in 1997 to make contact with community and minority groups in order to address these issues. The key aims of the initiative were to:
- promote and raise awareness of the 2001 Census and encourage participation of previously under-enumerated groups by underlining its uses, in particular, by attending

or organising meetings, presentations and exhibitions with national, community and minority groups;

- identify ways of providing assistance with form completion, including producing translations and transcriptions, liaising with ethnic and disability groups to establish best practice;
- identify ways of providing practical help to community and minority groups and how the groups, in turn, could provide input to the Census programme; and
- provide guidance and support to Census field staff.

73 The Census Office carried out the liaison at a national level but the work was continued by the regional census field managers, who worked closely with local community, ethnic and voluntary groups. These managers were provided with instructions and specialist training to assist them with this work, and this was cascaded down through the field force as part of the training package.

74 Enumerators issued translations of the Information leaflets in any one of 26 foreign languages including:

| | | | |
|------------------|----------|------------|------------|
| Albanian/Kosovan | German | Polish | Swahili |
| Arabic | Greek | Portuguese | Tamil |
| Bengali | Gujarati | Punjabi | Turkish |
| Cantonese | Hindi | Russian | Urdu |
| Croatian | Italian | Serbian | Vietnamese |
| Farsi/Persian | Japanese | Somali | |
| French | Korean | Spanish | |

and provided large print documents as necessary and dealt with specific requests for help, such as the need for an interpreter, when they delivered the Census forms. The provision of such assistance was seen as a major benefit of adopting a doorstep delivery strategy.

75 Response to the initiative was generally very positive and it was seen as a success in that it covered a very wide range of community and minority groups and national charities that represented many different sectors of the population. **The overall assessment is that a future census would clearly benefit from the repeat of a similar initiative. However, the 2001 experience also highlighted the need to initiate contacts much earlier. In particular, the Census Office will need to review its procedures for the enumeration of minority groups, particularly the disabled, working closely with relevant organisations to adopt a best practice approach and to ensure that no sections of the community are disadvantaged by the Census methodology, and that, in particular, provision is made for the availability of census forms in Braille or other formats suitable for completion by the disabled.**

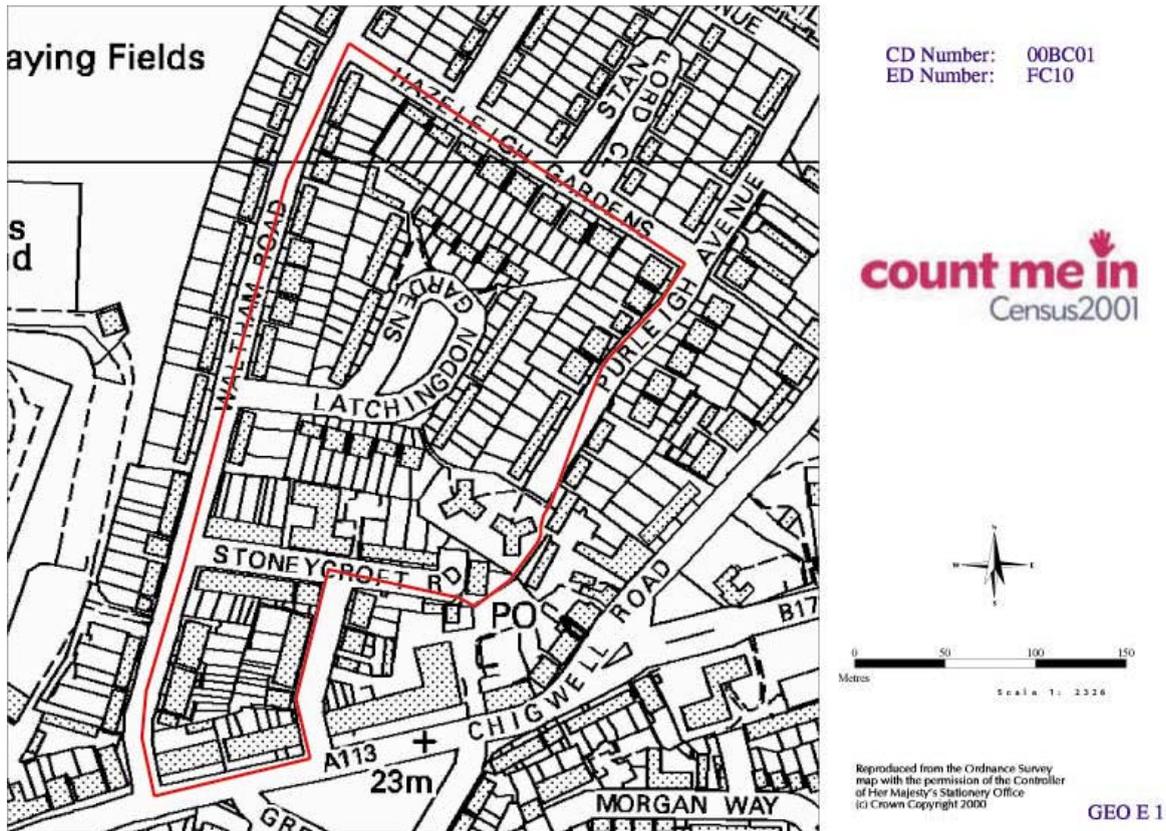
Planning Census Geographies

Input Geography: Enumeration Districts

76 The requirement for fewer enumerators in 2001 Census meant that Enumeration Districts (the data collection geography) had to be larger than in 1991, and that flexibility had to be introduced in order to enable enumerators to manage more than one district or for two or more enumerators to cover a number of districts jointly. Furthermore, advantage was taken of advances in digitisation and GIS technology to enable planning of the areas (and of subsequent Output Areas – see below) to be automated.

- 77 The principal consideration of Enumeration Districts (ED) design is to create geographical areas which facilitate efficient and accurate distribution and collection of census forms by enumerators, while attempting to equalise enumerator workload.
- 78 This requires knowledge of the location of residential addresses, together with some understanding of the likely difficulty of enumeration. Factors particularly increasing enumerator workload include: sub-divided properties such as bedsits; flats in which individual front doors are protected by entry-phone systems; and residents who may not have English as their first language. Relevant physical considerations also include the density of housing, particularly in rural areas where enumerators may have to travel long distances between addresses, reducing the number of households which can be enumerated within a standard workload. In planning for both 1991 and 2001 Censuses, the UK Census Offices attempted to trade-off the difficulty of enumeration resulting from such factors with the size of the ED. Furthermore, it is necessary to place ED boundaries in such a way that enumerators have responsibility for areas which 'make sense' on the ground. This generally requires that EDs do not straddle major roads, rivers, railway lines or extensive areas of open space.
- 79 Design of EDs for 2001 was implemented using a GIS-assisted methodology, in which ED planners undertook a design exercise very similar to the 1991 approach, but working within an entirely digital environment. The methodology has been reported elsewhere (Martin (1998) and Martin (2002)), and ED design was completed in less time and by fewer planners than in 1991. Once the entire country was complete, certain areas were re-planned using the latest available data enhanced by additional addresses discovered during the fieldwork so as to ensure that the ED geography reflected the actual residential structure on census day as accurately as possible. An innovation resulting from this approach has been that it was possible to issue enumerators with appropriately scaled single-sheet ED maps (see Figure 1) and also to produce record books with pre-printed address lists.
- 80 The methodology and technology used in designing EDs were major successes in the 2001 Census and will likely form the basis of planning any future census input geography. **It is expected that further automation of ED planning will be achievable, reaping the benefits from the creation of digitised boundaries and a complete and fully grid-referenced address database.**

Figure 1: Sample enumerator's ED map



Output Area design

- 81 Following on from prototype work undertaken at the time of the 1997 Census Test, a further innovation for the 2001 Census was the decision to create separate geographies for data collection and output throughout the UK (although this had been done in Scotland for the 1991 Census). As described above, ED design proceeded within a GIS environment to create a digital representation of the ED geography from which enumerators' maps and address lists were generated. Although the cost and complexity of setting ED planners to go back over the entire country drawing up another separate geography according to output considerations would have been prohibitive, advances in computing power, digital data infrastructure and the use of zone design algorithms made possible the creation of a separate output geography by automated means. Output geography design was based on the same GIS database that had already been created for ED design.
- 82 It is not within the scope of this paper to describe the methodology for creating Output Areas (OAs) (instead, see Martin (2002)). It is perhaps sufficient just to record here that one of the first principles in creating a purpose-specific output geography was that the OAs should, for disclosure control purposes, all be above a required minimum population and household threshold, thus avoiding the 1991 problem of suppressing output for some small areas.
- 83 A second significant factor was that it was necessary to create OAs, which were, as far as possible, constructed from whole unit postcodes, thus facilitating better integration between geographical information referenced by census and postcode geographies. Additional, but nevertheless important, considerations also included the standardisation of

OA population sizes, maximisation of internal social homogeneity and some control over the more irregular geographical shapes of the areas.

- 84 It is still too soon to evaluate the performance of the 2001 Census Output Areas, but from the initial favourable reaction of users to the recently released boundary sets (and from the experience of users in Scotland in the 1991 Census) it seems that the creation of purpose-built areas for statistical analysis better meets uses needs for small area geography. **Furthermore, the Census Output Areas have subsequently been adopted throughout the ONS as the basis for planning a standard small area geography for a wide range of social and demographic statistics – in effect enabling the creation of a stable UK geography at the ward/NUTS 5 area level.**

The One Number Census and the Census Coverage Survey

- 85 The One Number Census (ONC) Project aims to integrate the 2001 Census count with the estimated level of under-enumeration in the Census – that is, the number of households and residents not counted – the with aim of:
- (a) providing a new base for the annual mid-year population estimates down to the local authority district level; and
 - (b) adjusting the census database itself for the undercount so that all statistics add up to One Number – the national estimate of the population. This was a major user requirement, which lead many of the strategic changes in the 2001 Census.
- 86 The key element of the ONC methodology was the Census Coverage Survey, which was designed to enable the census counts to be adjusted for under-enumeration at the national, regional, local and small area levels. This was, perhaps, the most significant innovation in the 2001 UK Census.
- 87 The Census Coverage Survey (CCS) was the largest household survey ever run in the United Kingdom. The Survey was designed with the sole purpose of enabling ONS to identify those people and households missed by the Census. To this end, it had to be entirely independent from the Census, there had to be full geographical coverage, it had to allow population counts to be adjusted for under-enumeration down to small area level, and it had to allow population estimates to be made for all districts.
- 88 The sample size had to be sufficiently large to ensure that the accuracy of the subsequent annual population estimates, at both national and local area level, were acceptable. The larger the sample size, the more accurate the population estimates, but this must be balanced against the cost and practicalities of carrying out a larger CCS.
- 89 The ONC research programme examined the optimal sample size representing the best value for money in terms of accuracy. This research indicated an overall sample size of around 300,000 households for England and Wales would provide an acceptable level of accuracy. The final selected sample contained 320,000 households in 16,400 postcodes in England and Wales, (and 40,000 household in 2,400 postcodes in Scotland, and 10,000 households in 700 postcodes in Northern Ireland).
- 90 A project of this scale required some careful planning well in advance and continuous monitoring and assessment as it progressed. Overall the procedures developed for the CCS were successful - none failed completely and none of the difficulties encountered were prohibitive in ensuring the CCS delivered its aim sufficiently for the One Number Census

methodology to be applied. Inevitably, though, there are some lessons to be learned and evidence of a need for some improvements to be made in any similar future exercise.

91 In summary these are that:

- **Early planning on all aspects of a Census Coverage Survey is essential, and particularly the development of an IT system to facilitate better field management and reporting;**
- **The team-work ethos of the CCS was a real strength throughout the project;**
- At the stage of sample selection it became apparent that the average number of households in each postcode was higher than previously estimated, and that the random selection of postcodes meant that some very difficult places could occur in the sample. Difficulties included the distance required to travel, and potentially dangerous areas. Each needed careful local management at the time of the survey. **In future the sample needs to have a high degree of contingency built into it in terms of the number of households expected to be selected.**
- There were some difficulties in recruiting sufficient numbers and quality of interviewer. **Despite the statistical requirement to keep the CCS independent from the Census there would be benefit in running a combined recruitment campaign for the two exercises. This would ensure there was no confusion when the advertisements were placed, and would help in making sure there were enough applications to fill all the posts. It would be easier to run a single campaign to recruit 70,000 people than separate but overlapping campaigns to recruit 66,000 census enumerators and 4,000 interviewers.**

92 Detailed results from the 2001 Census across the UK are not yet complete, and the Evaluation Report of the One Number Census not yet prepared, but early indications from users are generally that the counts at the local area level, which sum to the One Number Census total and which will form the basis for future population estimates and projections, are acceptable.

93 This, however, is not the case universally, and a small number of local district authorities have expressed concerns about the validity of the ONC methodology and have disputed the accuracy of the 2001 Census counts in their areas. Their particular concerns stem from the effect of the alleged census under-count in these areas on the level of annual funding provided by central government, which is principally calculated the basis of the census figures.

94 In response ONS are co-operating in an independent matching exercise to compare the 2001 Census counts of the number of households with data from administrative records held by the authorities concerned. **Furthermore, a review of the methodology for estimating migration, particularly out migration, both internationally and at the local area level, is in progress with a view to improving the methods of estimating the annual mid-year populations.**

Outsourcing census operations

95 Another important change was the move to outsource elements of the Census operation on a greater scale than had been done previously. The major operation awarded to outside contractors was for the primary processing of the completed census forms, but many other contracts were set up and managed.

96 The greater use of outsourcing in 2001 brought:

- an influx of external expertise and knowledge;
- a lessening of risk for the Census Office;
- savings realised through a series of competitive bids, and

- a transfer of skills.

But with it also came a greater degree of contract management and effort required to achieve the standards and quality of services required for all three UK Census Offices.

- 97 In addition to the major contract for the primary processing of all census forms (awarded to Lockheed Martin), the other contracts/services managed by the project included:
- planning and handling the postback of completed forms (Royal Mail);
 - providing help with enquiries from the public through helplines during the recruitment phases and the Census itself (Cable & Wireless);
 - designing and printing all public forms (Lockheed Martin);
 - designing and printing all other documents and materials (Central Office for Information);
 - distribution and collection of all census forms and other materials, and subsequent disposal other than completed census forms (TNT);
 - planning and producing field staff training videos (Scottish Executive Video and TV Production Unit (SEVTPU) in collaboration with Mirage TV Productions;
 - the provision of Payroll Services (Chessington Computer Services Ltd); and
 - publicity campaign (MS Saatchi).
- 98 The use of external service providers for the 2001 Census brought mixed fortunes. On the positive side, the contracts generally delivered quality services, risk was been transferred from the Census Office, strong partnerships were formed, considerable goodwill on both sides resolved some very difficult issues, and skills have been transferred to the Census Office.
- 99 However, the effort required to manage the contracts was under-estimated and, in an operation as complex as the Census, there are bound to be problems which may not be as easily solved because of contractual and operational constraints (involving resources and costs) in a fast-moving arena.
- 100 Consequently, the effort required to manage several large contracts was a significant drain on resources, and at times it conflicted with the daily business of running field operations. Considerable effort was made to ensure the services were delivered, **but future outsourcing needs to consider more carefully the potential impact of external influences on requirements for service delivery.**
- 101 **Contract management awareness and understanding, through training and teamworking, should be continued and extended as necessary, so that operational managers continue to regard the crucial importance of contractual rights and obligations.** This will take on even greater significance if more outsourcing is envisaged in future. **But this approach can only be successful if the right resources are in place to give due and equal focus to both operational and contractual matters.**
- 102 There were to many outsourced activities to report on in full in this paper, but three in particular illustrate both the benefits and difficulties that contracted projects bring.

Public Enquiries Service

- 103 One of the consequences of the decision to adopt a postback methodology for form collection was the increased numbers of forms being received directly at the Census Office (about 450,000: a fourfold increase on 1991). In response to the unprecedented number of

calls to the main Helpline asking for forms, a dedicated emergency Call Centre was set whose sole purpose was to send out forms.

- 104 The Call Centre was operational from 27th April 2001. Completed forms were returned direct to ONS. Additionally, from 30th May, after the closure of the local collection points, forms were directed to ONS. Delays in householders returning forms and postal difficulties resulted in more form being returned after the cut-off point than had been anticipated. In both cases, these forms were specially routed through the mail system and were assimilated into the main stock of forms collected and made ready for processing.
- 105 The Public Enquiry Service itself was overwhelmed by the number of calls that far exceeded all expectations. This was tremendously encouraging, in as much as it demonstrated that the advertising campaign publicising the Helpline had succeeded, but it put pressure on the system, which at times was unable to cope. Additional resources and extra lines were, however, successfully added to ease the pressure.
- 106 The automated messages were not always easy to use, but there is evidence to suggest there is general unease with this sort of technology in the public domain. The Census, as a subject, was not the easiest to explain on a helpline facility.
- 107 The numbers of emails received indicated that there was a greater requirement to respond through this medium than had been anticipated. **Future plans need to assess the extent to which each form of enquiry service can be utilised, particularly through electronic media.**
- 108 The public helpline was inundated with calls from people concerned that their Census forms had not been delivered. In many cases, these calls came well before the end of the delivery phase. **To avoid public concern and undue strain on the helpline, any future publicity strategy should ensure the public is made more clearly aware of the timetable for delivering and posting back forms. Contingencies should be prepared well in advance for dealing with areas where the delivery timescale can not be achieved**
- 109 **Helplines need to be easy to use, and designers of automated response processes must make them as user-friendly as possible. However there should be an even greater synergy between the Helpline and with events as they occur during the Census .**

Data Processing

- 110 Traditionally, Census forms in the UK have always been processed in-house, but following the 1997 Census Test it was decided that better value for money could be obtained from contracting-out the main scanning and capture services for the 2001 Census. After a procurement project was undertaken in 1998 the processing contract was awarded to Lockheed Martin (LM). The latest technology was used to process 27 million census forms, with forms scanned at high speed, and responses coded electronically, using automated and computer-assisted methods.
- 111 The key benefits to outsourcing the processing operation were to:
- streamline the capture and coding process with a view to reducing staff costs and accommodation required;
 - improve quality in terms of accuracy and consistency of the captured and coded data;

- improve the speed of processing data over 1991 so that outputs could be delivered to a more prompt and reliable timetable; and
 - enable a trade-off between reduced costs and the coding all responses - only 10 per cent of responses to Industry, Occupation and Workplace Address questions had been coded in 1991 and previous censuses. Full coding of all questions was undertaken in 2001.
- 112 The overall aim of the Processing Project was to provide an efficient and effective processing operation that delivered a set of clean, consistent and complete data for input to the Edit and Imputation systems, which were separately managed by the Census Offices in-house. There were three key indicators that were used to measure the success of the Processing Project - quality, cost and timeliness.
- 113 All agreed quality standards were met or exceeded. These standards formed part of the contract and were measured continuously throughout processing. The final total costs (of around £60M) compared favourably with the original estimates. There were some timetabling issues, which ultimately led to a slippage of some five weeks in the final data delivery milestone, but this was still an excellent achievement in the context of a four-year programme.
- 114 Overall, the performance of the Project compares extremely favourably with equivalent Government Information Technology projects. The Census Offices' relationship with the Contractor was managed in a very professional manner and was the cornerstone of the Project's success. The Project provided a good model on which future similar initiatives can be based. There were, however, a number of lessons learned that can be carried forward.
- 115 The Project involved a number of separate contracts awarded for a range of services provided. This created number of problems resulting from the complex interface issues between the separate contracts. **To avoid these it may be beneficial in future to award all related and inter-dependent contracts to one supplier. The benefit of this needs, however, to be balanced with the risk that one supplier they may not be a specialist in every area within the contract.**
- 116 **To overcome this consequence, additional experts can be employed to advise. For example, concerns about the print aspects for 2001 Census Forms were mitigated by buying-in the services of an expert in forms design and printing for scanners/Optical Character Recognition (OCR), to support the Census Offices. Overall, the decision to place the onus for printing scannable forms with the contractor responsible for processing turned out to be sound and avoided a potentially difficult interface between two contracts. But wherever the cut-off is placed, the contractor needs to understand subsequent processes to ensure as seamless an interface as possible.**
- 117 **On balance, the UK Census Offices have concluded that a single contract could well have improved overall effectiveness. There is scope for combining processing, transportation, all printing, storage (and perhaps payroll) and call centre activities into one contract, dependant on a detailed assessment at the time. Arguably, the inclusion of the secondary processing activities, that is, those which managed the subsequent Edit, Imputation and One Number Census (ONC) processes, would have avoided an interface, but overriding factors mitigated against this.**
- 118 The lack of time between award of Contract and the Census Rehearsal (in April 1999), and the consequent delays, had a serious impact on the time available for the development and testing of systems for 2001 Census and ultimately led to a number of recommendations, **in particular that procurement must consider awarding the contract earlier in the planning cycle .**
- 119 It should also be recognised that governmental processes, such as the legislative programme, can be slow and may impact severely on the timetable. **Awarding the**

contract earlier would enable the Contractor to gain a better understanding of the emerging requirements and to build a solution by stages, rather than following a ‘big bang’ approach as was the case for the 1999 Rehearsal.

- 120 Any future contract must cater for managing the need to test changes between any Rehearsal and Census itself. When testing that the contractor’s system meets the necessary requirements, both parties need a clear understanding of the objectives of each test and to share a common understanding of what constitutes an acceptable result - acceptance criteria should be agreed up front and met before moving on to the next phase.
- 121 To achieve this, a large scale ‘live’ test should be run using the final version of the forms and using a much larger test base, at least 1 million forms. It should be the final test before ‘live’ operations. Prior to that there should be as many dry runs as possible using the forms collected from earlier field tests.

Publicity

- 122 The 2001 Census publicity campaign covered both paid and free media activity during the period October 2000 to May 2001. The main aims of the campaign were generating public awareness and promote a rapid return of forms, particularly by encouraging those who might be suspicious or resistant to taking part. This was achieved through:
- a successful advertising campaign;
 - using a co-ordinated approach to ensure that key public relations events were covered at both the national and regional level; and
 - the effective use of partnerships, most notably with BBC Local Radio and Public Relations agencies.
- 123 As well as the mainstream campaign effort, tactical initiatives were run to target those groups known to be especially liable to under-enumeration. Campaign activity promoted the key messages of confidentiality and the social benefits of the Census, which had been shown by research to be the most influential and important to the public.
- 124 Furthermore, the ‘Count Me In’ slogan (see example in Figure 1) in conjunction with the handprint logo and a raised hand motif provided a branding or unifying device which conveyed the underlying ethos of the Census and invoked the power of inclusion. These devices were successfully translated into a range of products that had a cumulative effect on the target audience through press, poster and television executions. The logo with its bright yellow background formed one of the most striking and recognisable aspects of Census publicity.
- 125 Despite the overall success of the campaign, the value of the innovative approaches adopted and the success of the partnerships with radio, television and the press, there were a number of issues that should be considered in the planning and implementation of future Census public relations activities. Key lessons include:
- **The difficulty of developing a creative concept for the Census which informs, convinces and yet at the same time entertains both mainstream and niche audiences;**
 - **The advertising campaign showed that a simple but positive idea delivered with frequency provides the most successful formula; timing must be well judged.** The early appointment of the selected Agency enabled the ONS to produce a core concept that could be introduced into key items like the Census form. At the same time the late

production of advertising meant that ONS were able to keep its options open and take account of current events as demonstrated by the inclusion of a Welsh scene in the television advertisement;

- **Researching facts and figures, historical information and points of special interest paid dividends throughout the campaign.** This helped ensure that much of the coverage was interesting and helpful to the Census Office rather than leaving an opportunity for the media to cover with negative angles. In addition by offering information broken down to a local level, regional coverage was increased considerably;
- **The value of identifying fixed events within the public relations calendar and the combinations of information, interest, fun and local people, stimulated a steady stream of positive coverage in advance of the Census, particularly at regional level;**
- There was insufficient time for the preparation of information to support individual news stories and as a result material for Census Area Managers was delayed. This hampered local media activities. **The campaign would have benefited from recruitment of key individuals at least two to three months earlier with more time to take forward detailed projects;**
- **The need to develop a positive post-census feature material was identified;** in the absence of such new material for the 2001 Census the weekend papers ran an exaggerated story about tens of thousands of undelivered forms;
- **The Census can become a focus for a wide variety of concerns and special interest campaigns and can be used to draw attention to a range of political issues.** For example, in 2001 in Wales it became a rallying point for the devolution campaign, and the Royal National Institute for the Blind used the Census to highlight their campaign to change the Disability Discrimination Act.

Edit and Imputation processes

- 126 Users of Census data in the UK asked the Census Offices, from the outset of consultation, to provide output from the 2001 Census that is *complete* and *consistent*. They did not wish to fill in gaps in tables containing ‘not known’ responses by having to make their own estimates for missing values. The pattern of non-responses is often different from that of reported data, and without access to the individual records users would not be able to correct for such non-response bias or accurately estimate values for derived variables that were based on more than one item. There would also be a danger that different users would make estimates for the complete population in different ways, creating inconsistencies between their results.
- 127 An Edit and Imputation strategy was therefore put in place with the aim of estimating for all missing data and resolving inconsistencies in responses for the people and households affected. The strategy adopted an number of principles:
- only changes that would improve the quality of the data would be made;
 - the number of changes to inconsistent data would be kept to a minimum;
 - as far as possible missing data would be imputed for all variables, so as to provide a complete and consistent database; and
 - the system had to be relatively easy to develop and capable of processing large amounts of data automatically within short timescales.

- 128 In developing a system for the 2001 Census the experience of the systems used in the 1991 Census was built on. Extensive tests were carried out on the use of neural networks, which can detect complex relationships in data without the need for using complex modelling techniques. However, the testing failed to impute results which were consistent with the edit rules and neural networks did not perform as well as the 1991 'hot deck' method.
- 129 A donor edit and imputation system was also trialled, whereby all missing or inconsistent values on a record would be adopted from another similar individual (the donor). However, it was decided that setting specific values in the editing routines rather than basing them on similar donors would be more operationally efficient (although possibly less statistically accurate).
- 130 Thus an Edit and Donor Imputation System (EDIS) was devised for the 2001 Census, which was applied to individual records once the data had been loaded by ONS. It was designed to fill in almost all the gaps in records for existing people and households. The One Number Census process (see above) imputed for whole households and people who were missed from the Census. EDIS modified values for individuals on returned census forms.
- 131 A set of **Edit** rules was applied to missing items or responses which appeared to be in error or inconsistent when compared with other data (such as married couples of the same sex, a child less than 13 years younger than its parents, or a married person under 16). These were known as hard checks. In determining how to resolve such inconsistencies, the Fellegi/Holt principle of making the minimum number of changes was followed as far as possible. Thus if a person was recorded as being aged under 16, married and had answered employment questions such as occupation, the 'age' would be set to missing, since the inconsistency could be resolved with the least change by imputing a value for age between 16 and 74.
- 132 The system also identified unlikely, but not impossible responses. In some cases rules were applied to eliminate these: for example, a purpose-built flat was considered unlikely to have more than 10 rooms, the value was set to 'missing' for subsequent imputation; in other cases, for example, where people aged under 35 were retired from paid work, no further action was taken. The number of these 'soft checks' was reported but the data were not changed as a result.
- 133 **Imputation** was applied when there was no answer on the Census form, or when a response failed the multi-tick rules or was invalid, or when the filter rules or Edit marked it for imputation to resolve an inconsistency.
- 134 The principle of a Donor Imputation System is to search for a single donor household to supply *all* the missing variables in a recipient household. Exceptions are imputation for postcode of usual address one year ago and postcode of workplace, which were carried out at a later stage than imputation for all the other variables.
- 135 Although the EDIS system performed fully to specification, and within planned running times some aspects might have worked better if there had been more opportunity for testing.
- 136 It did not prove possible to carry out a full run-through of the EDIS system on the data collected during the 1999 Dress Rehearsal. This would have afforded the opportunity of checking whether ideas that appeared sensible in theory would stand the test of practical application on live data. **Rehearsal data needs to be delivered more rapidly, at least on a small scale, or the Rehearsal itself should be brought forward so that it takes place more than two years before Census day.**

- 137 Inability to test the system meant that assumptions about how well the public would follow instructions, for example on answering or omitting certain questions, proved to be not entirely valid. This impacted most noticeably on imputation of age, although the extent of systematic error turned out to be very slight.
- 138 Within EDIS, a number of assumptions were based on age being correct rather than other items. However, year of birth was occasionally mis-stated, not scanned correctly or given a wrong value during processing. Particularly when there was an error in the next to last digit of the year, EDIS may have imputed for a range of items where no value was needed, or conversely set reported data to 'no code required'. Further checks could have been tested if more time had been available to investigate real data from the Rehearsal. Nevertheless, a few households contained multiple errors that would have been difficult to resolve accurately by any automatic editing system.
- 139 Late changes to the questionnaire design had some impact on EDIS. For example, splitting the qualifications questions into two parts, which occurred after the 1999 Rehearsal, meant that new rules had to be devised for 2001 to deal with the new section on professional qualifications. This turned out to be the question having the largest non-response rate as many people considered that it did not apply to them. Furthermore, as extra room had to be found on the form for the expanded qualifications question, the question on 'work last week' was consequently squashed into a smaller space. As a result, two of the bullet points which appeared separately on the Rehearsal form were combined into one, and it appears from the pattern of responses to this and later questions that some form-fillers misunderstood the question and answered 'No' when they were actually in work.
- 140 A resolution was found to this problem by amending the filter rules for the derivation of Activity Last Week, but a small number of answers may have been miscoded as a result of the extra complication which was introduced. **Late question changes should therefore be avoided in any future census.**
- 141 A single edit and imputation system was designed to deal with the Censuses in England, Wales, Scotland and Northern Ireland, which all had slightly different requirements. Variations in the design of the Census form and in editing requirements meant that great attention had to be devoted to ensuring that the processing for each country was carried out to the desired standards. **A single system may not be possible if there is further divergence between future census forms across the constituent parts of the UK.**

CONCLUSIONS

- 142 It is still too early yet to evaluate fully all aspects of the planning and execution of the 2001 UK Census. From the evaluations prepared to date performance is, in some respects as good as that in the 1991 Census, in some other respects it is better, but in others less so.
- 143 For example, looking at the components of the overall census response and coverage rates given in Table 1 below it can be noted that although the proportion of the population recorded on completed forms received fell from 96 per cent in 1991 to 94 per cent in 2001, the 2001 Census was better able to estimate the number of people in dwellings that were identified as occupied but from which no forms had been received (4 per cent compared with just 2 per cent in 1991). Furthermore, in 2001 the One Number Census process was then able to estimate the number of people not included on returned forms and in households that were entirely missed in the enumeration (a further 2 per cent), and to include these in the base population for all statistical outputs. Thus, while the 1991 Census

overall coverage rate was 98 per cent, the 2001 Census was able to produce output representing 100 per cent of the estimated UK population.

Table 1 Components of UK Census response and coverage rates (percentages) 1991 and 2001 Censuses

| | 1991 Census | 2001 Census |
|---|-------------|-------------|
| Form response rate | 96 | 94 |
| Other people identified in households | 2 | 4 |
| Overall census response rate | 98 | 98 |
| People not included on returned forms and in wholly missed households | 2 | 2 |
| Overall census coverage rates in results | 98 | 100 |

REFERENCES

Martin D (1998). 2001 Census output areas: from concept to prototype. *Population Trends* 94, pp 19-24.

Martin D (2002). Geography for the 2001 Census in England and Wales. *Population Trends* no. 108, pp 7-15.

SUMMARY

The 2001 UK Census was in many ways a traditional census. Planning of the 2001 Census built upon the strengths of the successes from previous censuses, as well as the lessons learned from the 1991 Census. There were, in particular, a number of issues arising from the 1991 Census and subsequently, which needed addressing. Several aspects of the planning of the 2001 Census attempted to rectify one or more of these problems. This, and the need to respond to a greater range of users' requirements for information, increased public concerns about confidentiality, the need to reduce the burden on the public, and to incorporate up-to-date technology to enable the Census to provide better value for money, gave rise to a number of major innovations in 2001 making it far less 'traditional' than had previously been the case.

The paper provides an overview of the evaluations of a number of the innovations introduced into the 2001 UK Census to address the several difficulties encountered in 1991, and highlights where these lead to significant successes and benefits as well as noting where further lessons can be learned for the planning of any future census. The paper focuses on: the date of the Census; question content and form design; the ethnic group question; data collection activities; community liaison; designing census geographies; the One Number Census methodology; outsourcing census activities; and the edit and imputation system.