

Information paper

Census strategic development review Alternatives to a Census: Review of international approaches

Contents

| | Page |
|---|-----------|
| 1 Introduction | 3 |
| 2 Traditional census | 3 |
| 2.1 Case study: USA (population: 281 million) | 3 |
| 2.1.1 History | 3 |
| 2.1.2 Census 2000 | 4 |
| 2.1.3 Future | 4 |
| 2.2 Other countries | 5 |
| 2.2.1 Australia (population: 19 million) | 5 |
| 2.2.2 Canada (population: 30 million) | 5 |
| 2.2.3 Israel (population: 5.6 million) | 5 |
| 2.2.4 Italy (population: 57 million) | 5 |
| 2.2.5 New Zealand (population: 3.7 million) | 6 |
| 2.2.6 Slovenia (population: 2 million) | 6 |
| 2.2.7 Spain (population: 41 million) | 6 |
| 3 Population register and other administrative sources | 7 |
| 3.1 Case study: Netherlands (population: 16 million) | 7 |
| 3.1.1 History | 7 |
| 3.1.2 The Social Statistics Database | 7 |
| 3.2 Other countries | 8 |
| 3.2.1 Belgium (population: 3.4 million) | 8 |
| 3.2.2 Denmark (population: 5.4 million) | 8 |
| 3.2.3 Finland (population: 5.2 million) | 8 |
| 3.2.4 Germany (population: 82 million) | 8 |
| 3.2.5 Norway (population: 4.5 million) | 9 |
| 3.2.6 Singapore (population: 4 million) | 9 |
| 3.2.7 Sweden (population: 9 million) | 9 |
| 4 Rolling census | 9 |
| 4.1 Case study: France (population: 60 million) | 9 |
| 4.1.1 History | 9 |
| 4.1.2 The new model | 10 |
| 5 Conclusions | 11 |

Census strategic development review

Alternatives to a Census: Review of international approaches

1 Introduction

For many years, population figures have been required by governments for a variety of purposes, from taxation and military service to planning of schools and hospitals. The simplest way to obtain these figures is to count people, which is the original derivation of the word 'census'. The modern census, however, is far more than a head count - it provides a wealth of information, at a low level of geographic detail. Topics on the most recent round of censuses were as wide-ranging as work, housing, health and religion.

The history of the modern census stretches back over 200 years, and most, if not all, countries in the world have conducted one at some point during that period. In many places, however, the 'count everybody' model is beginning to fall out of favour. There are a number of reasons behind this, including falling response rates, respondent burden, cost, and a desire for more accurate information in non-census years. This paper reviews, at a high level, the census designs being used in a number of countries. It does not attempt to cover the detail of population base and definition, coverage adjustment, outputs, costs etc. The term "household" is here used to mean a group of people enumerated on a single census form (or equivalent), although in practice a number of different definitions (e.g. households, housing units, dwellings) are used in different countries.

The designs fall into three broad categories – a traditional census, administrative sources (including population registers) and a rolling census. A case study is given for each category, followed in the first two cases by an overview of other countries.

2 Traditional census

The traditional census is still used by the vast majority of countries, especially outside Western Europe. In essence it consists of a complete enumeration of the population on a particular day, carried out periodically (usually every 5 or 10 years). However, within this structure there is wide scope for variation: for instance, enumeration may be carried out by interviewers

or a self-completion form; if the latter, forms may be delivered and collected by enumerators or via the postal service; the census may count people at their location on census day or at their usual place of residence; and the number and type of questions asked may vary extensively. In addition, there is the issue of under-enumeration - how, if at all, it should be measured, and whether or not published figures should be adjusted to take account of this.

2.1 Case study: USA (population: 281 million)

2.1.1 History

The requirement to hold a census every ten years is enshrined in the US constitution. This is due to the country's federal nature - since its foundation, up-to-date population figures have been required for each state to determine how many representatives each may send to Congress. The first census was held in 1790 - just over a year after George Washington was inaugurated as president - and the decennial sequence has continued unbroken since then. In 1902 a permanent Bureau of the Census was established, and has been responsible for the count ever since.

As recently as 1950, the census was conducted almost entirely by a doorstep interview - it was not until 1960 that self-enumeration became a major part of the process. Most questionnaires are posted to households rather than hand-delivered by an enumerator, although initially enumerators still called to collect the completed form. By 1990, some 94 per cent of the population was covered by post-out and post-back, but the postal response rate declined steadily, dropping as low as 65 per cent in 1990, leaving enumerators to visit over one-third of households. In the US census people are enumerated at their usual place of residence even if they are away on census night.

A distinctive feature of the US census is sampling, which was first used in 1940. Since then, there has always been a subset of questions that are asked only for a proportion of the population. The proportion of the population sampled, along with the sampling method, has varied over the years, as have the number and

types of questions asked, but in all cases, as with the main census, participation has been mandatory.

The overall net undercount in 1990, estimated by a post-enumeration survey, was 1.6 per cent. However, the figure varied widely between areas and demographic groups: for instance, it reached 5.0 per cent among people of Hispanic origin, 3.2 per cent for under-18s and 3.4 per cent in the Washington DC area. These are net figures, which include a significant overcount as well as non-response.

2.1.2 Census 2000

The USA's most recent census day was 1st April 2000. In early March, 120 million forms were delivered to households, 98 million of them by post (the remainder, mainly to dwellings in remote and rural areas, were hand-delivered by enumerators). People who did not receive a form were able to obtain one from one of a number of public locations. Census returns could be submitted over the internet or, for people who had difficulty reading or understanding the questionnaire, by telephone, as well as by post. After census day enumerators were sent into the field to retrieve outstanding forms.

The response rate excluding fieldwork was 67 per cent. This included postal, telephone and internet returns. It was 2 percentage points higher than in 1990 - ending a decline stretching back to 1970 - but was still 8 percentage points lower than in 1980. No undercount rate is available due to problems with the post-enumeration survey. This was known in 2000 as the Accuracy and Coverage Evaluation (ACE) survey, and was designed to estimate both undercount and overcount to arrive at a final net undercount figure. In the event, though, it was believed that the ACE failed to identify an overcount of approximately 3 million. The estimated population would therefore be too high by this amount, and the unadjusted counts were considered more reliable.

Sampling in 2000 was carried out by means of a 'long form', that is, a sample of the population were asked for more detailed information. The sampling rate nationally was about 1 in 6, but the need to provide information at a low geographic level meant that a larger proportion needed to be sampled in less densely populated areas.

The 'short form' collected only age, sex, race, tenure and relationship information, while additional topics on the long form included

marital status, education, work, income, household bills and number of rooms.

2.1.3 Future

Given the cost of the census, the problems in assessing coverage, and the large amount of administrative data available, the bureau is examining the feasibility of increased use of administrative sources in census operations. These range from their use as auxiliary data for assessing coverage levels to an entirely records-based census. Since the census is used for defining electoral boundaries for local assemblies, any method used needs to be able to provide accurate counts at the block level.

An administrative records experiment was carried out on a small subset of the population in conjunction with the 2000 Census. This was in two stages: in the first, administrative records (for example, tax and health records) were used to determine demographic details for individuals in the test areas. Following this, fieldwork was carried out at those addresses in the Bureau's master address file for which no record was found and to follow up apparently invalid addresses in the administrative sources. (In practice, fieldwork from the 2000 Census was used instead of carrying out additional fieldwork.)

In total, the administrative records experiment counted 99 per cent of the unadjusted census population (no information is available on how reliable the census was in the test areas). At county level, the proportions varied from 96-102 per cent. For under-18s the figure was 87 per cent, while for adults it was 103 per cent. The overcounts are believed to be due to deaths since the sources were collected (some were a year old on census day) and failure to eliminate duplicates. Deficiencies in the sources available meant that ethnicity figures were highly unreliable.

Because of the limitations of the 2000 experiment, including the limited number of administrative sources available and relatively small test sample (two sites, with a total population of approximately 2 million), it was not considered reasonable to replace the entire census with an administrative records-based survey in 2010. However, it was recommended that there should be a larger experiment in conjunction with the 2010 Census, of sufficient size to be able to generalise to the entire population. It is also possible that administrative

records may be used to provide data for households that do not respond to the census.

The 2010 Census itself will continue to be run on traditional lines, but the long form will no longer be used. The information previously gathered this way will now be obtained by the American Community Survey, which has an annual sample size of 3 million households. Over 10 years, therefore, approximately a quarter of the population will be covered, significantly more than had previously received a long form. In addition, it will eventually be possible to produce annual population estimates at a low level - reliable estimates for areas with a population of over 65,000 will be available after the first year of the national survey, for populations of 20,000 – 65,000 after three years, and for smaller areas when there is five years of data available.

2.2 Other countries

2.2.1 Australia (*population: 19 million*)

Legislation provides for a census to be taken every five years. The count is of the population present on census night, although temporary residents are also asked for their usual place of residence. A drop-off/pick-up methodology is used, without any use of the postal service. The level of undercount is determined by a post-enumeration survey. In 2001 the estimated net undercount was 1.8 per cent, up from 1.6 per cent in 1996. The 2001 Census included questions on the use of personal computers and the internet, and, for the first time, allowed respondents to tick an 'opt-in' box giving permission for the record to be released in 99 years - the form will otherwise be destroyed, as has previously been the practice.

A traditional census will once again be conducted in 2006. Two major changes are planned. For the first time, it will be possible to fill in the census form via the Internet. In addition, it is intended to use post-back in buildings, such as apartment blocks, where it is particularly difficult to obtain a response through pick-up.

2.2.2 Canada (*population: 30 million*)

The Canadian census, a count of usual residents, occurs every five years. A drop-off/post-back methodology has been used since 1971, although post-out is being considered for 2006. The short form contains seven questions, asking only basic demographic information. Additional questions on the long form, sent to one household in five, include parents' birthplace, single-sex

couples and a very detailed language question. Gross undercoverage rates have been steadily increasing, reaching 3.3 per cent in 1996.

Canada's next census, in 2006, will again be run along traditional lines. It is intended that two thirds of forms will be posted out, and production of a comprehensive address register is underway to facilitate this. It will also be possible to use the internet to complete the questionnaire.

2.2.3 Israel (*population: 5.6 million*)

A census has been held every 10-12 years since the establishment of the state in 1948, the most recent being in 1995. A drop-off/pick-up methodology was used. Israel has a population register, but this was not brought into play until the pick-up stage. At this point, enumerators checked each household's return against the register's record of residents. This helped to improve coverage at both individual and household level. Each person's unique identification number was appended to the form by the enumerator, and register information used in the edit and imputation operation. The short form contained five demographic questions, while the long form, delivered to one household in five, contained questions including fertility, income and housing conditions.

The 1995 Census was the last to be conducted along traditional lines. From 2006 it is intended to conduct an administrative sources census based on the population register and other sources. The register is known to have coverage problems - in particular, it does not cover permanent residents who are not Israeli citizens, and conversely it does include citizens who have permanently emigrated. For this reason a large-scale sample survey, covering 15-20 per cent of the population, will be carried out to assess both over- and undercoverage. It will also include questions on topics that are not covered by administrative data.

2.2.4 Italy (*population: 57 million*)

A census is normally held every ten years, going back to 1861 - shortly after the birth of the state. Italy also has municipal population registers, but as each municipality is responsible for its own register, and these vary highly in quality - some are not even computerised - and, since funding is partly dependent on the number of people on the register, there is little incentive to remove people who have died or moved. Traditionally, the census has been used to update the registers as well as for statistical purposes. The registers, in turn, are used to estimate undercount in the

census; in 1991 for example, the raw register figures gave an undercount of 5 per cent, but attempts to remove superfluous register data reduced this figure to 1 per cent.

There can be a significant difference between resident population and registered population in an Italian municipality. Reasons can include students who are registered in one municipality but live in another to study, people who are temporarily resident in another municipality (or out of the country) due to work, and people performing national military service. The 2001 Italian Census attempted to enumerate both the legal population and the population actually present.

The national statistical office, ISTAT, is co-ordinating an attempt to standardise the population registers and other administrative sources, with a view to conducting register-based censuses in the future. It is not clear when this is envisaged to commence.

2.2.5 New Zealand (*population: 3.7 million*)

The New Zealand census began in 1851 and is currently taken every five years, the most recent being in 2001. The methodology was drop-off/pick-up and the count was of the population present on census night, even to the extent of distributing forms to passengers on overnight trains (although householders were asked names and basic relationship information for residents who were absent on the night). A special effort was made to count the homeless via city missions and other community support services.

Each individual was given their own form, separate from the dwelling form, regardless of the type of accommodation in which they lived. Māori and Pacific peoples, who had been difficult to enumerate in the past, were specifically targeted in the publicity campaign, with bilingual Māori/English forms available on request. Questions asked on the individual form included income and an optional religion question, while the dwelling form included questions on amount of rent paid and the type of fuel used for heating. Questions on smoking and fertility, which had been asked in 1996, were dropped in 2001.

A post-enumeration survey to measure undercount was introduced in 1996. In 2001 the PES surveyed approximately 0.7 per cent of dwellings. The sample design was based on New Zealand's labour force survey, which divides primary sampling units into 120 strata.

The proportion of PSUs sampled in each stratum varied according to its ethnic and age characteristics. The net undercount increased from 1.6 per cent in 1996 to 2.2 per cent in 2001, consisting of an undercount of 2.8 per cent and an overcount of 0.6 per cent.

Another traditional census is to be conducted in 2006. New Zealand is believed to have a more technologically aware population than the average, and so the Internet is expected to be a major method of completion in 2006. No information is available on expected take-up rate, but a system is currently being tested. No other major changes are envisaged, with the emphasis being on continuity and improving data quality.

2.2.6 Slovenia (*population: 2 million*)

2002 saw Slovenia's first census as an independent nation, although regular censuses had been conducted when the area was part of Yugoslavia and earlier, going back to the mid-19th century. The 2002 Census used administrative sources extensively, including a central population register, and employment and benefits registers. Some data, including place of birth, marital status and employment details, were taken entirely from these sources. Others, including sex, address one year ago and place of education, were available for some individuals but not for others. Any information not available this way was collected by interview or self-enumeration; thus, some people were asked more questions than others were. No information is available about future census plans.

2.2.7 Spain (*population: 41 million*)

The 2001 Census in Spain was claimed to be the first in the world to allow all residents to complete their forms via the Internet. Central to this was the existence of municipal population registers. Paper questionnaires were pre-printed with information from the register, which they could correct if necessary. The pre-printed form also contained an identity code and password to enable access to the on-line version. The Internet form also required other authenticating information that was not printed on the form. All consistency checks were applied on-line in real time.

There is no information available on how heavily this option was promoted, but the final take-up rate, just 0.1 per cent of households, was a disappointing.

3 Population register and other administrative sources

Many countries have some form of population register, either at national or local level, and in principle this can completely replace a census to provide a snapshot of the population at any time and at any level of geographic detail. In practice its use is constrained by the quality and extent of the data contained in the register, and many countries combine their population register with one or more of a traditional census, sample surveys and other administrative sources to build up a complete and accurate picture. Modelling using sample surveys cannot, of course, provide information at the same level of geographic detail as a traditional census.

3.1 Case study: Netherlands (*population: 16 million*)

3.1.1 History

The Netherlands has not held a conventional census since 1971. Increasing public concerns about privacy, and an associated risk of low response rates, led to the 1981 Census at first being postponed, then followed by a complete abandonment of the traditional census approach. Since it was still necessary to produce census-type data in 1981 and 1991, Statistics Netherlands used the best information available to them from administrative registers and existing sample surveys. However, it was only possible to use administrative data at aggregate level, meaning that it was not possible to integrate them fully with survey data.

Changes in legislation now allow individual administrative records to be used for statistical purposes. Statistics Netherlands now tries to use such records to their full potential, in order to reduce both costs and respondent burden. All Dutch residents are expected to register with the local authorities, and the population register thus obtained is among the administrative sources available. It contains name, address and basic demographic data such as sex, date of birth and marital status. It is updated once a year, on 1 January.

3.1.2 The Social Statistics Database

In 1996, Statistics Netherlands began the process of building a Social Statistics Database (SSB in Dutch) which will contain all available information at an individual level. It aims to cover all those who were resident in the Netherlands during the year in question, based on the population register. Since the population register is only updated annually, a 'pooled population register' is created from the records

at the end of the year. This covers most of the target population, obvious exceptions being illegal immigrants and those who were not resident at the end of the reference year. Where people have moved during the year, their address as at the end of the year is used. Clearly, people who move during the year and die or emigrate before its end will not have their most recent address in the pooled register.

The population register is the spine from which the rest of the SSB hangs. Records from other administrative sources (such as benefits data, employee insurance schemes) and sample surveys (for example, the labour force survey, health interview survey) are matched to the register using variables such as sex, date of birth and address. The social-fiscal ('sofi') number appears on many administrative records, but was not included on the population register until 1997. At household level, such sources as the geographic address file, housing register and household budget survey are used. The matching process aims to match over 90 per cent of the records, with no more than 5 per cent false matches.

A coverage estimate of the population register is hard to come by, due to the lack of another measure of the population. Although there is no law making registration compulsory, it is generally considered that it is virtually impossible to function in Dutch society without being registered. For instance, anyone working without being registered would have to pay the top rate of income tax. The register is regularly checked against other sources, including the police's information on immigrants, and it is in the best interest of the individual municipalities to register all their resident population. At the time of the 1971 Census, the register was found to have an undercount of just 0.02 per cent, an overcount of 0.05 per cent, and 0.18 per cent registered in the wrong municipality. There may, of course, have been significant changes in the last thirty years.

Not everyone in the population register will be covered by the other administrative sources, and the sample surveys, by their nature, only cover a small fraction. Imputing the missing values is not a reliable method because of the small sample sizes, so a method of weighting is used to produce consistent tables. A weighting procedure is also used to reduce non-response bias. The 'virtual census' conducted every ten years is now purely for the purposes of fulfilling Statistics Netherlands' obligations to Eurostat.

At a domestic level, it is possible to be far more flexible with the frequency and type of data produced.

3.2 Other countries

3.2.1 Belgium (*population: 3.4 million*)

The last decennial census was conducted in 2001. The municipal population registers were centralised in 1985, and since this time it is this National Register that has been used to determine the official population figure. This is therefore no longer a function of the census, which was renamed the 'General Socio-Economic Survey' in 2001. In future, existing administrative records and specially developed databases will be used to determine census-type information. The specially developed databases are on dwellings and level of education, both of which will be created initially from information in the 2001 survey. To update the dwellings database information will be required from anyone registering a change of address on the population register, while the education database will be updated with information from educational establishments.

3.2.2 Denmark (*population: 5.4 million*)

The last conventional census in Denmark took place in 1970. Denmark has had municipal population registers since 1924, but in 1968 a nation-wide Central Population Register (CPR) was set up. This was based on the municipal registers (themselves originally taken from the results of the 1924 Census) and checked against the 1970 Census results. Registration of a change of address with the municipal authorities is compulsory, and these, along with birth and death registrations, are used to update the CPR.

Although the CPR only contains basic address and demographic information, most other administrative registers came into existence at a later date, and are linked to the CPR by the unique person identifier that it contains. A register-based population census took place in 1976, and following the establishment of a Register of Buildings and Dwellings, a register-based population and housing census in 1981. Neither of these included any form of fieldwork.

Denmark's Statistics Act obliges public bodies to make available any information required by Statistics Denmark for statistical purposes. It also requires them to consult Statistics Denmark when developing registers to ensure that they are appropriately organised for statistical use. Population statistics are now produced every year from the CPR and other administrative sources.

3.2.3 Finland (*population: 5.2 million*)

Finland has had a computerised Population Information System since the early 1970s, with information both on individuals and on the households and buildings in which they live. A decennial census continues to be held, but since 1990 it has been based purely on administrative data - there is no actual fieldwork carried out. A number of registers, including those holding details on taxation, degrees and businesses, are linked to the population register using unique codes allocated to individuals, addresses and businesses.

3.2.4 Germany (*population: 82 million*)

Traditionally censuses have been carried out at the operational level by the Länder (provinces), which provide agreed tables to the Federal Statistical Office. It is therefore difficult for the Office to check the quality of the data or produce any additional analysis. Population registers, associated with the issuing of identity cards, are held at municipality level. Previously (in West Germany) the census was used to update the registers, but a court case in connection with the scheduled 1983 Census ruled against this dual statistical and administrative use of the census. The census eventually took place in 1987, with population register information used in cases of household non-response.

The 1987 Census was unpopular with the general public (partly because they believed the information was already available from population registers) and costly, and for these reasons the Government has not authorised another since (each census requires a specific piece of legislation). A 1 per cent 'microcensus' is currently carried out each year, which is used to update census figures and incorporates other sample surveys such as the labour force survey. The sampling, however, is based on the 1987 Census, which is now significantly out of date.

Two models have been proposed for the future, both of which involve the use of the population and employment registers. The 'Federal' model involves tabulating employment data for small areas but not linking it to the population register at individual level. The microcensus would continue to be used to gather other census-type information. In the 'Länder' model, data from the two registers would be linked at individual level and a reduced census taken which would ask only housing information and the names of residents. The Federal model is by far the cheaper of the two.

3.2.5 Norway (*population: 4.5 million*)

The 2001 Norwegian Census combined registers and questionnaires, but it is expected that future censuses will be purely register-based. The Central Population Register was used to pre-print names and addresses, which household members were asked to check and amend if necessary. There were also questions relating to the dwelling. All other information was supplied from the registers. It was possible to complete the form via the Internet.

Prior to 2001, flats within a single block in Norway did not have a unique address. Since these are separate dwellings for census purposes, it was necessary for them to be uniquely identifiable for a register-based census to be carried out. As part of the enumeration process in 2001, therefore, all flats were identified and given a unique number within the building. They can now be listed separately on population and address registers, and data derived about each flat individually.

3.2.6 Singapore (*population: 4 million*)

Singapore conducted a traditional census, approximately every 10 years, from 1871 to 1990. During the 1980s a 'people hub' was set up, which gave every individual a Unique Identification Number (UIN) and contained basic demographic and address information. For the 1990 Census information from this and other administrative sources was pre-printed onto census forms, with individuals being asked to verify the details. The results were then used as the basis of a Household Registration Database (HRD), which was established in 1996 and has since been updated quarterly from registrations of births, deaths, marriages, changes of address and other information.

In 2000 a primarily register-based census was conducted, with basic information extracted from the HRD and the National Dwellings Database. A 20 per cent sample survey was conducted to gather information not included in the databases, including relationships, education and employment. There were three methods of data collection, of which the Internet was used by 15 per cent of households in the sample. Besides the removal of the need for field staff, advantages to this method included immediate consistency checks and automatic coding. Any households in the sample that had not made an Internet return within a reasonable period of time were contacted by telephone, and if this failed after several attempts, interviewers were sent into the field.

3.2.7 Sweden (*population: 9 million*)

Sweden has long held a number of good administrative registers, and since 1996 work has been ongoing to integrate these into a single system that is capable of being used for statistical purposes. Four linked registers have been developed from the sources available, namely registers of population, jobs (and other activities), real estate and businesses. It was originally intended that the 2000 Census should be entirely register-based, but since the registers do not contain up-to-date information on household membership, it was necessary to collect this information via a questionnaire for those households that had moved address since the previous census. It is intended, however, that future censuses will be purely register-based.

4 Rolling census

An alternative to the traditional census model is to enumerate the population by means of a continuous, or rolling, survey, covering the whole country over a period of time rather than on one particular day. There are a number of possible ways of selecting the sample to be enumerated each year, for instance by covering one region in its entirety or by taking a sample from every area. There is never a complete enumeration of the whole country at one time, but the data collected each year can be used to model a more accurate national estimate than the inter-censal estimates derived from the traditional model.

Only one country is currently known to use the rolling census model, namely France.

4.1 Case study: France (*population: 60 million*)

4.1.1 History

Until 1999, France conducted a traditional census. However, there is no enshrined constitutional requirement for a census to be taken and fresh legislation was needed each time, which contributed to a periodicity that was not as regular as for most countries - it has generally been conducted every 7 to 9 years, mainly to avoid clashing with France's electoral cycle (there are three separate national elections, which until recently were held every 3, 5 and 7 years). In addition, there has often been trouble obtaining funding from the government, which has led, on occasions, to the census being postponed.

Each year, the French statistical office (INSEE) is required to publish a 'legal' population count for each commune (an administrative area). Although a 'statistical' population count - a demographic estimate adjusted for births, deaths

and migration, similar to the method used in the UK - is also produced each year, the most recent census figure has until now been the one used for legal purposes. Clearly, this can result in public money being distributed on the basis of population figures as much as a decade out of date - a commune with a growing population can ask for an inter-censal recount, but this rarely occurs. This is one of the factors that led to INSEE moving away from the traditional model and towards a continuous survey - others included a government requirement to smooth out the peaks in spending, and a desire to be able to incorporate technological changes as they occur rather than redesign all systems before each census.

4.1.2 The new model

Strictly speaking, a rolling census would cover every household in the country over a given period of time (such as 5 or 10 years). The model proposed for France would perhaps be better described as a rolling survey, as a large proportion of the population would not be counted at all in the five-year cycle. At time of writing the continuous survey is expected to begin in 2004, with the first full results in 2008.

The main unit of local government in France is the commune. These vary enormously in population, from 350,000 to less than 50. The treatment of a commune under the French model depends on its size - a true census will be carried out only in those with a population of less than 10,000, each of which will be enumerated once every five years. In the larger communes, accounting for roughly half the population, a sample survey will be taken, although in this case any given commune will be sampled every year. The sampling fraction is 8 per cent, so in any given five-year period 40 per cent of the population will be enumerated in communes with over 10,000 residents and the entire population enumerated elsewhere.

The sample survey element of the process requires an accurate sampling frame, and a building register was created for this purpose. This includes geographical information as well as address and type of building, be it a residential property, a communal establishment or business premises. The 1999 Census was used to create the initial register, which is continually updated using information such as construction and demolition permits, along with other administrative sources and surveys.

France has 22 regions, and in each of these regions the smaller communes are stratified according to socio-demographic makeup. This is used to divide the communes into five rotation groups, each broadly representative of the region as a whole. A full census of one rotation group will be taken each year. The larger communes are stratified at a lower level, namely blocks of about 2,000 people. These are used to create five representative rotation groups within each commune, with one rotation group to be surveyed each year. A random sample will be drawn from within that year's rotation group.

The household enumeration exercise each year will take place during January and February. As has been the case for traditional censuses in France, the actual enumeration will be handled by the communes (or groups of communes in the case of those with a small population), with the drop-off/pick-up method being used. Communal establishments, however, will be enumerated separately. The sampling and data collection methodologies are essentially the same as for households, but the enumeration will be carried out centrally by INSEE and will take place in March. In addition, telephone interviewing may be used to update the population counts of the 80 per cent of communal establishments not enumerated in a given year.

One fifth of industrial buildings will be enumerated each year. The purpose of this exercise is to verify that they do not contain any residential dwellings, such as converted flats or caretaker's accommodation. Finally, all new residential buildings will be enumerated each year so that they can be placed into a rotation group.

The annual population estimate for each of the larger communes will be based on the sample survey taken that year. Rather than simply gross up the count according to the sampling fraction, households will be weighted according to the numbers on the buildings register. This should correct for households that do not return a form and for changes between the drawing of the sample and the most recent update of the register. For the smaller communes, an accurate population count can be made available for those communes sampled in the current year. Administrative records can be used to estimate undercount. Administrative records are also used to produce estimates for years when a census

is not taken in a particular commune. These estimates may subsequently be revised based on the following census.

Although this model is based primarily on a continuous survey, good administrative records are crucial to its success. In particular, it will be vital, for example, that the building register is kept up to date.

5 Conclusions

The research conducted for this review suggests that, outside of the developed world, the traditional census is the only model used. Even in developed countries, it is generally only within western Europe that a complete enumeration of the population is not always attempted, the major exception being Singapore with a register-based census supplemented by a 1 in 5 sample survey. Otherwise, it is principally in the north continental European countries that administrative records are the primary source of census-type data, with all the Scandinavian countries having good population registers that are already, or are about to be, the sole means of enumerating the population. Belgium is also moving in this direction (as, probably, is Germany), while the Netherlands has a unique model that makes good use of all the sources available - population registers, other administrative sources and sample surveys. France uses a rolling model which is part census and part sample survey, although this still depends on at least one good administrative source (a buildings register).

Of those developed nations that still use the traditional model, many are looking to make increased use of administrative sources in the future. Italy, for instance, already uses population registers to estimate the level of under-enumeration, and is looking to develop a more standardised system that can be used to conduct register-based censuses. The USA, meanwhile, is moving towards a 'head count plus survey' model. The USA is also conducting experiments into the possibility of a primarily records-based census, despite the fact that, unlike the countries mentioned above, the USA does not have a compulsory population register.

It seems that no two countries are exactly alike in the way they gather census information, even for those that still conduct a traditional census. Equally, not every model would work for every country - factors that may affect the choice of method include population size, quality of administrative records, population base issues,

public perception, political climate, funding, legislative and constitutional issues and the existence of a reliable population register, all of which will need to be taken into account when deciding on the way forward for census taking in the United Kingdom. Although no one model can be lifted wholesale from another country, we can assess whether each model is feasible in the particular circumstances of the UK and how it may need to be adapted.