

An overview of progress on the potential use of administrative data for census information in New Zealand

Census Transformation programme



Crown copyright ©

This work is licensed under the [Creative Commons Attribution 3.0 New Zealand](#) licence. You are free to copy, distribute, and adapt the work, as long as you attribute the work to Statistics NZ and abide by the other licence terms. Please note you may not use any departmental or governmental emblem, logo, or coat of arms in any way that infringes any provision of the [Flags, Emblems, and Names Protection Act 1981](#). Use the wording 'Statistics New Zealand' in your attribution, not the Statistics NZ logo.

Liability

While all care and diligence has been used in processing, analysing, and extracting data and information in this publication, Statistics New Zealand gives no warranty it is error free and will not be liable for any loss or damage suffered by the use directly, or indirectly, of the information in this publication.

Citation

Statistics New Zealand (2014). *An overview of progress on the potential use of administrative data for census information in New Zealand: Census Transformation programme*. Available from www.stats.govt.nz.

ISBN 978-0-478-42904-6 (online)

Published in July 2014 by

Statistics New Zealand
Tatauranga Aotearoa
Wellington, New Zealand

Contact

Statistics New Zealand Information Centre: info@stats.govt.nz
Phone toll-free 0508 525 525
Phone international +64 4 931 4610
www.stats.govt.nz



Contents

List of tables and figures	4
1 Purpose and summary	5
2 Introduction	6
Census transformation strategy	6
What is census information?	6
What is administrative data?	7
Why investigate administrative data for census?	7
Key challenges in producing census information from administrative data	8
International trends in using administrative data for census	8
3 Census Transformation programme: A phased and iterative approach	10
Developing options and assessment criteria	10
4 Results	11
Using administrative data to count the population and identify where people live	11
Using administrative data for information about the population (attributes)	15
How an administrative census could work in New Zealand	16
A national population register	16
5 Towards the future	18
Future pathway for the New Zealand census	18
Alignment with cross-government initiatives	19
6 Conclusion	21
Next steps	21
References	23



List of tables and figures

Figures by chapter

4 Results	11
Figure 1: National coverage of three administrative datasets compared to the Estimated Resident Population (ERP), males and females by single year of age, June 2012	11
Figure 2: National coverage of linked data (IDI tax, education, and external migration) compared to the Estimated Resident Population (ERP), by single year of age and sex, June 2010	12
Figure 3: Territorial authority areas: coverage of linked data (IDI tax, education, and external migration) compared to the Estimated Resident Population, June 2010.....	13
Figure 4: Selected Territorial Authority areas: coverage of linked data (IDI tax, education and external migration) compared to the Estimated Resident Population (ERP), by single year of age, June 2010.....	14
5 Towards the future	18
Figure 5: New Zealand census pathway to the future.....	18



1 Purpose and summary

This paper provides an update on the progress of Statistics New Zealand's investigation of administrative data sources as a means of providing census information in the future.

In March 2012 the New Zealand Government agreed to a census transformation strategy which has two strands: to focus in the short to medium term on modernising the current census model and creating efficiencies; and to investigate alternative ways of producing small-area population and socio-demographic statistics in the long term. This includes the possibility of changing the census frequency to every 10 years, and exploring the feasibility of producing census information from administrative data sources in the longer term.

Early investigations show that existing administrative data sources cannot at present act as a replacement for current census and population statistics, but have also indicated some good potential uses in the current census model. A future pathway has been identified that builds from a modernised digital form of the current census to a possible administrative census in future.

Two approaches appear to offer feasible solutions for an administrative census if barriers can be overcome:

- Creation of a **national population register** where the administrative function is closely aligned with the census need to count the population. This approach is the basis for successful international examples. However in the New Zealand context, other benefits for government would need to be identified and consideration given to privacy and cultural issues. It is Statistics NZ assumption that there will be no New Zealand national population register in the next 10 to 20 years.
- Alternatively, **linking multiple existing administrative data sources** shows promise as a statistical solution to produce population counts that would avoid the need to survey everyone. This approach is the current focus of our investigations.

In either case, a sample survey component would also be needed. A coverage survey is necessary to provide an independent validation of the administrative data. A large scale sample survey may be necessary for census information which cannot be obtained from administrative sources.

While there are major limitations in the administrative sources available at present, recent government initiatives within Better Public Services and the Government ICT Strategy offer real opportunities for enhancing the statistical value of administrative data and systems. To contribute significantly to the ability to use government data for statistical purposes in general, new developments would need to recognise the basic principle that statistical use is built in by design.

The work programme will continue through 2014 and 2015, with a further report in late 2015, sufficient to enable the Government to agree a preferred future direction for the New Zealand Census.



2 Introduction

Census transformation strategy

The census transformation strategy was agreed with the Government in March 2012. The strategy and considerations feeding into it are described in [Transforming the New Zealand Census of Population and Dwellings: Issues, options, and strategy](#) (Statistics NZ, 2012). In summary, the strategy calls for:

- a focus in the short to medium term on modernising the current census model and creating efficiencies
- a work programme to investigate alternative ways of producing small-area population and socio-demographic statistics in the long term. This includes the possibility of changing the census frequency to every 10 years and exploring the feasibility of producing census information from administrative data sources in the long term.

A phased approach through to the mid-2020s was proposed to manage the implementation of the census transformation strategy and enable the Government to make future investment choices as the options become clearer and the associated risks better understood.

For the first phase, a three-year work plan has been developed within Statistics NZ's Census Transformation programme. This information paper provides an overview of early progress made on the long-term strand of the strategy: investigations into the feasibility of using administrative sources to produce census information. These early findings have implications for the next census and the future census pathway, and have formed the basis for the first progress report to the Government.

What is census information?

The census is about counting everyone, once, only once, and in the right place. It is about counting people in the place where they live, and about creating a picture of communities. This means local communities in the geographic sense, and also small but distinct and important groups such as iwi, Pacific peoples, new migrants, single income or single parent families, and crowded households.

Census information can be split into two types: **counts** and **attributes**.

Counts: Includes counts of dwellings and people, down to small geographic areas.

Attributes: The census provides a range of socio-economic information about small sub-populations and for small geographic areas. Attributes include information about education, work and income, ethnicity and culture, health, housing, households, and families.

There is an enduring need for the finely grained information about small area population and socio-economic characteristics currently based on the periodic census.

The current census model in New Zealand is a full-enumeration survey of the population conducted every five years. A coverage survey (the Post-Enumeration Survey) is held immediately after the census to measure net census undercount.

The census counts form the basis for official population statistics¹. Population estimates are updated between censuses using measures of population change derived from administrative data from birth and death registrations and external migration. While several additional administrative sources inform the updated population estimates for geographic areas within New Zealand, these subnational population estimates are much more difficult to produce due to the lack of a direct measure of movements within New Zealand. Determining the change in population is especially important in New Zealand because of the high mobility of New Zealanders.

The census and the same administrative sources also form the basis for projections of future populations.

Producing census and population statistics information is now an integrated system that relies on a combination of a full census survey, sample survey, and administrative data. Investigations into an administrative census are about how this information might be produced in future, and to what quality. The question is to understand where the most effective balance of survey and administrative data lies.

What is administrative data?

Administrative data can be defined as data that is collected by government agencies or private organisations in the course of conducting their business or services. It is data that is not collected primarily for statistical purposes. Rather it is collected for operations such as delivering a 'service' (eg the health service, tax payments), or legal requirements to register events (eg births, deaths, marriages) or as a record of transactions or events (eg overseas travel journeys). The population (*who* is included in the data source) and data content (*what* information is collected) is defined by the organisation responsible and they have primary control of the methods by which the administrative data are collected and processed. As a result, administrative data differs in nature, scope, and quality from data collected through the census or surveys, where control of who is asked for what information is in the hands of the statistical agency.

A rapid increase in digitisation has meant that administrative data has become more abundant and more accessible. Combined with improvements in technology and methods of combining and analysing large datasets, this has resulted in more opportunities for using it for official statistics. The more recent 'tsunami' of big data sources and the new methods being applied by the commercial sector add another dimension.

Why investigate administrative data for census?

There has been a long history of using administrative data for official statistics. The benefits of making wider re-use of administrative data for statistical purposes are well-recognised internationally and [Statistics New Zealand Strategic Plan 2010–20](#) includes a strategic priority to maximise the use of administrative data in producing official statistics. The main benefits of using administrative data for official statistics are:

- reduced collection costs
- reduced respondent burden by using the same data multiple times
- improving quality and increasing the range of research possibilities and relevant statistical outputs.

The potential for a major step change in reducing the cost of census information is the main driver for considering a shift to an administrative census in the long term. Another

¹Population estimates or the estimated resident population (ERP), and population projections can be found at: http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections.aspx

advantage is the potential to increase the frequency of census information from what can be provided in a periodic census held every five years or every 10 years. There is also potential for selective use of administrative data to reduce costs, reduce respondent burden, and improve data quality in the current census model. At the same time there are constraints and risks associated with using administrative data. In the census context the main risks are loss of respondent trust, a lowering of data quality, a narrower range of variables, and the potential for disruption to supply.

Key challenges in producing census information from administrative data

In producing census information from administrative data, the key challenges are related to the comprehensive nature of census information and the high quality standards needed to meet critical uses. These census requirements can be in opposition to the nature of administrative data. While the specifics of what is essential census information are yet to be tested, it is clear that three core functions need to be met:

- achieve comprehensive coverage of the New Zealand population including core demographic variables (age, sex, ethnicity, Māori descent)
- ensure that people can be located where they usually live
- ensure the dataset has sufficient personal and household attributes to inform a range of programme development, investment, and resource allocation needs.

The first requirement of a census, to count everyone only once, is central. Difficulties arise because administrative data is not comprehensive by its nature (health data is a health service register, education data are mainly children and young adults) and may include too many people, while missing others.

In addition to coverage issues, if information is not important or necessary for an administrative purpose, data may not be collected or may be of poor quality. This affects the need to locate people where they live. For example, as agencies move services on-line, residential addresses become less important as a contact point and may not be captured well.

And it has a major effect on the extent and usefulness of personal and household information. For example, when ethnicity is not a determinant of programme participation, it may not be asked, or if it is asked, there is less incentive for an agency to ensure the ethnicity data collected is of good quality.

International trends in using administrative data for census

More specific illustration of the value of using administrative data for census can be found by considering overseas experience. Many countries, particularly in Europe, are seeking to reduce their census costs by obtaining census information directly from administrative registers held by government agencies (eg Tonder 2008, Valente 2010, United Nations 2007 and 2011).

Countries such as Denmark, Norway, and Sweden, that have made a successful transition to a fully register-based census, report costs for their census at around 10 percent or less of a full survey-based census, though this does not include the costs of maintaining the associated administrative systems. The transition has occurred over a period of 20–30 years, and they have had to establish the following infrastructure:

- a national population register
- a national address register (and building and dwelling registers in some cases)

- accurate linking between administrative systems (a unique personal identifier)
- compulsory notification of address change.

Political and public acceptance of the need for a population register and personal identity number, and accurate and timely capture of address changes are prerequisites for this model. It also demands close cooperation between the statistical agency, register authorities, and the public administration. Large cost savings have been achieved because the base registers have the same purpose as the census – namely maintaining an up-to-date list of everyone in the country and where they live. The national population registers are embedded in the fabric of government administration, and are a normal part of participation in everyday activities like work, housing, paying taxes, and receiving social benefits. The costs are met by government because of the administrative purpose.

The full register-based census of the Nordic countries is the most advanced example. More common is a combined approach, where administrative data is used where available (often a national population register and address register to support the basic population counts) as well as surveys. Sample surveys are used for the census population attribute information, and coverage surveys may be used to check and adjust the administrative population lists. Countries using a combined approach include Germany, Poland, Israel, and Switzerland. While these combined approaches do appear to achieve cost reductions, the savings are not as significant as the fully register-based census models (Table 4 UNECE 2009).

The Office for National Statistics (ONS) in the United Kingdom is actively investigating the feasibility of an administrative census without a national population register. The ONS experience is particularly relevant to New Zealand because of the similarity in legal, cultural, and historical contexts.

User consultation was undertaken on two possible options for their future censuses:

- An online decennial census replacing the traditional paper census (the census in 2011 for the first time allowed online responses for those who wished)
- Annual population statistics derived from the use and re-use of administrative data supplemented by household surveys.

The National Statistician has recommended a predominantly online census in 2021, supplemented by increased use of administrative data and surveys in order to enhance the statistics from the 2021 Census and improve annual statistics between censuses (Office for National Statistics, 2014).

The next section provides an outline of Statistics NZ's approach to the investigation of alternative data sources for census information. Section 4 shows some results from investigations of administrative data, and provides a brief description of two options for an administrative census. The final sections look to the future. Section 5 describe a future pathway, building on a modernised version of the current census, and discusses the need for a whole of government approach to exploring an administrative census. The paper concludes with the next steps in the work plan.



3 Census Transformation programme: A phased and iterative approach

Statistics NZ's investigation into the potential for an administrative census takes a phased and iterative approach. The first phase is designed to provide an evidence base that is sufficient to inform a decision in 2015 on the preferred direction for future development of the New Zealand census. The early focus is on developing an understanding of future census information requirements, and the ability of administrative sources to meet those requirements. The first iteration of work is designed with a broad scope, but limited depth. The findings will guide the direction of further effort and in-depth analysis. Developing the most likely approach is also providing focus to subsequent work on policy and legislative requirements.

Developing options and assessment criteria

An initial group of five conceptually based options for how an administrative census might work in New Zealand has been identified as a set of working assumptions on which to base technical investigations (Bycroft, 2013a). Three survey-based options (a five-yearly census, 10-yearly census and 10-yearly short-form census with annual survey) are also included as possible options for future censuses, but are not reported on in this paper.

Draft assessment criteria have been developed to assess options. The options and assessment criteria are being refined in an iterative process as investigations progress. Our current understanding of the two best administrative-based options is outlined below.

High-level assessment criteria include benefit, cost, and risk dimensions, but also specifically identify public acceptability, public burden, and technical feasibility (whether it can be done).

The main focus of investigations in the first year has been on assessing the technical feasibility of using administrative sources to produce census information. From the perspective of statistical use, administrative data is often 'noisy', incomplete, and subject to change. The first step is to understand how disparate sources can be brought together to produce population counts and socio-economic attributes. While not excluding the possibility of using commercial, social media, or other big data sources, the first investigations are based on administrative data generated through government activities.

The investigation of administrative sources to produce census information is complex and time-consuming. The most critical requirement to produce detailed counts of the whole population may sound simple, but is actually a difficult and technical issue. The wide range of attribute information collected on the census questionnaire imposes a corresponding volume of searching and analysis of many administrative sources. The initial investigations reported below have been targeted firstly to analysis of data that could provide the most important population counts and is already available to Statistics NZ, and secondly to a broad first cut of possible sources for attributes using information about the data source (metadata).

4 Results

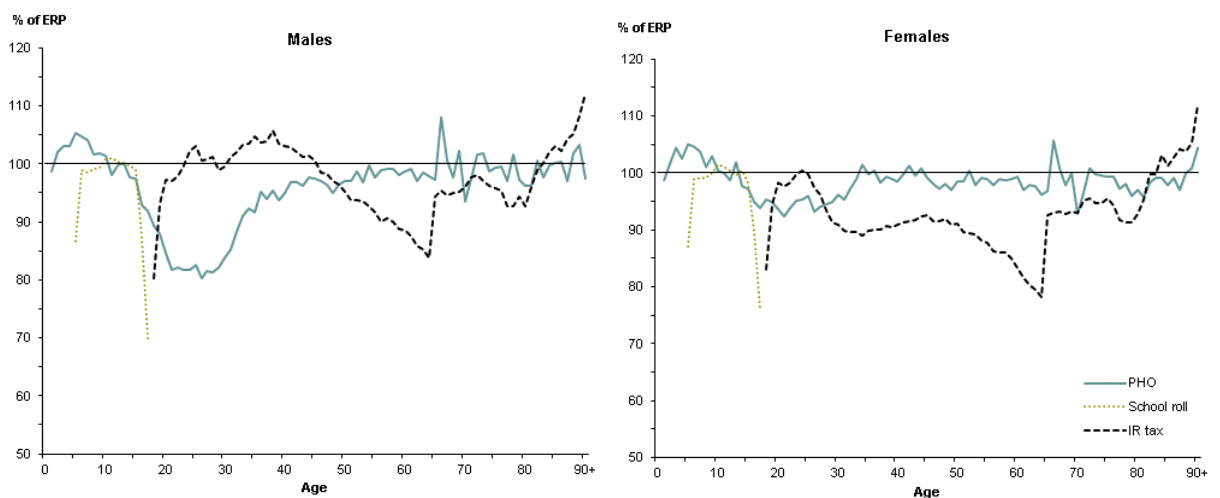
Using administrative data to count the population and identify where people live

Administrative data sources have been assessed to measure how close they are to official population statistics. Initial work has shown that currently available administrative sources could not produce population counts in place of the census in the near term. While early investigations show some promise that linked data can produce an adult population by age and sex that is reasonably close to the official estimated resident population, there are major limitations in the administrative sources available at present for producing more detailed sub-national and ethnic populations.

Investigation of single data sources showed that none would be suitable on their own (Statistics NZ, 2013). While individual sources have parts of the age-sex distribution where they individually have good coverage, they each miss significant sub-groups (figure 1).

Figure 1

National coverage of three administrative datasets compared to the Estimated Resident Population (ERP), males and females by single year of age, June 2012



Note: PHO is Primary Health Organisation data from the Ministry of Health.

Source: Statistics New Zealand

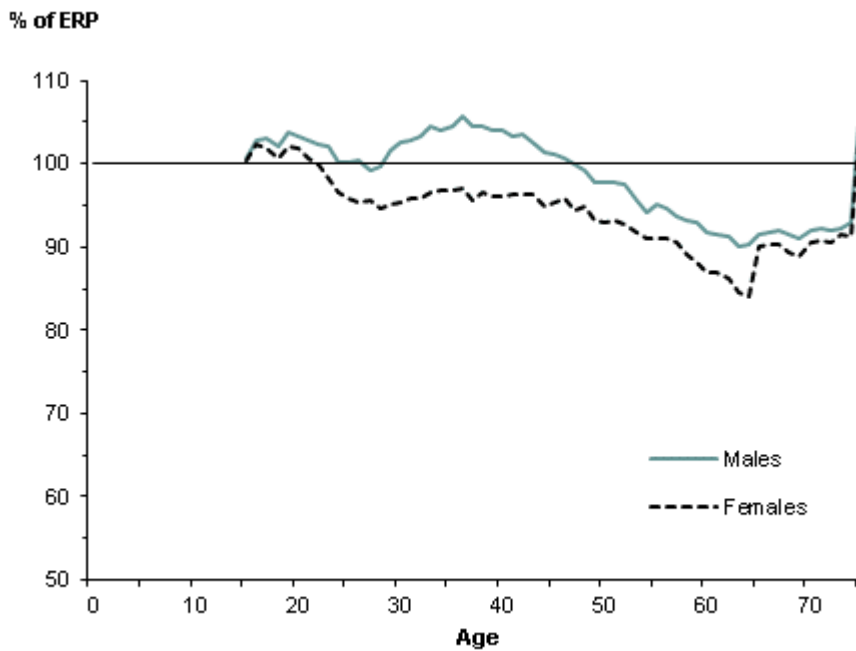
If single sources are not sufficient on their own, does linking data sources improve the situation?

Figures 2 to 4 shows results of investigating the linked administrative data available at present through the Statistics NZ Integrated Data Infrastructure environment (IDI) (Gibb and Shrobbree, 2014). In figure 2, a population derived from linked tax, education, and external migration data is compared to the official population figures. The administrative construct shows good agreement at the national level for ages 15 and over, although this varies between age groups, with some being systematically under-counted and other ages over-counted. Not enough information about children was available in the IDI to

attempt to measure ages under 15 years. Nor was there sufficient ethnic information available at the time of this investigation.

Figure 2

National coverage of linked data (IDI tax, education and external migration) compared to the estimated resident population (ERP), by single year of age and sex, June 2010

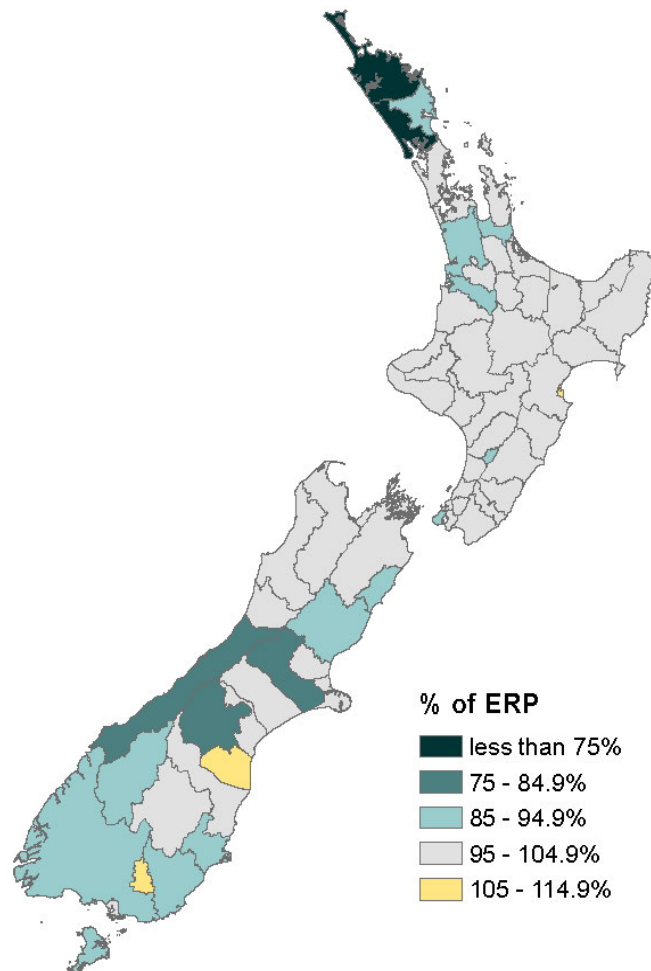


Source: Statistics New Zealand

Data sources that have a reasonable national distribution do not necessarily have adequate information about location (ie where people usually live). Figure 3 shows that subnational coverage for the IDI linked data construct is highly variable. While significant portions of the country are within 5 percent of the ERP (light grey areas), estimates of population size are very low in some territorial authority areas due to poor quality address data in the source.

Figure 3

Territorial authority areas: Coverage of linked data (IDI tax, education, and external migration) compared to the estimated resident population, June 2010

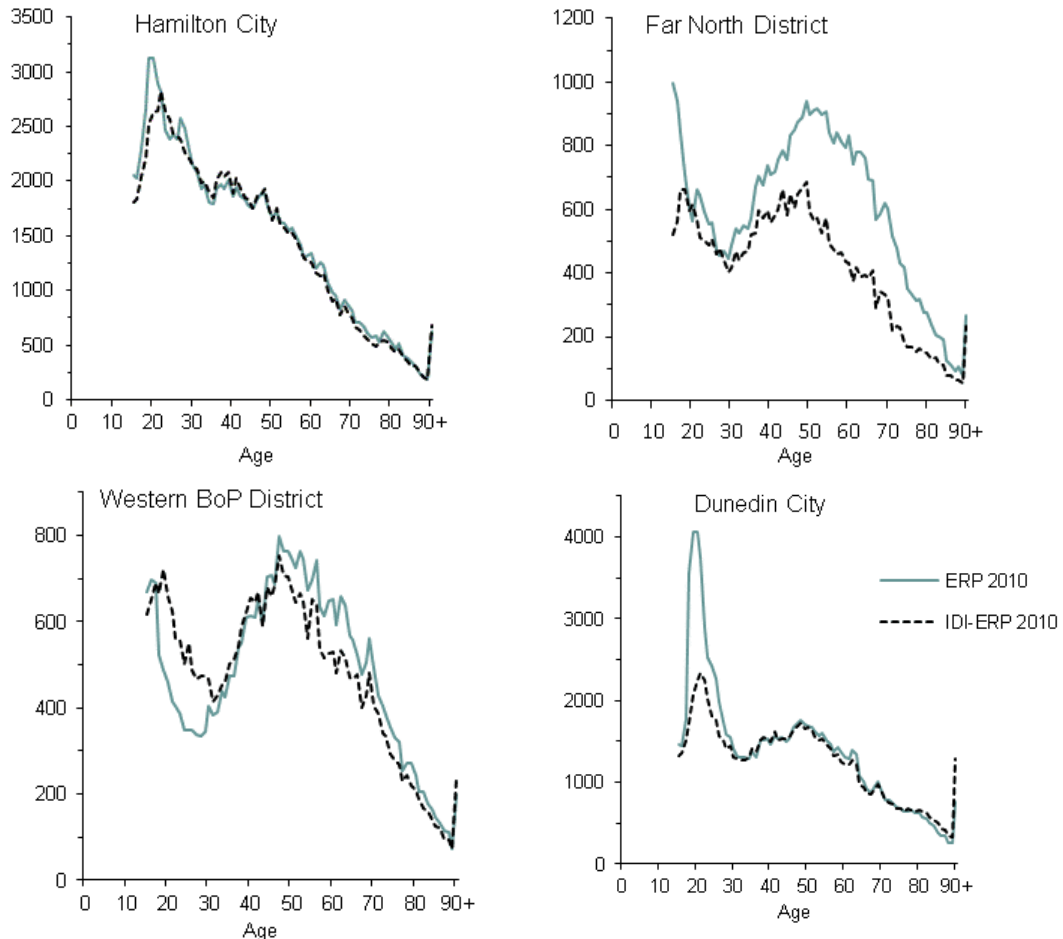


Source: Statistics New Zealand

More detailed analysis of the age-sex structure of sub-national territorial authority areas also shows a wide variation between areas. Figure 4 provides some examples. In some areas like Hamilton City, age structures show good agreement with the ERP. However major discrepancies are evident for other areas. For example, in the Western Bay of Plenty District, the total population agrees well, but the age distribution shows that undercounts in older adults are being balanced by an over-count of younger ages. In the Far North District most ages are well below the ERP, and in Dunedin City while ages over about 30 years correspond closely with the ERP, young adults are grossly under-represented in the estimates constructed from linked IDI data.

Figure 4

Selected territorial authority areas: Coverage of linked data (IDI tax, education, and external migration) compared to the estimated resident population (ERP), by single year of age, June 2010



Source: Statistics New Zealand

In terms of the availability of core demographic variables, linked administrative data sources available in the IDI at present do not support population statistics in the absence of a census, because:

- there is not enough information about children
- the quality of available address data is limited, which means that sub-national estimates for territorial authority total populations are inadequate
- ethnicity information is not adequate
- Māori descent information is not available
- linkage errors which are currently unable to be measured adequately may be distorting the results.

While there are obvious limitations, these early results are promising enough to continue. Some improvements could be made by including more datasets within the IDI. For example, the currently proposed addition of primary and secondary school enrolment data, birth registrations, and death registrations should improve the national coverage.

Ministry of Health data appears to offer considerable potential for use in a future administrative census as it includes people of all ages. Health data would need to be linked to other sources in an integrated administrative database. To investigate this further, agreement with the Ministry of Health for access to unit record data would be needed. Privacy and ethical concerns and retaining independence between the use of census data for health funding allocation and population estimates also need to be considered.

The electoral roll is another source of good data about adults. However current legislation (Electoral Act 1993) does not permit transfer of person-level electronic files.

With more data sources linked together with sufficient accuracy it might be possible to produce good national population estimates across all ages. However these national counts are only the first step. There is little to be gained from linked data sources unless they can also provide good quality information for the core demographic variables of ethnicity and Māori descent, and the accurate and up-to-date address information needed to show how people are distributed around the country.

Using administrative data for information about the population (attributes)

As well as providing population counts, the current census includes questions that collect attribute information about socio-economic characteristics such as language and culture, education, work and income, health and disability, families and households, and housing.

A first stage high-level assessment of the potential for administrative sources to replace the current census attribute information has been undertaken (O'Byrne et al, 2014). Quite a number of administrative sources have been identified that hold some relevant information, and several sources are relied on heavily². However the potential to replace census questions is fairly limited. Nine of 39 current topics appear to be 'likely' to be satisfied with administrative data sources. A further seven were assessed as 'possible' and over half the 39 topics were considered 'unlikely' to be met through administrative sources.

Reasons that administrative sources may not be suitable include:

- information is only available for parts of the population and misses significant groups (eg post-school qualifications are available for New Zealand graduates since 2004, but not for qualifications gained previously, or from overseas institutions)
- only some of the categories are available (eg legally registered marriages and civil unions are recorded, but not de facto relationships)
- the information is not collected in administrative systems (eg language spoken)
- there is not enough information to construct more complicated derived variables (eg household and family arrangements, household crowding).

This indication that administrative sources are likely to contribute only a small proportion of current census topics reinforces the importance of understanding the essential information requirements provided by a census for small areas and population sub-groups. On a more positive note, the results also indicate where administrative data is most likely to be useful in the current census model to improve quality where responses are missing, or possibly to replace questions. Examples include questions about participation in study, income, and income sources.

² In particular: Births, Death, and Marriage registrations, external migration arrival and departure cards, educational enrolment data, tax and income records, tenancy bond database and vehicle registrations.

How an administrative census could work in New Zealand

Based on these findings and other technical considerations, the initial five potential options for a census based on administrative data have been reviewed (Bycroft, 2013b) and two options remain:

- a national population register as the basis for population counts
- multiple linked existing administrative sources to produce a statistical population list.

The two options reflect international experience. The most direct route to an administrative census is through a formal New Zealand national population register. In the absence of a national population register, linking multiple sources of person-centred data is the next best approach to create a statistically derived list of the population. It is a more complex statistical approach, but similar to that investigated by the ONS (ONS, 2013).

The rejected administrative options are: Single source for population counts, an aggregate model and 'big data'.

Both remaining options would work as a system of linked administrative sources. The system relies on two 'base registers'. One is a population register, or else a statistically compiled list of everyone who lives in the country that would play the same role. The other base register is a national address register (a list of all valid addresses). The two base registers are connected through a record of the address where an identified person lives. These base person and address registers then serve as 'spines' that other data sources with information about the person or about the dwelling would be linked to.

A sample survey component would also be needed. A coverage survey is necessary to provide an independent validation of the administrative data. Without a population register, statistical methodologies are likely to play a key role in the estimation of the population. A large scale sample survey may be necessary for attribute information which cannot be obtained from administrative sources.

A national population register

Although New Zealand does not have a national population register, its inclusion as an option is necessary given that this is the successful international model. This is not to say that Statistics NZ advocates for a national population register. The position is a conditional one. If a national population register were to be introduced on its own merits, and proved to be of high quality, then this would be the best option to pursue for an administrative census.

While no specific costings have been carried out for establishing a New Zealand population register, international evidence indicates that a population register would be a high cost investment. Statistics Canada (2012), reports that a 1999 study estimated the cost of establishing a 'Common Client Identifier' supported by biometrics technology, at between \$1.1 billion and \$3.6 billion. And in 2003 start-up costs of a national identity card system were estimated at \$3 billion to \$5 billion. This compares to the approximately \$660 million cost of the Canadian census. In the UK, a voluntary identity card system was estimated at £5 billion to run over 10 years³ and was axed by the incoming coalition government in 2010. By comparison, the cost of the 2011 Census of England and Wales was around £500 million.

³ For example, The Independent, 12 May 2010 <http://www.independent.co.uk/news/uk/politics/national-identity-card-scheme-to-be-axed-1971897.html>

Statistical purposes are not likely to provide sufficient rationale for a population register, and so the cost would need to be justified by the benefits for administrative use. Privacy is likely to be a major concern, and a public debate would be required to determine acceptability, and this too would need to be informed by a clear statement of benefits. The Office of the Privacy Commissioner considers it unlikely that the benefits of a national population register would outweigh the privacy risks.

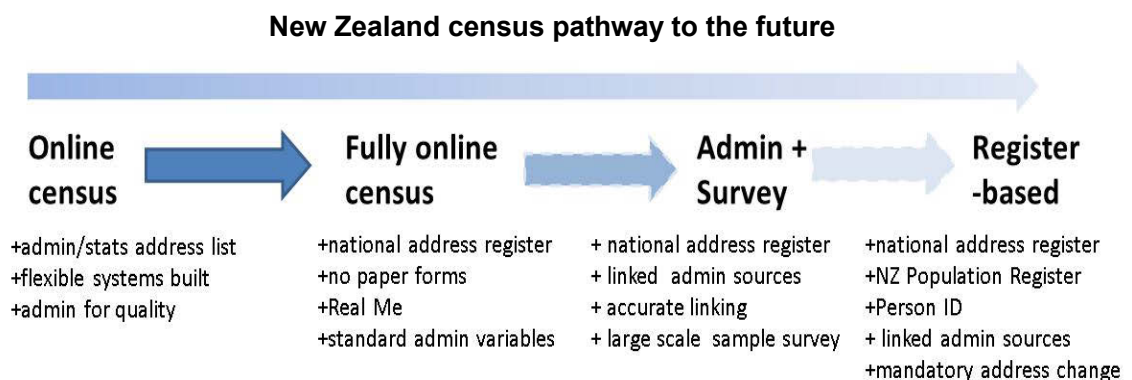
At present, there is no overt demand for a national population register in New Zealand, and even if a need were established, the time to implementation could be a decade or more. Given this uncertainty, Statistics NZ assumes that there will be no national population register established in New Zealand within the next 10 to 20 years, and therefore effort is placed on statistical solutions.

5 Towards the future

Future pathway for the New Zealand census

The next key part of the design process integrates the model for the development of person-centred administrative data with the development and investment activity in modernising the current census model. Figure 4 shows a likely future pathway for the New Zealand census. The barriers to moving to the next step become progressively higher and the model may settle at some point on the way. Equally, some steps might be missed, for example, if a mandatory population register were introduced.

Figure 5



Source: Statistics New Zealand

Modernising the current census takes a path of increasing use of technology working towards a fully web-based or 'online' census, reducing reliance on field staff and paper, and increasing use of administrative data and systems where this adds value.

Investing in the base registers suggested above will benefit both the efficient conduct of the current census and support a shift to an administrative census. Most importantly, the need for a comprehensive and accurate address list for mailing out census Internet codes is a forerunner to the basic requirement of a national address register in an administrative model. Use of administrative data for information missing from the census questionnaire, or for people missed by the census, could improve the current model, and starts to crystallise the best opportunities for using more administrative data.

Modernising the current census and development of alternative administrative approaches are inter-dependent. High quality censuses are essential in the short to medium term for testing the adequacy of administrative alternatives and understanding where improvements are needed. It is likely that two or more censuses may be needed before a decision to change the census approach could be made.

It is not yet clear how far the model can evolve. One possible outcome of investment in modernising the current census and the investigations of administrative data could be a hybrid model: a new combination of an efficient online census that maximises the use of administrative data, that relies on surveys only for essential information unavailable in administrative sources, and where field staff are concentrated on the hard-to-reach population, a group that is often the most important for policy.

Alignment with cross-government initiatives

To support the pathway outlined above, government administrative data sources and infrastructure will need to be strengthened to meet the statistical requirements for census and population statistics. Improvements can be grouped under requirements for:

- a national authoritative address listing
- accurate and up-to-date residential address collected by government
- high coverage of the population
- accurate linking of person-centred data sources
- standardised collection of core variables (eg ethnicity, Māori descent).

Future censuses will operate in a future administrative environment that will look different from today. Recent government initiatives within Better Public Services and the Government ICT Strategy⁴ are at the forefront of changing how government interacts with its citizens and offer real opportunities for enhancing the statistical value of administrative data and systems.

A workshop with key agencies and programmes was held in April 2013 to create awareness of the census transformation programme, and to identify common strategic goals and areas of strong alignment.

The Government ICT Strategy focus area on 'Information as an Asset' provides a supportive environment and an all-of-government view aimed at unlocking the value of government information.

The Government ICT Strategy focus area on 'Digital by Default' will drive significant changes to how government interacts with citizens, mainly through Better Public Services Result 10⁵. It is through the interactions with citizens that much 'information' is generated. If the connection can be made between digitising transactions with government and the value of that information for further re-use, then important quality gains could be achieved. As just one example, the Result 10 goal of 'easy to carry out', enabled by smart online forms and service design principles, could include harmonised collection of ethnicity to the appropriate standard. The adoption of data standards and a high uptake of RealMe is another avenue for improving the usability of some variables like ethnicity or current address.

Improving government holding of addresses is the single most important factor, and this is underpinned by work that Land Information New Zealand (LINZ)⁶ is doing to improve the quality, coverage, and accessibility of authoritative address data in New Zealand. Statistics NZ is working closely with LINZ on this development.

Analysis for Outcomes⁷ brings together key person-centred data from across a range of agencies through extension of Statistics NZ's IDI, and is aimed at research across government sectors, often in relation to specific service sub-populations. This is different from the population measurement function of census. It may also lead to greater demand for the detailed reference population data provided by census and population statistics. Analysis for Outcomes will determine priorities for the addition of new administrative datasets to IDI to support the analytics and insights work programme and this will also

⁴ Government ICT Strategy and Action Plan to 2017 <http://ict.govt.nz/strategy/introduction/>

⁵ Better Public Services - Result 10 <http://www.dia.govt.nz/Better-Public-Services>

⁶ New Zealand Geospatial Office and New Zealand Geospatial Strategy <http://www.linz.govt.nz/geospatial-office>

⁷ Analysis for Outcomes Cabinet paper <http://www.treasury.govt.nz/releases/2013-09-11>: Better Public Services next priorities <http://www.ssc.govt.nz/bps-next-priorities>.

support investigations into an administrative census. However improvements to the quality of source data are largely outside the scope of Analysis for Outcomes.

None of these cross-government initiatives are aimed at population measurement. Some like RealMe are voluntary and consent based, others are specific to a government service. It remains unclear whether any of these initiatives will be sufficient to eventually allow administrative data to replace the current census approach.

To contribute significantly to the ability to use government data for statistical purposes in general, new developments would need to recognise a basic principle that statistical use should be built in by design. It would be beneficial if agencies considered broader re-use when updating or developing data and systems.

Potential areas of exploration include:

- use of standard concepts, definitions and questions
- authoritative sources of information
- standard storage and documentation formats
- agencies publicly documenting data they hold
- agencies actively assuring the quality of their own data.

Improvements in the usability of person-centred data will have benefits well beyond possible use in the context of the census. Better management of data sources, standardisation of common input data and other factors as outlined under ICT Strategy Action Plans for managing information as an asset will add value to the agencies that own them and for cross-agency analysis of outcomes.

This desire to strengthen the strategic value of administrative data sources held by government agencies is evident internationally. For example, in Australia, Principle 1 of the High Level Principles for Data Integration⁸ states:

“Responsible agencies should treat data as a strategic resource and design and manage administrative data to support their wider statistical and research use.”

In the United Kingdom, the UK Statistics Authority released a monitoring report on ‘Creating official statistics from administrative data’ which offers some conclusions on the way forward to making more effective statistical use of administrative data. These include the need for government to find ways to make it easier for statisticians to identify and use administrative sources that have the potential to produce valuable statistics; harmonising definitions and standards in different administrative sources where possible; and that statisticians should be centrally involved in the design and operation of administrative systems that are used to produce statistics.

In February 2014, the United States Government’s Office of Management and Budget released a memo entitled ‘Guidance for providing and using administrative data for statistical purposes’ to help both programme and statistical agencies use administrative data more fully in a manner that respects privacy and protects confidentiality. Specifically, this guidance “will help program agencies manage their administrative data with statistical purposes in mind,” and identifies the benefits for agencies in doing so. For example, “the increased use of administrative data will enhance agencies’ ability to build evidence on which to evaluate the effectiveness of their programs and policies.”

⁸ [High level principles for data integration involving Commonwealth data for statistical and research purposes](#) Australian Bureau of Statistics website, Accessed 13/03/2014.

6 Conclusion

Strong inter-dependencies exist between government policies, administrative data, and census information. A model where administrative data is the main source of population counts requires suitable administrative sources and the ability to link them accurately, as well as an appropriate legal environment and public acceptability.

Making greater use of administrative sources requires a change in the relationship between Statistics NZ and government 'provider' agencies. At present, use of administrative data is largely opportunist and Statistics NZ would need to take a stronger influential role to take full advantage of opportunities emerging in the current fast-changing environment. Creating an administrative census would need to be a shared goal across multiple agencies

At the same time, achieving Better Public Services outcomes through targeted initiatives is creating greater demand for census and population information at detailed breakdowns and for small areas. These targeted interventions and greater use of administrative data as 'information' lead to a greater need to understand finely grained reference populations – information provided by the census. The aim of census transformation and investigations into the use of administrative data is to ensure that proposals for future censuses are the result of joined up cross-government thinking.

Because of the high value and importance of census information, any changes to the way the information is produced require open communications with stakeholders and a rigorous quality assurance process. During consultation undertaken in 2013 a number of agencies noted the importance of maintaining the current census model to protect the ongoing availability of valid population information, and stated that any replacement with administrative data should only be introduced when it has been robustly tested and able to reliably produce the necessary information.

The quality of census information is a key determinant of benefits. The first question is, 'How good do population statistics need to be?' Priorities for the coming year will be to develop specific quality criteria for core population statistics and for attribute information for small areas and small groups (how accurate, how often, at what geographic levels?). Determining essential census information requirements for Māori will also be a priority.

Users of census and population statistics are used to a range of census variables being available every five years, and will take time to work through the issues arising when they are challenged about what information they need, how often they need it, and how accurate it needs to be.

Public acceptability, especially around privacy, is a key determinant of whether any approach to census-taking is viable. Work to identify the privacy and legal issues for an administrative census is underway.

Next steps

The findings reported here are the results of initial investigations, with some avenues more advanced than others. The programme will continue with the technical and stakeholder engagement work already underway, and place more emphasis on developing legislative and policy implications. Next steps include:

- A technical stream involving wider investigation of administrative sources and research into statistical methodologies. The best constructed administrative files will be tested against the 2013 Census results and options will be refined, and assessed against transparent criteria. The focus is on developing the most feasible

form of an administrative census, and understanding the role of sample surveys and estimation methods.

- Part of the technical stream includes investigation of specific use of administrative data in a modernised census to reduce respondent burden, improve quality, or gain efficiencies.
- An external engagement stream, with a focus on working with key users of census information to develop specific quality criteria and to discuss trade-offs. Engagement includes work to understand public attitudes to issues raised by use of linked administrative data for census.
- A legislative and policy stream which will look at what changes and decisions across government would be necessary to implement an administrative census, and how best to leverage off and support government-wide initiatives for mutual benefit.

The extent of the future role of administrative data in producing census and population statistics information remains unclear, and timeframes are uncertain. Regardless of how census will change in future, investigation of administrative data and improvement of administrative sources to make person-centred administrative data more usable will have wide application and be beneficial in its own right.

References

- Bycroft, C (2013a). *Options for future New Zealand censuses: Census Transformation programme*. Available from www.stats.govt.nz.
- Bycroft, C (2013b). *Summary of progress on administrative census options*. June 2013. Unpublished Statistics NZ report
- Gibb, S & Shrosbree, E (2014). *Evaluating the potential of linked data sources for population estimates: the Integrated Data Infrastructure as an example*. Available from www.stats.govt.nz (forthcoming).
- Office for National Statistics (2013) [Beyond 2011 Producing Population Estimates Using Administrative Data: In Practice \(M7\)](#). ONS website.
- Office for National Statistics (2014). *The Census and Future Provision of Population Statistics in England and Wales: Recommendation from the National Statistician and Chief Executive of the UK Statistics Authority*. ONS website [National Statisticians Recommendation](#) March 2014.
- O'Byrne, E, Bycroft, C & Gibb, S (2014). *An initial investigation into the potential for administrative data to provide census long-form information*. Available from www.stats.govt.nz (forthcoming).
- Statistics Canada (2012). *Potential approaches for 2021 Census and beyond*. Available from <http://www12.statcan.gc.ca/strat/Beyond-Subsequents-eng.cfm>
- Statistics New Zealand (2012). *Transforming the New Zealand Census of Population and Dwellings: Issues, options, and strategy*. Wellington: Statistics New Zealand
- Statistics New Zealand (2013). *Evaluation of administrative data sources for subnational population estimates 2013*. Wellington: Statistics New Zealand.
- Tønder, J-K (2008). *The register-based statistical system – preconditions and processes*. Paper presented at the IAOS Conference, Shanghai. Available from <http://isi.cbs.nl/iaos>.
- UK Statistics Authority (2012). *Creating official statistics from administrative data*. Available from: <http://www.statisticsauthority.gov.uk/reports---correspondence/correspondence/letter-from-sir-michael-scholar-to-rt-hon-francis-maude--administrative-data---16032012.pdf>
- United Nations (2007). *Register-based statistics in the Nordic countries: review of best practices with focus on population and social statistics*. United Nations: New York and Geneva.
- United Nations (2011). Report on the Results of a Survey on Census Methods used by Countries in the 2010 Census Round Working Paper: UNSD/DSSB/1.
- United Nations Economic Commission for Europe (2009). Main results of the UNECE-UNSD survey on the 2010 round of population and housing censuses ECE/CES/GE.41/2009/25
- United States government Office of Management and Budget (2014). *Guidance for Providing and Using Administrative Data for Statistical Purposes*. <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2014/m-14-06.pdf>
- Valente, Paolo (2010). Census taking in Europe: how are populations counted in 2010? *Population & Societies*, No 467 May 2010.