Standard Report
on
Methods and Quality (v1)

For
Census of Population

This documentation applies to Census 2006

Last edited: 19 March 2009

CENTRAL STATISTICS OFFICE

Swords Business Campus, Balheary Road, Swords, Co Dublin.

Locall 1890 236787

www.cso.ie

Contents

Appendices

1   Overview

The first major census, using a household form, was the so-called Great Census of 1841. In recent times, a Census of Population of Ireland is held every 5 years. The most recent census was taken on Sunday 23 April 2006 in accordance with the Statistics (Census of Population) Order 2005 (S.I. No 786 of 2005).

The census, which is the largest statistical operation organised by the CSO, is a compulsory statistical enquiry. The Census Order specifies the information to be provided (i.e. the responses to the census questions) and the classes of persons who are legally obliged to provide this information (mainly householders and persons in charge of communal establishments).

For the 2006 Census, a team of approximately 4,800 temporary field staff were recruited and trained to carry out the census field operation. The team comprised 6 Liaison Officers, 40 Regional Supervisors, 400 Field Supervisors and 4,400 Census Enumerators.

The census dissemination program started on 19 July 2008 with the publication of the Preliminary results. The main schedule of final results ran from 29 March 2007 to 29 November 2007. This included two summary reports providing the main census results, primarily at national level, a further 13 reports providing more detailed data on individual subject matters and Small Area Population Statistics (SAPS).

The results of the Census are used at national level for planning the provision of health care, education, employment, etc. Regional and county figures are needed for developing regional policies and for the operation of regional and local authorities. The detailed population figures at local level are used to identify the need for schools and hospitals, areas of relatively high unemployment, the best location for new shops, etc.

The census is also the only means of accurately measuring the exact extent of migration and it is needed to carry out constituency reviews.

2   General Information

2.1   Statistical Category

Primary statistical survey

2.2   Area of Activity

Census of population and information on households.

2.3   Organisational Unit Responsible, Persons to Contact

Census of Population Division

Aidan Punch, Director	Tel: +353 1 4984316	Email: aidan.punch@cso.ie
Gerry Walker, Senior Statistician	Tel: +353 1 4981343	Email: gerry.walker@cso.ie
Deirdre Cullen, Senior Statistician,	Tel: +353 1 4981334	Email: deirdre.cullen@cso.ie
Noel Walsh, Principal Officer	Tel: +353 1 4981331	Email: noel.walsh@cso.ie
Brian King, Statistician	Tel: +353 1 4981335	Email: brian.king@cso.ie
Declan Smyth, Statistician	Tel: +353 1 4981345	Email: declan.smyth@cso.ie
Shaun McLaughlin, Statistician	Tel: +353 1 4981474	Email: shaun.mclaughlin@cso.ie
Dermot Corcoran, Statistician	Tel: +353 1 4981316	Email: dermot.corcoran@cso.ie
Cormac Halpin, Statistician	Tel: +353 1 4981355	Email: cormac.halpin@cso.ie
Evelyn Cronin, AP	Tel: +353 1 4981319	Email: evelyn.cronin @cso.ie
Pat Flynn, AP	Tel: +353 1 4981322	Email: pat.flynn@cso.ie
Mary Duffy, AP	Tel: +353 1 4981343	Email: mary.duffy@cso.ie
John Flanagan	Tel: +353 1 4981302	Email:john.flanagan@cso.ie
Enquiries	LOCALL 1890 236787	Email census@cso.ie

2.4   Objectives and Purpose; History

The purpose of the census is to accurately measure the population in Ireland, to provide information on the key demographic and socio economic characteristics of the population and on the housing stock in the country. The census involves the collection of information relating to persons and households, the capture and analysis of the data and the dissemination of a wide range of information through specialised publications, releases, web-tools and other media so that it is available for effective policy, planning and decision making purposes.

At national level current population statistics are essential for planning the provision of health care, education, employment, etc. Regional and county figures are critical for determining regional policy and for the operation of regional and local authorities. The greatest strength of the census is the provision of detailed population figures at local level. These help to identify likely demand for schools and hospitals, areas of relatively high unemployment, the best location for new shops, etc.

The census is also the only means of accurately measuring the exact extent of migration. By comparing the results of successive censuses, and taking account of the number of births and deaths that have occurred over the intervening period, an accurate measure of net migration (the difference between inward and outward migration) can be derived.

Furthermore, census data is required in order to carry out constituency reviews. These reviews take place once the results of the census have been published. Article 16.2 of the Constitution (Bunreacht na hÉireann) lays down that the total membership of Dáil Eireann depends on the population as measured by the census (i.e. 1 TD per 20,000 to 30,000 persons).

The following is a brief history of censuses in Ireland:

1841 to 1911: The first major census, using a household form, was the so-called Great Census of 1841. Following 1841, censuses were subsequently taken at 10-year intervals up to 1911.

1926 to 1946: The first census following the formation of the State was undertaken in 1926. This was followed by censuses in 1936 and 1946.

1951 to 2006: Commencing with 1951, censuses have been undertaken at five yearly intervals up to and including the most recent census in 2006. There were two 'interruptions' to this pattern:

- The census planned for **1976** was cancelled at a late stage as a Government economy measure. However, the need for up-to-date population figures resulted in a census being undertaken in **1979** with a restricted number of questions. The usual five yearly pattern resumed with the 1981 Census.
- The census due to take place in April **2001** was postponed until **2002** due to the threat of foot and mouth disease at that time.

2.5   Periodicity

The census is held at five year intervals.

2.6   Client

The principal client is the Government.

2.7   Users

There is a wide range of users of census data, including
- Central and local government
- International bodies such as Eurostat, UN, OECD
- Individuals, groups and organisations involved in planning at a local, regional and national level.
- Individuals and organisations involved in education and research.
- The Constituency Commission, which uses the results to re-draw constituency boundaries as laid down in Article 16.2 of the Constitution (Bunreacht na hÉireann).
- The public at large.

2.8   Legal basis

The legal basis for taking Census 2006 was the Statistics (Census of Population) Order, 2005 (S.I. number 786 of 2005) which was made under 25(1) of the Statistics Act 1993 (No. 21 of 1993).

3   Statistical Concepts, Methods

3.1   Subject of the Statistics

At the level of the individual person, the census asks questions on a variety of demographic and social topics such as age, sex, marital status, residence now and one year previously, nationality, religion and ethnicity. A number of questions on commuting are included along with questions on important socio-economic variables such as education, employment status, industry and occupation.

At household level, the questions asked relate to housing characteristics of private households such as accommodation type, year built, nature of occupancy, type of dwelling, heating, water, sewerage facilities and P.C. and internet access.

3.2   Units of Observation/Collection Units/Units of Presentation

The target collection is all persons present in the country on census night, in both private households and communal establishments. The collection units are households containing at least one person and individuals in communal establishments. A household is defined as a person living alone or a group of people who live together in the same dwelling with common housekeeping arrangements. Most of the key analyses relate to individuals; analyses are also produced on household composition and families; and on housing type, tenure, etc for occupied dwellings.

3.3   Data Sources

The source of data in the census is the returned completed census forms.

In exceptional cases, an imputed form will be constructed by the enumerator on behalf of a household where contact could not be made but reliable data could be gleaned from neighbours. In such cases, key demographic data only are recorded. In addition, in the case of completed census forms where some of the data is missing the relevant missing data is imputed for date of birth, marital status and principal economic status, using distributions from the previous census.

3.4   Reporting Unit/Respondents

The primary reporting unit is private households. The householder or any adult member of a private household is responsible for providing the information on the household and the individuals in the household that are present on census night and those temporarily absent. All of the data on individuals in private households, and the housing characteristics of private households are collected on a household census form. Data on individuals in communal establishments is collected on individual census forms; data on type of communal establishments is collected on listing forms, and data on vacant dwellings is collected by the enumerator in the field.

3.5   Type of Survey/Process

Complete enumeration involving the delivery and collection of census forms by enumerators.

3.6   Characteristics of the Sample/Process

3.6.1   Population and Sampling Frame

Sampling does not apply to the census. The census targets every household in the state. Enumerators are instructed to identify every dwelling with a possibility of habitation in their area, to call to all dwellings and to deliver a census form to all occupied dwellings.

The Census figures relate to the *de facto* population, i.e. the population recorded for each area represents the total number of persons present within its boundaries on the night of Sunday 23rd April, 2006, together with all persons who arrived in that area the following morning, not having been enumerated elsewhere. Persons on board ships in port are included with the population of adjacent areas. The figures therefore

include visitors present on census night as well as those in residence. Usual residents temporarily absent from the area are excluded from the census count.

A limited number of questions were asked in respect of usual residents who were temporarily absent from their households on census night.

3.6.2   Sampling Design

Not applicable.

3.7   Survey Technique/Data Transfer

A temporary field force of about 4,800 persons, of which 4,400 were enumerators, was recruited for Census 2006. The entire country was designated into manageable enumerator workloads of between 250 and 400 dwellings. Each enumerator was provided with a map, identifying all known dwellings in their area. The enumerators' job was to identify every dwelling capable of habitation in the assigned area, whether on the map or not, to establish if the dwelling would be occupied on census night, to make personal contact with the householder in every dwelling that was occupied, to provide a census form to every occupied household in her area and to collect the completed form(s) after census day.

The enumerator also recorded all calls to all dwellings in a record book. In cases where the complete household was elsewhere in the State on census night, the enumerator initiated a tracking procedure to establish where the occupants spent census night and to ensure they were enumerated at that address.

The enumerator, with the help of her supervisor, where necessary, also made contact with the manager in every communal establishment in her area to agree arrangements for the enumeration of everyone who spent Census Night on the premises. She provided the manager with individual forms to be given by the manager to each resident for completion on the night and collected the completed forms the following morning.

Census forms were completed by householders on census night. The day immediately following census day the enumerators began collecting the forms. They called to communal establishment first, particularly hotels and guesthouses. Over the following 4 weeks they called to every dwelling to which they had delivered census forms to collect the forms. Several calls to a dwelling may have been necessary before contact could be made with the householder. In the event that ultimately no contact could be made a reminder form was left. The reference number of the household, the regional office address, household address, the date and the householder's name were written on the form along with a replacement census form. The field supervisor was also informed.

The collected census forms were summarised in the field by the enumerators who recorded the number of males, females and total persons on the front of every form. These headcount details were used to compile the preliminary population count. The forms were sorted by number and placed in special archive quality boxes. The forms were returned to the field supervisor who checked that all was in order, before they were collected and returned to the census warehouse facility for processing. All boxes of forms were registered on return and missing boxes were identified immediately and an enquiry was sent back out to the field for follow-up.

3.8   Questionnaire (including explanations)

The census questionnaire covers a wide range of demographic, social, and economic topics. While the core questions on the questionnaire remain constant from census to census the full content of the questionnaire, including any new questions, is arrived at following public consultation and subsequent pilot testing in the field. The Government decides the content of the census questionnaire in advance of every census.

In 2006 the following questions were asked: name, sex, relationship to others in the household, date of birth, place of birth, usual residence, usual residence one year ago, ever lived abroad, marital status, number of children born alive, ability to speak the Irish language, religion, ethnicity, disability, means of travel to work, school or college, time of travel, distance travelled , time taken, provision of caring, age education ceased, level of education, third level qualifications, participation in voluntary activities, economic status, employment status, occupation, industry and name and address of place of work.

The full text of the questions in Census 2006 and explanations are in Appendix 1.

3.9   Participation in the Survey

The census is a statutory survey and under Section 26 of the Statistics Act, 1993, participation is compulsory. It was clear from previous censuses, the most recent being in 2002, and from other household surveys such as the Quarterly National Household Survey, that conducting field operations was undoubtedly becoming more difficult. CSO therefore had to plan to mitigate these difficulties. The first decision taken was to stick to the tried and trusted model of using enumerators to deliver and collect census forms.

As part of their duties the enumerators had to list in their record books every building capable of being occupied on census night and mark the building on an up to date map of their enumeration area (EA). They then had to make contact with householders over a nine week period in April/May, first to distribute blank census forms and then to collect the completed forms. Bearing in mind the increased complexity of Irish society with more people at work and greater mobility extensive use was made of calling cards in cases where contact with householders was not easy to achieve. Each of the enumerators was issued with a mobile phone and the phone numbers were entered on the calling cards to facilitate texting or phoning the best time for the enumerator to get the householder at home. Where contact proved very difficult forms could to be mailed back to CSO (there were 16,000 mailed back forms).

The enumerator's job was to make personal contact even if this entailed multiple visits. Management companies were contacted for access codes in the case of gated communities of private apartments – a growing phenomenon as indicated by other CSO surveys.

It is strongly contended that using enumerators to establish personal contact with householders is the best method of ensuring a complete count. Alternative methodologies such as mailing out census forms and/or mailing back completed forms or transmitting them over the internet will not achieve the same coverage.

The fact that enumerators have responsibility for their own EAs is a powerful motivating factor. They have to account for every listed dwelling in their EA. When a household is temporarily absent on census night the householder is obliged to indicate where the household will be and if the address is elsewhere in the State the relevant enumerator checks to see that the household is enumerated there. Otherwise they will be approached later to fill out their census form at their original address.

An intensive public awareness campaign on TV, radio, newspaper and outdoor was conducted during the month leading up to Census Day. The advertising commenced at the launch of the fieldwork on 29 March and reached a peak on Census Day. Its purpose was to encourage and remind people to fill out their census forms that night and "Be part of the bigger picture". The post Census Day advertising was to remind the public to have their completed census forms ready for collection by their enumerators. Information about the Census was made available in 16 foreign languages and extensive use was also made of translations of census forms with a view to facilitating the growing number of non-Irish nationals living in Ireland.

3.10   Characteristics of the Survey/Process and its Results

3.11   Classifications used

The following is a summary of the main classifications used:

Topic	Classification used
Type of accommodation	As identified in the questions
Year built	As identified in the questions
Nature of occupancy	As identified in the questions
Rent paid	8 categories beginning with 'under 25 euro' and then going up in brackets of fifty euro to a final category of '300 euro or over'.
Number of rooms	11 categories in all, going up on ones, and finishing with 10 rooms or more.
Central heating	Yes or no
Water supply	As identified in the questions

Topic	Classification used
Sewerage	As identified in the questions
Number of cars	5 categories stopping at four or more, and including none
Personal computer	Yes or no
Internet	Yes, split into broadband or other, and no connection.
Sex	Male or female
Relationship	11 separate categories distinguished ranging from husband / wife, to other related, and including unrelated. Used in assigning family codes for outputs.
Family unit	A family unit *or nucleus* is defined as: (1) a husband and wife or a cohabiting couple; or (2) a husband and wife or a cohabiting couple together with one or more usually resident never-married children (of any age); or (3) one parent together with one or more usually resident never-married children (of any age). Family members have to be usual residents of the relevant household. The determination of household and family composition is based on responses to the question on the census form dealing with relationships within the household. In the 2002 and previous censuses the question referred to the relationship of all persons in the household to the reference person in the household. The question used in the 2006 census took account of relationships to the first four persons in the household. This version of the question has facilitated the coding of more complex and atypical family types than heretofore. In particular it allows a more accurate assessment of multi-family households. Six categories are identified covering couples with and without children, married and not married, and lone mother and lone fathers.
Family cycle	Families are classified by age of youngest child, age of eldest child, number of children in the family from 0 to 7 or more, number of persons aged under 5 and under 15, number of persons aged 65 and over, age group of the reference person, families comprising husbands and wives by age of the husband and age of wife in relation to the husband. The following classification is used for family cycle: Pre-family: Family nucleus of husband and wife or cohabiting couple where wife is under 45 years; Empty-nest: Family nucleus of husband and wife or cohabiting couple where wife is aged between 45 and 64 years; Retired: Family nucleus of husband and wife or cohabiting couple where wife is aged 65 years and over; Pre-school: Family nucleus where oldest child is aged 0-4 years; Early-school: Family nucleus where oldest child is aged 5-9 years; Pre-adolescent: Family nucleus where oldest child is aged 10-14 years; Adolescent: Family nucleus where oldest child is aged 15-19 years; Adult: Family nucleus where oldest child is aged 20 years and over.
Elderly population living alone	Persons Living Alone in private households are classified by sex and age group. The age groups identified are 65 and over, 70 and over and 75 and over.

Topic	Classification used
Reference Person	The term "Reference Person" is used in place of "Head of Household", which was used in the 1996 and previous censuses. The reference person in each private household is the first usual resident in the household. This is the person with the lowest Person Number. Where there were no usual residents present in the household on census night, the first absent person was used as the reference person. The private households and number of persons in private households are classified by the age group of the reference person, distinguishing 10 age groups from under 25 years in 5-year blocks to 65 years and over.
Household composition	Thirteen categories covering single person, couples, married and not married, with and without children, with and without other persons, multiple family households and finishing with households containing unrelated persons.
Date of birth / age	Various classifications ranging from single year of age, five year age groups, fifteen year age groups and other non-standard groupings.
Place of birth	Classified by Province, County and Country (if outside Ireland).
Nationality	Input classification is based on the Eurostat Geo-nomenclature of countries. Output classification covers Irish, dual-Irish, British, EU15, EU 15 to 25, other European, American, African, Asian, other nationalities. Within these categories different individual countries are separated out depending on their importance in terms of numbers for example Polish, Nigerian.
Usual residence	A wide range of classifications covering aggregate urban rural, Regional Authority, County, Country, Towns, Dail Constituency, Electoral division, Gaeltacht areas, local electoral areas, census enumeration areas (for the cities and their suburbs only), Garda geographical areas, and Health Service Executive areas.
Usual residence one year ago	Classified by same address, same county, and other county or outside the state. Place of previous residence classified by Province, County and Country (if outside Ireland). Countries identified cover Northern Ireland, Scotland, England and Wales, the (then) remaining 23 members of the EU, the united states and 'other countries'.
Ever lived abroad	Input classification is based on the Eurostat Geo-nomenclature of countries. The output classification for country of previous residence is UK, other EU countries, USA and other countries. Total who ever lived abroad for a year or more are shown as a percentage of the total usually resident.
Year of taking up residence	The following 8 classifications are used: Before 1951, 1951-1960, 1961-1970, 1971 to 1980, 1981-1990, 1991-2000, 2001-2006 and 'not stated'.
Marital status	As identified in the question
Ability to speak Irish	Persons aged 3 years and over are classified according to whether they can speak Irish or not. The number of persons who can speak Irish are further sub-divided into the following five groups on the basis of frequency i.e. Daily, Weekly, Less Often, Never and Not Stated.
Religion	Input classification of 34 categories. Output classification of 27 categories which changes according to the results of the question i.e. predominant religions are separately identified.
Ethnicity	Input classification used the Geo-nomenclature of countries allow for identification with country of birth. It also included important ethnic groups that are not readily identified with a singe country like Kurdish, Muslim, and Gypsy. Output classification is based on the categories identified in the question.

Topic	Classification used
Number of children born alive	Age of mother in 5 year age groups and number of children ranging from 0 to 9 or more.
Disability	Individuals are classified as having a disability if they answered YES to <u>any</u> part of the two questions on disability, including in particular if they ticked YES to any of the parts of Q16 even though they may not have ticked YES to any of the parts of Q15.
	The published classifications are based on the categories identified in the questions.
Means of travel and time of leaving home	Based on the categories identified in the questions
Kilometres and minutes	Distances ranges from 1 km up to 50 or more and Length of time goes up in 15 minute segments to one and half hours and over
Carers	Based on the categories identified in the questions
Education ceased	'Yes' or 'no' with age education ceased classified by under 15 years, single year of age from 15 to 24 years and 25 and over.
Level of education	As identified in the questions with some collapsing of categories.
Third level qualifications	As identified in the questions
Voluntary activities	As identified in the questions
Principal economic status	Nine classifications identify the principal economic status of all persons aged 15 and over. These are At work, Looking for first regular job, Unemployed, Student, Looking after home or family, Retired, Unable to work due to permanent sickness or disability and Other.
Employment status	Input classification of employee, self-employed, and assisting relative are use to determine social class.
Industry	2006 Census codes industry using NACE the General Industrial Classification of Economic Activities within the European Communities. NACE Rev. 1 is a 4-digit activity classification that was drawn up in 1990 and is a revision of the version originally published by Eurostat in 1970.
	NACE is a hierarchical classification, with 60 2-digit codes, 222 codes at 3-digit level and 503 at 4-digit level. The most detailed level of industry distinguished in the present report is given in Appendix 2 while the intermediate level of industrial groups is given in Appendix 3 and the broad level is in Appendix 4.
	All persons aged 15 years and over who are at work are classified according to the branch of economic activity of the local unit in which the person is working.
Occupation	The occupation classification used in the census, is based on the UK Standard Occupational Classification, with modifications to reflect Irish labour market conditions. The classification was first used in the 1996 Census. This classification adheres to the international occupation classification ISCO Com (88) and provides a link to ISCO Com (88). The hierarchical structure provided within this classification is suited to accurate automatic coding.
	The code to which a person's occupation is classified is determined by the kind of work he or she performs in earning a living, irrespective of the place in which, or the purpose for which, it is performed. The detailed classification as used in Census 2006 is set out in Appendix 5, while the Intermediate level is set out in Appendix 6 and the Broad level of classification is in Appendix 7.

Topic	Classification used

Socio-economic Group	The socio-economic group (SEG) of all persons aged 15 years and over who are at work is determined by their occupation and additionally in some cases by their employment status.
	Unemployed or retired persons are classified by socio-economic group according to their former occupation.
	All other persons are classified according to the SEG of the person on whom they are deemed to be dependent. The codes are as follows:
	    A = Employers and managers
	    B = Higher professional
	    C = Lower professional
	    D = Non-manual
	    E = Manual skilled
	    F = Semi-skilled
	    G = Unskilled
	    H = Own account workers
	    I = Farmers
	    J = Agricultural workers
	    Z = All others gainfully occupied and unknown
	Appendix 8 lists the occupations included in each SEG.

Social Class	The social class of all persons aged 15 years and over is determined in a similar way to the socio-economic group. In contrast to socio-economic group, however, social class is based on an ordinal scale with 1' being the highest social class. The social class of family dependants is derived from the social class of the parent having the highest social class. The codes are as follows: 1 = Professional workers 2 = Managerial and technical 3 = Non-manual 4 = Skilled manual 5 = Semi-skilled 6 = Unskilled 7 = All others gainfully occupied See Appendix 5 for more detailed information...
Type of household	Private and non-private, with private spilt into permanent and temporary. A *private household* comprises either one person living alone or a group of people (not necessarily related) living at the same address with common housekeeping arrangements - that is, sharing at least one meal a day or sharing a living room or sitting room. A *permanent private household* is a private household occupying a permanent dwelling such as a house, flat, apartment or bed-sitter. A *temporary private household* is a private household occupying a caravan, mobile home or other temporary dwelling and includes travelling people and homeless persons living rough on census night.
Size of household	The number of persons in a private household consists of the total number of persons usually resident there on the night of Sunday, 23 April 2006, including those absent from the household for less than three months. Visitors present in the household on census night are excluded. Classifications range from 1 person households to households with 12 or more persons

3.12   Regional Breakdown of Results

The following Geography classifications are used:

- Province (4),
- Counties (34) ,
- Towns with legal boundaries (67), (Cities (5), Boroughs (5) and 75 other legal towns)
- Census Towns (664)
- Suburbs/environs of Cities, Boroughs or legal towns
- Towns with a population of 1,000 or over (233)
- Electoral Divisions (3440),
- Rural Districts/Areas (160)
- Enumeration Areas for the five cities and their suburbs only (1,559),
- Dáil Constituencies (43)
- Gaeltacht Areas (8)
- Health Service Executive Areas (4),
- Local Electoral Areas (180)
- Regional Authority Areas (8),
- Garda Regions (6), Divisions (25) and Districts (109)
- County and Borough Electoral Areas,
- Aggregate urban (34) and aggregate rural (34) areas of each county.

Appendix 9 sets out a description of the Administrative and Census Areas used. It also describes the criterion used by CSO for creating census towns and boundaries for suburbs/environs of legal towns.

4   Production of the Statistics, Data Processing, Quality Assurance

4.1   Data Capture

Since 2002, the census form has been specially designed to enable it to be scanned and to allow for the automatic capture of data on the form. Drop-out ink is used in the printing, and the position of all tick boxes, write-in boxes and general form layout allows scanning and optical character recognition.

The forms are returned to the CSO in archive quality boxes, usually three boxes of forms per enumerator. The boxes are registered into a tracking system and stored in pre-designated shelving. When selected for processing the boxes of forms are divided into batches of 15 forms, are guillotined and then scanned. The batch of scanned images then passes through the automatic processing system, and once captured, validated, and coded is output to a database. The process and rules governing how the data is captured off the form are very strict and clearly specified.

The census form is composed of tick box questions (such as marital status), write-in numeric questions (such as date of birth and distance travelled), and write-in alpha questions (such as place of birth and nationality). The rules for the capture and coding of all questions are as follows:

Tick Boxes

The tick box questions are assigned the value of the box that has been ticked. Only one tick is allowed for each question, with the exception of the following:

Question 3	Relationship to persons 2, 3 and 4
Question 12 - Part 2	Frequency of speaking Irish
Question 24	Third level subjects
Question 21	Voluntary work

In all other questions, if more than one box has a mark in it (multi-tick) the question is examined by an operator. The values allowed for each field are specified in advance. Validation is performed on all values to ensure that they are only from the correct range of codes available. Blank values are allowed. Some fields can be blank at the beginning of the process but cannot be blank at the end of the process. This is indicated in the section on imputation and in the edits. In addition to tick box questions the following two areas of the form are captured as 'Optical Mark Recognition', or OMR, which looks for any mark in the areas specified.

- Name – if there is a name written in the Name field a value of '1' is returned; if the Name field is blank '0' is returned.

- Area adjacent to the write-in box for kilometres (this is to identify where the word 'miles' has been written next to the KM field. In such cases an operator converts the miles to KM).

Write-in numeric questions

All write-in numeric fields are recognised by the OCR (Optical Character Recognition) engines; those containing unrecognised text are presented to the operator for repair. The write-in numeric fields are validated against a range of values. These range values take the form of an upper and lower parameter limit, or a list of possible values. County, EA, ED and Street townland are combined and have a single range list. All other fields have their own separate range.

Write-in alpha questions

All write-in text fields are captured using OCR, and either coded automatically, or repaired and manually coded, via a lookup index. The facility is available to update, alter, delete or add entries to the look-up lists for these fields, as the need arises. A field that has been assigned from a lookup table is referred to as an assigned code. If a code cannot be assigned automatically then the text is presented to an operator for computer assisted coding.

4.2   Coding

The data is captured and coded using automated forms processing software. For Census 2002 and Census 2006 the CSO used Eflow 3. The data is first scanned and then recognised using Optical Character Recognition technology. As the census forms pass through the Eflow software each data field is matched against a 'look-up' list to find an automatic code. For example the 'place of birth' fields are automatically matched against the 'allplaces' look-up list which contains every possible place of birth (all places in Ireland - be they cities, counties, towns, districts etc. along with mis-spellings of all such places). If a match is found the captured text string is automatically coded. If no match is found then the partially captured text string is presented to the operator who will match the field against the look-up manually.

Question	Present Persons	Absent Persons
Relationship	Relationship is captured for persons <u>in private households only.</u> No relationship data is captured for persons in Communal Establishments 'coded C'. No relationship data is captured for Person 1. For Person 2 and up: the relationship to person 1 is captured through a tick box with a possible value of 1 through to 11. Blank is allowed. For Person 3 and up: In addition to the above, the relationship to person 2 is captured through a tick box, Relationship tick 2, with a possible value of 1 through 11. Blank is allowed. For Person 4 and up: In addition to the above the relationship to Person 3, Relationship 3, is captured through a tick box, with a possible value of 1 through 11. Blank is allowed. For Person 5 and up or Persons on Individual Forms: In addition to the above the relationship to Person 4, Relationship 4, is captured through a tick box, with a possible value of 1 through 11. Blank is allowed.	If tick box 1-3 or 11 has been ticked then the Relationship code is assigned according to given rules. If text has been written in the text box then this text is captured and coded against a Relationship lookup.
Place of Birth	If the question is blank then a code of 99 is assigned. The text is automatically recognised and a code is automatically assigned using the ALLPLACES Index File. County is coded first, and Country is only coded if County is blank. If both are entered only County is coded.	Not applicable
Nationality	If tick box and text are both blank then a Nationality code of 99 is assigned. If the tick is 1 and no other box has been ticked and the text is blank then the code IE for Irish is assigned. If ticks 1 and 2 are both ticked then the string Irish is concatenated before the text to give a dual nationality e.g. 'Irish-American'. Other rules apply depending on the combination of ticks and text in the question.	Not applicable
Usual Residence	If tick box, County and Country fields are all blank then a UR code of 99 is assigned. If tick box = 1 and the rest of the question is blank a UR code of H is assigned. If the tick box is not 1 and text has been entered then a code is assigned automatically using the Allplaces lookup. Other rules apply depending on the combination of ticks and text in the question.	If A6 tick box is 1 then assign UR code "H " If A6 tick box is not 1 (i.e. is 2 or blank) then assign UR code "ZZ"

Question	Present Persons	Absent Persons
Usual residence 1 year ago	Same as usual residence	Not applicable
Religion	If the tick box and text are both blank then a code of 99 is assigned.	Not applicable
	The tick box values are mapped to Religion codes according to rules.	
	If tick box 6 has been ticked and there is an entry in the Religion text field then a code is automatically assigned using the Religion Index File.	
Ethnicity	If the tick box and text are both blank then a code of 99 is assigned.	Not applicable
	The tick box values are mapped to Ethnicity codes according to rules.	
Present Principal Status	If the tick box and text are both blank then a value is imputed. See imputation rules below.	Not applicable
	If a box has been ticked then the PES code is assigned the value of the tick box from 1 through to 7.	
	If more than one box has been ticked then the person is manually coded.	
Family Nucleus Coding	The family nucleus code effectively defines the family structure that exists in private households. It is assigned using the Relationship and Usual Residence codes.	Not applicable
	In order to be coded as part of a family a person must be 'usual resident' i.e. they must live in the house. A usual resident is defined as a person with a UR code of H, 99 (blank), 35 (Ireland) or 43 (Republic of Ireland).	
	While the majority of households contain only one family it is possible to have up to three families in the one house. The codes for the first family in a household are as follows:	
	• Parents are coded 1 and 2 • Children are coded 3 • Other relatives who live in the house are coded A. • Persons who are unrelated or who do not live in the house are coded 0.	
	There must be at least 2 persons in every 'family' i.e. you cannot have a one-person family.	
	A person in a single person household is coded FNC 0.	

Question	Present Persons	Absent Persons
Industry and Occupation	Industry and Occupation fields are captured, coded and validated in a precise process where strict rules apply.	Not applicable
	Both fields are validated against other socio-economic variables of age, employment status, PES, ceased education, age education ceased, hectares, level of education, third level qualifications, third level subjects.	
	All occupation and industry fields that are not blank are repaired and coded, in strict sequence.	
	First the Industry text is coded against Industry lookup.	
	Then the occupation text is coded against Occupation lookup.	
	Then the suite of industry edits is run against the person. Uncoded strings or any person record that fails an edit are passed to an operator for computer assisted coding and repair.	

4.3   Data Editing

The data is captured, coded and edited using automated forms processing software. For Census 2006 the CSO used Eflow 3. The data is first scanned, and then recognised using Optical Character Recognition technology. As the census forms pass through the Eflow 3 software each data field is validated using rules provided. If any data field fails an edit then the failed fields are presented to the operator on screen who repairs the data.

The general principle for applying the edits is as follows:
- All edits are run on a field by field basis and all failed edits are examined by an operator.
- Edits are basically rules that stress-check the data captured on the forms for consistency. After the values have been assigned to the individual fields, and validated against a set range, or coded against a given lookup file, the edits are run. The answer to one question is checked against the answer to another one, and depending on the result the edit is either passed or failed. If the edit is passed then no action is taken; if it fails then the data that failed the edit must be presented to the operator and action taken. The edits are run as the fields are validated and coded e.g. as the household questions on page 2 are validated the edits associated with those questions are run. All the questions that cause an edit to fail are assigned to the same group so that they can be presented on screen for the operator to make a decision on the correct action to be taken.
- The edits are either Warning or Fatal. If a warning edit fails then the operator, having checked that the data was captured correctly off the form, will accept the edit or in certain cases may change the answer if it seems implausible (if the respondent has clearly misunderstood the question e.g. a child ticks 'married'). If a fatal edit fails then the operator must correct the data before moving on. If the operator is unable to decide the correct action then the forms are directed to a supervisor for a decision.
- No data, neither original nor changed, is allowed to pass through the system without being validated. After any field has been changed all the edits associated with that field are re-run.

The processing system ensures that the fundamental principle of re-validating all data that has been changed with all the edits associated with it is always adhered to.

Once the data has been captured, validated, coded and output, further data cleaning takes place at a macro level. Examples are cleaning fields where age filters apply and assigning usual residence codes based on the geography on the front of the form.

4.4   Imputation (for Non-Response or Incomplete Data Sets)

The data is captured, coded and edited using automated forms processing software. For census 2006 the CSO used Eflow 3. Imputation was also carried out within the Eflow 3 software. The data is first scanned, and then recognised using Optical Character Recognition technology. As the census forms pass through the Eflow 3 software each data field is validated using rules provided. If a data field has been identified for imputation then this field is routed to the imputation station. This imputation station was custom built for Census 2006, within the Eflow 3 software.

Missing values for date of birth, marital status and principal economic status are imputed, in strict sequence during processing.

1.   Missing values for Date Of Birth are imputed
2.   Family nucleus codes are assigned
3.   Missing values for Marital Status are imputed (which uses the family nucleus codes)
4.   Missing values for Principal Economic Status are imputed.

The imputation rules check for certain conditions and depending on the result a value is assigned off a rotation file. The rotation files are lists of values that represent the distribution of values for a particular combination of demographic characteristics. The values are selected off the rotation files one by one and when the end of the file is reached it should start back at the beginning.

Data field	General rules
Date of birth	If the any of the DOB fields, DD, MM, and YYYY are blank a value is imputed. Imputed values are held in a derived 'DOB' field and 'age' is calculated using the values in the derived DOB field.
	Year is imputed first, followed by month, followed by day.
Marital Status	This is based on the Family Nucleus Code (FNC) which has been assigned automatically by the system.
	The process uses age, sex and relationship codes, along with the family codes of the person whose marital status is being imputed and also the family code of related persons in the household. For example a person may have ticked that they are the husband / wife of another person but may have left marital status blank and so would be imputed to be 'married'.
	If FNC has not been assigned automatically then marital status is not automatically imputed; it is assigned manually by the operator during a later stage of processing.
Principal Economic Status	This is imputed for all missing values for persons aged 15 or greater. Rules are used that look at captured text in the field (for example if the text is "housewife" then a code of 5 is assigned). Education, employment status and age are also used. If there is any entry in either the industry or occupation fields then no imputation is done – it is passed to an operator to be manually assigned.

4.5   Grossing and Weighting

No grossing or weighting applies.

4.6   Computation of Outputs, Estimation Methods Used

In addition to the data captured and coded on the census form additional values are computed post processing; these take the form of 'counts', flags, and derived variables and are computed / assigned according to strict rules using the existing data for the household.

Examples of counts are: number of usual residents, number of persons aged 65 and over, number of persons in the household who speak Irish, number of persons at work in the household.

Examples of flags: flag to denote reference person, flag to denote person is in a consensual union or same-sex couple.

Examples of derived codes are: usual residence, usual residence one year ago, place of birth, age of eldest child in the family.

Households are divided into Private household (containing at least 1 usual resident) and non-Private households (communal establishments) where there may or may not be usual residents

Socio-economic group (SEG) and Social class group (SC) are also assigned post-processing. See also section *3.11* on classifications.

The final census master file is output from the processing database to a CSO managed Sybase database. The derived variables, counts and flags are assigned using programs run in visual basic by the census team. The data is then output to a flat file and read into a SAS data set. Aggregation is performed using the proc tabulate procedure or the table producing language TPL.

4.7   Other Quality Assurance Techniques Used

Extensive analysis and checking is carried out at a macro level as the output becomes available. Tabulations are produced and examined throughout the process. The analysis falls into different categories:

1. The results across all variables are checked against previous census results. This is carried out by a statistician and usually takes the form of closely examining the data and looking for anomalies.
2. Census variables are checked against other variables, for example place of birth against nationality, number of children born against age of mother, nationality against ethnicity. Where unusual patterns are found in the data a sample of the records are extracted and examined in detail. Where necessary changes are made.
3. Tables are produced that show the distribution of codes assigned by operator, by day. These tables are examined to look for unusual patterns, and where found are corrected.
4. The percentage of not-stated on each variable is checked against those of the previous census.

5   Quality

5.1   Relevance

In order to ensure the relevance of the Census to current demands for data, a public consultation is held in advance of every census seeking input from the public and interested parties. Input is sought on both questions to be asked on the form, and on the outputs to be produced. Changes to the layout of the form, changes to existing questions, new questions and changes to field methodology are all tested in a pilot census held -two years prior to census. An expert group is convened to discuss the submissions received and agree what should be tested. In 2004, in advance of the 2006 census, a sub-group on disability was set-up to agree a re-formulation of the disability questions. The question on Ethnicity was formulated with the assistance of the Equality Authority, the National Consultative Committee on Racism and Inter-culturalism and Pavee Point.

In 2004 the pilot test was carried out in 32 areas and covered in the region of 8,000 households. The results were analysed on the basis of response rates, with previous census figures (where there was a change to an existing question) and feedback from householders.

Further consultation takes place prior to every census on the outputs to be produced.

5.2   Accuracy and Reliability

5.2.1   Sampling Effects, Representativity

This is not relevant to the census.

5.2.2   Non-Sampling Effects

5.2.2.1   Quality of the Data Sources used

Not applicable.

5.2.2.2   Register Coverage

The census field operation is carried out from first principles. No register information feeds into the process apart from the use of the GeoDirectory – the national address database. The points indicated on the map are GeoDirectory points. The enumerator checks these and assigns a numeric code (D number) to each household visited. Where GeoDirectory points do not exist and the enumerator discovers a household she annotates the map with a dot and allocates a D number. The census does not set out to assess the coverage of the GeoDirectory.

5.2.2.3   Non-response (Unit and Item)

The preliminary results published on 19th July 2006 confirmed the accuracy of the population estimates published since the previous census in 2002. It also underscored CSO's belief that the methodology used in the field operation - the personal delivery and collection of census forms - is the most effective approach which could have been adopted in ensuring the comprehensive coverage of the census. To put the matter in perspective, the census field management system picked up a relatively small number of individuals who initially refused to cooperate in the census. The majority of these were persuaded of the need to comply. In a small number of cases field staff were instructed to provide direct estimates of the age, sex and nationality of individuals whom they knew to be present but who went out of their way to avoid enumeration. Some of these individuals were pursued with a view to prosecution by the Courts.

The following table provides estimates of item non-response for the individual topics in the census.

Question	Total	Number 'not stated'	Not stated as % of Total
Type of accommodation	1,462,296	31,803	2.2
Year built	1,462,296	69,210	4.7
Nature of accommodation	1,462,296	47,344	3.2
Rent paid	301,306	12,003	4.0
Number of rooms	1,462,296	45,444	3.1
Central heating	1,462,296	37,880	2.6
Type of piped water supply	1,462,296	47,791	3.3
Sewerage facility	1,462,296	47,181	3.2
Number of cars or vans		n/a/	n/a
PC	1,462,296	34,523	2.4
Internet access	1,462,296	75,744	5.2
Name		n/a	n/a
Sex		n/a	n/a
Relationship		n/a	n/a
Date of birth	4,239,848	24,436	0.6
Place of birth	4,239,848	52,347	1.2
Nationality	4,172,013	44,279	1.1
Usual residence now	4,239,848	130,699	3.1
Usual residence one year ago	4,111,333	87,263	2.1
Lived outside the Republic of Ireland for one year or more	4,111,333	167,979	4.1
Year of taking up residence in Ireland	778,577	63,823	8.2
Marital status	3,375,399	47,841	1.4
Number of children born alive	2,086,821	266,629	12.8
Ability to speak Irish	4,057,646	100,682	2.5
Frequency of speaking Irish	1,656,790	26,998	1.6
Religion	4,239,848	70,322	1.7
Ethnicity/ cultural background	4,172,013	72,303	1.7
Disability - Long lasting conditions			
Disability - Difficulty in doing activities	4,239,848	116,347	2.7
Means of travel to work, school, college	2,794,225	46,555	1.7
Time leaving home to travel to work, school, college	2,674,307	122,146	4.6
Distance travelled to work, school, college	2,794,225	463,706	16.6
Time to travel to work, school, college	2,674,307	244,452	9.1
Carer		n/a	n/a
Has full-time education ceased?	3,375,399	209,447	6.2
Age education ceased	2,850,333	258,077	9.1
Highest level of education completed	2,850,333	130,237	4.6

Question	Total	Number 'not stated'	Not stated as % of Total
Third level qualifications	820,760	10,286	1.3
Voluntary work	3,375,399	251,996	7.5
Principal economic status	3,375,399	132,111	3.9
Employment status	2,458,053	277,903	11.3
Occupation	2,109,498	230,918	10.9
Employers name		n/a	n/a
Employers address		n/a	n/a
Business of employer (industry)	1,930,042	132,910	6.9

5.2.2.4   Measurement Errors

There is no post-census coverage survey conducted following the census in Ireland, and so no measurement errors for the census as a whole are available.

The core set of demographic and socio-economic questions on the questionnaire are carried over and repeated from census to census. New questions and changes to existing questions are introduced and pilot tested between censuses. For the 2006 census a Pilot survey of 8000 households was carried out during which several questions were tested. Changes to existing questions are tested by way of comparison to the results to previous question formats. New questions are analysed by examination of the results by detailed age breakdown, sex, urban / rural responses and other relevant parameters. For example a new question on voluntary activity was introduced and tested in 2004. The results (participation by age, sex, type of voluntary activity, by socio-economic group) were deemed to be reasonable apart from the omission of a 'no voluntary activity' tick box which was learned through householder feedback, and the analysis of non-response.

The responses to the census questionnaires are provided by individuals in private households and communal establishments throughout the country. When the enumerator collects the completed census forms she is instructed to check that the information provided is accurate. A critical check is to ensure that every person who spent census night in the household or establishment is properly enumerated. The enumerator is also instructed to check through the form to discover areas of potential error and to have them corrected with the agreement of the householder. In no case can an enumerator alter the information provided on the census form.

The Field Supervisors also check a sample of the collected forms to ensure that the enumerator is carrying out her duties in a thorough way.

5.2.2.5   Processing Errors

Processing errors can arise in a number of different ways. Examples are:

(i)   False positives on written alpha characters

In any automatic OCR system false positives will occur. The software used to process the Irish census EFlow, estimated the maximum false positive rate to be 2 per cent. Actual rates of false positives are not available. False positives result in a string not being coded automatically. All text strings captured on the census form are matched against a look-up to find a code. If no match is found then the captured string is presented to an operator for assisted coding. Strings that contain false positives will either be coded automatically if that particular false positive was inserted into the look-up file (OUBLIN), or will require operator intervention. They never result in errors in the final output.

(ii)   False positives on written numeric characters

The estimated rate of false positives using EFlow software was 1 per cent. Actual rates are not available. False positives on numeric characters require special vigilance.

Date of birth: All captured date of birth values are validated against a number of different data fields on the form namely, 'Means of travel' (can't drive a car if under a certain age), 'Marital status' (under 15

years), 'Occupation' (high level of occupation for young person), other member of the family (age of mother / father), age of children.

Despite these checks false positives on date of birth can occur. To ensure that any errors that do get through the processed data are kept to a minimum distribution tables of age cohorts are compared against the population estimates and against the previous census.

Numeric codes in the front panel. All front panel codes are checked against a look-up list of valid codes so all false positives are immediately identified and corrected.

Kilometres and minutes. These are checked against means of travel and against each other (speed). Some errors do get through. Once all the data is processed tables are produced of time and distance travelled and anomalies can be spotted. These errors are then re-processed and the data repaired.

(iii)   Tick boxes

Processing errors with tick boxes can occur through a tick being missed, or 'false' tick being identified through dirt / dust on the form or through a strike-through on the form.

Dust and dirt do cause 'ticks' to be incorrectly picked up. Keeping the scanners dust free through regular (twice daily) cleaning is very important. Poorly scanned images are first identified at the scanning stage by the scanning operators and the images are re-scanned (they are examined with the human eye). The images must then be matched against a form type and this gives the operators a second chance to spot poorly scanned or dirty images. Nonetheless false marks can get through on the forms.

On a single tick question if a second tick is identified then the system will stop and the operator can correct the error. If it is a single tick question and the 'false' tick is the only answer on the form then this answer will be validated in the normal way. If the answer passes the edits then it will be captured as final. If it is a multiple tick question – like third level qualifications – then it is very difficult to identify. Vigilance must be maintained by the operators. However, because the majority of the questions require only a single tick, if multiple ticks were common due to dirt or dust this would become immediately apparent to the operators, the processing would be stopped and all poor images would be re-scanned.

A tick can be missed completely on the form if the 'Region of Interest' (ROI) on the form (the area where the system is to look for a mark) is set incorrectly, or if the image is slightly skewed. Poorly defined ROI in the 2006 processing system caused serious processing errors that were only identified mid-way through processing. The error was identified as the level of not-stated on certain questions was significantly higher than in 2002. The error was so widespread that all questions captured as blank (output as not stated) were re-processed once the ROI had been fixed. In all six million images were re-processed.

5.2.2.6   Model-related Effects

The Irish census model is enumerator drop-off and collect, followed by scanning and automatic data capture. The effects of scanning and automatic capture are covered under 5.2.2.5 Processing Errors.

Model effects can come about through the enumerators failing to identify one or more dwellings in their area, and then through failing to deliver and collect a census form to one or more dwellings, or finally through failing to enumerate all persons in that dwelling.

Every enumerator is given a map marked with all the known dwellings in their area. They are required to visit every dwelling and mark the appropriate dwelling number, the D number, against the dot on the map. Furthermore, dwellings identified during the course of enumeration which are not on the maps are marked by the enumerator and allocated a D number. The maps are checked on return and the correspondence between the front of the census form and the enumerator record book (ERB) thoroughly checked. Before going out into the field every enumerator is also given a detailed table of the population of their area at a very small geographic level (street/townland). On completion of enumeration the enumerators must summarise the headcount population for their area at a similar geographic level, and importantly account for all significant differences since the previous census whether they are positive or negative. This acts as a strong check for the enumerators for both summary errors and missed enumerations.

The enumerators are paid reasonably well and receive comprehensive training. They are instructed to visit every dwelling as often as it takes to make direct contact with the householder. They are paid for every form collected so there is a strong incentive to identify and visit 'new' dwellings not already on the map. As stated above, a critical check is to ensure that every person who spent census night in the household or establishment is properly enumerated. However, there is a limit to what - enumerators can do and ultimately they will depend on the honesty of the person completing the form.

5.3 Timeliness and Punctuality

5.3.1 Provisional Results

Provisional results from the 2006 census were released on the 19 July, 2006 just under three months after census day. These preliminary results are based on summary tables produced by the enumerators in the field. They consist of headcount data of total males, total females and total persons at Electoral Division (LAU 2) geographic level.

5.3.2 Final Results

The report, Principal Demographic Results, setting out the first set of final results of the 2006 Census was published on 29 March 2007, 11 months after census day. The report brings together in one publication the main tables relating to the following topics: age, sex, marital status, living arrangements, usual residence, migration, nationality, ethnic or cultural background, Irish language, religion and housing. It also contains a detailed commentary on the topics presented and presents some additional analysis in the form of summary tables and graphs and retrospective comparisons.

A second report titled Principal Socio-Economic Results, published in 28 June 2007, presents a summary of the final results for the remaining topics covered on the census form, namely principal economic status, industry, occupation, travel, education and housing.

More detailed results, by individual topic, are released in thirteen separate volumes. All results are released simultaneously on the CSO web site in interactive tables using the Beyond 20/20 software.

The publication schedule for Census 2006 was as follows:

Description	Publication Date
Preliminary Report	19 July 2006
Principal Demographic Results	29 March 2007
Volume 1 - Population Classified by Area	26 April 2007
Volume 2 - Ages and Marital Status	10 May 2007
Volume 3 - Household Composition, Family Units and Fertility	31 May 2007
*Small Area Population Statistics (SAPS) - demographic variables	21 June 2007
Principal Socio-economic Results	28 June 2007
Volume 4 - Usual Residence, Migration, Birthplaces and Nationalities	12 July 2007
Volume 5 - Ethnic or Cultural Background (including the Irish Traveller Community)	26 July 2007
Volume 6 – Housing	16 August 2007
Volume 7 - Principal Economic Status and Industries	06 September 2007
Volume 8 – Occupations	20 September 2007
Volume 9 - Irish Language	04 October 2007
Volume 10 - Education and Qualifications	18 October 2007
*Small Areas Population Statistics (SAPS) - all variables	25 October 2007
Volume 11 - Disability, Carers and Voluntary Activities	01 November 2007
Volume 12 - Travel to Work, School and College	15 November 2007
Volume 13 – Religion	29 November 2007
*Small Area Population Statistics (SAPS) were made available as interactive tables, free of charge on the CSO website	

5.4 Coherence

The primary coherence check of the census results is against the annual mid-April population estimates. In 2006 the final count of population was just under 5,000 more than the comparable population estimate. This was a very high level of coherence, though the differences for 15 year age cohorts were as follows:

Age group	Annual mid-April population estimates (2006)	Census 2006 results	Difference
0-14	864,572	864,449	-123
15-29	1,020,091	1,005,810	-14,281
30-44	952,207	972,795	20,588
45-59	745,018	747,141	2,123
60and over	652,984	649,653	-3,331
Total	4,234,872	4,239,848	4,976

A further comparison of births by county over the previous number of years is carried out although it is important to bear in mind that migration movements of families with young children may impact on this comparison.

5.5   Comparability

The model used in the field (visual enumeration followed by enumerator delivery and collection) for the Irish census is a long-standing one. Factors that can influence comparability over time are therefore changes to the way questions are asked, changes to the way the data is processed and changes in classifications.

Census questions

There is a constant tension between comparability over time on the one hand and the representativeness and continued relevance of a question wording on the other. The balance between the two has to be carefully monitored and managed. It is well known that changes to question wordings will have an impact on the comparability of results from census to census. The purpose of the census pilot is primarily to assess whether the revised question works and leads to believable results. A recent example of where a discontinuity was introduced was in relation to the Irish language question where an existing category was in the frequency of speaking the language (i.e. daily) was split into daily within and outside the education system. The outcome was a total lack of comparability in all frequency categories between the 2002 and 2006 censuses.

Changes in processing

The refinement of existing editing rules can have an impact on results. An example of this is the treatment of 'age education ceased' in 2006. The age filter on the form asked persons aged 15 years or over to skip to the end of the questions for that person. A higher than expected number of 15-year olds also skipped to the end of the remaining questions. During our macro editing we used a rule that said - If education ceased was blank and principal economic status was not 'student' then education ceased was set to 'yes'. This resulted in a higher that expected number of persons aged 15 being imputed as having ceased their education. All newly introduced edits and amendments to existing edits must be carefully managed to ensure no adverse impact on results.

Coding schemes

The introduction of new coding systems can sometimes give rise to discontinuities which it may not be possible to overcome without possibly coding on a current and previous basis. It may not be possible to do this because of the resource cost and because it may not be feasible from an operational viewpoint.

5.6   Accessibility and Clarity

5.6.1   *Assistance to Users, Special Analyses*

The census division has a dedicated enquiries section devoted to assisting users to access the census results. Census Enquiries provides a telephone and email service to customers Monday – Friday, 8am – 6pm.

The nature of enquiries generally fall into two categories:
- general enquiries on published census demographic data and
- requests for special census tabulation

The first category involves administrative staff giving assistance to customers with accessing online census information - most commonly recent (2006) or comparatively recent (2002) over the phone. Occasional queries for older census data also occur. The most popular census requests are for Small Area Population Statistics (SAPS). For the first time in the 2006 SAPS data was made freely available on the CSO website to download through the Beyond 20/20 software system. (See also section *5.6.3.3 Internet*)

From experience, Beyond 20/20 takes a little getting used to and so staff tend to spend a fair amount of time (typically ten to fifteen minutes) talking customers through its functionality. A supporting 'screen-grab' guide showing the steps for downloading SAPS is on the SAPS homepage for follow-up self help.

The second category of Special Tabulations involves the building of customised tables that have not been published before. Such requests can involve a number of variables for particular geographies. These requests tend to come from commercial bodies or local authorities. Examples include:

- place of work by occupation
- characteristics of family units that speak Irish (social class etc.)
- nationality by place of birth
- occupancy by education by location
- number of households with school going children that have PC's by socio-economic divide
- economic status of accession country migrants

This process typically involves a Statistician or dedicated tabulation section producing the tables from PC SAS or TPL programmes and administrative staff checking the results. Special tabulations are normally charged out on a cost recovery basis.

For all queries received over the phone or by email, use is made of an in-house recording/tracking system which details the treatment/progression of customer requests so that they are managed within the agreed timeframes (in the case of special tabulations). Special tabulations normally take about two to three weeks to deliver depending on complexity while all other queries are usually handled on the spot or on the same day.

The following data files are also made available on a restricted basis to *bona fida* researchers to enable them to carry out research projects:

Sample of Annonymised Records (SARS)

A 5% anonymised sample of the population covering demographic and socio-economic variables and associated housing characteristics has been lodged with the Irish Social Science Data Archive (ISSDA). ISSDA may authorise users who are bona fide students, staff or research personnel to have access to the data for the purposes of economic and social research. Use of the data and/or any results obtained from use of the data for any other purposes is prohibited without the express permission of the CSO.

Place of Work Census of Anonymised Records (POWCAR)

As part of the Census 2006 processing programme the place of work details of all employed persons who undertook a journey to work were geo-coded. A detailed file containing the demographic and socio-economic characteristics of these workers along with information on the origin and destination of their journeys to work has been made available for analysis.

The POWCAR is only available under strict conditions to bone fide researchers who are approved by CSO. Access is governed by CSO's policy on research microdata files (see http://www.cso.ie/aboutus/policy_access_res_microdata.htm).

5.6.2   Revisions

Revisions to census results can come about through errors made during data capture and coding. It is the policy of census division to correct any errors as soon as they become apparent.

In Census 2006, the following revisions were made to published data:

Geography coding error involving two adjacent Electoral Divisions (EDs)

363 households with a combined population of 940 persons, which should have been coded to the ED of Tallaght-Jobstown, were incorrectly coded to the neighbouring ED of Saggart in the Census 2006 file. Both EDs are in the administrative county of South Dublin.

Table 3 in Census 2006 - Principal Demographic Results and tables 5, 6 and 8 in Census 2006 Volume 1, which were affected by the error, were corrected in September 2008 and the amended versions of these tables were uploaded to the CSO website. The Small Area Population Statistics (SAPS) for the relevant EDs were also amended. These include the ED, HSE and Constituency files. Revised versions of the POWCAR were forwarded to accredited users before the end of September 2008.

5.6.3   Publications

5.6.3.1   Releases, Regular Publications

See full list of reports published at *5.6.3.2* below.

5.6.3.2   Statistical Reports

The following is a full list of reports from the 2006 Census:

Name of Report	Date published
Preliminary Report	19 Jul 06
Principal Demographic Results	29 Mar 07
Volume 1 - Population Classified by Area	26 Apr 07
Volume 2 - Ages and Marital Status	10 May 07
Volume 3 - Household Composition, Family Units and Fertility	31 May 07
Principal Socio-economic Results	28 June 07
Volume 4 - Usual Residence, Migration, Birthplaces and Nationalities	12 July 07
Volume 5 - Ethnic or Cultural Background (including the Irish Traveller Community)	26 July 07
Volume 6 - Housing	16 Aug 07
Volume 7 - Principal Economic Status and Industries	06 Sep 07
Volume 8 - Occupations	20 Sep 07
Volume 9 - Irish Language	04 Oct 07
Volume 10 - Education and Qualifications	18 Oct 07
Volume 11 - Disability, Carers and Voluntary Activities	01 Nov 07
Volume 12 - Travel to Work, School and College	15 Nov 07
Volume 13 - Religion	29 Nov 07
Non-Irish Nationals Living in Ireland	30 June 08

5.6.3.3   Internet

All census results are released simultaneously on the CSO web site at
http://www.cso.ie/census/default.htm

The following are the main formats in which data is provided on the website:

- PDF versions of the 16 census reports are available online at
  http://www.cso.ie/census/Census2006Results.htm.
- All published tables are also available in interactive format at
  http://beyond2020.cso.ie/census/ReportFolders/ReportFolders.aspx?CS_referer=&CS_ChosenLang=en
- Small Area Population Statistics (SAPS): 68 tables arranged over 15 themes are available for a wide range of geographic areas at http://www.cso.ie/census/SAPs.htm.
- 33 historical series, many going back as far as 1841, on a range of census topics are available on the census of population section of the CSO databank at

http://www.cso.ie/px/pxeirestat/database/eirestat/Census%20of%20Population/Census%20of%20Population.asp

The Census 2006 Boundary Files are also available on the CSO website at http://beyond2020.cso.ie/censusasp/saps/boundaries/census2006_boundaries.htm. The files can be utilised by users who have access to software which will enable them to perform their own analysis of census data within a **GIS** (Geographical Information Systems) environment.

The boundary files are in ERSI Shape (SHP) format which is a proprietary format created by the Environmental System Research Institute. The ESRI Shape (SHP) file format is a vector format. SHP files support point, multi-point, polygon, polyline and multi-patches. The ERSI file format contains three files **Main file: *.shp, Index file: *.shx** and **DBase file: *.dbf** for each boundary. The three files must be downloaded. Most GIS applications can read SHP files.

5.6.4   Confidentiality

All information collected in the census is treated as strictly confidential by the CSO and used only for statistical purposes. This is guaranteed under the Statistics Act, 1993.

Under the Statistics Act (sections 38, 42 and 44), the penalty for any breaches of statistical confidentiality can be as high as €25,394 (20,000 Irish 'punts'). All CSO headquarters staff and field staff are signed up as Officers of Statistics under Section 20 in a format specified in Section 21 of the Statistics Act. They are made fully aware of their legal obligations regarding confidentiality.

When census data are being captured, the names of individuals on census returns are not entered onto any computer database.

The involuntary disclosure of confidential data is avoided by taking care to combine up small areas where the number of households and persons is quite small with contiguous areas in such a way that disclosure is minimised.

6   Additional documentation and publications

http://www.cso.ie/census/documents/censusform_2006.pdf

http://www.cso.ie/census/documents/SAR_User_Guide_2006.pdf

http://www.cso.ie/aboutus/policy_access_res_microdata.htm

http://epp.eurostat.ec.europa.eu/portal/page?_pageid=1996,45323734&_dad=portal&_schema=PORTAL&screen=welcomeref&open=/&product=REF_TB_population&depth=2

http://www.statcentral.ie/viewstats.asp?type=Population