

**THE NORTHERN IRELAND CENSUS OF POPULATION AND HOUSING:  
LOOKING BACK AND LOOKING FORWARD**

NORMAN CAVEN, TREVOR EVANS AND ROBERT BEATTY  
Northern Ireland Statistics and Research Agency

(read before the Society, 12 March 1998)

---

**1. INTRODUCTION**

A census of population and housing, even in a country as small as Northern Ireland, is a major undertaking and much time, effort and resources are devoted to it. At a national level it is the largest single statistical exercise carried out by government - indeed at one stage during the 1991 census around 150,000 people were employed on it in the United Kingdom as a whole. Data collected in a census have many uses - foremost perhaps being the provision of an accurate population count at local level. In addition, the rich detail of census data provides essential information to plan and deliver services to the population as a whole and various subgroups within it, including those most in need.

**2. HISTORY**

Population censuses have a long history. It is recorded in the Bible that the fighting strength of the Children of Israel at the Exodus was ascertained by a count of all males of 20 years and upward made by enumerators appointed for each clan. However, it was the Romans who both gave us the word and placed the act of counting the people on a regular footing for reasons both of administration and taxation.

The first census in the British Isles took place in Great Britain in 1801 - an event which owed much to the work of John Rickman, Clerk to the House of Commons, who in 1800 outlined a list of arguments for such an activity. These ranged from reasons related to the general good of the country to the more detailed, such as ensuring that militia conscription reflected an area's population, how much corn needed to be planted to feed the population, and, perhaps initially somewhat more surprising, the need for information to stimulate the life insurance industry. Rickman persuaded a Member of Parliament to introduce a Private Member's Bill which was subsequently enacted, its passage eased by references to the need for a census to assess food supplies required after the disastrous harvest of 1800. In this context, Catherine Marsh<sup>1</sup> has noted that the general climate created by the then recent publication of Thomas Malthus' essay on population, (in which he outlined his

proposition that populations had an inherently greater capacity to grow than did food supplies), provided the backdrop for the first Great Britain census.

Official census taking came later to Ireland although social historians rightly point to the rich source of demographic and socio-economic data available from earlier work such as that of Sir William Petty in the Down Survey<sup>2</sup>. (Petty had estimated the population of Ireland at 1.1 million in 1672). However, the first official attempt at modern enumeration came in 1813 twelve years after the first census in Great Britain. Sadly, it was not a success, the supervision of the enumeration having been passed to the Grand Juries of the various counties who were not adequately structured to accomplish the task. In some counties no steps were taken whilst in others the work completed was unsatisfactory. After two years the attempt was abandoned<sup>3</sup>. Both the censuses of 1821 and 1831 also experienced certain difficulties - and it was not until 1841 when Ordnance Survey Maps were available and the police force provided a corps of enumerators that a totally comprehensive enumeration was achieved.

From 1821 censuses were taken at ten yearly intervals until 1911. Subsequently, events in Ireland resulted in the postponement of the 1921 census, 'the disturbed state of the country' being cited as the reason. Thus resulted a break with Great Britain census timing - one which was not restored until 1951. It was not until 1925 that the Northern Ireland Minister of Finance introduced the second reading on a Census Bill. He indicated that it was nearly 15 years since the last census in Ireland and that intervening events such as the Great War and the 1918-19 flu epidemic required a census to be carried out. It was accordingly held on 18 April 1926 (the day also of a census in the rest of Ireland).

In order to bridge the gap between the census of 1926 and the planned census of 1941 a census of more limited scope (e.g. omitting questions regarding occupation and industry) was taken in 1937. The outbreak of war and the subsequent paper shortage led to a restricted publication programme and later, inevitably, to the abandonment of plans for the 1941 census. In 1951 NI came back into line with the rest of the United Kingdom (or vice versa) and since then censuses have been held at ten yearly intervals with the exception of the additional 1966 mid term census.

The areas of inquiry at each enumeration since 1926 have an element of variation. As Paul Compton describes in detail in an article on the NI censuses in the 1991 Census Users Guide:

*"Elements common to all include relationship to head of household, age, sex, birthplace and marital status. Most censuses have also enquired into current residence, religion, employment/occupation and education although the emphasis in the case of the last mentioned has changed from school attendance in the early enumerations, to qualifications obtained in more recent censuses. A characteristic feature, particularly since 1961, has been the extension of the*

*scope of successive censuses to include information on dwellings, migration, fertility, journey to work, car ownership, long-term illness and the Irish language”<sup>4</sup>.*

As a compulsory activity the census has a legal basis. Here, prior to 1969, censuses were taken under the authority of separate Acts of the Parliament of Northern Ireland. Enumerations post 1966 have all been carried out under the authority of the Northern Ireland Census Act, 1969. This Act makes provision for a census to be taken by Order in Council at any time but with the proviso that at least five years must have elapsed since the last enumeration. As in earlier legislation, the Act confers a duty on the Registrar General to make such arrangements and do all things as he deems necessary for the enumeration. A schedule to the Act outlines the matters in which particulars may be required, viz:

1. Names, sex, age.
2. Occupation, profession, trade or employment.
3. Nationality, birthplace, race, language.
4. Place of abode and character of dwelling.
5. Condition as to marriage, relation to head of family, issue.
6. Education, professional and technical qualifications.
7. Religion.
8. Any other matters with respect to which it is desirable to obtain statistical information with a view to ascertaining the social condition of the population.

The last point provides considerable scope to include questions relevant to user needs at particular points in time although of course the proposed census topics decided on are subject to parliamentary scrutiny. As in Great Britain, the legislation also provides latitude to the Registrar General as to the number and form of abstracts and reports to be prepared thus enabling alterations to be made which reflect changing needs and the potential of technological developments to disseminate results.

### 3. THE 1991 CENSUS

#### *Planning the 1991 Census*

In 1986, following the Government's announcement that there would not be a mid-term census, work began on planning the 1991 Census. Users were asked about their data needs and from such contacts three categories of questions arose for consideration; previous questions that prima facie had sufficient support for inclusion, proposed amendments to 1981 questions to elicit additional information, and proposed new questions. The latter included questions on long-term illness, term-time address of students and schoolchildren and Irish language. Questions on hours worked and the number of children born alive to ever married women, previously asked in the 1961 Census, were also included. Particular attention was paid to ensuring public awareness of the census and a contract was awarded to an external public relations agency. The publicity campaign adopted focussed initially on highlighting public policy questions that a census of population was best placed to answer using billboard and later television advertising. A second stage involved screening a countdown of days to Census Day with encouragement to the public to complete what was considered an essential exercise and a final stage over the last ten days highlighted the actual date of the census.

#### *Fieldwork*

Fieldwork for the 1991 Census was organised in a similar manner to 1981 and census enumerators were responsible for enumerating all persons present or resident within each enumeration district on census night. These districts are the areas allocated to individual enumerators for the delivery and collection of census forms and they vary in size from 50 to 300 households according to the settlement geography of the area. The enumeration districts were drawn up using Ordnance Survey maps supplemented by Valuation Lists of rateable properties. The boundaries of the districts were constrained to follow those of townlands, electoral wards and district councils. 1981 enumeration districts were used as the building blocks for 1991, although where changes had taken place to statutory boundaries or to the housing stock either by clearance or new development this was not possible and new boundaries were drawn up.

#### *Published Reports*

As regards output, a Preliminary Report based on returns from census enumerator record books and a comprehensive Summary Report were supplemented by topic reports on subjects such as Economic Activity, Transport to Work, Housing and Household Composition, Migration, Education, Irish Language and Religion. These published reports describe the main characteristics of the population at the Northern Ireland and district council levels. Census results for smaller areas - electoral wards

and grid squares (100 metre in urban areas and 1 km in rural areas) - are available on request, in hard copy or in machine readable form.

Such output was available in 1981. What was new in 1991 was that under an Economic and Social Research Council (ESRC) 1991 Census Initiative, Small Area Statistics (SAS) at enumeration district level were produced for Northern Ireland. Modelled on the Great Britain equivalent, this data set comprises cross-tabulations of data for 75 pre-specified tables available at a wide range of geographical units. Again the data sets are available in machine readable format and cover all census topics. A major advantage with ED SAS is that users are not limited to single variable counts.

Also under the ESRC Initiative, two Samples of Anonymised Records (SARs) were produced. These are samples of records of individual data from the 1991 Census suitably anonymised. One is a two per cent sample of individuals in households and communal establishments, the other is a one per cent sample of households and individuals in these households. The samples give users considerable scope for preparing tables not otherwise issued or commissioned and offer greater scope for statistical analyses. The lowest geographical level in the individual sample (2%) is taken from areas comprised of combinations of local government districts with a minimum resident population of 120,000. The household sample (1%) is produced at Northern Ireland level only.

#### **4. PLANNING FOR THE 2001 CENSUS**

Earlier in the decade the Government decided that planning should proceed on the basis that the next census in the UK would take place in 2001. Although the three Census Offices, (the Office for National Statistics, the General Register Office Scotland and the Northern Ireland Statistics and Research Agency) operate under separate legal and administrative processes, work towards 2001 is proceeding in a co-ordinated way overseen by a United Kingdom Census Committee comprised of senior officials from each of the offices. The Committee agreed a detailed development programme the key points of which are outlined in Table 1.

To date preparation for the 2001 Census has involved consideration of the basic structure of the census questionnaire, cognitive and wording testing of possible new questions and consultation with users. In Northern Ireland a Census Advisory Group, representative of business, academia and the voluntary sector was established in 1995 and meets on a regular basis. Census developments are also reported to the independent Northern Ireland Statistics Advisory Committee which has a broader remit on local statistical issues. In 1997 there was a major landmark in the development programme with the conduct of a Census Test which involved the enumeration of about 9,000 households in a number of areas in Northern Ireland. There was a simultaneous Census Test in Great Britain with broadly similar objectives.

**Table 1 Detailed Development Programme**

Year	Month	Action
1997	March	United Nations Principles and Recommendations
	15 June	1997 Census Test
		European Union Guidelines Agreed
1998	Early 1998	Complete evaluation of 1997 Census Test
		Develop strategy for Data Processing
		Agree output strategy
	Mid 1998	Agree topics for census questions
	Autumn 1998	Publish proposals for 2001 Census (White Paper)
1999	Early	1999 Census Dress Rehearsal (Field Work)
	End	1999 Census Dress Rehearsal (Data Capture)
		Census Order laid before parliament
2000	Early	Make Census Regulations
		Appoint Senior Field Staff
2001	April	Census Day

*The NI 1997 Census Test*

Objectives

The Census Test examined three main variables.

1. Recognising that a large proportion of the cost of a census is accounted for by the enumeration process of delivery and collection of forms, the Census Test was designed to examine if returning forms by post would be acceptable to the public and would not affect the quality of response. This involved a comparison of the response rates to the Census Test using traditional enumerator collection and a post-back system.
2. Form design is an important part of any census and in the test two basic designs were used - a matrix form and page per person form. In the matrix form, each question forms a row of a matrix, while the replies for each person form a column. In a page per person form, the complete set of questions for the first person in the household precede the questions for the next person and so forth. The Test was to measure the acceptability of the two form designs.
3. Users had indicated a desire to see an income question included on the census. Accordingly, a banded income question was developed for the Census Test to examine the acceptability of the inclusion of such a question.

This primarily involved a comparison of the response rates to the Census Test with the inclusion and non-inclusion of an income question.

The test was also designed to examine the acceptability of these different approaches in areas with different socio-economic backgrounds. These backgrounds were defined in terms of urban/rural, level of relative disadvantage and religion. In addition, the test provided an opportunity to examine the logistics of the census operation involving, among other things, the designation of enumeration areas, the production of maps and address lists for enumerators, the control of forms returned to Census Office and entering the information on the Census Test forms onto computer.

#### The Design of the Northern Ireland Census Test

As indicated the Census Test was based on a number of enumeration areas. Within any enumeration area each household received the same combination of the three design variables. The enumeration areas were designed to be small and consequently more likely to be socio-economically homogeneous. The socio-economic background of each was defined in terms of its urban / rural character, its level of relative disadvantage and the predominant community background (religion), as follows:

Urban	Part of a town
Rural	other areas
Deprived	A degree of deprivation score exceeding zero, the Northern Ireland average <sup>5</sup> .
Non-deprived	other areas
Protestant	At least 80 per cent of the local population were either Protestant or other non-Roman Catholic religion in 1991.
Roman Catholic	At least 80 per cent of the local population were Roman Catholic in 1991.
Mixed	Other areas

We thus have the elements of a  $2^5 \times 3$  factorial experiment. A fully saturated version of this would require the selection of 96 enumeration areas for the Census Test, and would result in a model from which all main effects and their interactions could be determined. Cost was however an important consideration; accordingly, a half-replicate which allowed unbiased estimation of main effects and two-way interactions was used. The Census Test thus required  $2^5 \times 3/2$  - a total of 48 enumeration areas.

The statistical design was also orthogonal, so all main effects and two-way interactions could be estimated independently. The power of the design was such

that there was an 80 per cent chance that true differences of 3.3 per cent in the response rates between the levels of two-level factors will be detected as statistically significantly ( $p < 0.05$ ) different. For religion, true differences of 4 per cent between any two of the three levels will similarly be detected

The 3,729 Enumeration Districts from the 1991 Census were used as the basis for enumeration areas for the Census Test. The aim was to achieve similar precision for all estimates so, where necessary, the boundaries of the 1991 enumeration districts were amended so that each enumeration area contained broadly similar numbers of households, approximately 200. This was necessary because, in general, urban enumeration districts contain more households than rural enumeration districts.

The resulting 2,685 enumeration areas were placed into groups, representing the 12 different combinations of socio-economic conditions. From each group, 4 enumeration areas were selected giving the 48 enumeration areas required for the Census Test.

The design of the test was thus based on 48 enumeration areas, balanced across three socio-economic variables and with balanced representation of the three design variables; collection method, form type, and inclusion / non-inclusion of an income question. Enumerators were then set the task of completely enumerating each enumeration area. Apart from its voluntary nature, the Census Test was conducted as if it were a normal Census of Population.

### Results

The 48 enumeration areas contained 8,845 occupied households and the response rates (measured by the actual return of a census form) in individual enumeration areas ranged from 29 per cent (a university area dominated by student flats) to 98 per cent (a rural part of County Armagh). Given the wide range of response rates, for analysis purposes the response rates have been transformed using the root arc sine transformation which will give similar sampling variances within each enumeration area. The unweighted average response rate across the 48 enumeration areas was 73 per cent.

The average response rates for the main effects are shown in Table 2. A number of statistically significant ( $p < 0.05$ ) differences were detected amongst the 3 main design variables. For example the response rate for the conventional collection methodology was higher than that for postback, and the response rate where an income question was included was lower. Despite the statistical significance of these differences, a noteworthy substantive finding is the overall similarity of the response rates among the three design variables and the wider disparity in response rates when socio-economic differences are examined.

**Table 2 Observed Average Response Rates**

Factor	Response rate (%)
Conventional collection	74.2
Postback	72.1
Matrix	74.5
Page per person	71.7
Income question	71.9
No income question	74.3
Urban	65.3
Rural	80.3
Deprived	69.6
Non-deprived	76.5
Catholic	71.2
Mixed	68.0
Protestant	79.7

A model has been developed which brings together the effect of the design variables and socio-economic backgrounds on the response rate of an enumeration area. In developing the model the 3 levels of the religion variable were transformed into 2 orthogonal contrasts representing Protestant areas versus others and Roman Catholic areas versus others. The model is

$$\begin{aligned} \text{Transformed response rate} &= 1.026 \\ &+ [0.117 * \text{Protestant}] \\ &+ [0.085 * \text{Rural}] \\ &+ [0.039 * \text{Non-disadvantaged}] \\ &+ [0.099 * \text{Protestant x Form type PPP}] \\ &+ [0.038 * \text{Non-disadvantaged x Form type PPP}] \\ &- [0.086 * \text{Protestant x Non-disadvantaged}] \end{aligned}$$

$$r^2 = 0.53$$

$$\text{Adjusted } r^2 = 0.46$$

$$\text{SE of estimate} = 0.134$$

Thus, for example, the expected response rate in an urban, Protestant, non-disadvantaged area using a person per page form would be

$$\begin{aligned} & 1.026 + 0.117 * 2/3 \\ & + 0.085 * (-1) \\ & + 0.039 * 1 \\ & + 0.099 * (2/3) * 1 \\ & + 0.038 * 1 * 1 \\ & - 0.086 * (2/3) * 1 \\ & = 1.1047 \end{aligned}$$

which back-transforms to an expected response rate of just under 80 per cent. The expected response rates varied from 48 per cent (urban, deprived, Roman Catholic or mixed areas using a page per person form) to 91 per cent (rural, non-deprived, Protestant areas using a person per page form). It is interesting to note that the page per person form is predicted to give both the highest and lowest response rates.

The balanced statistical design of the test means that the coefficients in the above model can each be interpreted in isolation, unlike most multiple regressions. Apart from the form type appearing in interaction terms, the design variables did not appear in the model, reiterating the point that the response rates to the design variables were broadly similar; they vary more with the socio-economic background of an area. The interpretation of the main effects coefficients are that the response rate should increase as;

- areas become more Protestant
- areas become more rural
- areas become less disadvantaged.

An interpretation of the interaction coefficients is that:

- response rates in Protestant areas are higher than in mixed and Roman Catholic areas. The effect differs by form type, being more pronounced with a page per person form;
- response rates are higher in less deprived areas. Again the effect differs by form type, being more pronounced with a page per person form;
- within Roman Catholic and mixed areas, the response rate is higher in non-deprived areas. Within Protestant areas the response rate is higher but is similar for both deprived and non-deprived areas.

The above analyses were based on the act of returning a census form to a voluntary census test, and take no account of the extent to which the form was completed. A

small number of blank forms were returned, but the vast majority of forms had been at least partially completed. This was examined further in two ways:

- a minimum information set (age, sex and marital status) for at least one person was provided in just under 98 per cent of returned household forms;
- a measure of form completion was derived using the proportion of questions that each person had completed, with regard to the number of questions that they should have completed. On average, respondents answered 88 per cent of the questions which they were asked.

These two measures did not display any strong patterns of variation with the design variables or socio-economic background. This consistency is interpreted to mean that simple response rate provides a proxy for the acceptability of the Census Test.

The model thus has practical utility in informing approaches to the next census. If, for example we were to assume that a post-back methodology is adopted, using a page per person form, the expected response rates would be:

	Roman Catholic/mixed	Protestant
Rural Deprived	65	89
Rural Non-deprived	83	91
Urban Deprived	48	76
Urban Non-deprived	68	80

Whilst responses overall would be expected to be higher in the 2001 Census due to its compulsory nature, lower response rates might be expected in mixed and Roman Catholic areas, especially in urban deprived areas. As a postback methodology would involve follow up by enumerators to non-responding households, this would have implications for the potential workloads of enumerators in different types of areas.

#### Individual questions

The inclusion of an income question did not materially affect the response rate to the Census Test, but this takes no account of whether a response was actually provided to the income question. Of those who returned a Census Test form containing the income question and to whom the question was relevant, 75 per cent had completed the income question. This is placed in context in Table 3 which shows that this level of voluntary response is comparable to that for questions on economic activity.

**Table 3 Response rates to individual questions**

Question	Average response rate*
Age	96
Religion	95
Ethnic group	95
Irish language	93
Qualifications	86
<b>Income</b>	<b>75</b>
Work address	70
Number of jobs	69
Workplace size	60
Industry	55
Secondary economic activity – (unemployed)	51
Work postcode	27

\* The number responding to a given question expressed as a percentage of those who returned a form and to whom the question was relevant.

The response rate to the question on activities undertaken in addition to paid employment was relatively low and will require further examination. The level of response to the questions relating to employment, such as industry, size of establishment and type of work was also low. The lowest response rate was for postcode of workplace, possibly reflecting a lack of knowledge of the information.

Completion rates for individual questions were analysed by the design variables and the socio-economic factors. The collection methodology and the inclusion of an income question did not affect the response rates to any individual questions. The response rates did vary with form design for a number of questions, mostly related to economic activity such as description of main job, hours worked and secondary economic activity; the response rate was higher for the page per person type of form. Such differences, while they require further examination, suggest that all else being equal, page per person may be a more appropriate design to consider for the next census.

#### Census Test Evaluation Survey (CTES)

Consideration of individual questions raises the notion of the validity of the responses in the Census Test and the Census of Population. A census is a self-completion exercise and, accordingly, the answers given by respondents are accepted. It is important that questions in the census are worded clearly and cannot be interpreted ambiguously. To investigate this, a follow-up evaluation survey was

conducted. A sample of about 1,000 households was included in a Census Test Evaluation Survey.

### Design

The Census Test Evaluation Survey (CTES) involved revisiting 1,000 households who had responded to the Census Test. The 1,000 households were located in a subset of 24 of the enumeration areas from the Census Test, balanced by the three design variables and socio-economic backgrounds. An equal number of households were selected in each enumeration area. Selected households were visited by survey interviewers who had access to the household's response in the Census Test, such that differing responses to questions in the Test and the Evaluation Survey could be probed.

### Results

Given that the households had already responded to the voluntary Census Test, it is perhaps not surprising that the response rate to the evaluation survey was a high 82 per cent. After allowing for some households which had become vacant or where the occupants had moved, this yielded a responding sample of just under 800 households on which the following section is based.

#### *Attitude to the Census Test and overall opinion*

While the most frequently given description of the form was "easy to understand" (49 per cent of all respondents), there were sizeable minorities who thought it was "difficult to understand" (14 per cent) or "too long" (27 per cent). The pattern of responses on views of the form were similar for the matrix and page per person forms. There was a tendency for people from predominantly Roman Catholic and Protestant areas to be more likely to consider the form "difficult to understand".

Some 15 per cent of respondents indicated that they found some questions difficult, with the new question on the relationship between household members being named most frequently (24 per cent of those who had difficulties mentioned this question). In addition, although not identifying a particular question, there was a large group of people who said that they had problems with at least parts of the employment section.

When asked if they were unhappy about any of the questions, 18 per cent said they were, with the income question being cited most often - about 12 per cent of all those who were asked the income question. When asked directly if they minded being asked their religion only 5 per cent said they did.

### *Reproducibility of responses*

The Census Test household questions, and individual questions pertaining to the respondent, were then repeated to see how reproducible they were. The overall reproducibility scores for individual questions ranged from virtually 100 per cent on access to a toilet, bath or shower to 60 per cent for number of hours worked. Among those questions with lower reproducibility levels was 'health over the last twelve months', a subjective assessment where less than three-quarters of respondents (73 per cent) gave the same answer on both occasions. Although the proportion of the sample who claimed to have long-term illness was similar in CTES (30 per cent) and the Census Test (27 per cent), some 15 per cent of all respondents switched from one response to the other.

Respondents to the Census Test gave a number of hours worked in the week, to the nearest hour. Even when the responses were collapsed into 10 hour spans (0-9 hours, 10-19, and so forth) only 60 per cent of CTES responses were in agreement with the Census Test, with a distinct pattern of long working hours claimed in the Census Test being reduced in the CTES.

Analysis of the reproducibility rates by the design variables and socio-economic factors indicates that there are no statistically significant differences observed for the three design variables, that is the reproducibility is unaffected by the design variables. However, there were some differences in socio-economic characteristics, with the level of reproducibility being higher for some questions in urban areas. The difference was most pronounced on the question on any paid work where the level of reproducibility was 92 per cent in urban areas and 80 per cent in rural areas.

### *Interpretation*

Work is continuing to examine the results of the test which will be completed in the near future. The following points are emerging for detailed consideration on the way ahead:

- An income question. Whilst the presence or absence of an income question made little difference to the response rates, the CTES showed that there was dislike of such a question from a minority of respondents;
- Collection methodology. The response rate to the postal collection method was, in practical terms, similar to the conventional collection, suggesting that the method has potential;
- Form layout. The overall response rate was unaffected by the form layout but for individual questions, the person per page style yielded higher levels of response for a number of questions;

- Protestant areas, rural areas and non-disadvantaged areas tended to have higher response rates. Any adoption of a postback methodology will have to take this into account in planning work loads for enumerators.

Overall, the level of response to the voluntary Census Test was encouraging and has widened the options for Census enumeration in 2001. Further work remains to be undertaken to determine the costs and benefits of the different possible approaches in the wider context.

## 5. TIMETABLE TO 2001

The remainder of this paper highlights a number of other key activities in the period before 2001.

### *Business Cases for 2001 Census Questions*

Central to any census is consideration of the number and form of the questions included. A balance has to be struck between the need for information and the burden on the public which if mishandled can lead to major problems in response rates. As regards census content a Census User Needs survey undertaken in the mid 1990s involving both Government and public users informed the 1997 Census Test. As indicated the test findings are currently being evaluated in tandem with business cases being put forward by users for potential topics. Users have thus been circulated with proposals for the census content and asked to identify the uses made of each question. They have also been asked whether certain questions should be dropped or additional questions included.

The finalised list of desired questions will be assessed against the following criteria:

- there is a demonstrated need for the information by a substantial number of users;
- the information is not available from any other source in the detail or geographical level required;
- it is practicable to collect good enough data to meet the needs of census users; and
- the topic itself and the form of question are acceptable to the public;

and further considerations such as:

- acceptability of the topic in the 1997 test;

- quality of the statistics produced to meet user requirements;
- any limit on the overall cost of the census;
- overall burden of completion on the public.

The final assessment of business cases will be made later this Spring before feeding into the Government's proposals for the 2001 Census to be published as a White Paper in the Autumn. The final decision on topic content will be taken by Parliament.

*Census Geography and Output Strategy*

Critical to the whole census operation is the efficiency of the fieldwork operation to collect the information. In the past this has proved a major exercise using hand drawn enumeration districts on topographical maps which often prove difficult for enumerators to handle. A system to plan collection areas radically different from that used in 1991 was developed for the 1997 Census Test. In summary, the system will use Ordnance Survey of Northern Ireland (OSNI) computer generated maps centred on each enumeration district with an address list and accompanying twelve digit map reference for each dwelling.

Historically the enumeration districts of the collection geography have determined the output geography creating a lack of flexibility and relevance that users wish to see improved. Options for output from 2001 currently being considered are:

1. enumeration districts;
2. postcode units;
3. addresses with unique grid references; or
4. a combination of two or more of these.

At this stage the separation of output geography from collection geography through building output areas from unit postcodes is seen as a leading option. Of course with all options consideration will have to be given to risks of disclosure of information about identifiable individuals.

Finally, as regards the actual release of data, initial thinking is to process data for the whole of NI before releasing any results. This would mean no preliminary counts of the types produced from previous censuses.

Such an output strategy would:

- allow the provision of province wide results as a priority;
- allow improved quality checks on the data to be made and the results of this to be released with the statistical output;
- allow for free-flow processing (with the inclusion of late returns from the field); and
- allow for the introduction of a one-number census approach if this proves feasible (see below).

#### *Census Processing*

In actually processing the census - from returned questionnaires to a database ready for analysis - technological change continues to open up new opportunities. All 1991 basic processing systems are being reviewed with the aim of developing more cost effective ways of capturing and coding the data. In this context a percentage of the forms collected during the 1997 census test have been processed using automated data capture and coding technology. This involves scanning, optical mark reading, optical character recognition techniques and automatic and computer assisted coding of write in responses. The results so far have been encouraging and the three United Kingdom Census Offices are currently engaged in examining options to provide a workable system for the 1999 Census Dress Rehearsal. This rehearsal which will be held in a number of areas in Great Britain and Northern Ireland will provide an opportunity to test finalised questions alongside the main census processes of form distribution, collection, processing and data base construction.

#### *Towards a 'One Number' Census*

Finally, and of particular interest to statisticians is the methodological work going on to examine approaches to dealing with any under coverage in the census. In recent censuses no coverage adjustment has been made to the enumerated output. However, in view of the estimated undercount in Great Britain in 1991 - particularly the wide variation between population subgroups - the Office for National Statistics has undertaken to explore the feasibility of adjusting all the census counts. It is considered that this would be best achieved by supplementing the database of enumerated households and people by imputing those estimated to have been missed. A key source in estimating the size of any undercount will be an independent post - enumeration survey whose sample size must be large enough to give precise estimates at the required geographical level. Any such survey, which would go into the field soon after Census Day, would take the form of a short interview to check on the coverage of households and people within households and collect basic

demographic data. Demographic analyses would also be carried out to provide an independent check of the population counts at national level. This process would take some time but the trade-off is that higher quality estimates would be produced. Early indications are that the major customers would find such a timetable acceptable - providing the methodology works. A final decision on an 'one number' census approach will be taken after the 1999 Census Dress Rehearsal.

## 6. CONCLUSION

The Census Offices in working towards 2001 are thus aiming to reduce both development and operational costs utilising new technology and to combine this with better value products directed at a wider customer base.

Looking back over nearly 200 years it is tempting to say that census taking has come a long way since the first halting steps of the early part of the nineteenth century. Yet all along the signposts to the future have lain embedded in the landscape of the past. Each census has learnt from and built upon the previous one. There is thus much to admire in the early work of those who struggled with inadequate resources to plan, undertake and report the census in the early part of the 19th century, through to the dedication of those who worked to complete a census enumeration in Northern Ireland in the difficult times of 1971 and 1981. Recognisable throughout also is that spirit of enquiry epitomised in the report of the 1841 Census Commissioners.

*"We have only now to add that everyone of our inquiries might, with advantage have been extended and that with more time, many interesting combinations and comparison might have been obtained from the existing returns"<sup>6</sup>.*

## Footnotes

1. **Marsh, C., 1993.** “An overview”, in Dale A, and Marsh, C., *The 1991 Census User's Guide*, London: Her Majesty's Stationery Office.
2. **Petty, W., 1691.** *The Political Anatomy of Ireland*. Dublin.
3. **Report of the Commissioners, 1902.** *The Census of Ireland. 1901*, Dublin: His Majesty's Stationery Office.
4. **Compton, P.A.,** “Population Censuses in Northern Ireland”, in Dale, A. and Marsh, C., *The 1991 Census User's Guide*, London: Her Majesty's Stationery Office.
5. **Policy Planning and Research Unit, 1994.** “Relative Deprivation in Northern Ireland”, Occasional Paper Number 28, September.
6. **Report of the Commissioners, 1843.** *The Census of Ireland 1841*, Dublin: Her Majesty's Stationery Office.

## DISCUSSION

**Opening remarks by Sir George Quigley:** It is a great pleasure and honour for the Northern Ireland Economic Council to be involved in your meeting today and to host a reception afterwards in honour of your longevity. You have been celebrating your 150th anniversary. Compared with that, the Council is a mere stripling, although I must remind you that another organisation with which I am involved -Ulster Bank - was already a growing boy - all of 11 years old - when your Society emerged from the womb.

I think it is generally accepted that your progenitor was the Whately Professor of Political Economy in Trinity, William Neilson Hancock. How appropriate that your President should also bear that proud title. Hancock contributed around 100 papers to the Society's proceedings - a record, I believe. You will note the qualifications I carefully introduced into that sentence - 'around', 'I believe' - to avoid being hoist on my own statistical petard.

Whately had no mean opinion of economists. The world, he said immodestly, must be governed, has been governed and will be governed by political economists. But, more contritely, he added, 'though many of them are bad ones'. In 1860 alone he presented 7 papers, one of which was 'A plan for obviating the Identification of luggage at Kingstown and Holyhead, and so accelerating the through traffic between London and Dublin'. In her admirable History of the Society, which I read with enormous pleasure, Mary Daly comments somewhat acidly that this 'suggests diminishing intellectual returns'.

I was interested to note that amongst your first vice-presidents was Sir Robert Kane, whose *Industrial Resources of Ireland*, originally delivered in the form of Lectures at the Royal Dublin Society, was published three years previously, in 1844.

So far as my researches go, Kane's profound analysis was largely ignored in his own day and he is also - with honourable exceptions - a much neglected figure in the history of political economy in Ireland. His vision was of a populous island, holding a pivotal position in Atlantic trade, building up a strong indigenous manufacturing base using local raw materials and Irish capital, and gradually extending its scope to include manufacturing using foreign raw materials. No real attempt was made to pursue Kane's insight that - primarily by technical education and technology transfer - Ireland must in some way replicate the innovative milieu which existed in the central areas of England. In today's jargon you could say that Kane advocated a strong national system of innovation.

Kane was in no doubt about the route to prosperity for Ireland. For him, knowledge was power but he was clear about its nature: "practical knowledge; for power is essentially practical". He quoted, with approval, words of M. Briavionne, a Belgian minister, who summed up in two words - "Industrial Knowledge" - the reason for the

difference between the condition of England and Ireland. The fault, said Kane, “is not in the country but in ourselves; the absence of successful enterprise is owing to the fact that we do not know how to succeed”. Education for industrial pursuits was underrated. Too much youth was misspent in learning how one should have spoken if one had lived 3000 years ago. So much for a classical education!

It is interesting that Kane only contributed two papers to your Society - in 1851 and 1867 - both being the addresses at the opening of the 5th and the 20th Sessions. I would be interested to know whether in either address there was any resonance with his brave ideas of 1844.

The establishment of your Society, like that of others around that time, reflected a desire to bring political economy into policy-making and, as the London Society put it, to end politics by obliterating conflicts of principle. The latter at least, I fear, was a misplaced hope. But what I find fascinating is the way in which, over the years, the Society has played a major role in creating a policy community. In your midst, academics have mingled with civil servants. Many of the large issues which underlay the Whitaker *démarche* of the 1950s and 60s, for example, were debated within the Society.

Judging by Mary Daly’s book, civil servants were - untypically - prone to cast off in your company at least a few of the 7 veils. Some, indeed, committed the ultimate sin of diverging from the official Department of Finance line as expounded by the then Secretary, the by all accounts very formidable J.J. McElligott.

Mary Daly provides clear refutation of the calumny that statistical papers are dry as dust. I liked the description of constructing formal economic models as a process of ‘setting down one’s prejudices in algebraic formulae’. Even if the author had not been identified, I might have detected the pen of Maurice Doyle in words which warned of the danger of building an inverted pyramid of mathematical expertise on a foundation of statistics which were not designed to take the load’.

The sea change in the parameters of debate in Ireland is mirrored in the emphasis in your papers on economic development issues in the last 40 years, whereas, in the previous 40, only 4 papers related to manufacturing industry.

The North moves in and out of your story. A Belfast Society *preceded* the Dublin one by 10 years. I am not sure what its fate was but, as your own Society developed, people with Belfast addresses became involved. However, when World War I arrived, they had disappeared and the North featured little until the 1950s. I am delighted that it was Sir Charles Carter, who much later became the first Chairman of the NI Economic Council, who played a major role in the establishment of a Belfast branch. Plaudits are due to Professor Bob Black for his sterling role in keeping the flame alive in the difficult days of the late 60s and 70s.

Interestingly, I read that in the 1950s, attempts to draw comparisons between the two Irish economies were deprecated, even by Roy Geary, because of differences in policies and economic circumstances. Today, I am glad to say, from Tralee to Tandragee, they talk of little else.

And speaking of rural areas, there is the wonderful story of the Barrington Lecturer who found himself without an audience in Tramore in the early 1960s because the lecture clashed with a showing of *The Quiet Man*. Much more surprising is the fact that an audience of 100 could *normally* be expected. In the days before television, such lectures were prime entertainment in rural Ireland. What a sad decline we have witnessed!

I hope that, over its next 150 years, the Society will continue to display from time to time the pungency which enlivens Mary Daly's pages. There is Geary, for example, offering economists a place on the statistical society bandwagon but issuing the challenge: if you continue to sulk in your tents, we must travel alone. There is Norman Gibson surmising, with (I suspect) some satisfaction, that perhaps most of the audience had disagreed with what he had said. There is Geary - again - tempting fate by cheekily remarking, in the course of his own Presidential address, that there is no tradition of sacrosanctity for Presidential addresses.

But, humour aside, the major - the serious - themes are there all the time, generated, as your very creation was, by the pressure of problems imperatively demanding attention. The motto of the Dublin Society - Our Pole Star is Truth, - has been the hallmark of all your efforts, tirelessly conducted in the spirit of earnest inquiry. That spirit remains entirely relevant today when, despite all the tremendous progress which has been made North and South, there are still intellectual dragons to be slain and new domains of policy to be explored, as we continue to pursue Kane's vision of a strong national system of innovation.

Thank you very much for allowing me to intervene.

\*\*\*

**Proposal of Vote of thanks by Aidan Punch:** I am delighted to note that population matters feature prominently in this year's programme of papers celebrating 150 years of the Society's existence. In addition to this afternoon's paper in Belfast, which looks at Northern Ireland Censuses of Population, past and future, the next paper in Dublin will consider the economic and social implication of population change. Indeed population issues and Censuses of Population have historically been well catered for by the Society, as the index to the recently published history of the Society "Spirit of Earnest Inquiry" by Professor Mary Daly attests.

Today's paper is also opportune in its timing, coming as it does in the lead up to the next census - the first of the new millennium. In my contribution this afternoon my

sights will be firmly fixed in a forward-looking direction. Mr Tom Linehan's December 1991 paper to the Society entitled "The History of Population Censuses in Ireland" is a comprehensive account of the historical situation with regard to census-taking in Ireland. Some useful additional insights were added by Mr Linehan in his paper earlier this session on the development of official Irish statistics. Given his presence here this afternoon I am content to leave the historical dimension to him and concentrate instead in the preparations for the 2001 census and beyond.

Firstly let me say how envious I am of the ten year census cycle afforded to the authors. We in the South have to contend with the pressure of five yearly censuses. Indeed at one stage we must have held the world record for traditional census taking when we held four in a period of twelve years. The period I am referring to is 1979-1991. While my counterparts in the North have the luxury of census tests and dress rehearsals, as explained in Section 4 of the paper, we in Dublin have, of necessity, a much shorter planning horizon. One positive aspect is that we get plenty of practice at staging censuses. However, the scope for innovation is limited as the resources committed to finishing off one census cannot be put to work on long-term planning for the next.

In that regard we have more than a passing interest in the results of the Census test carried out by NISRA last year. The more interesting results are quoted in the paper. I will comment on some of these in turn.

Given that the fieldwork costs associated with carrying out a traditional census can be as high as 50 per cent of the total census budget this is an obvious area for achieving efficiencies. The census test design compared the conventional method whereby enumerators collect completed census forms with the option of getting householders to mail back their completed census questionnaires. The expected response rates in the case of non-deprived areas were impressive and would certainly lead one to conclude that the mail-back option is a likely candidate for these areas in 2001. However, while there would undoubtedly be savings in field costs there would also appear to be important drawbacks.

One of these is the issue of quality of the information on the forms. Using enumerators to collect completed census forms, while expensive in terms of the effort involved in the field, does however allow emphasis to be put on checking the completeness and accuracy of the information collected. This is done at the door step and is effectively the first scrutiny carried out on the data. During the 1996 and previous censuses in the South we placed a heavy emphasis on this aspect of the work in the training of the fieldforce. Where the completed (or partially completed) forms are mailed back by respondents this important quality control element is missing and has to be substituted by statistical techniques which impute for missing persons and/or households and for particular questions which may not have been answered. It is accepted of course that enumerators will be used to mop up non-respondents in the field. Notwithstanding this the use of the mail back procedure will

amount to a trade-off between coverage and quality of response on the one hand and cost of collection of the data on the other.

A further consequence of the mail back option is the fact that the completed forms will not be arranged by enumerators in logical geographical order i.e. street/townland, ward, enumerator district. As pointed out in the paper this will effectively rule out the publication of provisional or early results based on enumerator counts which have traditionally been published within two to three months of census day. Instead the forms will be processed in a free-flow fashion according as they arrive at a central collection point. The underlying assumption is that the capture and processing of the data is going to be done at far greater speed than achieved to date using techniques such as scanning, optical mark reading, optical character recognition and automatic coding of literal strings. It is relevant to ask what the likely timescales are going to be.

Our recent census in Dublin which could be considered to have been a modest success used the traditional fieldforce model (enumerators dropping off and collecting forms), keyboard data entry and automatic coding where feasible (mainly occupation, place of birth and country of last previous residence). The processing concluded after 21 months and all outputs are planned to be delivered by 32 months after census day. Using technology the 2001 census will have to show vast improvements on this performance. Take data capture for instance. The results claimed to have been recently achieved in Turkey for their November 1997 census, if they are to be believed, are truly astounding. All the key data (date of birth, place of residence) in respect of 20 million households were captured in 35 days (½ million forms per day). The remaining data will be captured over a more extended period. Translating this into an Irish context, the key data could be captured over a weekend! Obviously the number of machines and operators has a major bearing on the speed of capture.

Practically speaking if scanning, OMR, OCR etc. are to be embraced then this aspect of the operation will have to be outsourced - a major departure from previous practice. Clearly a number of major decisions with far-reaching consequences will have to be taken before the blueprint for the 2001 census is written. The emphasis will have to be firmly placed on speedy data capture supplemented by statistical editing techniques and output products which will meet users' requirements. Technology is also playing a major role in shaping different types of outputs. We have recently concluded an agreement with a Canadian software company to produce all our 1996 census outputs on one CD-ROM with bundled user-friendly table manipulation software. We are currently negotiating with a commercial developer to make our 1996 census output available on the Internet. By the time the 2001 census results are available I have no doubt that technology will have moved on even further.

Turning to the Census Test Evaluation Survey (CTES) mentioned in the paper, this appears to have been particularly informative in a number of respects. The one that interested me most was the long-term illness question. Pressure has been mounting in the South to include such a question in our 2001 census. However, the fact that 15 per cent of respondents had a different response in the census test and CTES in what is a dichotomous question would seriously call into question the usefulness of the results for policy formulation in the health sector.

The final topic in the paper to which I will allude is the so-called “one number census”. Every census, no matter how well it is carried out, will be subject to some error. Normally this takes the form of undercounting particular population sub-groups which may be difficult to enumerate. An example is young transient males. Some double counting of usual residents may not be ruled out either. A post-census validation survey is designed to adjust the enumerated count for these short-comings. However, being a sample inquiry this is also subject to error, both sampling and non-sampling. The adjusted totals at national level are then compared with population estimates which themselves may be erroneous because of the out of dateness of the base and the difficulty of estimating annual migration flows. Is it possible to say a priori which of these estimates is likely to be the most accurate in determining the magic “one-number”. This is a matter in which the statistical community and users in general will have a keen interest.

The final point I wish to raise in my contribution this afternoon is the future of Censuses of Population beyond 2001. Discussions with colleagues from other countries, carried out admittedly in hostels as far afield as Brussels, Luxembourg and Strasbourg has led us to coin the term “Cyber Census”. This relates to the census questionnaire in electronic format being downloaded from the Internet by the respondent, completed on-screen and transferred to a mailbox belonging to the statistical office. For those who do not have access to a personal link the local post office, school or public institution could provide the facility.

When I spoke on returning census data using electronic means while responding to Mr Linehan’s 1991 paper I expressed myself somewhat dubious about the practicalities. I now have to admit some six years on to being a lot less dubious. Technology is changing rapidly and official statisticians if they are to retain credibility will have to strive to use this technology to the maximum. This is the challenge that faces those of us who will be engaged in conducting censuses in the dawn of the new millennium.

In conclusion I would like on behalf of the Society to propose a vote of thanks to the authors for their stimulating paper.

**Reply by Norman Caven:** I must first thank Mr Aidan Punch for his comments. It is both interesting and helpful for us to hear his thoughts particularly given his considerable experience in census taking. Those of us involved in census work have some appreciation of the pressing demands created by a quinquennial census cycle. In the UK the current ten year cycle has entailed detailed consideration of the approach to census taking which has brought its own demands.

In his comments Aidan makes a number of interesting points on data collection and processing. Regarding data collection Aidan has noted that using enumerators to collect completed census forms, while expensive in terms of the effort involved in the field, allows emphasis to be put on checking the completeness and accuracy of the information collected. Whilst enumerators may check to see whether all of the questions have been completed, this does not necessarily mean that the accuracy of the information collected is assured. Northern Ireland has investigated the issue of data quality in its 1997 Census Test by looking at the reproducibility of the responses made to the Test questions. Statistical analysis of the responses made in both the conventional collection and postback options did not show any significant difference in data reproducibility. Assistance with form completion will still be available in Northern Ireland through the operation of a telephone query number. With a postback methodology it is thus not intended that there should be a trade off between coverage and quality of response and the cost of collection. Indeed, a "Cyber Census", as described, would also involve a form of postback methodology.

It is correct to note that the mail back option precludes the publication of provisional results based on enumerators' records. However, the Census Office in Northern Ireland has found that where such information has been made available in the past it has not been widely used. Preliminary counts of this nature can vary quite significantly from the final results (0.5 per cent difference in 1991) and producing more than one population count can create confusion for users. Accordingly, Census Office have proposed that a provisional count will not be produced for the 2001 Census.

I would like to conclude by thanking the Society for inviting us to give this paper.