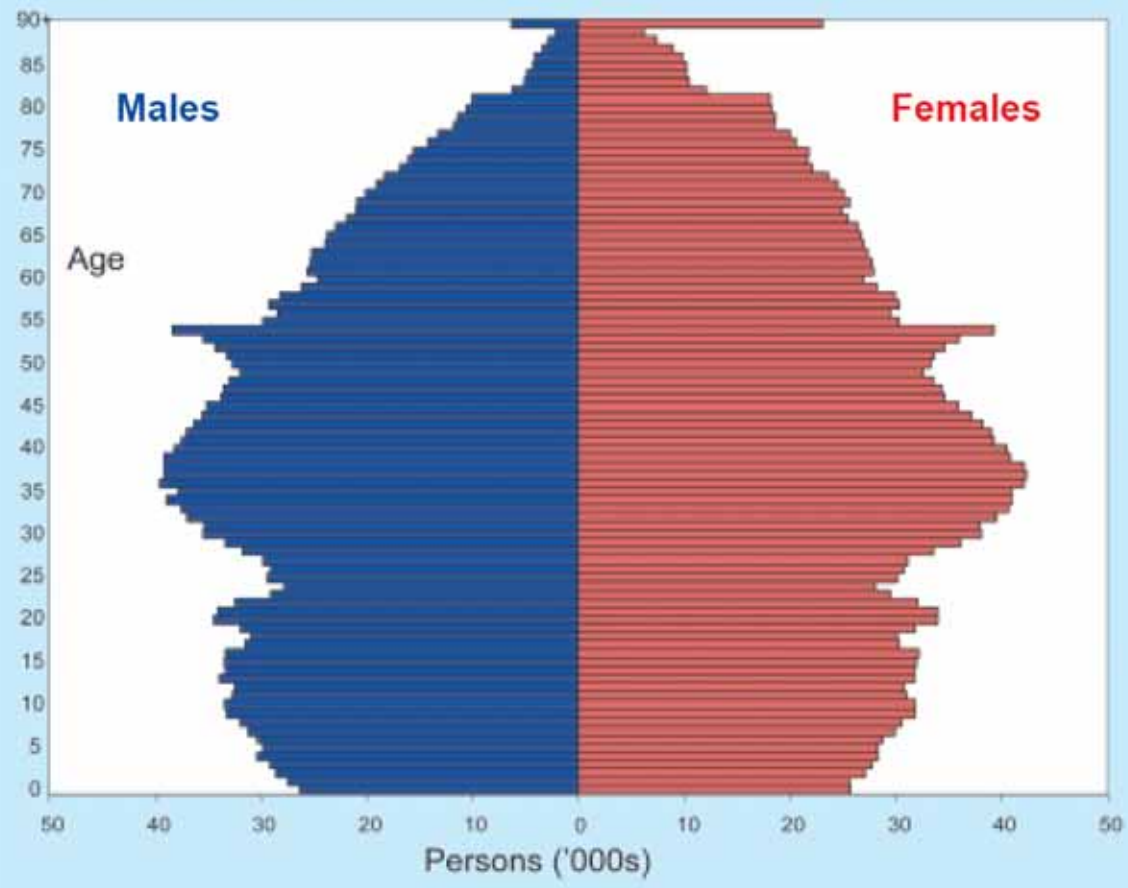
Scotland's Census 2011

Scotland Counts

Autumn 2004 Census Consultation

Estimated population by age and sex, 30 June 2001



General Register Office
for
SCOTLAND
information about Scotland's people

Table of contents **Page**

1 Consultation Plan and Timetable

1. The purpose of this document is to set out the proposals of the General Register Office for Scotland (GROS) for the 2011 Census in Scotland and to seek the opinions of Census users on all aspects of the Census, but particularly in the following areas:

   - Should the Census count base be 'Usually Resident' or is there a better population base?;
   - How should we prevent disclosure of personal information?;
   - How should we correct for of under and over enumeration?; and
   - When should the results be published?

2. Our guiding principle in preparing for the 2011 Census is:

   "*to provide high quality population statistics as required by key users, 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*".

3. This document has 5 sections:

   1. **Consultation plan and timetable;**
   2. **Evaluation of the 2001 Census in Scotland;**
   3. **Statistical strategy and design for the 2011 Census in Scotland;**
   4. **2006 Census Test design; and**
   5. **Summary and conclusions.**

4. We must ensure that the 2011 Census takes account of:

   - Lessons learnt from 2001;
   - Changes in demand for data;
   - Developments in data collection; and
   - Administrative data availability.

5. The views of Census users and other interested parties are vital. This document is the first step in the consultation process which will continue throughout the period of preparation (Table 1). Within the document those points on which we particularly welcome feedback from users are highlighted. Those points are printed in bold and are also listed in Annex B. More general views are also welcome.

6. In preparation for the 2011 Census, we plan to hold a Census Test in 2006 and a Census Rehearsal in 2008. We have to strike a balance when finalising particular aspects of the Census design because we must be sufficiently flexible to take account of changes in government policy, the demand for data and developments in data collection and technology. We must also leave sufficient time between the finalisation of our design and the date of the Census so that we can prepare adequately and ensure our plans offer good value for money

We invite users' views on the outline consultation timetable and methods (CP1).

If you respond to this consultation please ensure that you make clear the following:

- Whether you are responding as an individual or on behalf of a group or organisation;
- Whether you agree to your response being made available to the public;
- Whether you wish your response to be made public but anonymously; and
- Whether you are content for the General Register Office for Scotland to contact you again in the future in relation to this consultation response.

We will share your response internally with the other 3 Census offices. As GROS address the issues you have raised, we may wish to contact you again in the future, but we require your permission to do so. Could you also please give such permission?

Table 1 Consultation Plan and Timetable

Topic	Form of consultation	Timing
Question consultation	Web-based	3 monthly updates, on going
Formal Consultation on all Census issues	Meeting, with 12 week response period. Publication of findings	Nov 2004 Mar 2005
2006 Census Test questions	Final e-mail consultation	Aug-Sept 2005
Census Test	Survey of some 60,000 households (in Glasgow, Killin-Ballachulish area, and West Dunbartonshire)	April 2006
Formal Consultation on all Census issues in light of Test results.	Meeting with 12 week response period - particularly to consider output details with population bases and under enumeration Publication of findings	Nov 2006 Mar 2007
2008 Census Rehearsal questions	Final e-mail consultation	Aug-Sept 2007
Census Rehearsal	Survey of households	April 2008
Formal Consultation on all Census issues following Rehearsal	Meeting with 12 week response period. Publication of findings	Nov 2008 Mar 2009
2011 Census questions	Finalised by e-mail consultation	Aug-Sept 2009
Census		April 2011

In addition to the consultation with users and the wider public described above, the Scottish Parliament will be consulted at appropriate stages, culminating in the approval of the Order and Regulations under which the Census is conducted in Scotland.

2 Evaluation of the 2001 Census in Scotland

2.1 Introduction

7. This section summarises the GROS evaluation of the 2001 Census. We want to build on our evaluation to create a better Census in 2011. We also refer, where relevant, to the findings of reports on the Census in England and Wales by the Treasury Select Committee (TSC)[1] and the Statistics Commission[2]. Brief references are made to the strategies we plan to use to overcome specific difficulties encountered with the 2001 Census. These strategies are described in greater detail in later sections.

8. The 8 key objectives for the 2001 Census were:

1. To carry out the 2001 Census fieldwork in April/May 2001, with suitable publicity.
2. To ensure 2001 Census form data capture and coding is complete by December 2001.
3. To ensure that the procedures to capture and clean data from forms and create databases for analysis comply with predefined quality standards.
4. To decide, after consultation, the products and services to be delivered using 2001 Census data.
5. To produce counts of the population by age, sex and area by August 2002 taking account of any Census undercount.
6. To produce predefined outputs from the Census for all areas and covering all topics by March 2003.
7. To maintain a Geographical Information System to assist Census planning and meet other GROS geography needs.
8. To evaluate and review the 2001 Census programme and make recommendations for the future.

We consider that these key objectives were met. Do you agree? If not, which of these objectives do you believe were not met and what lessons can be learned? (CP2)

[1] Select Committee on Treasury, First Report, 2001 Census in England and Wales. http://www.publications.parliament.uk/pa/cm200102/cmselect/cmtreasy/310/31002.htm
[2] http://www.statscom.org.uk/media_pdfs/reports/Census%202001.pdf

2.2 Overview

9. The 2001 Census took place on 29 April 2001. It had many novel features compared to previous censuses:

- We counted the number of "usual residents" as in 1991 but included students at their term time address;
- More information was obtained about household relationships than in 1991. The 1991 Census asked for the relationship of each person to the head of the household; the 2001 Census asked for the relationship of each person to every other person;
- Questions on religion were included for the first time;
- The question on ethnicity, first asked in 1991, had more tick-box categories and write-in options;
- There were questions on general health and carers;
- The Gaelic question also included 'understanding';
- The travel-to-work question included travel-to-study. It gives daytime populations and links home and school;
- Census forms were returned by post, instead of being collected by Census enumerators;
- Optical Character and Mark Recognition (OCR and OMR) allowed the capture and coding of all the data on the Census forms, including write-in responses (instead of a sample for some questions, as in 1991);
- The Census database was adjusted for under enumeration at an individual and household level to create a 'One Number Census', with the results for small areas consistent with those for Scotland;
- Results were made available on the internet through the Scottish Census Results On Line (SCROL) website[3]; and
- Results were also available on CD and DVD.

10. We believe that this combination of new features improved the scope, quality and accessibility of the 2001 Census data.

[3] www.scrol.gov.uk

2.3 General Problems

11. This is not to say that the Census was without problems. Some of them were external and could not have been foreseen. For example, the February 2001 outbreak of foot and mouth disease restricted access to farm communities in parts of Scotland. As a result, we changed enumeration and follow-up methods. This sudden adaptation – mainly to use postal delivery in rural areas - was largely successful.

12. Some problems on the other hand could be foreseen, allowing us to work toward tackling them - such as reduced response rates among parts of the population.

13. However many problems were caused by the nature of the Census operation itself – a one-off large scale operation, generally larger than other data collections in the UK. For example, the 2001 UK Census employed more than 80,000 field staff over a short period. The pay system was complicated, as it had to reflect the different stages and types of work. While most field staff were paid on time, a significant number were not, with adverse effects on those people, enumerator morale and staff stress.

14. An area which also caused major problems was ramping up form processing from the numbers in the 1999 Census Rehearsal to the full requirement in 2001.

15. Our main response to these general problems is to bring forward aspects of the Census development by 1 year. This is in line with international recommendations[4]. It also allows us more time to develop final processing between 2008 and 2011.

16. However, not all of the changes we are proposing for 2011 are connected with failings in 2001. Many are in response to changes in data infrastructure in Scotland.

[4] http://www.unece.org/stats/documents/ces/2003/28.e.pdf (Page 4, Para 3)

2.4 Pre-Census Exercises

17. As with all Censuses, it was necessary, for 2001, to test the effectiveness of field methods. There were 2 large scale exercises: the 1997 Census Test and the 1999 Census Rehearsal. We propose, for the 2011 census cycle in Scotland, to bring these exercises forward 1 year, to 2006 and 2008. That will give us time after the 2008 Rehearsal to evaluate and implement lessons and to rehearse processing as fully as possible. The other UK Census Offices are proposing to test in 2007 and 2009. This will increase the scope for learning and sharing lessons.

2.4.1 1997 Census Test

18. The 1997 Test in Scotland was carried out, as part of a UK-wide exercise, in parts of Glasgow and SW Argyll. It covered about 16,000 households. The main purposes were to assess new questions (including religion and income), to evaluate new enumeration strategies such as post-out/ post-back and to check the page-per-person form design.

19. At the UK level, inclusion of an income question seemed to reduce the response rate, though in Scotland the results were not decisive. The religion and ethnicity questions seemed to be acceptable. The page-per-person form was deemed easier to answer than the matrix design. Post-out/post-back performed poorly in Scotland - though there were financial savings.

20. As a result of the UK experience, the income question was dropped from the 2001 Census form. For 2011, we propose to develop and test an income question as previous consultations have shown that users want it. We therefore plan to trial this question in the 2006 Census Test.

2.4.2 1999 Census Dress Rehearsal

21. The Rehearsal covered some 24,000 households in parts of Angus and Dundee. Its main purpose was to check the

procedures to be used in the Census and the follow-up coverage survey. It was also intended as a final test of the questions although, in the event, the questions on religion were not included in the Rehearsal but added subsequently.

22. We plan to bring forward the Census Rehearsal by 1 year to 2008 to allow a full rehearsal of processing. This move reflects international evaluations mentioned previously [5] and the lack of time we experienced before 2001 to complete a processing rehearsal and implement improvements.

2.5 2001 Census Evaluation

2.5.1 Form Design

23. In the 2001 Census, there were 10 household and 34 individual questions. In Scotland, the main household form was designed for up to five people. The form was 20 pages long. It was estimated that it would typically take about 10 minutes to complete the form. We think the page-per-person form design was successful.

24. Our current view is that we may remove some of the 10 2001 **household questions** either because the information is no longer useful (e.g. questions on bath and shower access have few "no" responses and therefore are no longer a good measure of deprivation), or is available from another source (e.g. dwelling type and number of rooms from the Scottish Assessors Association Portal[6] - although there are questions to resolve about data quality and comparability between the 14 Assessor areas). It would probably still make sense to have 1 page of household questions, which gives us the opportunity to ask new household questions (e.g. on internet access).

25. The expanded question about the relationship between people in the household worked[7]. Although seemingly

[5] http://www.unece.org/stats/documents/ces/2003/28.e.pdf (page 4, Para 3)
[6] http://www.saa.gov.uk./
[7] http://www.unece.org/stats/documents/2003/05/census/wp.5.e.pdf &

complicated, it did capture household relationships successfully. GROS propose to keep it.

26. Although we have no concrete plans for **individual questions** yet, some data may be available from other sources (see Annex 1). Potentially the individual question pages could be reduced from 3 to 2 pages. If the standard form, as in 2001, has space for responses from 5 people, this could reduce the form by 5 pages. Since printing processes dictate that the form is in multiples of 4 pages, this may allow some visitor information to be collected within a 16 page form, or could provide space for 6 people and visitors within a 20 page format.

27. The issues of population bases and visitor information are covered in more detail in Section 3.7. Question-specific issues are covered in Section 4.2.

28. We also need to consider other effects of form length. In 2001 the Canadian Census[8] used a mix of long and short forms. The short form had a 5% higher response rate, although follow-up procedures reduced this to 1%. Therefore reducing the length of the form may not be critical to response rate. However, because form length does have a direct effect on the cost of printing, distributing and processing, we must take this into account when considering adding questions.

2.5.2 Accessibility Issues

29. Coverage is paramount, as the Census is used to plan, among other things, improved accessibility to services. In Scotland, 193,000, 43,000 and 235,000 people receive incapacity benefit, severe disablement allowance and disability living allowance respectively[9]. So we must take account of the needs of people with disabilities, or who find it difficult to read or understand written English or have

---

I. Máté and G. Miller. The UK 2001 Census Question on Within Household Relationships, Statistical Journal of the UNECE, V 20, 1, 27-37.

[8] http://www.statcan.ca/english/sdds/3901.htm#InfoSurvey

[9] http://www.scotland.gov.uk/library3/society/equality/esd-00.asp

other characteristics which complicate completion of a Census form.

30. Due to the difficulties of adapting form processing systems, in 2001 the Census organisations did not adapt Census forms for people with a visual impairment or who did not read English. However, GROS did provide other support mechanisms for people who found it difficult to fill in a Census form including:

- Language interpreters accessed by Census staff through local authority Census Liaison Officers;
- Leaflets translating the Census questions into 7 languages (Gaelic, Arabic, Bengali, Chinese, Hindi, Punjabi and Urdu), together with information leaflets;
- Braille version of information leaflet;
- Large print version of the information leaflet;
- Tapes, distributed via the Talking Newspaper Network, giving a general overview of the purpose of the Census and the questions, and details of help available for those with visual impairments;
- Promotional video with subtitles; and
- Dedicated help lines for people with hearing impairments and those whose first language was not English.

31. Currently, there are some 37,000 people in Scotland registered as blind or partially sighted[10]. This is the minimum number of people who have eyesight difficulties which would make it difficult or impossible to access explanatory documentation and complete a Census form autonomously; in addition there are many people with sight problems who are not formally registered.

32. The RNIB criticised the Census because of the form text size and the lack of contrast in the form design. Some text was orange and therefore did not show up well against the form's pale red background. This made it difficult for people with a visual impairment to read and complete.

[10] www.rnib.org.uk/xpedio/groups/public/documents/code/public_rnib003462.hcsp

33. The RNIB state that 12 pt text can be read easily by only 50% of the population, whereas 14 pt can be read by 95%. Of course, an increase in text size results in a longer form and greater form production and processing costs. We will experiment with a larger text size form.

34. We also propose to review our form colour design. In 2001, optical scanning was a new technology for us. We were advised by the scanning contractor on which colour contrasts were best for OCR and optical mark recognition OMR. In 2011, we intend that the design will also better reflect form-filler needs.

35. In 2001, a standard form had to be completed by or on behalf of people with disabilities and language difficulties. As a result, these people did not have the same level of confidentiality as other form fillers. We will consider processing alternative format forms manually.

36. One of the proposed areas for the 2006 Census Test is in Glasgow, with relatively high percentages of different ethnic communities. That will allow us, in co-operation with the Racial Equality Scheme Implementation Group (RESIG), to test ethnicity questions which are acceptable to all ethnic communities and which allow the monitoring of racial equality.

2.5.3 Field Recruitment

37. Field staff recruitment and retention was more difficult in 2001 than in 1991, especially in urban areas where staff reserves were quickly used up due to resignations. The fact that 7,000 enumerators were employed, compared to 12,000 in 1991 did ease the pressure on recruitment somewhat. We expect to use about the same number of field staff in 2011 as in 2001.

38. In 2006, the Census Test will include some areas where recruitment was difficult. While, because of the relatively small size of the Census Test, we do not expect recruitment

problems and so will not be able to evaluate recruitment fully, we will plan from the start for enumerators to do double workloads if required and also consolidate the bonus and incentive schemes into the basic fee where possible to make the headline lump sum more attractive.

39. In 2001, we followed Scottish Executive guidelines on fair recruitment. However, only 56 enumerators (0.8%) and 1 Census District Manager (0.6%) were recruited from ethnic minorities. The follow-up survey also had 1 Team Manager and three interviewers from ethnic minorities.

40. Our ideal is to increase the number of people from ethnic minorities working in the 2006 Census Test, especially since we have chosen for the Test areas in Glasgow with relatively high percentages of different ethnic communities.

2.5.4 Pay

41. In 2001, the contracted-out field staff pay operations went wrong. There were problems ramping up an operation for the whole UK field staff of about 80,000 people, made more complicated by the short periods of employment of some staff and the complexity of pay forms and rates of pay.

42. We plan to develop an in-house pay system and test it in 2006. Such a system was used successfully in 1991. The few difficulties encountered then should no longer apply because of technological improvements, including better software and automatic bank account payment systems. We will assess, as part of the Test, the value for money of the system.

2.5.5 Communications

43. In the 1997 Census Test, as part of our strategy to inform the public about the Census, the Royal Mail delivered information leaflets in advance of the Census form. Sometimes the leaflet did not arrive until after the Census form had been delivered. We will try to create a firmer contractual relationship with the postal service if we use

post-out/post-back, or the enumerator will deliver the information leaflet and form.

44. In 2001, the enumerator had to try to make contact with all households within 2 visits. On the first visit, if unsuccessful, a “no contact” leaflet was delivered. This leaflet stated that the enumerator would call again ”within the next few days”. Since forms were delivered over an 18 day period, this phrase was misleading and the Census Helpline was over-whelmed by calls from householders who thought their Census form was not being delivered.

45. The TSC noted that the number of calls to the Helpline had been vastly under-estimated. In addition to seeking to improve the estimate of call volumes and the strategies to deal with high call volumes, it recommended that, in order to reduce the pressure on the Helpline, ways should be sought to reduce the amount of help required by the public.

46. We also hope to learn from the Census in Ireland. They noted the misleading text on our Census leaflet and took steps to avoid the problem. We hope to use their methods and estimate our call centre burden from their 2006 Census. We will re-plan our pre-Census day approach.

47. Some enumerators left a “Reminder” leaflet in place of a “No-contact” leaflet, an error which was attributed to the similarity of their appearance. This was a problem since the Reminder leaflet implied non-compliance on the part of the household and therefore risked confusing and antagonising householders. We will try to be more aware of the potential of colour coding leaflets etc to avoid confusion.

2.5.6 Form Collection/Return

48. In 2001, for the first time in a full Census, householders were asked to post back their completed forms. At the planning stage, we assumed that at least 70% of forms would be returned by post without additional action by enumerators. In practice, the figure was 72%.

49. The Royal Mail in Scotland experienced difficulties handling the 2 million plus items of mail over a short period. The resulting log jam caused three main problems:

- There were instances of “dead time” when enumerators had no forms to check;
- When the log jam cleared there was a deluge of forms; and
- When enumerators followed-up missing forms, they could not be sure whether a form had not been returned or whether it had been posted but not yet circulated back to them. This risked de-motivating enumerators as well as annoying the public and reducing confidence in the Census.

50. The TSC recommended that the procedures for form delivery and collection be looked at again in the light of the problems experienced in 2001 – and that, if postal form return was used, a tight service level agreement should be made with the service provider to ensure that the Census was not impeded by the quality of service.

51. The TSC also commented that postback methodology seemed to have had an adverse effect on question response rate. They suggested that the Census Offices should consider what effect delivery and collection methods had on the response rate to particular questions. Response rates are shown in Table 7 toward the end of this document.

52. Post-out/post-back is the cheapest enumeration method because enumerator costs can be reduced. While delivery/ collection retains most control over data quality it is probably the most expensive. Changes, such as a single delivery round rather than 2 attempts to contact households, would reduce costs. Delivery/post-back combines robust address and household validation with reduced collection costs. Post-out/post-back and delivery/ post-back both relinquish control of the Census collection to the postal service. Improvements to each method need to be evaluated for cost, quality, risk and practicality.

53. It is not yet clear how to overcome potential mail problems. We plan to trial local postback procedures (in which Census

forms returned by post go back to the enumerator) which may be the most efficient arrangement. But to improve response rates and data quality, we plan also to strengthen enumeration; the enumerator will be available to help complete the form and, if this is easier for the respondent, to collect it. This will reduce pressure on the postal system.

54. Ideally, to go down the postout postback route, we would want to agree barcode designs with the service provider and want them to introduce automated handling of C4 (229 x 325 mm) envelopes. We think Royal Mail is trialling this.

55. In short, in the 2006 Test, we plan to use a number of different form distribution and collection methods and should thereby get a clearer idea of which method, or combination of methods, should be used for the 2011 Census.

2.5.7 Adjustments for Under-Enumeration

56. The Census Validation Survey, which measured under enumeration in 1991, covered some 20,000 households across the UK. In 2001, the similar Census Coverage Survey (CCS) covered some 40,000 households across Scotland. The response rate from known addresses was 95%. This was above the design requirement of at least 90% and was achieved in all council areas. The sample fraction was larger than the design required.

57. People found by the CCS were compared to people found by the Census in the same postcodes. Matching identified:

- 'Wholly missed households' (which had not returned a Census form at all); and
- 'Missed people' in 'Census households' (households which had returned a Census form but not included all the household members).

The level of 'missingness' was extrapolated across the whole country. The survey method is more fully described

on the GROS website[11].

58. The estimated under-enumeration in 2001 across the entire population was approximately 4%. However, in some areas, and for certain demographic groups, this figure rose to 20%. In order to compensate for these missing people, synthetic people and households were imputed.

59. The number of people imputed relied on a Dual System Estimation procedure[12]. The method assumed that the two data 'capture' operations (Census and CCS) were independent - that is, that if someone is missed by the Census they have as good a chance of being 'captured' by the CCS as anyone else.

60. However, comparison with demographic analysis showed that this assumption was wrong. Therefore, we adapted the original under-enumeration method.

61. For the total number of occupied households, we used the enumerator count of occupied household spaces. We imputed households into all the occupied spaces for which there was no Census return based on the CCS characterisation of missed households and using local 'donors'. We assumed that the enumerator correctly reported whether or not the dwelling space was occupied. During 2006 testing we plan to revisit this assumption.

62. There was also a problem of how to estimate the number of people missed within Census households. In the matched CCS data the average ratio of missed people in wholly missed households and missed people in Census households was about 2:1. We assumed that, if this ratio was higher in an area, the CCS had not worked well there and so increased the number of imputed people in Census households until the ratio was 2:1. Thus, in Glasgow, the ratio from the field work, matching and dual system estimate had been 4:1. We then reduced the ratio to 2:1 by

[11] http://www.gro-scotland.gov.uk/grosweb/grosweb.nsf/pages/cenop7
[12] http://www.statistics.gov.uk/census2001/IntroOneNumber.asp

increasing the number of synthetic people in enumerated Census households.

63. Having re-estimated the required additional synthetic people by adding the two missed population components, we could not, with the imputation system as designed, constrain the system to produce a predetermined number of synthetic households, simply an overall number of synthetic people. In 2011, for any similar approach, we would develop a system which allowed us to constrain the number of households and the number of people by enumerated and missed households.

64. We necessarily made one further assumption - that the people missed by the Census and the CCS had the same overall characteristics as people missed by the Census but found by the CCS. This assumption is almost certainly wrong but we had no evidence from the Census and CCS system to build other assumptions on. We need another way of checking Census coverage.

65. The Statistics Commission recognised that the ONC methodology was the best which could have been used for a conventional Census in 2001 and in general it worked well, although it was unable to cope with the most extreme of circumstances[13]. Circumstances appear to have been more severe in England than in Scotland, especially in Manchester and Westminster[14] where the Census-based population estimates were subsequently adjusted upwards by about 26,000 and 17,000 people respectively.

66. GROS have not adjusted any Census-based estimates and do not currently believe this is necessary. However, we are working jointly with Glasgow City Council to investigate whether the Census underestimated Glasgow's population by 5-10,000 and we will make an adjustment to population estimates if there is proof of significant underestimation.

[13] Add in website reference
[14] http://www.statistics.gov.uk/downloads/theme_population/2001CENSUSLAPOPULATIONSTUDIES.pdf

67. The adjustments in the population estimates for Manchester, Westminster, and to a lesser extent other parts of England and Wales, were based on detailed studies which brought to light some causes of error in the ONC.

68. Some areas of Manchester had extremely poor Census response rates (60% or less). Response rates in Scotland were nowhere as poor as this. It is suggested that these extremely low response rates resulted from the large number of new houses which were not included on enumerators' address lists and then not identified by enumerators in the field.

69. In Scotland we were able to use more up-to-date address lists, issued 3 months before the Census. Our postcode boundaries are revised quarterly to reflect any changes and the maps provided to enumerators reflected these boundaries, as well as features on the ground. This meant that the enumerators could locate the postcode and enumeration district (ED) boundaries relatively easily.

70. The English studies also found that the characteristics of many areas changed during the inter-Census period, so that 2001 CCS coverage fitted in well with the hard-to-count index derived from 1991 data, but not with that derived from 2001 data. Re-stratification of 2001 results would have resulted in further delays in data publication. In Glasgow, GROS did enhance the CCS sample to cover particularly 'hard to count' areas, but it is not clear that this was completely successful, as those areas themselves were particularly prone to redevelopment.

71. The English studies also found that the degree of dependency between the CCS and the Census was much higher than expected. However, in Scotland, we adjusted for dependency within both Census households and wholly missed households, as far as the Census evidence would allow and using only Census evidence.

72. The English studies also found forms which had not been processed. We are not aware of any returned forms from

Scotland not being processed (other than late returns). Nonetheless, we feel we must strengthen form control procedures to ensure the correct identification of households and to contact under-enumerated groups.

73. Failure to complete a Census form is a criminal offence and offenders risk a fine. The TSC noted that, despite this, the number of prosecutions for non-compliance with the Census was very low (of 8 cases sent to the Procurator Fiscal in Scotland, 5 were prosecuted successfully). The TSC were concerned about the message this sent out about the legal duty to complete a Census form and recommended that the reasons for this low rate of prosecution be examined.

74. We took the view that we would rather have a completed Census form than prosecute. Consequently procedures were geared more towards the former than the latter. The prosecution process was slow, complex, inefficient and expensive. Although information leaflets stated that the fine was up to £1,000, none of the eventual prosecutions levied fines near to that level. We plan to discuss with the Crown Office whether a different approach is warranted in 2011.

2.5.8 Data Capture and Coding

75. In 2001, data capture and coding was out-sourced. It cost £5.77m, or £2.69 per household form - much cheaper than the estimated cost for an in-house process of £3.72 per form. Data capture had 2 elements: OCR for capturing write-in responses and OMR for tick-box responses. The response to some questions was captured by both methods.

76. Capture and coding of data was not completed until March 2002, 4 months behind schedule. This was due to problems in exporting data from the processing centre.

77. A full evaluation report on data capture and processing is being drawn up by the 3 UK Census Offices. One lesson that has been drawn is that we can strengthen enumeration quality by pre-printing addresses onto Census forms, thereby reducing processing costs and potential for errors.

78. There were also problems with the accuracy of some of the scanned data. Table 2 gives the data capture accuracy rates and Table 3 gives the coding accuracy for some topics. In most cases expected standards were met. The exceptions are in bold in the tables.

Table 2: Census Form Data Capture Accuracy

Scotland – Capture Accuracy (ONS Analysis)				
Data Category	Standard	Average Achieved	Maximum Achieved	Minimum Achieved
OMR	99.3	99.6	100.0	99.2
OCR Alphabetic	96.0	99.1	99.6	95.2
OCR Alphanumeric	95.0	99.1	100.0	98.2
OCR Numeric	98.0	99.5	100.0	98.9
Date of Birth	99.5	99.9	100.0	99.6
Form Identity	100.0	**98.8**	100.0	79.8

Table 3: Census Form Data Coding Accuracy

Scotland – Coding Accuracy (Lockheed Martin Analysis)					
Data Category	Standard	Av. Achieved	Max Achieved	Min Achieved	ONS Assessment
Country of Birth	96.0	99.9	100.0	82.4	99.8
Ethnic Group	96.0	99.0	100.0	66.7	96.5
Industry	88.0	88.5	94.6	83.7	88.2
Occupation	88.0	91.4	100.0	86.8	**87.3**
Enumeration Address	100.0	**99.2**	100.0	75.0	**97.9**
Address – 1 year Ago	96.5	**93.0**	100.0	60.0	**92.1**
Address – Workplace	94.5	**92.5**	100.0	85.2	**85.7**

79. One particular area of concern was form identity. Form IDs were handwritten by enumerators and some were incorrect or illegible, hence the comparatively poor accuracy. This had serious implications for data quality since bias, a function of handwriting quality by ED, was introduced, e.g. if an enumerator always wrote a poorly formed 7 which was then interpreted as a 1, there was a bias. Errors were rectified by cross-checking ED and postcode to identify potential capture errors and allow manual correction.

80. In the 2006 Test we will assess the costs of insisting on 100% form completion. Although we recognise the difficulty of doing this within a voluntary test, the information leaflets will stress the importance of 100% form completion and enumerators will assess the feasibility of follow-up and doorstep completion.

2.5.9 Media Strategy

81. The Census media campaign was launched by a press conference on 8 March 2001. The publicity campaign was multi-faceted, with paid-for advertising in print and broadcast media, together with coverage generated by news releases. A comprehensive media pack was distributed to every local newspaper, TV and radio station in Scotland, all business publications and every journalist who expressed an interest.

82. The overall purpose of engaging the media was to communicate 6 key messages, which are listed below with an evaluation of how effectively each was communicated.

- Census results are vital to improving Scotland: This was repeated often in extensive media coverage.

- First Census of the Scottish Parliament: It was not highlighted by the media. It did not significantly enhance the core message of the Census - the importance of filling in and returning a form.

- The impartiality of the Census: This message was stressed throughout and illustrated by all-party support for the Census.

- Strict confidentiality of the Census: The media used it repeatedly.

- Legal obligation to complete a Census form: This message was only used by the media when Census forms began to be delivered. A “softly-softly” approach was used initially,

followed by a harder tone post-Census day. The media stressed the risk of prosecution. We think this had an impact because many callers to the Helpline asking for forms stated that they had no wish to be fined.

- Existence of help in filling in a Census form: Widespread coverage in the media, which probably increased the number of callers to the Census Helpline.

2.5.10 Cost

83. The total cost of the 2001 Census (spread over several years) was £254m (£4.32 per head) in England and Wales, and £32m (£6.32 per head) in Scotland. The Scottish figure was higher largely because of lack of economy of scale and also the remoteness and inaccessibility of many areas.

84. The cost of the Census was criticised across the UK: for instance, the TSC did not believe that the level of expenditure for England could be justified and recommended that a rigorous cost-benefit analysis of the 2001 Census should be carried out, and that the cost of future Censuses should be similarly justified.

85. GROS will carry out a cost-benefit analysis of the 2011 Census in Scotland. Initial analysis indicates that the cost of the Census is dwarfed by possible misallocation of government expenditure e.g. for the Scottish Health Service, we estimate that £284m would have been misallocated if pre-2001 Census, rather than 2001 figures, were used to allocate resources – although misallocated resources are not necessarily wasted.

86. In the design of the 2006 Test, as in other preparations for the 2011 Census, we will pay close attention to value for money, consistent with seeking to achieve high data quality.

2.5.11 Publication timetable

87. The first data from the 2001 Census was available in August 2002. In line with our pre-announced timetable, full results were published in March 2003, some 23 months after Census day. The TSC noted that by the time full 2001 Census results were available, the previous 1991 Census data, upon which local authority spending assessments are based, was 12 years old and that the possibility of a shorter processing time should be considered. Though we would like to publish results more quickly, it is not clear how strong the need is nor, to us, where the balance lies between high quality and quick publication. This is considered in Section 3.

3 Statistical Strategy and Design for the 2011 Census in Scotland

3.1 Introduction

88. This section describes the strategic and statistical framework for the 2011 Census in Scotland, focuses on the key statistical aims and principles and identifies the high priority areas of statistical research and development over the coming years.

89. Our main aims for the 2011 Census are to:

a. Build effective partnerships with other organisations, including local authorities and community representatives, in planning and executing the enumeration;
b. Build public confidence and encourage participation in the Census;
c. Manage and control all aspects of the Census operation effectively;
d. Maximise overall response rates;
e. Minimise differential non-response rates across areas and population sub-groups;
f. Ensure that key minority groups are well defined and enumerated;
g. Protect, and be seen to protect, confidential personal census information;
h. Build user confidence in the final results;
i. Provide value for money; and
j. Provide high quality statistics that meet user needs;

90. We intend to achieve these objectives by the following:

3.1.1 External and research inputs

- Learn lessons from the 2001 Census;
- Make best use of all relevant information known before the Census;
- Take a rigorous and open approach to identifying user requirements through effective consultation processes;

- Research and identify hard-to-count groups/areas and develop procedures to improve their enumeration;
- Develop collection methods to facilitate response from all communities; and
- Be explicit about the trade-offs between cost, quality and time.

3.1.2 Field work

- Provide informative, timely and responsive publicity to actively encourage public participation;
- Minimise the problems associated with recruitment, pay and retention of the field staff;
- Provide rapid, accurate information to support operational decisions by monitoring household forms; and
- Effectively assess live coverage and quality of the collection operation.

3.1.3 Processing

- Make every effort to protect the confidentiality of personal data, through all census processes;
- Define, measure and manage data and process quality; and
- Strive for earlier delivery of quality outputs.

91. Harmonisation with the Census Offices responsible for the Censuses in England and Wales and Northern Ireland (ONS and the NISRA respectively) to achieve consistent and comparable Census outputs across the UK is a guiding principle of the GROS 2011 Census programme.

3.2 Context

92. This section focuses on the key statistical aims and principles and identifies the high priority areas of statistical research and development over the coming years.

93. The 2011 Census context will be different from previous censuses. Technology and society changed markedly

between 1991 and 2001 and we expect the rate of change to increase between 2001 and 2011. The design of the 2011 Census must therefore be flexible so that it (and we) can respond to change more quickly and more radically.

94. The percentage of residents counted in the last 3 Censuses has fallen from 97.3% in 1981 to 96.9% in 1991 to 96.1% in 2001. Differential response – where areas and population subgroups had response rates significantly lower than the national average - increased. These trends are not unique to the Census; they are also evident in surveys and election turnouts.

95. People in Scotland are becoming more mobile and their living patterns more complex. More people have more than one address. This provides problems for both data collection and outputs. We must have a clear definition of where people should be counted by the Census and population definitions need to be easily understood.

96. Increased availability of administrative data and improved record linkage should mean wider and more frequent availability of some census-type data. Annex A is a brief description of current administrative databases which could be used in place of the Census or some of its questions. Such data may also be used to assess person and item non-response and help adjust the Census at an individual level.

97. The development of population and address registers may further enhance our understanding of the population. We must take advantage of them. Also, assuming highly accurate address lists, we could move to a postout method by 2011 and marshal resources into effective follow-up.

98. We do not expect radical changes to the form design, except to make it easier for visually handicapped people to read, although the questions will of course change.

99. At this stage, we have no final operational design for the 2011 Census. But, by strengthening field methods to improve quality, we hope to speed up processing,

particularly editing and data quality checks and so improve accuracy of output.

3.3 Maximising Response

100. Over the next few years, we must understand barriers to response and address them through enumeration strategies, definitions, questionnaire design, publicity and follow-up.

101. The explanation of who should be included on a questionnaire must be clear and easily applied by all. The questionnaire design, content and respondent burden must be acceptable to the public.

102. We will also take pains to ensure every question on the form is answered. In the 2006 Test we will be assessing the cost of ensuring that forms are complete for all questions by making it clear to householders that enumerators will return if forms are not fully completed.

3.4 Address Registers

103. The development of a good national address register is critical to potential improvements to the 2011 Census. We are keeping in touch with implementation, by the Scottish Executive and local councils, of the 'Definitive National Address' (DNA) list. It is too early to say exactly which address registers will be available in 2011. The hope is that there will be a single agreed address register as in Northern Ireland – agreed by the local councils, Royal Mail and Ordnance Survey. Currently, we are comparing two external address registers, Royal Mail's Postcode Address File (PAF) and the Scottish Assessors Association (SAA) Portal[15].

104. However, we cannot assume that any national address register will have 100% coverage. Therefore, in the 2006 Test, we will introduce a new role of 'Address Checker'. They will check:

[15] www.saa.gov.uk

- Address differences between PAF and SAA lists before the Census Test enumerator work starts;
- Addresses which the enumerator deems to be unoccupied; and during enumeration
- Confirm the enumerators' non-occupancy status.

105. We can use the information to agree with local authorities a list of properties before Census Test day. The Test will take place in Glasgow, West Dunbartonshire, Perth and Kinross, Stirling, Highland and Argyll and Bute. From Census returns, would then flow a total number of households from whom we would expect to receive a form with the number and type of vacant properties. The resulting information could also be used to improve the address register.

106. This may become simpler if confidence in the address register increases. An improved address register will also facilitate the imputation of wholly missed people. This strategy follows the Statistics Commission's recommendation to use locally-based administrative sources to assist in the conduct of the Census. We will also take the Commission's advice that any quality assurance of local authority address lists and household population must be independent of the local authority since there is a potential conflict of interest.

107. However, even with an up-to-date accurate address register, enumerators may be needed to carry out Census form quality checks and address classification. They could also deliver forms if there were problems with the postal service, and the enumerator delivery round is the activity which gives 'ownership' and knowledge about the area and so helps improve data quality.

108. While a national address register may be an accurate list of dwellings, we may also need data about the number of households at each address. Additionally, there is a requirement to ensure that the appropriate questionnaires and accompanying information are sent to each address. Ideally, the address register would distinguish between

communal establishments and private households. Some work has now started on getting better information pre-Census on communal establishments.

3.5 Targeted Enumeration

109. Even if we take great care to get questionnaires to and from households, some people will not respond. In order to tackle this element of non-response, GROS have supported research to identify hard-to-count groups and areas.

110. Practical reasons for non-response are disability, language or literacy difficulties (particularly for those whose first language is not English). We will, as Section 2.5.2 envisages, thoroughly review ways to allow such people to participate fully in the Census.

111. Many people are worried about supplying the detailed personal information which the Census requires. Others do not see the relevance of the Census to their everyday lives. It is therefore vital that we communicate the importance of completing a Census form, that the information provided is used to improve local and national services, that there are many safeguards restricting what Census information can be used for and who can access it and that the penalties for misuse or inappropriate disclosure of data are severe.

112. Research in this area involves interacting with people who do not respond – in order to develop enumeration strategies, select the population definition for enumeration and improve questionnaire design. This research is similar to that being conducted internationally to tackle non-response to social surveys and censuses. We will take full account of this wider work.

3.6 Quality and Outputs

113. We consider these aspects to be closely intertwined. The response rate affects data quality, not only because it is below 100%, but also because is it spread unevenly across demographic groups and geographic areas. Attempts to

compensate for non-response will, in general, delay the final publication of output data and will normally require a second data source.

3.6.1 Coverage Assessment

114. We must have a means of assessing coverage which is sophisticated enough to provide an estimate of differential response down to small areas and population sub-groups. We will review the 2001 strategy of using the household classification given by enumerators and the CCS held shortly after the Census itself (see Section 2.5.7).

115. To adjust for under-enumeration, an estimate of the number of missing people must be made. This task is made more difficult if there is dependency between the Census and any follow-up survey. There is some evidence for such dependency in the 2001 Census; people missed by both the Census and the CCS had different overall characteristics from those included in at least one of these exercises.

116. We believe that the use of quality assured address lists will reduce the number of wholly missed households. Administrative data about individuals may, by 2011, allow us to measure and characterise to some extent within-household missing people. We will seek to use and improve such data for coverage assessment.

117. We will collect evidence of over-enumeration in 2001 and report on it to Census users in 2007 so that we can decide whether we need an over-enumeration strategy in 2011. For example, there is anecdotal evidence that children of divorced parents may sometimes be included at both parents' addresses.

3.6.2 Output

118. We need to decide whether to publish the 2011 Census results more quickly than in 2001, as recommended by TSC. But we also need to carefully adjust the Census results for under- or over-enumeration. In 2001, too little time was

allowed for rigorous quality assurance and the time was squeezed by the slow return of data.

119. We are pursuing a number of ways of resolving the speed/quality dilemma. A more stringent field quality target could potentially speed up processing and reduce processing costs but may be more expensive. By strengthening enumeration and form completeness, we may reduce the time spent fixing errors. With a better address list and pre-addressed forms linked to a forms control database, imputation can be reduced and carried out more quickly.

120. A cost-benefit analysis focussing on the processing time may help to assess both the demand/need for quicker dissemination and the willingness of Census users to tolerate the consequent increased costs.

Could users indicate how they would benefit from earlier publication of full results (say a year after Census day instead of 18 months) in a way which allows us to assess how much extra it would be worth spending to meet that timetable? (CP3)

121. Another way to help resolve the quality/speed dilemma might be to release unadjusted Census results quickly – within a year of Census day – followed by adjusted results once we were completely confident of their accuracy – perhaps 18 months after Census day. Both sets of results would give small area data adding to a Scottish total. This is a solution which we would hopefully explore if users found it helpful – though we would have to be sure that neither set of results (nor comparison between them) would breach confidentiality of information about individuals.

Would users welcome a two stage publication of results (provisional followed by fully-adjusted)? (CP4)

122. Our aim is to provide flexible outputs that are freely available over the internet. However, we will follow the over-riding rule that the confidentiality of personal data

must be protected throughout all Census processes, from form collection to production of outputs.

123. We are researching, with the other UK Census Offices, ways of controlling disclosure while maximising the utility of the information produced.

3.6.3 Quality

124. We will produce a comprehensive quality strategy, which will define the quality standards for each stage of the Census operation, from data collection, through processing to the final outputs. Here we will outline how we will measure quality and ensure quality standards are met and recommend how data quality is reported to users.

125. Part of the quality work for 2011 will focus on meaningful measures of data quality to accompany outputs. Local authority area level confidence intervals are not helpful to users who wish to compare small areas within a local authority. Levels of imputation are misleading as a measure of response quality if there is evidence that a question was misinterpreted by some types of respondents.

126. We will examine other approaches to the problem of quality assurance, with a view to their applicability to Scotland. For example, the New Zealand Census uses field information to produce early provisional estimates that are compared with previous population estimates. Where they find large differences, Census staff work in partnership with local experts to attempt to explain those differences prior to publication of the Census results.

127. This approach would allow any necessary adjustments to be made to the Census results prior to publication should problems with the Census operation be identified.

128. Alternatively, if no adjustments are necessary, there is increased time to understand and document differences prior to publication of results. This increases public

confidence in results that may at first glance appear unusual.

3.7 Population Bases and Definitions

129. We must decide whether outputs on two population bases or combinations of these are required or sufficient for user needs.

130. One of the most fundamental aspects of the Census is the population base which is to be enumerated. The changing nature of the population over time means that the population bases for enumeration and output must be reviewed: we need to know the limitations on population bases that the Census can successfully enumerate.

131. In addition to the enumeration base, on which the Census is carried out, other bases for output purposes may be generated from appropriate question(s) in the Census form. The final decision of which base to use for enumeration and which additional output bases to allow for rests on several competing factors:

- User requirements;
- Easiness to understand the questions and their purpose;
- Effect on accuracy and response rate;
- Data quality; and
- Technological and logistical limitations.

3.7.1 Possible population bases

132. The UK Census Offices are considering four main options for the enumeration base[16]:

1. Population Present

133. 'Population present' counts everyone at the address they are at on Census night. This has the advantage of being easily understood. However, in order to generate the

[16] Consultation with Users on Population Base for the 2011 Census (ONS June 2004)

‘usually resident’ population, a question asking for usual address would have to be included, so an individual could be ’transferred back’ to their usual address. We could not therefore guarantee the minimum geographical level at which this population base could be produced, since this would be dependent on the quality of the data generated by the question on usual address.

134. People who are temporarily away from Scotland would not be included at all in a ‘population present’ base enumeration. Conversely, people who are temporarily present in Scotland would be included. Clearly, neither of these circumstances is desirable. Some household relationship information is also lost.

2. Population usually resident

135. Enumerating on a ‘usually resident’ population base (the approach taken by all recent UK Censuses) removes the difficulties encountered with ‘population present’. No information is collected on any visitors and those who have no usual address are fully enumerated at the address they are at on Census night. However, a clear definition of what is meant by ’usually resident’ is needed.

3. Population present plus usual residents temporarily absent

136. A modification to the ‘population present’ base collects some information on those usual residents who are temporarily absent. The amount of information collected, including full enumeration, would depend on user requirements. There is also an issue for those who are temporarily absent from their usual address but are still within the country. A decision must be taken as to which address should be used for their enumeration, ranging from full enumeration at one address only, full enumeration at one and partial enumeration at the other or full enumeration at both.

137. Whilst collecting at least some information at both addresses could improve coverage and data quality, this places an increased burden on temporarily absent usual residents. We must consider what effect this increased burden might have on response rate. A lower response rate from temporarily absent usual residents would introduce bias into the Census and could nullify the benefits of double enumeration. To avoid double counting we must also consider how to identify temporarily absent residents if they are at another address in the country on Census night.

4. Population usually resident plus visitors

138. For this population base, usual residents would be fully enumerated and any visitors on Census night partially enumerated. As for the previous population base, this could be extended to fully enumerate everyone including visitors and we need to consider, as before, whether visitors and temporarily absent usual residents should be required to fill in two Census forms and decide how to identify these people to avoid double counting.

139. In order to yield useful and consistent data from the population bases discussed here, in addition to 'usually resident', we need to decide on a definition of the terms 'temporarily absent' and 'visitors'.

3.7.2 Current Population Bases

140. In recent Censuses, users have preferred 'usual resident' population as the enumeration base. In 2001 the workplace population, the school population, and the daytime population were all derived from information on the Census form. We expect to continue to have these output bases for 2011. The Statistics Commission recommended that alternatives to 'usual residence' measures of population base should be developed.

141. To allow cross-UK comparisons, we intend to agree the main population base with the other UK Census Offices. At present the favoured base is 'usually resident' with some,

as yet undefined, information about visitors – perhaps enough to provide a usual address and matching data.

142. Additional base(s) would require a significant demand from users. Through consultation, ONS has found that users would also like information on 'population present', to take account of those who have no usual address.

Could users give their opinions of the four population enumeration bases (paragraphs 133-139) which we are considering? If changes from the 2001 Census are needed, please explain why. (CP5)

3.8 Other Definitions

143. We will also consider requirements for outputs on households, families, household spaces, the household reference person, dwellings and communal establishments, workplaces and educational establishments.

Could users give their opinions on other aspects of the Census which they feel we may not have fully considered? (CP6)

4 The 2006 Census Test Design

4.1 Background

144. This section describes the proposed 2006 Census Test which is an important stepping stone to 2011, and on which we will need to take decisions soon.

145. The Census Test is traditionally a test of new features, as distinct from a rehearsal of the main methodological content of the main Census. The Test seeks to generate enough data to allow us to develop and evaluate processing options. It also allows us to:

- Train staff for the Census proper;
- Trial new technologies which impinge on field methods;
- Trial new questions; and
- Evaluate quality of results against cost.

4.2 Test questions

146. In September 2004, we launched an internet consultation on questions to be included in the Test[17]. We will publish summarised responses at 3 monthly intervals on the website. The first summary of responses is now lodged there. The next summary will be lodged by February 2005. That will include a digest of responses so far and we will use that as a basis for a question consultation document.

147. We have not as yet drawn any conclusions about questions to be included in the 2006 Census Test (except income and ethnicity – see below). However the Test form will probably reflect the current question set. Table 7 lists the 2001 Census question topics and compares results in Scotland and England and Wales. The Table also contains:

[17] http://www.gro-scotland.gov.uk/grosweb/grosweb.nsf/pages/cenconsult

- The non-response rate, where appropriate, for GROS and ONS;
- The difference between response rates; this may be a guide to the reliability of the data collected by the questions in the 2001 Census;
- Possible alternative sources of the data which was or would be collected by the question;
- A rank indicating the relative importance of a question; and
- Justification of, or comments on, a question's importance and proposed inclusion, removal or modification.

148. Figure 1 (towards the end of the document) shows the difference between question non-response rates for GROS and ONS. Non-response to a question is the result of one of 3 possible actions by the form-filler – where the form filler:

- Believes that the question does not apply to them;
- Understands the questions but for another reason makes a conscious decision not to answer the question, e.g. because of confidentiality concerns; and/or
- Does not understand the question and decides not to answer or answers incorrectly.

149. Although the GROS, ONS and NISRA Census forms and questions were similar, differences in the response rates may show strengths or weaknesses in the different presentation of a question. Generally non-response rates between ONS and GROS for significant shared questions were similar (most differences less than 1%). GROS had higher non-response rates for address/ postcode of workplace, travel to work and study, landlord and marital status. ONS had a higher non-response rate for size of workplace.

150. Two questions, on ethnicity (see paragraph 36 above) and income, have already been highlighted for inclusion and assessment in the 2006 Census Test. Consultation before the 2001 Census showed that users would find the Census more valuable if it included an income question. This view is

shared by the TSC[18]. They pointed out that most people answered the religion question even though it was voluntary and suggested that compulsion is not necessarily an essential element in securing a high level of responses to specific questions. They recommended that further consultation should be carried out on whether an income question should be included in any future Census, and whether it should be voluntary or compulsory.

151. If a question on income is to be included we must work to determine:

- Wording of the question;
- Explanatory text to encourage completion;
- Income banding;
- Whether achievable accuracy is acceptable;
- Whether a non-compulsory question would work;
- Whether it should be an individual or household question;
- Whether income source is also required or is enough; and
- Whether there is proxy information (e.g. dwelling Council Tax banding) or alternative sources which are acceptable.

We would welcome input on all aspects of an income question for the 2006 Census Test. (CP7)

4.3 Test Design

152. The Test is designed to evaluate a number of features (Table 4). We will compare the application of 2 methods of enumeration in 3 types of areas, with and without an income question.

153. In previous Census Test designs, an enumerator has only had to trial one enumeration methodology. The Census Office in Ireland had mixed methods of enumeration by a single enumerator in their recent Census pilot. We propose to do the same, allowing us to compare response rates of closely similar areas with the same enumerator.

[18] http://www.publications.parament.uk/pa/cm200102/cmselect/cmtreasy/310/31002.htm

Table 4: Proposed 2006 Test Design

Requirement	Comment	
Income question	On 50% of forms	
Ethnicity question	Ethnically rich and diverse area - no split design required but may be many question variants for follow up purposes.	
Cost	Postout/Postback-Collection Single round delivery/ postback collection	Rural Suburban Urban
Data Quality	Enumerator check back. Cost estimate	
Field staff pay	Cost estimate of in-house field staff pay system	
'Address Checkers'	Check address lists before Census - Compare PAF & DNA.	Rural Suburban Urban
	Confirm occupancy status of empty dwellings in Census	
Processing	Allow prototype processing; Allow data quality to be assessed;	

154. To meet the aims of the Test design, we plan to hold the Test in (see map provided separately):
- An ethnically diverse area of Glasgow;
- public sector estates with high unemployment in West Dunbartonshire; and
- a rural area with holiday homes between Killin and Ballachulish, taking in parts of Highland, Argyll & Bute, Stirling and Perth & Kinross Council areas.

155. The areas chosen in Glasgow are non-contiguous. The other two Test areas are formed from contiguous areas.

156. We will trial an address check procedure. Address Checkers will check about 4,500 addresses on the ground. We hope to liaise with the local authorities in the Test areas on our findings from the Autumn 2005 Address Check (see Section 3.4).

157. The check is also in preparation for pre-printing form and envelope addresses for postout and for enumerator delivery. Our aim is to make delivery as accurate as possible by correcting address anomalies before enumerators start work, and then concentrating our effort on collection.

158. The enumerator is there to help households complete the form if needed. The householders will also have the choice to post or get the form picked up. This has 5 potential advantages:

- We retain the delivery/collect option if the postal service has problems;
- We retain the collect option for all households with difficulty completing the form or posting it;
- It gives an enumerator a sense of ownership of an area;
- We can divert enumerator resources to checking form quality; and
- We may achieve savings over traditional enumeration.

159. We plan that an enumerator would cover about 500 households in urban and 400 households in rural areas. With 120 enumerators, we would cover about 57,000 households. Field staff by area is shown in Table 5.

Table 5: 2006 Test Field Staff and Form Count by Area.

Staff Grade	City centre	Peripheral	Holiday area	Totals
Regional Census Managers	1	1	1	3
Local Census Managers	2	1	1	4
Local Census Manager Assistants	6	3	3	12
Enumerators	60	30	30	120
Total no of households	30,000	15,000	12,000	57,000
Est no of forms returned (60% response rate)	18,000	9,000	7,200	34,200

160. With this size of sample, assuming a 60% rate of return, each treatment block (income by enumeration method) would have at least 8,000 - 9,000 returned forms (Table 6) - and 2% differences would be statistically significant.

Table 6: Planned number of forms issued and estimated number of forms returned

Variant	Forms issued	Forms returned
Post out/postback	33,000	19,800
Delivery/Postback	24,000	14,400
Urban	30,000	18,000
Suburban	15,000	9,000
Rural	12,000	7,200
Income question	28,500	21,735
No income question	28,500	21,735

4.5 New Technology and Data Collection

161. We considered whether to test electronic form completion by households – by internet, telephone or interactive TV. We are concerned at the relatively high cost (especially to ensure security, while linking the form to the household) and the scope for cost escalation during the project and technology may change significantly by 2011. Nor do we have evidence that it would guarantee enhanced response rates from historically under-enumerated sectors of the population. We do not therefore intend to include a full electronic option in the Test. We plan however to include a test question about whether people would prefer to complete the form on-line. We might also adapt the method we use for the web-based question consultation to allow on-line form completion in test areas, making clear however that we do not guarantee confidentiality.

We would welcome your opinions of how useful/relevant alternative methods of Census data collection would be. (CP8)

5 Conclusions/Summary

162. To achieve our aim of providing high quality population statistics as required by key users on a consistent and comparable basis for small areas and small population groups, there are 5 main steps we need to take for the 2011 Census.

163. First, we must ensure we ask the right questions. We will take maximum advantage of information from other surveys and administrative sources to discard some of the traditional Census questions. But users will want to add new questions and we will consult them carefully and check the practicality and costs and benefits of their suggestions.

164. Second, we must ensure that everyone gets a Census form which they can fill in. Better address lists will be a great help. But we need to pay careful attention to form design and the provision of special help for people with disabilities and language difficulties.

165. Third, we must strive to achieve a very high response rate. The role of the enumerator, careful tracking of forms and sensitive enforcement of the obligation to complete the Census form are the key elements, but the design of these elements needs to be based on a better understanding of why people do not return their forms.

166. Fourth, the processing of forms must be rapid and accurate. We will do more work on the right balance between cost and speed. We must be able to identify and compensate for under- and over-enumeration.

167. Finally, we must publish results in an accessible form which preserves confidentiality of individuals' information – at a speed which accords with user demand yet does not sacrifice accuracy or cost-effectiveness.

168. That is the challenge of the 2011 Census (and of the planned 2006 Test and 2008 Rehearsal). We would welcome your help in meeting that challenge.

Table 7: Topics and questions for potential inclusion in the 2011 Census

Question	In 2001 Census?	2001 Non-Response %			Alternative Sources (See Annex A)	Question Priority*	Justification / Comments
		(GROS)	(ONS)	Diff (ONS - GROS)			
Household Accommodation							
Household tenure	Yes	3.7	3.4	0.3	None	1	Strong user need
Accommodation type	Yes	3.2	3.0	0.2	Scottish Assessors	1	Strong user need
Number of rooms	Yes	4.7	5.4	-0.7	Scottish Assessors	1	Strong user need
Number of vehicles	Yes	3.2	2.7	0.5	DVLC.	1	Strong user need. Question may be removed for 2011. DVLC can give car owners by postcode and gender.
Self-contained accommodation	Yes	3.0	3.9	-0.9	None	3	Low no. of “Yes” responses
Bath/shower and toilet access	Yes	1.9	2.5	-0.6	None	3	Low no. of 'No' responses
Central heating	Yes	3.4	2.2	2.2	None	0	Low no. of 'No' responses. Remove?
Lowest floor level	Yes	3.3	4.0	-0.7	None	1	ONS consider less useful than in the past
Landlord	Yes	7.6	2.9	4.7	CPS?	1	Even if not in the CPS, research suggests this info could come from other surveys

Question	In 2001 Census?	2001 Non-Response %			Alternative Sources (See Annex A)	Question Priority*	Justification / Comments
		(GROS)	(ONS)	Diff (ONS - GROS)			
Furnished or unfurnished Accommodation	Yes	10.1	N/A	N/A		1	May be poorly understood by people who rent.
Demography							
Household relationships	Yes	3.6	3.5	0.1	None	1	Strong user need
Sex	Yes	0.3	0.4	-0.1	None	1	Strong user need
Name	Yes	N/A	N/A	N/A	None	1	Strong user need
Date of birth	Yes	0.6	0.5	0.1	None	1	Strong user need
Marital status	Yes	3.2	0.8	2.4	None	2	Strong user need, possible change with civil partnerships
Culture							
Ethnicity	Yes	4.5	2.9	1.6	None	0	Strong user need
Country of birth	Yes	2.6	2.5	0.1	None	1	Strong user need
National identity	No	N/A	N/A	N/A	None	3	Strong user need
Religion (raised)	Yes	8.0	N/A	N/A			Need to establish use and user need.
Religion (practising)	Yes	5.4	7.6	-2.2	None	0	
Gaelic language	Yes	4.8	5.5	-0.7	None	1	Strong user need
Main language	No	N/A	N/A	N/A	CPS?	3	Some user need, but not asked before. Problematic to include.
Qualifications							
Academic qualifications	Yes	6.4	6.2	0.2	ScotXed	0	Alternative sources may be suitable
Professional	Yes (ONS)	N/A	17.2	N/A	DfES records?	0	Didn't work well in 2011 in England and Wales. Wasn’t used

Question	In 2001 Census?	2001 Non-Response %			Alternative Sources (See Annex A)	Question Priority*	Justification / Comments
		(GROS)	(ONS)	Diff (ONS - GROS)			
qualifications							
Health							
Attitude to own health	Yes	3.7	3.1	0.6		0	Strong user need
Carer	Yes	6.5	6.1	0.4	Admin data?	0	Strong user need
Disability/Long-term illness	Yes	4.7	3.9	0.8	Admin data?	1	Strong user need
Labour Market/Employment							
Occupation	Yes	2.7	3.2	-0.5		0	Response rate relates only to people then working
Industry	Yes	8.3	7.8			0	Response rate relates only to people then working
Work last week?	Yes	0.5	2.1	-1.6	DWP	0	
Name of employer	Yes	N/A	N/A	N/A	DWP	0	
Work activity	Yes	N/A	N/A	N/A		0	
Hours worked last week	Yes	7.1	8.0	-0.9		0	Didn't work well
Looking for work	Yes	N/A	N/A	N/A	DWP	0	
Could have taken job	Yes	N/A	N/A	N/A		0	
Non-working status	Yes	N/A	N/A	N/A	DWP	0	
Waiting to start job	Yes	N/A	N/A	N/A		0	
Employment	Yes	6.5	6.5		DWP	0	

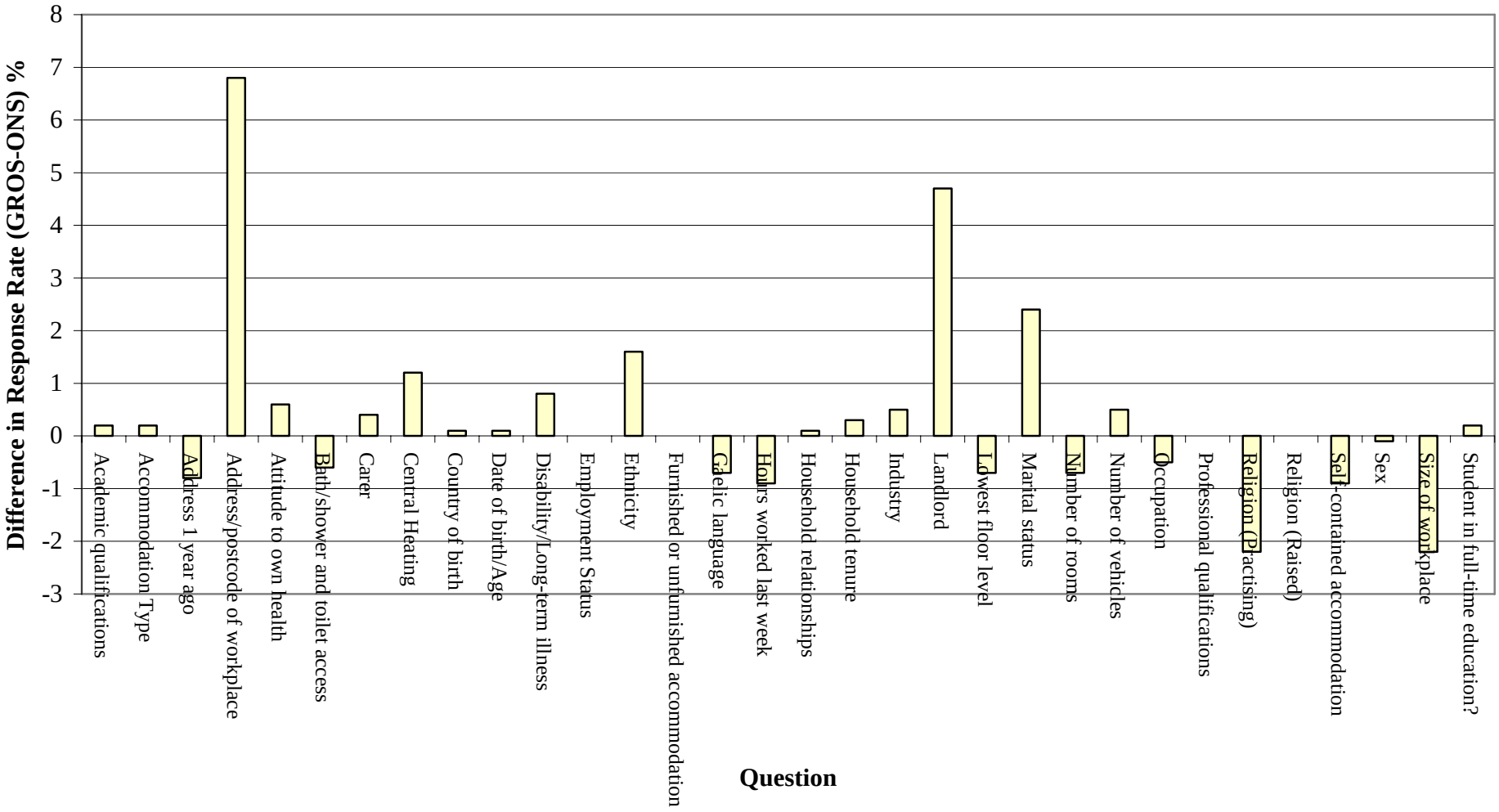
Question	In 2001 Census?	2001 Non-Response %			Alternative Sources (See Annex A)	Question Priority*	Justification / Comments
		(GROS)	(ONS)	Diff (ONS - GROS)			
Status							
Ever worked	Yes	N/A	N/A	N/A	DWP	0	
Size of workplace	Yes	11.7	13.9	-2.2		0	
Supervisor?	Yes	6.9	6.8	0.1		0	
Transport							
Travel to work	Yes	12.6	6.3	6.3	CPS?	1	Strong user need
Travel to study	Yes	12.6	N/A	N/A	CPS?	2	Potential user need
Students							
Student in full-time education?	Yes	1.5	1.3	0.2	None	1	Strong user need
Migration							
Address 1 year ago	Yes	4.6	5.4	-0.8	None	1	Strong user need
Year of entry into UK	No	N/A	N/A	N/A	None	2	Strong user need but not included before
Income and Wealth							
Income	No	N/A	N/A	N/A	IR data?	2	Strong user need, but data quality issues and not included before
Sources of income	No	N/A	N/A	N/A	CPS?	2	Strong user need, but data quality issues and not included before

Question	In 2001 Census?	2001 Non-Response %			Alternative Sources	Question Priority*	Justification / Comments
Population Bases							
Workplace address or postcode	Yes	14.6	7.8	6.8	None	1	Strong user need
Term-time address	Yes	N/A	N/A	N/A	None	1	Strong user need
Usual residence	Yes	N/A	N/A	N/A	None	1	Strong user need
Address of study place	Yes	N/A	N/A	N/A	DfES records?	2	A significant user need may emerge
Second residence	No	N/A	N/A	N/A	None	2	Strong user need but never been asked before

*** Priority Categories**

0 No view as yet;
1 Probably included in the 2011 Census;
2 Maybe included in the 2011 Census but more research required;
3 Probably excluded from the 2011 Census.

Figure 1: % Difference between response GROS and ONS Question Non-Response Rates



Annex A: Alternative Data Sources

169. Alternative data sources may ultimately obviate the need for some questions in the Census. In this section we describe some possible alternative data sources (mentioned in Table 7) and their potential effect on question content.

5.2.1 Scottish Assessors Association Portal

170. The Assessors Portal[19] is the first phase of the Definitive National Addressing initiative for Scotland. It contains the combined property addresses held by the 14 Scottish regional assessors. The Portal is a potential source to us and to Census users of address lists, accommodation data (accommodation type, number of rooms (not standardised yet) and presence of a garage) and council tax banding.

5.2.2 The Driver and Vehicle Licensing Agency (DVLA)

171. The DVLA has a UK-wide database which lists every vehicle licensed for use on public roads. The sex and postcode of the registered keeper for each vehicle is also stored. It might therefore be possible to use the DVLA data, which is updated on a monthly basis, to usurp the Census question on cars or vans available for use for each household. We propose to compare the DVLA database for 2001 to the Census results to assess its quality and usefulness.

5.2.3 ScotXed

172. ScotXed (Scottish eXchange of Education Data) is an initiative for the collection of individual data on primary and secondary school pupils and is therefore a possible source of demographic data for school pupils.

173. It is also a potential data source for the Scottish school level qualifications held by those born from about 1978 onwards. The qualifications data to which ScotXed is linked

[19] http://www.saa.gov.uk./

dates from 1994 and uses a unique Scottish Candidate Number assigned to each candidate for SQA examinations.

174. Data is collected annually from state and, to some extent, private schools. ScotXed data would not entirely remove the need for a Census question on qualifications – it does not cover qualifications above school level – but it may nonetheless help to make such a question less cumbersome than was the case in 2001. Ultimately, it may simply provide an administrative source for quality assuring the Census data.

5.2.4 The Department for Work and Pensions (DWP)

175. The DWP is a potential source of data on current and past employment status with monthly updates. Combining its data on state benefit claimants with income tax data held by the Inland Revenue would also allow data on income to be collected for the vast majority of the resident population along with more detail on employment type. Indeed, a Lifetime Database initiative is currently being progressed up with a view to marrying IR and DWP databases together.

176. Potentially the database could remove the need for at least some of the employment questions.

5.2.5 Other Sources

177. There are other administrative sources which could be used to supply Census-type data. For example, data used to administer disability and carer benefits could be used to give data on the extent of disabilities and people with caring responsibilities. However, as with the DWP's Lifetime Database, there are confidentiality issues.

Annex B: Consultation Points

Para 6: We invite users' views on the outline consultation timetable and methods (CP1).

Para 8: We consider that the key objectives of the 2001 Census were met. Do you agree? If not, which of these objectives do you believe were not met and what lessons can be learned? (CP2)

Para 120: Could users indicate how they would benefit from earlier publication of full results (say a year after Census day instead of 18 months) in a way which allows us to assess how much extra it would be worth spending to meet that timetable? (CP3)

Para 121: Would users welcome a two stage publication of results (provisional followed by fully-adjusted)? (CP4)

Para 142: Could users give their opinions of the four population enumeration bases which we are considering? If changes from the 2001 Census are needed, please explain why. (CP5)

Para 143: Could users give their opinions on other aspects of the Census which they feel we may not have fully considered? (CP6)

Para 152: We would welcome input on the design aspects of an income question for the 2006 Census Test. (CP7)

Para 161: We would welcome your opinions of how useful/relevant alternative methods of Census data collection would be. (CP8)

Consultation Events:

17th November 2004 – Meridian Court, 5 Cadogan Street, Glasgow G2 6AT
30th November 2004 – COSLA Conference Centre, Rosebery House, 9 Haymarket Terrace, Edinburgh EH12 5XZ
3rd December 2004 – Edinburgh University, Main Library, George Square, Edinburgh EH8 9LJ

Contact Details:

Statistics Division, General Register Office for Scotland, Ladywell House, Ladywell Road, Edinburgh EH12 7TF
Ian.Mate@gro-scotland.gov.uk Tel: 0131 214 4295
Brian.Paget@gro-scotland.gov.uk Tel: 0131 314 4319