

Censuses in a register-based statistical system: Norwegian experiences

Andersen, Espen

Statistics Norway, Department of Social Statistics

2225 Kongsvinger, Norway

E-mail: esa@ssb.no

Utne, Harald

Statistics Norway, Department of Social Statistics

Pb.8131 Dep 0033 Oslo, Norway

E-mail: hut@ssb.no

The last fully traditional census in Norway was conducted in 1960. In 1964 the Central Population Register (CPR) was established, introducing a unique personal identification number. In the following years a number of administrative registers were established, all using the same identification number. Statistics Norway developed register-based statistics in several fields, and the role of register data in censuses increased over the years. In 2001 only information on households and dwellings was collected from census forms. The 2011 Population and Housing Census in Norway will be totally register-based.

1. The Census strategy of Statistics Norway

During the last decades it has been a major concern for Statistics Norway to rationalize data collection and to reduce the response burden. To reach these goals an increased use of administrative data has been a significant element of the strategy. When it comes to the Census and use of administrative data sources there are two main trends in the development of register-based censuses in Norway. The *first* one is that censuses have been used quite extensively in establishing registers. The *second* trend is the step-by-step development. In the first steps subject matter statistics are tested and published in different areas. Register-based variables are introduced in the census as soon as the quality is considered sufficient. When statistics has been developed for all areas relevant for censuses, a totally register-based census can be conducted. These same trends have been observed in other countries developing register-based censuses.

The first Norwegian population census dates back more than 200 years. Up until the 1960s, all censuses were traditional. As mentioned, in 1964 the Central Population Register (CPR) was established, and an official, personal identification number was introduced. The 1960 Census provided the basis for allocating the personal identification number. In 1970 the census was used to establish the Register of Education. In the 1980 Census, information on demography, education and income, as well as geographical characteristics, was collected from registers. Administrative registers were also used to rationalise data collection, and data capture by postal mail only was introduced. In the 2001 Census all data on persons (place of residence, demography, education, labour market, income etc.) were based entirely on registers. Data on families, households and housing was based on a full coverage survey with a questionnaire for every household.

To conduct a fully register-based census a complete dwelling register was needed. Statistics Norway has played an active role in arguing for such a register since the 1980's. In 1999 it was decided to

allocate a unique address¹ to all dwellings in the Cadastre. A joint project involving the authorities responsible for the Cadastre and the CPR and Statistics Norway was established in year 2000 and Statistics Norway was given the role as project manager. The project was indeed included as an important part of the 2001 Census project and the supervisor of the project was also the leader of the Census. The census forms were used to establish links between resident persons and their dwellings (in the CPR) and to update information on dwellings in the Cadastre.

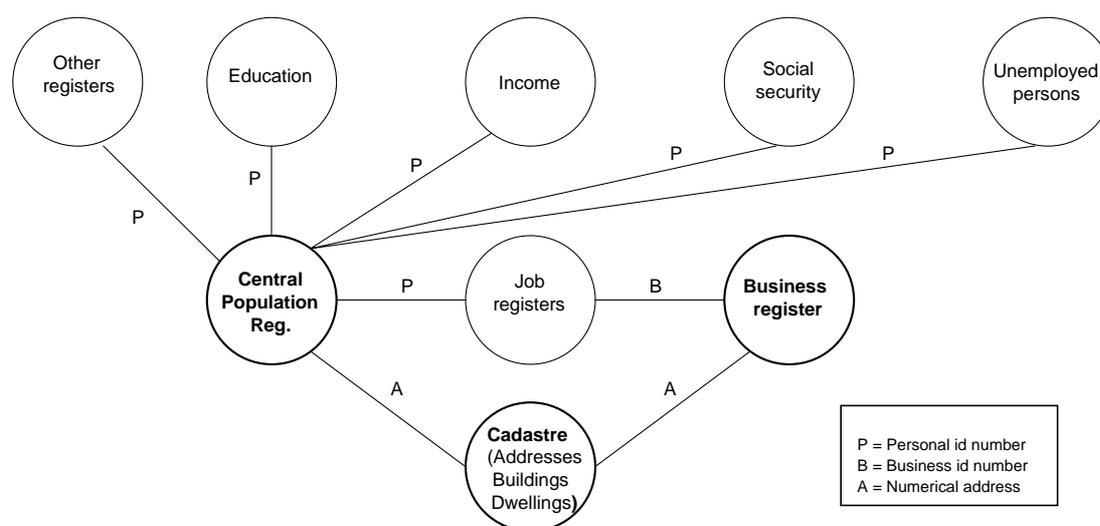
2. The Norwegian Census system

To conduct a register-based census a solid legal base is needed. In Norway the Statistics Act (Act No. 54 of 16 June 1989) gives Statistics Norway the right to use administrative data as basis for official statistics. Furthermore, Statistics Norway must be notified when a new administrative register is established or when major modifications of an existing register are being planned. Statistics Norway must also be notified as early as possible to make it possible to suggest changes in proposed plans. The Statistics Act gives Statistics Norway substantial influence on design (and redesign) of administrative registers.

In Norway a number of statistical registers have been created by processing data from an even larger number of administrative registers. A major objective has been to integrate these statistical registers into *one* system. One statistical register may be used as a source in different subject matter statistics. For instance, demographic variables are produced in the statistical population register and are used in all kinds of social statistics. Population and housing censuses are an integrated part of this system of register-based information. The same statistical registers are used as sources for censuses and for subject matter statistics.

Data for the 2011 Census will be extracted from a register-based statistical system that is common for a number of statistics. The system is built around three basic registers: The Central Population Register, the Business Register and the Cadastre. In addition to the information contained in the basic registers, a number of other registers are used in the census, for example registers on employment and unemployment, education and income. In total about 30 major register sources will be used and in addition a number of minor registers. All statistical units relevant for a census are linked to one another by using common identifiers: identification numbers for persons and businesses and numerical addresses.

Figure 1. The Norwegian census system



¹ For an address to be unique in a multi-dwelling building it must comprise a dwelling number in addition to the street address.

Figure 1 provides a simplified description of the census system. Persons, families and households in the population register are linked to dwellings and buildings by the numerical address. Through job registers employed persons are linked to their place of work in the business register and the different registers related to persons are linked to the CPR.

Even though the census is based on data from registers, most census variables are not directly available from administrative sources. It is necessary for Statistics Norway to combine data from different administrative registers, and data are further processed to obtain the variables required for censuses. For instance, to construct the variable *employed persons*, data from 11 different registers are used. This is partly because no single register covers all kinds of employment, partly because it is necessary to combine several sources to obtain sufficient data quality and coverage.

2.1 Advantages

The main advantages of register-based censuses are lower costs and reduced response burden. A register-based Census is based on the same sources as annual subject matter statistics. Consequently, the marginal costs of producing a census are low compared to a traditional census. The reduction of response burden is substantial since a census is covering the whole population.

Another advantage of a register-based approach is that census statistics and annual statistics are directly comparable, or to put it another way: Census type of statistics can be produced currently, for most variables annually. This has been considered a major advantage by many users.

2.2 Disadvantages

In a register-based census only variables that are included in registers or could be derived from register variables, are available. In many countries the population census is an important tool for data collection that is used to meet emerging information needs. This flexibility is lost when data are no longer collected by means of questionnaires. Consequently, this type of information will only be available from sample surveys.

Additionally, some restrictions are imposed on the definitions of units and variables. In a register-based census a private household is defined as all persons living in the same dwelling (*dwelling household*). *Housekeeping units* (persons living in the same dwelling with joint board) cannot be included. The population is normally counted according to *legal* place of residence (according to the population register) and not *de facto* place of residence. For some users, housekeeping units based on *de facto* place of residence would be more appropriate for their analyses.

3. The 2011 Census, organization and time-schedule

A traditional census is a huge operation. A lot of resources are needed for designing and testing questionnaires, collecting and processing data and for public information. When using register-based data, census costs are considerably reduced. This is not to say that register-based statistics are "free of charge". To transform administrative data into statistical data involves quite a lot of work. However, since different subject matter statistics as well as census statistics are based on the same data, the additional costs of producing a register-based census are not very high.

3.1 Organization

Unlike the previous censuses in Norway, the 2011 Population and Housing Census is not organized as a separate unit, but as a project in Division for Population Statistics. The project group consists of 2-3 persons only. The size of the project group is reflecting the role of the Census in the register-based census system as a coordinator of different census related projects conducted by different units within

Statistics Norway. The total budget of the Census is 1.5 mil. Euros. This is the additional costs and it's about 10 per cent of the costs of the 2001 Census.

The Census project is divided into four target areas. Firstly, to collect data from the different statistical registers and develop an *IT system* to combine these data into a census file. Secondly, to *develop the topics* required in the EU-regulation and to improve the quality of some the topics already existing. Thirdly, to measure and describe quality, including preparations for the *quality reports* and *metadata* descriptions required in the EU-regulation. Fourthly, *dissemination*, the most important task is to make the *hyper cubes*² to be published by EUROSTAT's census hub system. Census dissemination at a national level will be harmonized with releases from annual statistics. A main focus for the Census will be small area statistics and GIS-based products.

3.2 Time-schedule

To take the step from a combined approach in 2001 to a fully register based approach in 2011 several improvements have been implemented or are about to be implemented. The most important has been to improve the quality of data as recorded in administrative registers. This is the responsibility of the register owners, but Statistics Norway also plays an important role both formally (with reference to the Statistics Act) and practically by everyday cooperation. Furthermore, the processing of data in Statistics Norway, transforming administrative data into statistical data, is being improved in several ways. New derived variables are created by combining data from several sources.

In the time up to the Census the main focus is on improving the data to be used. In addition a test will be conducted by using data from 2010. The Census day is set to 19. November 2011, which means data from the different registers will refer to this day³. However, the actual extraction of data from the registers will take place somewhat later. This time lag depends on the procedures for updating the different administrative register; from a few weeks for the population register to more than a year for taxation data. Furthermore, data processing in the statistical registers takes some time. This means that the first data will be available for the Census by April 2012. The rest of the data will be available throughout the 2012 and early 2013 (income data). The main activity throughout 2012 is the development of the census files. In 2013 the main activities will be development of the hyper cubes and quality reports. In addition it will be published statistics from the Census at national level for some selected areas throughout the period 2012 – 2013. The project is planned to be finished by the end of March 2014.

4. The 2011 Census, challenges

In 2008 the European Parliament approved a regulation on Population and Housing Censuses for the European Union⁴. The regulation is based on the CES⁵ recommendations. However, replacing the former “gentlemen’s agreement” by a European law means that the importance of fulfilling all requirements for census data is more important than before. The regulation provides for the specification of outputs, the means of submission of these to EUROSTAT, and the requirement to provide metadata and quality reports.

In a register-based statistical system most of the census topics are already produced and published by subject matter statistics. This makes the international perspective, which is to meet the requirements of the EU Regulation, the main challenge in conducting a Population and Housing Census in Norway.

The CES Recommendations identifies the following dimensions of quality in statistics: Relevance, accuracy, timeliness, accessibility interpretability and coherence. In this paper we will mainly discuss

²A “hypercube” is a multidimensional cross tabulation of breakdowns which contains a cell value for the measurement of each category of each breakdown.

³Some data will have a different reference date, for instance data on education is pr. October 1.

⁴The regulation is also binding for Norway as a member of the European Economic Area.

⁵Conference of European Statisticians

the dimensions of relevance and accuracy when using administrative registers as the source for census data.

When conducting a traditional census, using census forms, the statistical office may ask questions directly corresponding to the definitions given in international recommendations. In this way the census data will relate to the concepts that it is desired to describe, in other words the *relevance* is high. When using register-based data it is always important to bear in mind that the data is originally collected for another purpose than making statistics. This may pose some restrictions on definitions. To obtain a high degree of *accuracy* is a challenge independent of data collection methods used, but the type of problems may be somewhat different when using register-based data. Topics that are important for the administrative agencies responsible for the registers are normally recorded with a high degree of accuracy.

To secure, or improve, the quality of administrative data close cooperation with the different register owners is vital. Statistics Norway plays an active role in giving support to administrative registers: technical support (reporting quality problems) as well as political support (call for resources if needed). It is also important that both parties benefit of the cooperation. Statistics Norway has also been an initiator of the establishment of different forums where register owners and users meet to discuss relevant issues related to administrative registers on a regular basis.

4.1 Defining the resident population

Since the establishment of the CPR in 1964 the CPR has been the only source when it comes to defining the resident population and the place of residence in Censuses. This also means that in defining the statistical population, we are fully dependent on the rules for registration in the CPR. The population register and the dwelling register constitute the back-bone of a register-based population and housing census. The fact that we use the Central Population Register (CPR) to enumerate the resident population has a great impact on most of the census topics.

Place of usual residence defines the resident population of each country and each region within the country. According to the EU regulation *place of usual residence* is the place of living at the time of census (de facto), but registered or legal residence (de jure) can be used if the circumstances defining place of usual residence can not be established.

In a register-based census the legal residence is the most obvious option. For most people the legal residential address are the same as the de facto address, but for some groups there may be deviations. According to the EU-regulation only persons who have lived in their place of usual residence for a continuous period of at least twelve months (or if not, have the intension of staying there for at least a year), should be regarded as usual resident.

In the instructions for updating the Norwegian CPR there are similar requirements with regard to where a person spends most of his/her daily rest. However, the corresponding rule of duration is six months. This of course has some consequences for the count of the resident population in Norway. Furthermore, there are some administrative rules and practices that do not fully comply with the regulation. Students may, according to instructions for population registration, choose where they want to be registered. It is rather common for students to be registered at their parents' home address even if they are actually living in a different household, and in fact most students do so. According to the EU-regulation it is acceptable to consider students family home as their usual residence, but this is not the case for persons studying abroad.

There are similar types of deviations when it comes to people living in institutions. According to the CPR instructions spouses should normally be registered at the same address in the CPR even if one of them actually is living in an institution such as homes for the elderly. Persons who are not married are normally registered as residents in the institution.

Using the concept of legal residence of course implies that persons staying illegally in the country will not be counted. However, this group will be difficult to count precisely in traditional censuses as well. But even some persons staying legally in the country are not registered as residents in the CPR, for instance most asylum seekers.

In the 2011 Census comparability between countries is a main issue. Harmonisation of the definitions of resident population according to the definition of place of usual residence is very important. To achieve this Statistics Norway consider using several additional data sources.

4.1.1 De facto population

In Norway it will be possible to use additional data sources (registers) to give a better picture of the de facto resident population. This information will however, only be available for certain groups of the population and the quality of the registers may vary. Statistics Norway has investigated and is still investigating different data sources with the objective to produce data on de facto residence for some groups, mainly students and persons resident in homes for the elderly. By using other administrative data sources in addition to the CPR, it could in principle also be possible to include asylum seekers and exclude persons studying abroad from the resident population. After a comprehensive investigation it is decided that the Census population will be the resident population according to the CPR (legal residents). This means that for instance asylum seekers will not be included, mainly because the data quality for this group is not satisfactory. It is decided to use de facto address for students studying within Norway and persons resident in homes for the elderly if the data quality is satisfactory by the time of the Census in November 2011.

The decision to use additional sources to define the population for certain groups in the 2011 Census brings up new issues. Moving towards “de facto” population will improve the relevance of the address data. However, these “new” register sources give less accurate data than the CPR. For many persons we will have the street address only. These addresses do not point to a unique dwelling as is the case with the address data from the CPR. In order to establish households when including this data, some kind of imputing or estimation will have to be used. Another challenge is related to updating. Statistics Norway gets updates from the CPR on a daily basis. This is not the case for the other registers that can be used to define the census population. Despite these problems Statistics Norway are still working towards the goal to publish “de facto” population for selected groups in the 2011 Census.

Since data from additional sources are substantially delayed compared to data from CPR, the annual and quarterly population statistics will have to be based on the CPR only, at least in the near future. Using other sources in the census therefore will give population counts different from the population annual statistics. This means that improving the relevance in this case will have a negative impact on comparability between statistics. It will also be a break in the time series compared with previous censuses.

4.2 Household and housing

The lack of full coordination of the addresses in the CPR and the Cadastre has made it impossible for Statistics Norway to conduct a fully register-based Census until 2011. The link between these two registers is decisive to derive statistics on households, which is an essential unit of every census. The main challenge has been related to people living in multi-dwelling buildings. As mentioned in chapter 1 the 2001 Census was used to establish this link for people living in multi-dwelling buildings through a joint project between Tax Norway (responsible for the CPR), The Norwegian Mapping Authorities (responsible for the Cadastre) and Statistics Norway. Statistics Norway was in charge of the project and the project was included as an important part of the Census 2001 project. There were several reasons for Statistics Norway to take the role as leader of this project. First, Statistics Norway had historically been the driving force in establishing a complete dwelling register. On this basis, it was not unnatural that Statistics Norway took this role to reach the goal of updating the Cadastre to include unique addresses for every dwelling in Norway. Second, there was also an economic component that supported a project led by Statistics Norway. The cost was estimated to be

significantly lower since the project was a part of the 2001 Census. The main reason of the cost reduction was related to postage since the census was used for data collection.

The introduction of a unique address to all dwellings made it possible to establish links between dwellings and resident persons in the CPR with the address as the key variable. The first period after the introduction of the unique address for all residents the quality was not satisfactory in the registers (CPR and Cadastre). This made it difficult to develop a register-based statistics on households. We have experienced a major improvement of the quality especially over the last two years. Today 95 per cent of the residents living in multi-dwelling buildings are registered with a unique address in the CPR.⁶ This means that approximately 99 per cent of the whole population is registered with a unique address in the register. This is a good starting point for a register-based census, but good coverage in a register does not automatically mean that we can produce register-based statistics with satisfactory quality.

A complete register-based statistics on families and households was published for the first time for the year 2005. Due to the quality problems described in the previous section, information from the 2001 Census and weighting in addition to the administrative registers had to be used to produce household statistics of good quality.

A register-based dwelling statistics was first published for the year 2006. The dwelling statistics has gradually been extended and will by 2012 cover all the core housing topics. One challenge still to be solved in 2011 is to classify dwellings as occupied or vacant. To solve this challenge we need to harmonize the dwelling statistics and the statistics on families and households. As mentioned above the present household statistics is not based on a register count only. Statistics Norway is currently developing a method for harmonization between household data and data on dwellings. According to the definitions, the number of (dwelling) households must be the same as the number of occupied dwellings. This work will be finished in time for the 2011 Census. This harmonization can result in a break in the time series for households.

In the Norwegian CPR most people are registered as living in conventional dwellings. Only very few persons are registered as *homeless* (with “no fixed abode”) in the CPR. Most de facto homeless persons is in fact registered at a legal address, for instance with friends or relatives. They may have some sort of connection to this household, but they do not spend most of their daily rests there. Since there is no available register sources to cover this topic a survey carried out by the Norwegian State Housing Bank will be used to give an estimate on the number of homeless people. This estimate will be included in the quality reports. No one will be classified as homeless in the Census, but persons registered with “no fixed abode” in the CPR may be classified as “not stated” in some tables.

When conducting a register-based census almost all persons will be classified as occupants of conventional dwellings even if they de facto are living in other types of housing arrangements. The group *living in other housing units* (that is other than conventional dwellings) is almost non-existing when using CPR as the main source. According to the body of rules for registration in the CPR it should not be possible to register people in second homes (for instance summer cottage, boats/piers, cabins, caravans and tents), but in practice this is done if this is the actual place of residence and there are no alternatives. In the Census we can retrieve housing information for some of the cottages in the Cadastre. It is also decided to classify cottages where people are registered as resident as conventional dwellings. The majority of these dwellings have a standard equivalent to a conventional dwelling.

4.3 Other topics

A national register of employees was established in 1978, but *occupation* was not included in the register before 2001. *Occupation* was planned as a register-based topic in the 2001 Census, but the

⁶ In 2001 less than 50 per cent of the population registered in a multi-dwelling building was registered with a unique address in the CPR

coverage was too low and it was not possible to include information on occupation in the census. The coverage has improved and occupation will be included in the 2011 Census.

A well known problem when using register data to produce statistics is the lack of data describing the situation before the register in question was established. In the census this is in particular associated with topics like *place of birth*. In Norway we do not have this kind of information for the period before 1964 when the CPR was established. We have the same problem with the topic *year of arrival in the country* but this is not a problem in the census since reporting is limited to events after 1980. In the Cadastre we had much of the same situation concerning buildings and dwellings built before the register was established (1983), but the 2001 Census was used to update the statistical version of the register. This type of imperfections is gradually getting less important as years goes by.

Some variables are not available directly from any register, but could be derived by combining different kinds of register information. In Norway there exists no “cohabiting register”. To determine whether two persons living in the same household are cohabitants or not, we use information on the resident address, whether the persons have a child in common, if the persons are relatives, the mutual age difference and the date of moving into the dwelling. This method makes it possible to identify cohabiting couples of opposite sex in a reliable way. However the method does not allow for identifying same sex cohabiting couples⁷. This example shows one of the problems of identifying relationships between members of a household. We do, of course, not have a register that cover people’s emotional relationship to each other.

A problem area is education taken by immigrants before they entered Norway. In 1990 and 1999 Statistics Norway carried out surveys among immigrants asking about their education. The register-based information in this field is still not satisfactory. Statistics Norway has received founding to conduct a survey in 2011 to collect data for immigrants arrived after 1999. In a longer perspective we will hopefully be able to collect this information using administrative sources.

4.4 Topics difficult to cover in a register-based Census system

It is not possible, and in some cases not even desirable for the society, to have administrative registers covering all topics relevant for statistical purposes. We have already mentioned the example of *cohabitation*, and other examples are *homemakers* and *family workers without contract of employment*. For this type of variables we do however have some relevant register information, and imputation or estimation of values will, at least in principle, be possible. Other topics are related to human behavior and attitudes, examples being *mode of transport to work/school*, *ethnicity*, and *type of energy used for heating*. This kind of topics will probably never be covered by register data, and combining data from sample surveys and registers will be the only option. None of these topics will be developed for the 2011 Census, but Statistics Norway is investigating the possibilities to develop some of these topics (primarily *type of energy used for heating*) by using administrative data in a longer perspective.

4.5 Use of additional data in the 2011 Census

For the 2011 Census, as for previous censuses, Statistics Norway is planning for a quality assurance programme. This program will comprise quality monitoring in all parts of the process. One important part is a quality assessment of census data, comparing register-based data with data from other sources. As in the 2001 Census we will use data from existing sample surveys (as the Labour Force Survey and EU-SILC) as well as survey data collected especially for the census. The plan is, however, to use this data only for quality assessments and not for improving census data.

⁷Same sex registered partners and married couples are identified in the CPR

5. Summary

The 2011 Population and Housing Census in Norway will be totally register-based. This is a result of a long term strategy to rationalize data collection by increasing the use of administrative data. This strategy is also reflected in censuses where the use of administrative data has increased gradually since 1970. The 2011 Census project group is small and consists of 2-3 persons. This reflects the role of a Census in the Norwegian census system; a coordinator of different census related projects which is carried out by other units within Statistics Norway. Since most of the Census topics already are covered by subject matter statistics, the international perspective and how to meet the requirements of the EU-regulation on Population and Housing Censuses is the most important challenge. When it comes to dissemination, the construction of the census hubs is the most important task. Dissemination at the national level will have the main focus on small area statistics and GIS-based products. There are several challenges related to meet the requirements using a fully register-based approach. In Norway the greatest challenge has been to establish households which are coordinated with the dwelling statistics. Another challenge is defining the resident population since the administrative rules and practice do not fully comply with the regulation. To give a better picture of the de facto population, Statistics Norway plan to use additional sources to the CPR to define the census population for residents. Despite some challenges Statistics Norway is ready to conduct a register-based Population and Housing Census in 2011.

References

Andersen E and Utne H (2007): The first register-based Census in Norway in 2011: How to comply with international recommendations? Invited paper: Registers in Statistics -methodology and quality 21 - 23 May, 2007 Helsinki

Longva S., Thomsen I. and Severeide P.I. (1998): Reducing Costs of Censuses in Norway through Use of Administrative Registers. Reprint from International Statistical Review, Vol. 66, 223-234.

UNECE (2006): Conference of European Statisticians Recommendation for the 2010 Censuses of Population and Housing.

UNECE (2007): Register-based statistics in the Nordic countries. Review of best practices with focus on population and social statistics.