Report of 2004 Census Pilot Survey

1. Introduction

The Government in its decision of 23 July 2003 authorised the taking of a Census of Population in 2006. It also agreed that CSO should

- commence consultations with users on the questions to be included in the 2006 census; and
- test the methodology to be used and any new or revised questions in a census pilot survey.

In response to the Government decision, a wide scale public consultation took place in late 2003. As part of this consultation process CSO set up a broadly based Census Advisory Group to assist it in its work. The census pilot survey was carried out on 25 April 2004. The final sample covered approximately 7,500 households in 32 enumeration areas located throughout the country. These were representative of urban and rural and deprived and non-deprived areas.

The present report describes

- the consultation process which took place;
- the objectives and design features of the census pilot survey; and
- the results achieved; and
- a summary of recommendations made.

Each of these topics is considered in the remainder of this report.

2. The consultation process

Consultation on the pilot survey took place between November 2003 and January 2004. Submissions were invited from the public and interested bodies in November 2003 on the content of the census questionnaire[1]. Over 80 submissions were received in response to this public consultation. These were from Government Departments, Local Authorities, Government agencies, the research and academic communities and private individuals. Over 25 different topics were covered in the various submissions.

Two meetings of the Census Advisory Group[2] were held in December and January. A sub-group on disability, which was representative of the various interest groups in that area, met to agree on a revised format for the two disability questions. A sub-group representing the Equality Authority, Pavee Point, and the National Council for Racism and Interculturalism met with CSO to agree the layout of the Ethnicity question.

The content of the Census Pilot questionnaire was decided on the basis of the consultations and discussions which took place. It was decided to adhere to the 24-page layout used in the 2002 census. The topics included were those where there was an expressed departmental policy interest or where there was public support. A

---

[1] See notice in Appendix 1.
[2] The composition of the Advisory Group is given in Appendix 2.

number of questions included in the 2002 census were excluded from the pilot questionnaire in order to facilitate the inclusion of the relevant new and revised questions. These were:

- Usual residence one year ago;
- If the person lived outside the country for a continuous period of one year or more, the year of taking up residence in Ireland and the country of last previous residence;
- Religion;
- Business of employer (tick box); and
- ILO suite of questions.

The exclusion of these questions does not compromise the possibility of including them in the 2006 census form.

3. The objectives and design features of the Pilot Survey

The following were the primary objectives of the pilot survey:

Mail back: To assess the use of mail back as an alternative to the traditional enumerator collection procedures and the impact this would have on response rates and data quality. This aspect was included to provide a fall back position for the CSO in the event that it might be unable to recruit a sufficient number of suitable candidates for the enumerator jobs.

Geo-directory: To assess the usefulness of the GeoDirectory (the national address database developed by An Post/Ordnance Survey Ireland) to assist the census field work phase.

New questions:
- The number of children born alive to women
- Ethnic group
- The number of hours of unpaid work looking after the home or family by persons over 15 years and
- The types of voluntary activity undertaken by persons over 15 years.

Changes to existing questions:
- Type of accommodation
- Year house built
- Nature of occupancy
- Type of sewerage facility
- Internet access
- The Irish language question
- The disability question
- Travel to work – time of leaving home and distance travelled in kilometres
- Education – third level subjects studied.

The income question: Half of the households selected were administered a form which included an income question. The forms used in the remaining households excluded the income question. This methodology was used to test the acceptability of the question and the impact of its inclusion on response to the questionnaire as a whole.

The relationship question: In order to meet the growing demand for information on non-nuclear families and the number of second families within the same household, the question used in the pilot survey asked persons 2 to 5 to tick their relationship to all persons previously listed on the form and persons 6 and higher to tick their relationship to persons 1 to 4.

Survey Design

The survey design involved selecting a nationally representative sample of households in which to test the pilot questionnaire and to subsequently analyse responses across a number of response factors. The sample used in the pilot survey consisted of 8,616 dwellings in 32 enumeration areas (EAs) throughout the country. Twenty of the EAs were in urban areas while the remaining twelve were in rural areas. This reflected the urban/rural breakdown of the population from the 2002 census.

The sampling frame from which the sample was selected was constructed from the EAs used in the 2002 census. This provided a sampling frame of 3,975 primary sampling units. The sample was selected using a multi-stage stratified design to ensure as wide a coverage as possible and to minimise the administrative costs by restricting the sampled EAs to selected counties. Prior to the selection of the sample, each primary unit was classified into one of six strata based on level of urbanisation. The composition of the strata and the number of units selected from each are presented in Table 1.

Table 1 Strata used in Pilot Survey

Stratum	Number of EAs	Description of Stratum
1	4	Dublin City
2	4	Cork City and Suburbs
3	4	Suburbs of Dublin contained in the administrative Counties of Fingal, South Dublin and Dún Laoghaire-Rathdown
4	4	Large Towns with a population of 10,000 or more
5	4	Other towns
6	12	Rural areas

For administrative purposes a further level of stratification based on administrative counties was introduced into strata 3-6. Units in Strata 1-2 were sub-stratified by level of deprivation thereby allowing the effect of deprivation on response factors to be assessed.

The first phase of sample selection involved, where relevant, the random selection of an administrative county. In Strata 4-5 a town or towns were then selected at random. The required number of EAs was then selected by taking a random sample within the unit selected.

To facilitate the examination of the response factors relating to method of collection and income (i.e. forms with and without the income question), each EA was divided approximately into four equal parts based on numbers of households. This allowed both of these factors to be balanced across all EAs.

4. Survey Results

Mail back versus traditional collection procedures

There were three aspects to the evaluation of the mail back of census questionnaires:

- overall response rates of households in mail back zones versus enumerator collection zones;
- the difference in data quality (measured as the average number of questions answered per person); and
- the difference in average household size.

In order to test the difference in response rates each EA was split into a mail back zone and an enumerator collection zone based on street and townland codes. In the mail back zones the enumerators marked the forms MAIL before giving them to the householder along with a return envelope marked with the census logo addressed to a local P.O. box no. The householder was instructed to mail the form back as soon as possible after census pilot day.

A two week period was allowed to elapse before the enumerator called to collect any outstanding forms. The co-ordinator date stamped and recorded forms as soon as they were received. Towards the end of the second week all the forms received up to that point were returned to the Enumerator, who marked up the forms received in their Enumerator Record Books (ERSs) before going out to collect outstanding forms. The quality of the forms received by post was not assessed at this point and there was no follow-up on poorly completed forms or forms returned blank.

A total of 8,616 dwellings were identified during enumeration, of which 836 were vacant and a further 236 had no persons present on Pilot Day. The remaining 7,544 households had persons present and this is taken as the starting point for analysing response rates. A response was defined as a form with at least three out of the four key demographic questions (sex, date of birth, usual residence and marital status answered) answered for person 1 on the form. The response rate varied from 63.1 per cent in a mail back area in Louth to 100 percent in a collect area in Wexford.

The overall average response rate was 92.7 per cent, of which 89.7 per cent was in mail back zones and 95.7 per cent was in collection zones (see Table 2). Given that participation in the survey was not compulsory the overall response rate achieved can be viewed as satisfactory.

The data in Table 2 indicate that just over 30 per cent of the 3,815 forms delivered in mail back zones were posted back directly without the need for enumerator intervention. A further 24 per cent required one or more visits from an enumerator while a significant 35 per cent had to be eventually collected by an enumerator.

The post-pilot feedback revealed an extremely negative reaction by householders when the enumerator called to collect the forms not posted. A mitigating factor may have been that the forms had been delivered too early (in some cases four weeks before Pilot Day) and that many householders forgot about completing and mailing the completed form back. Many of the forms had to be replaced after households had mislaid them.

Table 2 Mail back versus collection response rates

	Collection zone	Mail back zone	Total
Forms delivered	3,729	3,815	7,544
Forms returned	3,574	3,423	6,997
Response rate %	95.8%	89.7%	92.7%
Method of return:			
Collected	3,517 (94.3%)	1,335 (35.0%)	4,852 (64.3%)
Posted	57 (1.5%)	2,088 (54.7%)	2,145 (28.4%)
of which:			
posted directly	-	30.5%	-
after one visit	-	16.3%	-
after two or more visits	-	7.8%	-

The main explanation for the 6 per cent differential in response rates between mail back and collection zones was the insistence by householders that they had returned a form, despite the fact that the form was never received.

While is must be acknowledged that the overall response rates would be considerably higher in a live census given the statutory nature of the census itself and given that a high profile public awareness campaign would be undertaken, the differential response rate between mail back and collect observed in the pilot would be a cause for concern if carried through to a full census .

A further negative feature of the mail back approach is the increased complexity and loss of control over the field operation which it would entail. Using the traditional method each enumerator is responsible for ensuring that every household in her EA provides her with completed form(s) in respect of all persons who spent census night in the relevant household. This enables, inter alia, accurate preliminary headcount figures to be provided within a few months of census day. The widespread use of mail back would rule out this option or at least seriously compromise the confidence that can be placed in the current approach.

The overall quality of the forms returned was measured as the ratio of the average number of questions answered per person on the form to the number of questions expected to be answered, taking question filters into account. Using this yardstick there was little difference between mail back (91.2%) and collection (92.8%) - this despite the fact that enumerators are instructed to check forms on collection, to ensure that all relevant questions have been completed. However, given that only 30.5 per cent of householders (the most compliant ones) mailed their forms back without the need for a reminder the closeness of the results is hardly surprising.

Average household size was almost identical in both mail back (2.86) and collection (2.83) areas. This compares with 2.94 in 2002.

In summary, the use of mail back would give rise to some negative side effects – the main one being a fall in response rates. However, on the positive side, mailing census forms back to a central collection point(s) proved to be a viable alternative to the traditional method of using enumerators to collect completed census forms which could be used as a fall-back method in cases where, for example, there might be difficulty in recruiting a sufficient number of enumerators. On the grounds of ensuring as complete a count of the population as possible it is proposed to adhere to the tried and trusted method of getting enumerators to deliver and collect census forms assuming a sufficient number of enumerators can be recruited.

GeoDirectory

Feedback from users of census data suggests that the results would be greatly enhanced if individual households were spatially referenced thereby enabling output areas to be used which are independent of current statutory boundaries such as Electoral Divisions (EDs)[3]. The large variability in the size of EDs mean that they are not ideal from a statistical output perspective. The launch of the An Post/OSI GeoDirectory product, which contains a unique building ID and spatial reference for every postal address in the State, provides a potentially useful source for classifying results at an agreed small area level.

One of the objectives of the pilot survey was to examine how the GeoDirectory building ID code could be used to facilitate the census field work . An initial examination of the Geo-Directory identified two problems - the lack of unique addresses in rural areas and the number of digits used to define the building ID code. A small number of townlands in rural EAs was used to assess the impact of the non-uniqueness of the addresses covered.

In the remaining enumeration areas two methods of capture were tested:

- printing a reduced version of the building ID code on the enumerator maps;
- providing enumerators with address lists containing the ID codes from the GeoDirectory.

In both cases the enumerators were instructed to record the building ID code on the census form from the information provided.

On completion of the pilot survey, debriefing meetings were held with the enumerators. Among the topics discussed was the use of the GeoDirectory in the field operation. The enumerators were asked for their opinions on the practical aspects of using the product and in particular the problems encountered using it in the field. Problems highlighted included:

- In some areas road names on listings were different from those on enumeration maps;

---

[3] Each form is coded to street/townland at present.

- Out of date nature of the GeoDirectory in certain areas;
- Lack of house numbers on street listings, in particular streets in the smaller towns;
- In cases of multi-occupancy buildings only one building code supplied; and
- Difficulties printing ID numbers on maps within the scales used by the census field operation.

Overall it was felt that the GeoDirectory contained flaws, which at the present time prevent it from being of practical use in a full census.

New questions and changes to existing questions

The following table summaries the main points in relation to new questions and changes to existing questions. The relationship matrix and income question are considered separately in later sections of the report.

Table 3 Recommendations on Questions

New questions

Q nr.	Question heading	Consistency of results	Comments at the back of form	Recommendation
Q 9	Number of children born alive A version of this question was asked in 1961, 1971, and 1981, "No of children born alive to the present marriage".	The average number of children born per woman varied in line with expectations. The response rate was a satisfactory 94.3%.	This question caused upset to a small number of persons (10) Examples of some of the comments received were: "Wording is insensitive and insulting; it does not acknowledge stillborn babies." "This question undermines a person's previous life." "You should ask how many children men have fathered." "Question is confusing."	This question is a possibility for inclusion.
Q11	Ethnic Group	The question appeared to have worked reasonably well. The categories Irish and Irish Traveller together accounted for 93.7%. This compares well with the 92.9% who classified themselves as Irish in the nationality question in the 2002 census. 'Any other white background' was the second largest category at 3%. 13 people mistakenly ticked 2 categories instead of availing of the residual category 'other including mixed background'. The level of 'not stated' was very low at 2.1%.	17 comments made in total. These included: "The number of ethnic classifications were too limited, they do not reflect the ethnic mix in Ireland at present." "Terms such as 'white' and 'black' are racist." "Irish travellers are not an ethnic group, it is offensive to them that they are singled out."	A high level of acceptance on the part of the public suggests that this question is a possibility for inclusion.
Q21	Unpaid work in the home	Four out of ten of the women aged 15 years and over who answered this question ticked the 'no' box i.e. they did not do any unpaid work looking after the home or family. Of those who answered the question 62% of women who work for payment or profit stated that they did some unpaid housework compared with 74.7% of women who look after the home/family.	This question caused confusion. Some people were not sure if child minding for others was to be included. Others were not clear if question was intended simply for those who are at home full-time or for everyone. In total there were 29 comments on the question, including:	It is recommended that the question not be asked in the 2006 census given the confusion which the wording seems to have generated.

Q nr.	Question heading	Consistency of results	Comments at the back of form	Recommendation
		There was an overall response rate of 89% to this question with 53% of women and 38% of men indicating that they do some unpaid work in the home.	"Question was badly worded, difficult to understand." "Don't understand the need for this question, it is intrusive. Difficult to estimate how long a person works in their own home." "Should the work a husband does be included?" "Should be asked of under 15s as many of them take care of others and have to do these tasks".	
Q23	Voluntary work	A total of 2,812 persons representing 18% of respondents indicated that they do some voluntary work. Involvement in sport figured highest among males, while social and charitable work was the most popular form of female voluntary activity. Nearly 14% of those who undertook voluntary activities were engaged in more than one activity. The residual category 'any other voluntary activity' accounted for a third of all those undertaking voluntary work.	11 comments were made including: "There should be a no box." "Youth work should have a category of its own."	The drafting of the question did not distinguish between those who did not respond and those who genuinely did no voluntary work. Even though the question is more appropriate to a time use type survey an amended version of the question containing a "No" box is a possibility for inclusion.

Amended housing questions

Q nr.	Question heading	Consistency of results	Comments at the back of form	Recommendation
Q H1	Type of accommodation Duplex moved to terraced category and bed-sits a category on their own.	Results in line with 2002 despite categories changing slightly.	"Difficult to answer." "Question not useful, concentrate on how many people need houses." "What does a purpose built block cover?"	Include new categories
Q H2	Year house built 1996-2000 added.	The category 2001 or later accounted for 6.6% of housing units.	"How do I know the age of the house?"	Include new category
Q H3	Nature of occupancy A new category 'Rented from a voluntary body' was added.	The new category accounted for 2.1% of the total.	"Option 2 could be interpreted that the owner is in arrears to the Bank or Building Society." "Question not relevant." "What is a voluntary body?"	Include new category
Q H8	Type of sewerage New category (individual treatment system other than a septic tank) added.	96.7% of those who ticked the new category were households in rural areas, in line with expectations.	"Can't answer this as we rent the property."	Include new category
Q H11	Internet access Internet connection divided into with and without broadband.	9.2% of households have a broadband connection. 9.2% of householders did not answer the question.	"Will the Government provide PCs for all the households that don't own one? Or connect up those who have no access to the Internet?"	Include new category

Amended individual questions

Q nr.	Question heading	Consistency of results	Comments at the back of form	Recommendation
Q3	Relationship matrix	See separate section in report.		
Q10	Irish Language The question used in the 1996 and 2002 censuses related to ability to speak Irish and frequency of speaking it. The current version covered the competencies: understand, read, write and speak with a tick box for no knowledge of Irish. An additional category was added to the frequency part of the question covering 'daily, within the educational system'.	Nearly 39%[4] of persons aged 3 years and over indicated they could speak Irish compared with 41.9% in 2002. The version of the question used in 2002 was introduced in the 1996 census and gave rise to a discontinuity at that stage. Over a third of persons aged 3 and over indicated that they had no knowledge of Irish. The corresponding fraction for persons of school going age (5-19) was one in six. The results indicate declining ability according to the competencies understand, read, write and speak. This is in line with expectations. The inclusion of the additional category 'Speak within the educational system' resulted in the number of daily speakers increasing from 21.6% in 2002 to 38.7% in 2004. This would appear to be a question effect. Only 1.2% of children of school going age who can speak Irish indicated that they spoke it daily outside the educational system. The revised question shows declining ability to speak Irish compared with 2002. This may be plausible despite the discontinuity in the question.	This question caused by far the most comment among householders. A total of 58 comments were made. "The question does not allow for identifying the level of Irish understood." "There is no box to indicate basic Irish i.e. no box for 'cúpla focal'." "Impossible to answer correctly, I have some ability."	The absence of a category "Some limited knowledge of Irish" in the competencies part of the question would appear to have inflated the numbers ticking the "No knowledge of Irish" box. Distinguishing daily speakers of Irish into those within and outside the education system provides useful additional data. Revising the question so soon after introducing the previous version makes the interpretation of long term trends very difficult.
Qs12 & 13	Disability Three new categories added to the question on long-lasting condition. One new category added to the activities question.	The proportion of persons with disabilities according to the pilot was 10.6% of the population. Without the new categories this falls to 9.2%, up marginally from the 8.3% reported in 2002.	"Question 13 should not be asked of small babies." "Not sure if arthritis/blood pressure qualifies as a chronic illness." "Q12 & 13 not detailed enough." "Saying no to Q12 means I can't answer Q13." "Getting a more accurate picture will do nothing to alleviate the situation for hard working carers."	Despite the complexity of the two questions, the expansion of the categories appears to have worked reasonably well.
Q15	Time of leaving home An extra category 'Before 6:30' was added.	The new category was ticked by 6% of those who travelled to work. Proportionately more people are travelling before 7am, in line with expected trends.	"No account is taken of shift work or home based employment with irregular travel patterns." 'Not applicable' is not the right answer in these cases. "Times are different when the schools are closed." "Does this question cover mothers dropping children to school."	Include new category.

---

[4] A total of 642 persons did not tick any of the competencies boxes in the first part of the Irish question but indicated that they spoke Irish either daily, weekly or less frequently. A further 2015 persons did not tick the 'speak' box but ticked at least one of the other competencies boxes (i.e. understand, read, write) and also indicated that they spoke Irish either daily, weekly or less frequently. Both categories are added to the 5301 persons who indicated that they spoke Irish to make up the total of 7352 which is used to compute the percentage of Irish speakers.

Q nr.	Question heading	Consistency of results	Comments at the back of form	Recommendation
Q16	Distance travelled Miles changed to kilometres.	This was asked in Kilometres for the first time. 'Not stated' was 23% compared with 12.8% in 2002. If we compare the distribution of the distances travelled for the 32 pilot EA's in 2002, the results were broadly comparable.	"Was difficult to answer as I work in different places." "Van driver no fixed mileage." "Sales rep, distance travelled varies". "There should be a not applicable option." "Should ask how long a journey takes and ask homeward journeys too." "Using minutes is a little confusing." "Cars still use mileage and older people are used to it, so why use kilometres."	Change to kilometres. The question of whether we include the conversion (i.e. 1 mile = 1.6km approx.) is an open one as it gives rise to an increase in the number of decimal points on the Census form.
Q20	Education – third level qualifications Some categories were collapsed while one category was separated out (Social Services).	Results were consistent with 2002 census.	"Both questions 19 and 20 are confusing, do both parts have to be answered?" "Question 20 seems to be repeating question 19." "Question 20 does not make sense. IT is missing." "Question does not cater for those over 60 who went through a different education system."	Include revised question.
Q32	Income	See separate section in report		

The relationship question

In the 2002 census all persons present in the household on Census Night were asked to identify their relationship to Person 1. Those temporarily absent on Census Night were also asked to respond to this question. The purpose of the question was to distinguish and code family units within a household. While this has been adequate to cater for the majority of standard family types it fails to take account of all second families within the one household e.g. a daughter living with her own child in her parents' house.

There is a growing demand for information on non-nuclear families and the numbers of second families within the one household, as identified during the consultation process. To enable this more complex family coding to be done, every person should ideally be asked his or her relationship to every other person on the form. Unfortunately space constraints do not allow this. Instead, under the new format of the question tested every person up to person 5 was asked to identify their relationships to every person preceding them on the form. For households consisting of 6 or more persons, persons 6 and higher were asked to identify their relationship to persons 1-4 only, for space reasons. The question layout used in the pilot survey is given on the next page.

The types of relationships listed were selected to reflect the most common relationships identified in 2002. The relationship write-in text box used in 2002 was removed. An example of how the question should be answered was put on the back of the form to facilitate respondents. Furthermore, the enumerators were instructed to

familiarise themselves with the question in order to assist householders in answering it.

The evaluation of the question was based on:

- Overall response to the question;
- The number of second families coded within the same household;
- The percentage of grandchildren coded as other than 'other relative'; and
- An analysis of the number of households that answered the question correctly.

An analysis of the 'Relationship to Person 1' returned in the survey showed that the breakdown of these relationships was similar to 2002 and that there was no significant difference in the overall response to the question as measured by the amount of relationships not stated (i.e. left blank)[5]. This indicates that the more complex format of the question did not interfere with the basic relationship data as collected in previous censuses. Table 4 compares the data drawn from the 2004 pilot survey with the corresponding data from the 2002 census.

The criteria for coding family nucleus, and crucially identifying whether a particular household can be coded automatically or not, were the same as those used in Census 2002 apart from the following distinction. All households containing a 'grandchild' were manually coded in the pilot, and where the additional relationship data allowed, a second family unit was formed. The overall profile of the family nucleus codes assigned therefore show some significant differences, as expected, as more second families were identified. Less than ½ per cent of persons were coded to second families in 2002 compared with almost 1½ per cent in the pilot. This would imply we could expect to get over three times as many persons coded to second families using the new relationship matrix question[6]. In the case of grandchildren just under 30 per

---

[5] The non-response to this question by persons other than Person 1, who is not required to answer the question, was just under 1% but comparable data for 2002 is not available.
[6] In 2002 there were 5714 households containing two or more family units.

cent of them had their parent or parents identified in 2002 and coded accordingly whereas in the pilot this figure rises to nearly 75 per cent.

Table 4 Relationship to Person 1

Relationship to Person 1	2002 % of persons	2004 % of persons
Not applicable – Person 1	33.7	34.7*
Husband or wife	17.9	17.8
Partner	2.1	2.0
Son or daughter	37.9	38.3
Mother or father	1.2	1.2
Sister or brother	1.5	1.2
Son-in-law or daughter-in-law	0.2	0.2
Grandchild	1.1	1.2
All other related	1.0	1.1
Not related	3.5	2.3

* 'Not stated' accounts for 0.9% of this figure. The complete figure for 2002
  is not available.

Feedback from the field force indicated that there was a mixed reaction to this new version of the relationship question. Some enumerators reported difficulties with householders needing a lot of assistance in completing the question, especially in deprived areas and in larger households, whereas in other areas it was felt the question was well understood and there was no major difficulty. All households were examined to see if they answered the matrix question correctly or not. A household was deemed to have answered the question incorrectly if any part of the question was answered wrongly (as indicated by the other relationships) or if any part of the question was left blank.

Overall 81 per cent of all households answered all parts of the question correctly though this falls to below 60 per cent in households with 5 or more persons. However the majority of these 'incorrect' answers are due to the householder failing to answer a part of the question. Households were also examined to see if the new format of the question had provided additional useful information that would enable more correct family codes to be assigned and this was found to be the case in over 60 per cent of households with 3 or more persons. The completion rates for separate relationships were quite high; over 80 per cent of persons 3 and higher gave their relationship to person 2 and similar results were obtained for relationships to person 3 and person 4.

The conclusion drawn from the survey results was that the new format of the question did not impact on the overall response rate to the question, nor did it impact on the more basic relationship data collected previously. It did however provide useful additional data in households where it is needed that will enable more accurate and complex family coding to be done in the future. On balance it is felt that the extra

burden placed on respondents is outweighed by the additional policy relevant data which will be made available[7].

The income question

In order to assess the impact of an income question on response rates and average household size, half of all households received a form with the income question and half without. Every enumeration area was split into an 'income' and 'without income' zone, based on street and townland codes. The feedback from the enumerators indicated that this question received a mixed reaction on the doorstep with the wealthier and the poorer households thought to be the most distrustful.

The overall response rate from households that received a form without the income question was 93.5 per cent, only marginally higher than from those that received an 'income' form (92%), indicating that the presence of an income question alone would not appear to affect response. However an income question combined with other factors does appear to have had an impact: income forms in mail back areas had a response rate of 88.8 per cent, while income forms in deprived mail back areas had a response rate of 85.3 per cent. Some enumerators did report households refusing to complete the census questionnaire at all because of the presence of the income question.

There was no difference in average household size between those who received the income form and those who did not.

The non-response to the income question itself varied from 38.2 per cent in deprived areas to 27.6 per cent in non-deprived areas, but varied very little between urban and rural, or mail back and collect areas. When examined by present principal status 90 per cent of those who indicated they were working for payment or profit answered the income question whereas less than half of students and only two thirds of homemakers answered it.

The distribution of incomes for those who did answer the question followed expected patterns with more persons in the higher social classes having a higher income bracket and vice versa. Less than 3 per cent of persons in deprived areas reported an income in excess of €40,000 compared with nearly 9 per cent in non-deprived areas.

The main purpose of the inclusion of an income question is to provide information on the location of deprived households throughout the State. The high non-response rate to the question in deprived areas (38%) is therefore a cause for some concern. Given that there appears to be agreement among users that the underlying dimensions of social disadvantage can be accurately modelled, both spatially and over time, using other census variables (e.g. demographic decline, labour market deprivation and social class deprivation)[8] the need for a direct question on income may not now be as acute as it was formerly. In addition, the availability of comprehensive information on income, albeit at a regional level only, from the EU-SILC coupled with the likely

---

[7] Form design and scanability implications of the revised layout have yet to be fully assessed.
[8] See Deprivation and its Spatial Articulation in the Republic of Ireland, Trutz Haase and Jonathan Pratschke, 2004.

inclusion of an income measure in the QNHS will also reduce the information deficit in this area to a significant extent.

The single biggest issue however is the danger of a negative impact on the public response to the census if a question on income were included. While attitudes in relation to the inclusion of income questions are softening it is nonetheless considered that the time is not yet ripe to recommend that an income question should be included in the 2006 census form.

5. Summary of recommendations made

- The use of mail back would give rise to some negative side effects – the main one being a fall in response rates. However, on the positive side, mailing census forms back to a central collection point(s) proved to be a viable alternative to the traditional method of using enumerators to collect completed census forms which could be used as a fall-back method just in case there is difficulty in recruiting a sufficient number of enumerators. On the grounds of ensuring as complete a count of the population as possible it is proposed to adhere to the tried and trusted method of getting enumerators to deliver and collect census forms assuming a sufficient number of enumerators can be recruited.

- The lack of unique addressing in rural areas and the limitations in terms of up to dateness in certain urban areas imply that the national address database "GeoDirectory" is not a suitable vehicle on which to base the census enumeration. However, it is an invaluable tool in helping to highlight areas of major housing development for census planning purposes. It will also assist the coding of business addresses at the processing stage.

- Four new questions were asked: number of children born alive, ethnic group, unpaid work in the home and voluntary work. The first two questions were seen to work well and given their importance are candidates for inclusion in the 2006 census form.
- The question wording used for the question on unpaid work in the home proved difficult for respondents to understand. It is therefore proposed that it should not be included.

- The drafting of the question on voluntary work was not ideal (i.e. no "No voluntary work" box) and would probably be more suited to a time use survey rather than a census.

- All of the proposed amendments to the housing questions worked well.

- The revised Irish language question yielded more useful information for policy formulation than its predecessor. However, revising the question so soon after introducing the previous version would make the interpretation of long term trends very difficult.

- The expansions to the disability questions give a better basis for the introduction of a specific disability survey based on census returns than the former versions of the questions.

- Both the expansion of categories in the question on time of leaving home and the substitution of kilometres for miles worked well, although in the latter case there was an increase in the number of respondents who did not answer the question.

- The streamlining of the categories in the question on third level qualifications was also successful.

- A major revision to the relationship question was tested to enable different types of families to be distinguished and to allow multi-family households to be identified better than heretofore. The revised version of the question did not impact negatively on the more basic relationship data collected previously i.e. relationship to person 1. The revised question will allow more accurate and complex family coding to be done in the future. Given the additional policy relevant data which will flow from the data it is proposed to adopt it for the 2006 census.

- A question in relation to current gross income was tested for all persons over 15 years. The non-response to the question was 27.6 per cent in non-deprived areas compared with 38.2 per cent in deprived areas. The high non-response, especially in deprived areas where the results of the question are most needed, would compromise the accuracy of the inferences which could be drawn. The acceptance of more easily understood proxy measures of deprivation, mainly drawn from census sources, lessens the need to have an income question included. If an income question were to be included in the census form it would undoubtedly give rise to negative publicity for the census as a whole. This would result in greater difficulties in ensuring an overall comprehensive census count. It is therefore proposed that the income question not be included.

Appendix 1 Notice of Public Consultation

2006 CENSUS OF POPULATION
CONSULTATION ON
QUESTIONNAIRE CONTENT

The CSO invites submissions on the questionnaire content for the 2006 Census of Population and on the outputs to be produced.

Proposals for new questions or changes to existing questions should include

- a detailed justification for each proposal

- a suggested question wording, with emphasis on ease of completion and on avoiding ambiguity in interpretation

- a summary, where relevant, of corresponding analyses required.

In considering submissions, the following will be taken into account

- the need to ask core questions in a census

- data already available from existing sources

- practicality (particularly the need to keep questions simple and unambiguous)

- cost and the burden of response on householders.

The 2002 Census questionnaire is available on the CSO website ***http://www.cso.ie*** or in the published 2002 Census reports.

CSO plans to conduct a census pilot survey in selected areas in April 2004. The survey will test the census enumeration procedures and the proposed content of the questionnaire. The final content of the 2006 Census questionnaire will be subject to Government approval.

Please send submissions by email, FAX or post to:

Patricia Thornberry	**email pilot@cso.ie**
Central Statistics Office	**FAX (01) 497 2360**
Ardee Road	**LoCall 1890 313 414**
Rathmines	**(Extns. 4251; 4281)**
Dublin 6	

by Friday 28 November 2003 at the latest.

Appendix 2 Members of the Advisory Group (apart from CSO representatives)

Government Departments	Nominated Representative
An Taoiseach	Gerry Cribbin
Finance	Barra Ó'Murchadha
Community, Rural and Gaeltacht Affairs	Denise Grogan
Education and Science	Mary Dunne
Enterprise, Trade & Employment	Marie Dempsey
Environment and Local Government	Norita Griffin
Health and Children	Hugh Magee
Social and Family Affairs	Paul Morrin
Transport	Edel O'Dea Kenny

Other bodies

ESRI	Prof. Gerry Hughes
Equality Authority	Laurence Bond
IBEC	David Croughan
ICTU	Paula Carey
Regional Authorities Directors' Association	James Stone
Regional Planning Guidelines Review (GDA)	Mary Darley
Cork City Council, Community and Enterprise	Pat Ledwidge
Dublin Transportation Office	Owen Shinkwin

Universities

UCD	Prof. Pat Clancy
NUI Maynooth	Prof. James Walsh