

POPULATION AND MIGRATION STATISTICS COMMITTEE (SCOTLAND)

FUTURE CENSUS OPTIONS - 2011 PROGRESS AND PLANS

PAMS (03) 04 provided a strategic review of options for producing Census-type statistics in Scotland to 2011. This paper covers the detail of work toward a conventional Census in 2011. I seek members' views on the proposed consultation process and the 2006 census Test.

General

1. Statistics Division has been reorganised into the following branches
 - 1 2011 Census and Finance - Ian Máté: Traditional census data collection, consultation and central divisional support function.
 - 2 Information Systems and Geography - Norman Gillies: IT support for Statistics Division, 2011 processing, IT security and Geographic systems, particularly the modified PAF, postcode indexes and postcode boundaries, and SCROL
 - 3 Dissemination, 2001 Census Analysis - Frank Thomas: This includes future website dissemination, 2001 evaluation and customer services.
 - 4 Demography - Cecilia Macintyre - Population Projections and Estimates and Migration
 - 5 Vital Events and NHSCR - Graham Jackson (unchanged)
 - 6 Census Alternatives - Vacant - Including DPA, SLS and CHD matching projects and interfaces with administrative sources.
 - 7 A proposed small area housing estimates branch which would arrange assembly of occupied household data from council tax billing records

2. This structure should better able to meet the dual approach GROS has toward the 2011 Census - promoting alternative sources by seeking to replace Census type data from administrative and survey sources and planning for a traditional census.

Traditional Census

3. From April 1st Branch 1 expanded to cover 2011 consultation and a 2006 Census Test. International evaluations¹ and UK Census Offices experience was that 1999 rehearsal processing and pay testing either concatenated with, or were not completed before, the 2001 Census. Therefore, and with the dual challenge of devolution and rebuilding staff skills, planned testing in Scotland has been brought forward by a year in the census cycle to 2006 and a rehearsal in 2008. The consultation is also being started a year earlier.

¹ <http://www.unecce.org/stats/documents/ces/2003/28.e.pdf>

Consultation strategy

4. The consultation process will be 'continuous', that is we will endeavour to up date web statements of where we are at regular intervals (3 monthly?) and notify census users and interest groups of changes to our current thinking. This will be possible using the GROS website.
5. The consultation process will attempt to operate at a number of levels:
 - With local authorities as partners in developing methodology and agreeing address lists;
 - With all census customers as users ;
 - With under enumerated groups to promote an inclusive census, e.g. single parents and school leavers;
 - With groups of people about whom the census collects data - e.g. carers and minority ethnic groups;
6. At the same time it will attempt to draw other official bodies into the Census process, such as Parliamentarians and SE statistical bodies such as RESIG, Scotstat and PAMS.
7. The Census will try to integrate new data sources to allow a reduction in census questions and ease the burden on the public. Therefore the consultation process will influence census data users to appreciate operational costs and difficulties and to consider alternative options fully.
8. In all of this we will record our consultation process fully and update the lists of people consulted, the results of that consultation and, importantly, who any views expressed represent.
9. Small scale testing and Census Test generally will seek partners from our users and from representatives of the various interested communities.
10. This process will end, as far as possible, with parliamentary legislation. We are currently initiating a project to consider devolved legislation requirements.

Consultation practice

11. The RG will meet our Minister soon to discuss the future provision of statistics including initiatives elsewhere which will impact on census taking and social statistics and alert ministers to the consultation process, legislation and input from the Scottish Parliament.
12. We will build a question content consultation web presence. This will develop to include:
 - The 2001 question content, univariate results initially and later error levels;
 - The questions which were strong contenders in 2001 but not included;
 - A paragraph on the current status of each question for 2011;
 - An e-mail response mechanism;
 - Submissions received about each question.
13. We acknowledge the strength of the business case for an income question while noting the presence of DWP data on income. We will seek views on the design of an income question. It is currently proposed that a form of the income question will be tested in 2006.

14. We will work with Racial Equality Strategy Implementation Group to design and test an ethnic question and create an inclusive process with members of minority ethnic communities.
15. Rather than rely on browsing, the consultation team will build up a comprehensive list of contacts, initially from people and organisations involved in the 2011 round of consultation and Customer Services census up-date lists. At the same time we will seek to extend the list by adding new contacts. We will be constructing a process to quantify the extent and depth of the consultation.
16. There will be a separate initiative to work with the 32 councils to create an agreed address list. This means looking carefully at the Census timetable to update and agree address lists prior to a census field operation.
17. We will seek to involve disability groups to inform our enumeration, questionnaire design team and publicity.
18. We have initiated contacts with Schools Advisors and some headmasters to promote the use of the Census in School Projects. It is very early yet we intend to focus on schools in the more deprived areas of Scotland to try to reach people in traditional under enumeration groups before they age into them. This may require curriculum and school support.
19. We are currently proposing to hold a forum in the autumn. We would seek views on what such a public meeting should cover however provisionally I intend it to cover:
 - The consultation programme;
 - A consideration of 2011 Question Content with:
 - Results from the 2001 Census;
 - 2001 Evaluation;
 - Data Quality;
 - Harmonisation across the UK Census Offices.
 - Population bases
 - Under enumeration
 - The Alternative Census programme with:
 - ONS Initiatives on the Continuous Population Survey and Integrated Statistics System;
 - Administrative data sources;
 - Comprehensive and appropriately specified Address list initiatives;
 - Small Area Population Estimates improvements;
 - Neighbourhood Statistics.
 - The 2011 Traditional Census Development Programme.
20. Having completed a round of Consultation by March 2005, we would seek to inform Ministers and Parliamentarians of the outcome and our draft proposals for question content in the 2006 Census test in March 2005. We would have until September 2005 to conclude this stage of the work. At that point the questionnaire must go to the printer for the 2006 Test.
21. For the 2001 Census a business case was submitted for each question. This was a very onerous task for our users. Therefore we would like to use the 2001 business cases as a basis for 2011 as

much as possible. The method of assessing business cases of new questions (and 2001 questions which had no business case) against 2001 questions with business cases has not been decided, but proponents of any new question would have to provide a case.

22. Further developments in the availability of administrative data may undermine the case for 2001 questions, but considerations of burden on the public, form size and cost will necessarily constrain the question content.

I would like PAMS members views on:

1. **The content of the proposed census workshops**
2. **The web-based question content methodology**
3. **Methods for assessing business cases for census question content**
4. **Topics for inclusion in the 2006 Census Test (see below for design)**
5. **Ways of taking forward the Alternative agenda to reduce the need for a census in the future.**

The Census Test

23. The Census Test is traditionally a test of new options compared to previous methods and will be carried out in the light of the 2001 evaluation. Therefore the test will cover a wide range of methods and seek to generate enough data to evaluate field methods and some questions and allow processing options to be developed. We will:
- Trial new technologies which impinge on field methods;
 - Acts as a method for training staff for the Census proper;
 - Trial new questions; and
 - Evaluate results against cost so that choice of method is determined on quality and cost.
24. The 2001 Census has its critics. There was close scrutiny of the Census in England & Wales by the Treasury Select Committee (TSC)². I have selected some TSC recommendations and criticisms which highlight the perceived problems which, though aimed at ONS, are also relevant to GROS.

Cost

25. *'ONS were unable to supply us with robust evidence to justify expenditure of over £250 million on the Census. We recommend that a rigorous cost-benefit analysis should be carried out of the 2001 Census and published in time to inform the need for, and timing of, any future Census. We consider that any future Census should also be justified in cost-benefit terms.'*
24. This paragraph is taken as a reflection of concern over costs but at a strategic level. I understand that the cost benefit analysis is unlikely to be affected by marginal changes in costs caused by choice of methodology. This implies that we are either free to choose a method which best balances quality against cost against publication timetable or we should not do a Census at all. ONS have produced a draft analysis document which we are currently considering.

Quality

25. *'The Market Research Society noted from preliminary information published on the ONS website that the new data collection procedure appears to have had an adverse impact on question response rates. For the majority of questions asked in both 1991 and 2001, the non-response rate had increased from less than 1 per cent in 1991 to 1 to 5 per cent in 2001.(See Table A below comparing 2001 and 1991)'*
26. *'We recommend that in their evaluation of the 2001 Census, ONS review the balance of resources devoted to enumeration in the best performing areas and those devoted to the worst and consider what changes may be necessary to the Census in the light of the response rates to individual questions. In view of the comments of the Market Research Society, ONS might also usefully evaluate whether returning forms by post, rather than through enumerators, had any impact on the response rate to particular questions.'*

^{2 2} Select Committee on Treasury, First Report, 2001 Census in England and Wales.
<http://www.publications.parliament.uk/pa/cm200102/cmselect/cmtreasy/310/31002.htm>

Table A: Comparison of England and Wales 2001 Non-response and imputed percentages with GB 1991 missing, invalid and inconsistent responses percentages³

Variable	Total (including imputed) 000s	Non- response 000s	Imputed 000s	Non- response %	Imputed %	1991 missing, invalid and inconsistent %
Age	49,359	262	278	0.53	0.56	0.5
Sex	49,359	199	219	0.4	0.44	0.2
Marital status	49,359	372	158	0.76	0.32	2.5
Country of birth	48,848	1,211	829	2.48	1.7	0.4
Ethnic group	48,848	1,405	1,421	2.88	2.91	0.9
Welsh language	2,754	153	153	5.54	5.57	2.3
Long-term illness	48,848	1,899	1,915	3.89	3.92	0.9
Address one year ago	48,848	2,198	2,213	4.5	4.53	1.5
Activity last week	35,367	-	1,301	-	3.69	8.5

27. I interpret this as encouraging a 3-factor evaluation of enumeration methodology - by data quality, under-enumeration and cost. Postout/Postback is the cheapest method. Delivery/Collection retains the most control over data quality but is probably most expensive (However changes to Delivery/Collection, such as a single delivery round rather than 2 attempts to contact, will reduce costs). Delivery/Postback combines robust address and household validation with reduced collection costs. Postout/Postback and Delivery/Postback relinquish control of key elements of data collection to the postal system (see Para 31). Improvements to each of the methods need to be evaluated in terms of cost, quality and practicability.

Publication timetable

28. *'We note that the first results from the 2001 Census will not be available until August 2002 and that the main results will not be available until the first half of 2003, when data on which local authority spending assessments are based will be 12 years old. We recommend that ONS review the trade-off in cost benefit terms of the Census results being available earlier for users and public resource planning against the additional cost of doing so, and publish the results of this exercise. In undertaking this work, ONS should take account of the requirements of all data users, and not just the resource allocation round for local government, which seems at present to be the primary determinant of the timetable.'*
29. Many find the census publication timetable slow. I am concerned that any adjustment to the Census results for under-enumeration in 2011 will either ensure that data could not be published more quickly or reduce the time available for quality assurance. In 2001 quality assurance was more difficult because the delivery timetable slipped. Arguably, a more stringent but expensive field quality control method could potentially speed up processing and reduce processing costs.

³ <http://www.statistics.gov.uk/census2001/editimputevrep.asp>
and 1991 Census General Report for Great Britain, 1995, p 124.

30. In 2001, the Census ONC adjustment relied on Enumerator evaluation of non-respondent household spaces (no-contacts, refusals, second homes, derelicts etc). This was not part of the original plan and not evaluated although GROS had confidence in enumerator accuracy and the quality checks in place. Relying on Enumerator assessment of household spaces, potentially simplifies household imputation methodology in 2011. Therefore we will try to evaluate the reliability of Enumerator assessments of household status under different conditions.

Reliance on outside agencies

31. *'In view of the difficulties experienced with the return of Census forms, we recommend that for any future Census ONS evaluate the benefits of postal return versus enumerator or other means of returning forms. If a future Census is to be based on postal return, ONS should conclude a service level agreement with the service provider aimed at ensuring that the conduct of the Census is not impeded by the quality of service received.'*
32. The Census is a one-off operation. We therefore reduce risk and build in contingencies as part of our planning. In some areas, the use of outside contractors increases risk. This came to the fore with our postal and pay contractors. We have to weigh the risks involved with methods which rely on contractors or service providers without any performance agreement and see whether in-house methods, over which we intrinsically have more control, are to be preferred.
33. There are many current address initiatives which should lead to a Definitive National Address database. It is planned that the 2011 Census use an address list agreed with each council.
34. The hardest areas to enumerate lie in our major cities. Post out/Post back is the cheapest method of delivering the Census; it has also been stated that it allows the field force to be used more flexibly (although it significantly alters the nature of the work from what an enumerator may expect and requires a different skill set). However a component of a Post out/Post back method (which could also be valuable in a traditional census) is to have regional offices in major cities. We are assessing this option.
35. There has been speculation about the place of internet data capture, telephone data capture and interactive TV in the traditional Census methodology. It is a government requirement to be able to provide services over the internet but the Census is not a service and Traditional Census Data Collection is not planning an internet option and GROS is not funded to trial this in 2006.

I would be interested in PAMS members views on an internet option for 2011, accepting that ONS will have an internet Census completion option in 2011.

36. As noted GROS will test ethnic/cultural questions developed by RESIG and an income question⁴.

Test Design

37. From this it follows that the Test design should meet a number of requirements. They are laid out below (Table 1). What is not covered in this paper are testing requirements which follow from the proposed consultation in Scotland which would include for example new questions and

⁴ Select Committee on Treasury, First Report, 2001 Census in England and Wales.
<http://www.publications.parliament.uk/pa/cm200102/cmselect/cmtreasy/310/31002.htm>

potentially new population bases. This paper does not include a detailed evaluation plan of the Test. This would follow when we finally knew which questions we would trial in the Test. Acceptance of this plan allows the work of the branch to move forward in the main areas of preparation including strategic plans for recruitment, training, pay etc.

Table 1: Requirements of the 2006 Test

Requirement	Comment	
Income question	In 50% of forms	
Ethnicity question	Ethnically rich and diverse area - Glasgow - no split design required but may be many variants of ethnic/cultural questions for follow up purposes.	
Cost	Postout/Postback Delivery/Postback Delivery/Collection	Rural Suburban Urban
Cost	Single round delivery Triple round delivery	Rural Urban
Data Quality	Postout/Postback Delivery/Postback Delivery/Collection	No check back Minimum information Doorstep check
Reliance	Postout/Postback Delivery/Collection	
In house and external pay	Develop an in house pay system and compare to private sector costs – each for 50% of the Test.	
Regional Offices	1 in Glasgow	
Agreed Address lists 'Address Checkers'	Postout/postback.	Rural Suburban Urban
Enumerator identification of vacant property	Holiday homes area Partially vacated tower block	
Processing	Allow prototype processing; Allow data quality to be assessed;	

38. The location of the 2006 Census Test is partially determined from these requirements and superficially I would suggest an ethnically diverse area of Glasgow, a rundown suburb - selected on levels of unemployment and public sector housing, and an area of holiday homes perhaps around Ballachulish, Crianlarich and Killin
39. For the 1999 Census Rehearsal a number of other criteria were used to select the Rehearsal Areas, but then we were also part of the design of the greater test for the UK. Table 2 contains the relevant criteria from that design.
40. EC1 constrains the design by requiring particular inter-staff grade ratios. However I have taken it that the ratio between Regional Census Managers and Local Census Managers would not be tested and so can vary to suit our budget.

Table 2: Criteria used to select the 1999 Census Rehearsal Area

Source	Criteria	2006 Requirement
Coverage Survey need	Ambiguous household residency	Rural area Good vs. Bad Urban Area. Students/transients workers Tenements
Data Collection costs	Relative costs (Table 3 below shows costs from the 1997 test)	Rural, suburban, urban against Postout/Postback (popb), Delivery/Postback (dpb) and Delivery/Collection (dc) methods
Essential Criteria (EC)1	3 Realistic field organisations See Table below for smaller variants	e.g. 2 Regional Offices with 3 staff 2 Regional Census Managers, 12 Local Census Managers, 36 Census Assistants, 324 Enumerators 60 Address-list Checkers
EC 2	A student hall of residence	University City
EC 3	High Ethnicity	Glasgow
EC 4	Residents with qualification	Not relevant
EC 5	Widespread types of occupations	Not relevant but would happen in any case

41. Comparing the application of 3 methods of enumeration in 3 types of areas seems a central requisite. Such a comparison was made in 1997 using data from across the UK. The results from 1997 are shown in Table 3.

Table 3: 1997 comparison of costs by 3 enumeration methods.

Quality	Post out/ Post back	Delivery/ post back	Delivery/ collection (SW Argyll)	Delivery/ collection (not SW Argyll)
Person Qs	5.2% < delivery collect	2% > delivery/collect		
Household Qs	3.1% < delivery collect	1.3% > delivery/collect		
% hhld response rates	55.5*	52.9**	81.6	60.5
Cost/hhld	£2.56	£2.49	£3.71	£2.58

* influenced by outlier low rate in Brent B.

** influenced by lack of effective follow-up, [but some costs of this are included].

42. A repeat of this exercise is proposed because:
- Some reduction in costs is available on Delivery/Collection by having a single delivery round;
 - We can improve address lists substantially and wish to trial the concept of Address-list Checkers;
 - We want to assess data quality against each method;
 - A hybrid enumeration scheme has been suggested. We would cover much of Scotland by postout/postback and so concentrate Enumerator resource on hard-to-count areas;
 - The 1997 exercise was distorted because most of the Test volume was in England.

43. This leaves us with a 3x3 design which, with income/no income, becomes a 3x3x2 design. The size of the Test follows from this and the staff required is shown in Table 4.

Table 4: Suggested Test Field Staff Count.

Staff Estimates					Minimum Cost Option		
Staff grade	Min	Med	Max	Ratio between staff	Cost (£) per staff including expenses in 2001 and 2006		Cost (£) in 2006
Regional Offices (1 or 2)	3	6			na	5,000	5,000
Regional Census Managers	2	3	4		nk	10,000	20,000
Local Census Managers	6	9	12	3	4,100	5,800	34,800
Assistant Managers	18	27	36	3	2,000	2,320	41,760
Enumerators	162	243	324	9	600	700	113,400
Address checkers	13	20	27	1: 4 popb Enumerator	600	700	9,100
Total staff	204	308	403		Est Cost (£)		226,000
Enumerators per variant	9	13-14	18				

44. A minimum-size design is shown in Table 5.

Table 5: 'Minimum' design for 2006 Census Test showing number of Enumerators by treatment.

Enumeration Method	Urban		Suburban		Rural		Total
	Income	No Income	Income	No Income	Income	No Income	
Popb	9	9	9	9	9	9	54
Dpb	9	9	9	9	9	9	54
Dc	9	9	9	9	9	9	54
Total	27	27	27	27	27	27	162

45. It is proposed that popb, dpb and dc Enumerators would cover 600, 350 and 200 households respectively. Multiplying up by the number of Enumerators gives the number of households per treatment (Table 6).

Table 6: 'Minimum' design for 2006 Census Test showing estimated number of delivered forms by treatment.

Enumeration Method	Urban		Suburban		Rural		Total
	Income	No Income	Income	No Income	Income	No Income	
Popb	5,400	5,400	5,400	5,400	5,400	5,400	32,400
Dpb	3,150	3,150	3,150	3,150	3,150	3,150	18,900
Dc	1,800	1,800	1,800	1,800	1,800	1,800	10,800
Total	10,350	10,350	10,350	10,350	10,350	10,350	62,100

46. Table 7 aggregates the forms by each treatment. The important thing is that because of the size of this sample we would get statistically significant results with a 1% difference in response rate between any of the treatments - even at lower response levels. A 1% difference in response rates would start to cease to be significant at about the 20,000 level - where a greater difference would be required. For the dp vs other treatments a 2% response rate marks the boundary of significant difference at the 5% significance level (for a two tailed test).

Table 7: Planned number of forms issued and estimated number of forms returned

	Forms issued	Forms returned
Popb	32,400	22,680
Dpb	18,900	13,230
Dc	10,800	7,560
Urban	20,700	14,490
Suburban	20,700	14,490
Rural	20,700	14,490
Income	31,050	21,735
No income	31,050	21,735

47. These figures suggest that from a statistical test point of view we could reduce the size of the test by about a further 1/3rd. However Enumerator and other management factors could then come into play. If we had to remove outliers and failed areas etc leaving us to differentiate between income and no income for urban delivery collection, the difference in response rates required, to be significant, increase to about 4-5%.
48. However, in any final design, an enumerator could implement more than 1 type of enumeration which would give a more even number of forms across each treatment and ensure that areas were more alike and the enumerator effect was minimised. In any case income and no income treatments would be mixed across EDs in a pre-determined distribution. Originally we viewed mixed enumeration as problematic but it is being carried out successfully in Ireland currently.
49. It is important to note that a worse response to the income question would not automatically preclude its presence in the 2011 Census. Other designs etc would still be pursued.

Where?

- 50. We have provisionally decided that ethnically rich areas, a poor suburban outer city estate and a third area of holiday homes would meet the criteria set for the Test. It was a further consideration that training could all be concentrated in Glasgow and that travel reduced as far as possible. Therefore we will choose areas from inner city Glasgow and the area between Ballachulish and Killin. A third area of Public sector housing and unemployment and relatively high household imputation in 2001 has not been identified yet.
- 51. The dual aims of having a high proportion of members of minority ethnic groups and a rich variety of ethnicity in one area are not compatible. Analysis shows that Glasgow is quite ghettoised. This is shown in Figure 1 where an ethnic diversity index is plotted against minority ethnic group percentage.
- 52. Therefore the areas chosen in Glasgow will have to be spread across the inner city, while it is expected that the other two test areas would be formed from contiguous postcode sectors.

I would welcome PAMS members comments on:

The design and location of the Census test

The balance between cost, data quality and coverage.

Views on household vs individual income and designs to minimise refusals.

Fig 1: Diversity vs % non-White Ethnic Population in Glasgow by Postcode Sector

