

Chair

Cabinet Economic Growth and Infrastructure Committee

## **CENSUS TRANSFORMATION – A PROMISING FUTURE**

### **Proposal**

1. This paper presents progress on the investigation of future options for transforming the New Zealand Census, beyond the planned modernised 2018 Census. It recommends that Statistics NZ be directed to actively work towards a future census based primarily on government's administrative data, supported by redevelopment of its household surveys.

### **EXECUTIVE SUMMARY**

2. There is significant opportunity to exploit the data revolution to transform the New Zealand census. The Government's focus on seeking opportunities to gain greater value from data provides a conducive environment to consider alternative approaches to producing the data provided by census.
3. Accurate population estimates and quality social and economic information at detailed levels underpin, for example, electorate boundary setting, Gross Domestic Product, and billions of dollars of investment by business, government and communities. This information, currently produced from the census, is critical to New Zealand, with an estimated net present value over 25 years of close to \$1.4 billion dollars (the 2018 Census will cost \$110M).
4. It is getting harder to run a traditional census. Costs continue to increase as population increases, and declining response rates are making it more difficult to meet critical customer information needs.
5. A promising future direction is a census based on administrative data, supported by a redevelopment of Statistics NZ's household survey programme. This aligns well with the Government's drive to increase the value gained from its data. It will reinforce and assist cross-government initiatives, such as the Social Sector Investment Change programme, and would result in an estimated 70% reduction in compliance burden on the public. However, a census based on administrative data is not yet possible.
6. Achieving an administrative data future for census will require overcoming barriers and challenges which are a microcosm of the issues facing government as a whole in realising the full potential of its administrative data. Working to overcome these barriers for census transformation will have wider benefits for realising Government's data vision.
7. Further work and testing is required before a recommendation could confidently be made to move away from the current modernised census. This includes further investigation of electoral implications, as any decision to move away from the five yearly full enumeration census will require changes to the Electoral Act 1993.
8. Statistics NZ seeks direction from Cabinet to actively work towards an administrative data census, by leading a programme of work across government to:
  - Improve the quality of key administrative data across government;
  - Resolve legislative issues in the Electoral Act 1993 and the Statistics Act 1975;
  - Further develop the necessary innovative statistical methods, publish an experimental series of population estimates based solely on administrative data, and test the results through and against the 2018 Census; and
  - Explore new data sources to meet critical customer information needs, including scoping the redevelopment of Statistics NZ's household survey programme.

## BACKGROUND

9. On 19 February 2014 the Cabinet Economic Growth and Infrastructure (EGI) Committee considered progress on census transformation investigations and options for the next census [EGI (14) 9]. Cabinet agreed to conduct a modernised 2018 Census, using the full enumeration model, and:
  - a) directed Statistics NZ, in consultation with other government agencies, to undertake a programme of work, including technical, user engagement and legislative and policy work streams, on census transformation;
  - b) directed Statistics NZ and the Ministry of Justice to lead work on identifying options to address the electoral implications of changes to the census model;
  - c) directed Statistics NZ to report back to Cabinet by 31 October 2015 with a final report on census transformation investigations, drawing on the work in recommendations a) and b) above, including the potential for census models fundamentally based on administrative data, sufficient to enable Cabinet to agree a preferred future direction for the New Zealand Census; and
  - d) noted that \$2.5 million over 18 months, to complete the further Census Transformation work programme (outlined above), was included in the Budget 2014 Business Case for The 2018 Census and in the related appropriations.[EGI Min (14) 2/11]

### The current census: what, why, how?

10. Put simply, census is the vehicle for collecting important demographic and other information about the population of New Zealand. This information falls into two broad categories: counts and attributes.

Counts: Includes counts of dwellings and people (by ethnicity, age and sex), down to small geographic areas. These counts form the basis of New Zealand's official population estimates – the estimated resident population (ERP).

Attributes: The census provides a range of social and economic information about small sub-population groups and for small geographic areas. Attributes include information about education and training, work and income, ethnicity and culture, health and disability, transport, housing, households and families.

11. The census produces the authoritative, and in some cases the only, comprehensive source of population and dwellings statistics on local communities and small populations (e.g. Māori, iwi, children, migrants, families).
12. The high quality information produced from the census is critical to New Zealand, with an estimated net present value over the next 25 years of close to \$1.4 billion<sup>1</sup>. By comparison, the next five yearly census, a modernised census in 2018, will cost around \$110M in operating costs.
13. Surveying every person in New Zealand on census night, known as full enumeration, means census data can be relied on to draw accurate conclusions on the whole population and allow multi-variate analysis at detailed levels (e.g. small areas and small population groups, by age and sex etc). Statistics NZ also runs a programme of household surveys, which provide a deeper understanding of key topics. However, the sample size of these other surveys do not allow the detailed analysis possible from census.

### What are the critical information needs?

14. Continuing to meet critical information needs must underpin decisions on the future of the census. This does not mean fulfilling everyone's wishlist - needs must be prioritised and costs, amongst other factors, taken into account.

---

<sup>1</sup> Bakker, Carl. *Valuing the Census: A report prepared for Statistics New Zealand which quantifies the benefits to New Zealand from the use of census and population information*. (Wellington: Taylor Duignan Barry, April 2013).

15. Population statistics are the most important requirement for a census to deliver. Accurate population figures underpin the setting of electorate boundaries as well as key economic and social measures (e.g. per capita Gross Domestic Product and Greenhouse Gas Emissions). They inform a wide variety of planning and investment decisions, by business and central and local government, such as the billions of dollars New Zealand invests in critical infrastructure, housing, health and education. At a minimum, census must be able to produce general population counts and Māori descent population counts at a meshblock level (approx. 60 – 110 people) to set electorate boundaries, and meet other critical information needs.
16. Statistics NZ has undertaken specific recent consultation with census users on their quality requirements for population statistics, on essential information needs for and about Māori, and on the content for the 2018 Census.
17. Essential information needs for and about Māori, identified in consultation with key customers, are: Māori descent, Māori ethnic group, iwi, and te reo Māori.
18. Users have continued to identify their need for a range of social and economic attribute information across the main topic areas (see paragraph 10), with information needed for a range of “unit types” (i.e. individuals, dwellings, families and households) and available for small population sub-groups and small areas.

### **WHY CHANGE THE CENSUS?**

19. It is getting harder to run a traditional census in an efficient manner. The costs of a full enumeration census continue to rise as the population increases. Declining census response rates<sup>2</sup>, which impact on the quality of data especially for small sub-populations and small areas, are an ongoing trend for New Zealand and internationally.
20. The modernisation of the planned 2018 Census responds to these drivers in many ways. The “digital first” strategy, with an online census completion target of 70%, will result in savings (compared to 2013 costs) through significant reductions in field staff and a model less sensitive to population growth. Re-use of administrative data, to support collection and processing of the census, and greater use of technology will also reduce costs. The business case for the 2018 Census also predicted increased online completion and other additional savings for a following modernised census (i.e. the census following 2018), surpassing the 5% efficiency target set for these two census cycles.

### **Maximising the value of government data**

21. There are increasing opportunities to create greater value by re-using administrative data, and increasing advances in technology and methods that support re-use. Government has clearly signalled its intent to grasp these opportunities and understands the power of administrative data to support decision making and evaluation. Census Transformation is one of a number of cross-government initiatives driving greater use of administrative data, including the Government ICT Strategy and Action Plan, the New Zealand Data Futures Partnership programme, Analysis for Outcomes and the Privacy Reform work programme<sup>3</sup>. The Social Sector Investment Change Programme is considering how Government can systematically use administrative data to understand what works for whom, at what cost, to guide future investment decisions.
22. Increased use of administrative data can result in savings from efficiencies, in improved quality of official statistics and in reduced compliance burden on the public. Like many national statistics agencies internationally, Statistics NZ is increasing its use of administrative data in the production of official statistics. The United Kingdom, Italy, Spain, Israel and

---

<sup>2</sup> 1996 Census achieved a response rate of 95.6% which declined to 94.8% in 2006 and to 92.9% for the 2013 Census. These response rates are significantly lower for some groups, including Māori. Some questions also have traditionally lower response rates e.g. income.

<sup>3</sup> Other complementary activity includes the GCIO Partnership Framework and the cross-agency Information Group, LINZ’s National Address Strategy, Government Enterprise Architecture for New Zealand (GEA-NZ) v3.1, and the Open Government Information and Data Programme

Canada are amongst other national statistics agencies looking to use administrative data to supplement or replace full enumeration census surveys.

23. Government data will increase in quality over time as agencies invest to meet business objectives and take up new opportunities to ensure that the public can complete their transactions with the government easily.
24. There will be increasing opportunities to make further use of commercial and local government data sources, as digital engagement with citizens' increases and the value of data is increasingly realised. There are numerous emerging examples which have the potential to provide data with greater detail and flexibility than is currently provided by census (e.g. new Bluetooth traffic sensors compared to the current census "travel to work" question).
25. Statistics NZ has a significant role in increasing the opportunities to gain greater value from the data ecosystem, and encourage and support agencies to improve data quality for effective re-use in official statistics, research and evidence-based decision making.

### Potential future census models

26. Six future census models have been investigated, all of which would increase use of administrative data:

| Census Model | Population Statistics | Attributes |
|--------------|-----------------------|------------|
|--------------|-----------------------|------------|

#### Survey-based censuses with administrative data support

|                                 |                                     |                      |
|---------------------------------|-------------------------------------|----------------------|
| 1. Five yearly (modernised)     | 100% survey, long form <sup>4</sup> |                      |
| 2. 10 yearly                    | 100% survey, long form              |                      |
| 3. 10 yearly with sample survey | 100% survey, short-form             | Annual sample survey |

#### Administrative-based censuses with survey data support

|  |                        |                           |
|--|------------------------|---------------------------|
| 4. Admin data supported by sample survey | Linked admin sources   | Annual sample survey      |
| 5. Admin data only                       | Linked admin sources   | Linked admin sources only |
| 6. Population register                   | NZ Population Register | Linked admin sources only |

27. In survey and annual sample survey models (Models 1 – 4), increased use of administrative data will mean, over time, replacing some questions with administrative data.
28. Further detail on these models and the assessment of their potential is contained in Appendix 1. Comparison of these models shows that:

|   |
|---|
| <b>Model 1:</b> A further modernised five yearly model would deliver the full range of information but has a high public burden, is less sustainable over time and has high ongoing cost.   |
| <b>Model 2:</b> A full 10 yearly census retains the range of information but provides it too infrequently to meet critical customer needs, is less sustainable over time and has a medium level cost.   |
| <b>Model 3:</b> The 10 yearly short form census supported by surveying provides more frequent information but otherwise has few benefits when compared to other models – the cost remains high and the level of detailed information for small areas and small population groups is reduced.                          |
| <b>Model 4:</b> An administrative data census supported by surveying would provide information more frequently than the survey based models, has a lower public burden, is more sustainable over time and has a medium level ongoing cost (around the cost of the full 10 yearly model 2)                             |
| <b>Model 5:</b> An administrative data census alone does compare favourably on frequency, public burden, sustainability and cost. However, given the range and quality of administrative data available, it would currently produce a very limited range of information which would not meet critical customer needs. |

<sup>4</sup> Long form is the current census model, comprising approximately 46 individual and 21 household questions.

**Model 6:** A compulsory population register, while potentially less costly and more sustainable for census, would have high system cost, provide limited information and may be less acceptable to the public.

29. Further progression of the population register model (Model 6) is not recommended, at least in the foreseeable future. The population register model, which would require assignment of a unique identifier to every individual, was thoroughly discussed in the Census Transformation Interagency Working Group (see agencies listed in paragraph 82). While a population register could provide a solid base for population counts, it would still require the bulk of attribute information to come from other administrative sources and be supplemented by surveys. There was consensus that a population register would be unlikely to deliver significant service provision benefits beyond those that can be gained from further exploiting the use of current administrative sources. Officials also agreed that a population register would put an unnecessarily high compliance burden on the New Zealand public.
30. On balance, the administrative data census models appear to be strong sustainable contenders for the future, aligning well with Government's strategic direction and increasing the value that is gained from investment in data. These models would result in a significant reduction in the burden on the public. Even the administrative data census model which relies on a survey component would result in an estimated 70 percent reduction in the compliance costs placed on citizens.
31. While a promising direction for the future, a census based on administrative data is not currently viable. There are a number of barriers, including data currently being of insufficient quality to meet critical customer needs.

## **CAN ADMINISTRATIVE DATA REPLACE CENSUS NOW?**

### **For population estimates – “Getting better but not yet”**

32. Statistics NZ has evaluated the potential for linked administrative data to produce population estimates<sup>5</sup>, using the Integrated Data Infrastructure (IDI). There is clear improvement in recent population estimates produced from the IDI. However, some limitations remain (see Appendix 2 for further information):

*a) The population coverage of administrative data is not yet good enough*

A population derived from administrative data misses some people in the population (e.g. people who are not working or receiving a benefit/superannuation and are not enrolled with a doctor) and includes others who have left the country.

*b) Difficulty locating people where they usually live*

Accurate and up-to-date residential addresses are essential to census to ensure people are counted in their place of usual residence and to understand population mobility. New Zealand has a highly mobile population, with around 50% of people moving between five yearly censuses. More than 20% of the 2013 Census population had a different address in 2013 administrative data. This means administrative data is currently not good enough for purposes such as setting electorate boundaries.

*c) Ethnicity data not good enough for population counts*

Ethnic breakdowns of populations and communities are a fundamental and widely used aspect of census information, for example in allocating population-based health funding. The collection of ethnicity across different administrative sources lacks consistency and is particularly problematic when people have multiple ethnicities. If population estimates were to be created from current integrated administrative data, without further methodological adjustment, the Māori population would be around 25% higher than the current official Māori estimated resident population of New Zealand.

---

<sup>5</sup> A population count can only be taken by a full enumeration census. Any other method of arriving at a number of people in the population is an estimate – hence the term “administrative population estimates”, not counts.

## **For attribute information – “No, it gives only part of the picture”**

33. Statistics NZ has completed a number of in-depth investigations into the potential for administrative data to provide the attribute information currently provided by census, specifically for the topics of education, income, workplace, Māori descent and iwi affiliation, households and families, and housing.
34. These investigations have concluded that:
- a) *Some administrative data is more precise and accurate than census*

For example, total income from Inland Revenue and MSD benefit data provides a more accurate picture than the five yearly census data. Administrative data can also be provided more frequently than five yearly census data (e.g. annually) and the data is longitudinal, meaning an individual record can be analysed over time.

- b) *However, administrative data doesn't provide sufficient population coverage*

Often there is good information in administrative sources about large parts of the population, but no information for specific groups. For example, administrative data doesn't record events that occur overseas (e.g. New Zealanders who gained their highest qualifications overseas). Recently arrived migrants, who have no personal history in New Zealand records, are particularly absent from administrative data.

- c) *And it doesn't contain information similar to that provided by census.*

For example, languages spoken, including te reo Māori, is not information usually required by government agencies, and therefore is not collected. Iwi information is also not broadly collected. Census information on families and households is relied on for purposes such as forecasting benefit expenditure and the Families and Whānau wellbeing indicators<sup>6</sup>. However, families and households cannot currently be constructed from current linked administrative data.

## **REALISING THE FULL POTENTIAL OF ADMINISTRATIVE DATA**

35. Work on census transformation has shed light on the barriers and challenges which are holding back unleashing the full potential of administrative data. Statistics NZ will use transformation of the census as a catalyst to lead a programme of work across government to overcome these barriers. This work will have benefits not just “for census” or “for Statistics NZ” but will also have significant benefits to all data users, including the agencies collecting the data, and to government as a whole, as the value of government data is realised.

### **Driving improvements in the quality of administrative data across government**

36. Meeting future information needs through increasing re-use of government data requires improvement in the overall quality of administrative data and standardisation of data across government. The Productivity Commission's recent report on social services also recommended increased standardisation of data across government<sup>7</sup>. Improved quality and consistency of administrative data will improve the quality of advice Ministers receive and provide more accurate information for service targeting and evaluation purposes. Quality improvements driven from the Census Transformation work also support other government investment in data such as the Integrated Data Infrastructure, the Social Sector Investment Change programme and the Government ICT Strategy and Action Plan to 2017.
37. The key administrative data sources for an administrative census have been identified currently as Health, Education, Benefits, Inland Revenue, Immigration, Border Movements, and Births, Deaths and Marriages. The agencies responsible for these crucial sources are referred to throughout this paper as the “key administrative data agencies”. They are Ministry of Health, Ministry of Education, Ministry of Social Development, Inland Revenue, Ministry of

---

<sup>6</sup> The Families and Whānau wellbeing indicators are published by the Social Policy Evaluation and Research Unit (Superu) as part of its legislative requirement to annually publish a report that measures and monitors the wellbeing of New Zealand families (Families Commission Act 2003 and the Families Commission Amendment Act 2014).

<sup>7</sup> The New Zealand Productivity Commission. *More Effective Social Services* (August 2015)

Business, Innovation and Employment, and Department of Internal Affairs. These agencies are working with Statistics NZ, the Government Chief Information Office (GCIO) and Land Information New Zealand (LINZ) on understanding the impacts of census transformation.

38. These key administrative data agencies have agreed, in principle, to a series of actions aimed at improving the overall quality of administrative data. These actions include:
- a) a concerted cross-government strategy to improve consistency of data collection and data entry
  - b) improving the recording of basic demographic information, including ethnicity, Māori descent and iwi (where these are collected), by moving to standard data classifications;
  - c) improving the accuracy and timeliness of residential address data, where it is collected, including moving to validated addressing systems; and
  - d) ensuring data is recorded with sufficient data description and time reference to enable efficient and effective data re-use.
39. These changes are necessary in order to assess the realistic potential of administrative data to replace census data, and to inform future decisions by Cabinet. The key administrative data agencies have agreed in principle to take all reasonable steps to implement the agreed changes as part of planned data system upgrades or rebuilds in coming years, either within existing baselines, or as part of budgeted data system changes. These changes to data systems in each agency will take some years to implement and further time to then become apparent through the data.
40. As far as is practicable, agencies will aim to implement the agreed changes as part of planned upgrades to allow 2018 Census results to be tested against administrative data benchmarked to the same date. Statistics NZ will, in conjunction with GCIO & LINZ, work with agencies to develop plans for these actions and will report back to Cabinet by 31 October 2016 on any implications arising from those plans, including financial implications.

## **Enabling transformation through legislation**

### *Decoupling the electoral system from the census*

41. Any decision to move away from the five yearly full enumeration census will require changes to the Electoral Act 1993. This is because under the Electoral Act the census is the ‘trigger’ for the electorate boundary review and the Māori Electoral Option (MEO); and is also the data source for setting electorate boundaries:
- The completion of a quinquennial<sup>8</sup> census currently triggers the start of a 4 month MEO period.
  - At the completion of the MEO, census data is used along with Māori electoral data to produce data to review electorate boundaries and calculate the number of electorates.
  - The completion of a periodical census currently triggers the calling together of the Representation Commission to undertake the electorate boundary review.
42. A change in the census model will therefore require Government to make decisions on the preferred timing and triggers for the MEO and electorate boundary review (which are held after each census), and a preferred data source for calculating electoral populations and setting electorate boundaries.
43. Statistics NZ and the Ministry of Justice, in consultation with the Electoral Commission, have considered a range of options for resolving the electoral implications of census transformation (Appendix 3). There is agreement between the three agencies that the most practical options for further investigation, in light of census transformation, involve decoupling the electoral elements (the MEO trigger, the boundary review trigger and the data source for boundary reviews) from the census. This would allow both the electoral and the statistical

---

<sup>8</sup> The terms “quinquennial census” and “periodical census” are used in different sections of the Electoral Act.

systems to continue to evolve without placing undue restrictions on each other in future. The most practical options for further investigations include:

| <b>Boundary review trigger:</b>  | <b>MEO trigger:</b>  | <b>Data source:</b>  |
|--|--|--|
| align the boundary review with an electoral cycle based timeframe (e.g. 6 or possibly 9 yearly). | a) retain a defined period (e.g. 4 months), within which Māori electors can change rolls but this period is decoupled from the census (e.g. occurs every 5 or 6 years); or<br><br>b) a continuous system which allows Māori electors to change rolls once during an electoral cycle. | a) use the Estimated Resident Population, NZ's official population count; or<br><br>b) allow the Government Statistician to decide the best data source at the point in time required (using a Certificate of Population as per the local electoral system). |

44. The processes for setting electorate boundaries and the delivery of the MEO have also been raised in submissions to the Justice and Electoral Select Committee inquiry into the 2014 general election currently underway. The Select Committee may make recommendations on these issues.
45. Ensuring the accuracy of population estimates at the detail level required for electoral purposes is a fundamental requirement which must be met before any decisions can be taken to move away from the current census model.

*Enabling greater flexibility and freedom to evolve*

46. Any departure from the current five yearly full enumeration census model will require change to the Statistics Act 1975. The Statistics Act requires that the census be held in 2013 and every five years thereafter - 2018, 2023 etc (Section 23). It also requires that the census be a full enumeration census, with census forms completed for every person (Sections 25 – 27).
47. The Statistics Act is highly prescriptive with regard to census operations, as was common in legislation of that time. A planned review of the Statistics Act, beginning in 2015/16, will consider ways in which future legislation can allow for greater flexibility in the sources of data and modes of collection for a census and for production of other official statistics. This would allow the freedom for census to evolve over time.

*Ensuring the supply and assuring the quality of administrative data to Statistics NZ*

48. A move to greater use of administrative data for census purposes will also require new powers for the Government Statistician in relation to government's administrative data. These changes are necessary to ensure reliable and ongoing access to administrative data for statistical purposes, to ensure supply is not disrupted by unexpected changes to the source data, and to assure the quality of administrative data for critical statistical purposes. This is in line with international legislative developments.
49. Therefore, future legislative changes should give consideration to options for including:
- a) A right for the Government Statistician to access, for statistical purposes, administrative data held by government agencies, and
  - b) A right for the Government Statistician to be consulted on changes to administrative data systems, and make recommendations on these systems, for statistical purposes.
50. Māori descent is collected in electoral enrolments. However, under the Electoral Act, the electoral roll data is not currently available to Statistics NZ at an individual record level. The Electoral Act would require amendment to allow roll data to be supplied to Statistics NZ.

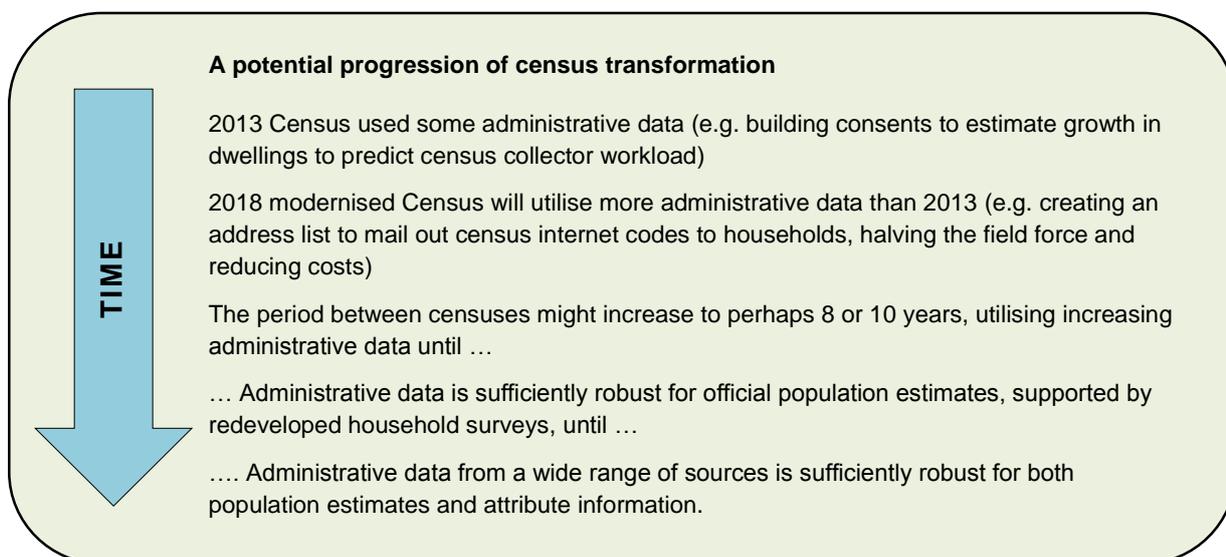
**Innovative statistical modelling and experimental testing**

51. In 2016, Statistics NZ will begin production of an experimental series of administratively based population statistics which can be compared to official population estimates. Publishing this series will clearly demonstrate progress towards an administrative data census and provide a vehicle for ongoing input, on methods and results, from data users, external experts and the public.

52. A number of statistical methods must be employed to assess the accuracy of the population coverage of administrative data and adjust it to closer reflect the actual population, so we can draw conclusions about the whole population. Amongst other methods, coverage surveys will be used to provide a level of assurance that administrative based population estimates accurately reflect the population, similar to the coverage surveys conducted after current censuses. The methods investigated by Statistics NZ to date are promising, and in line with international developments, but further work is required.
53. To further understand the viability of an administrative data-based census, Statistics NZ will utilise the modernised 2018 Census and its development. This will include a test of the coverage of administrative data sources being undertaken as part of the regular post-censal coverage survey in 2018.
54. The resulting administrative population estimates will then be compared to the 2018 Census population count and the resulting official estimated resident population, released in June 2019. This will be the first point at which a robust assessment can be made of the viability of creating population estimates from administrative data.

### **SUPPLEMENTING ADMINISTRATIVE DATA TO MEET CRITICAL INFORMATION NEEDS**

55. Investigations to date demonstrate that government’s administrative data alone does not provide a sufficient range of attribute information to meet customer needs. While aspiring to rely on administrative data alone in the longer term, the medium term goal that Statistics NZ proposes to work towards is a primarily administrative data census, supported by a redeveloped household survey programme and use of data from alternative sources (Model 4). This enables the benefits of re-using administrative data to be realised, while also continuing to provide the broad range of information required by customers.
56. This is a key stepping stone to a longer term administrative data census (Model 5). An indicative progression can be illustrated as follows:



### *Exploiting alternative data sources*

57. The potential for commercial “big data” to be used in official statistics is being explored worldwide. Success varies widely between different data sources and different countries and there are significant challenges and limitations to be overcome.
58. Often alternative data sources do not collect information that would meet census users’ needs. Local authority rating data, for example, cannot provide information on the nature of a property: whether it is a usual residence, a vacant house, a holiday home or a shed; or whether there is a mortgage over the property. Thought must also be given to ensuring commercial data supply and quality can be relied on for critical statistical purposes.
59. Statistics NZ will actively explore the potential for alternative data sources to provide census-type information in the next phase of work. Statistics NZ will also explore with key

government agencies the potential to collect new or different information from their customers in future.

60. New Zealanders with RealMe verified identity accounts could, with a change to regulation, have the ability to volunteer their identity information to Statistics NZ as an additional source of administrative data. However this is a voluntary service so, under current legislation, the information contained in RealMe could not be provided to Statistics NZ for statistical re-use without the specific consent of every person. Over time, better management and a greater use of authoritative sources of identity information will lead to improvements in the quality of administrative data held by participating agencies.

#### *Redeveloping Statistics NZ Household Survey Programme*

61. Where administrative and alternative sources cannot meet critical information needs, a large scale survey will be needed to fill the gaps. A large scale attribute survey would become part of a redeveloped suite of Statistics NZ household surveys.
62. Statistics NZ's existing household survey programme includes the Household Labour Force Survey, the Household Economic Survey, the General Social Survey, the Disability Survey and Te Kupenga (a survey of Māori wellbeing). These surveys approach around 30,000 households, or around 1.5% of the population.
63. While these surveys provide timely, detailed information on enduring topics of interest, they are unable to provide information at fine geographic levels or for small population groups. To provide the accuracy and level of detail of information currently collected by the census, it would be necessary to survey 5% of the population each year or 25% of the population over five years. Therefore, if an administratively based census is to be pursued, it is proposed that the existing programme of surveys be redeveloped to achieve greater coverage and meet some of the information needs currently met by census.
64. However, while having potential to collect much of the attribute information currently available from the census to supplement administrative data, redevelopment of the household survey programme will not allow accurate analysis at the very detailed level currently available from the census for small areas and small population groups.
65. For example, drawing conclusions about low income working sole parents would be difficult at local levels. Family status is not currently available from administrative data, so this would rely on survey data, and conducting multi-variate analysis (their age, sex, income and location) on such a small population would have limited accuracy. Similar issues would arise with, for example, multi-variate analysis for most iwi. For example, regional breakdowns could only be provided, to reasonable accuracy, for large iwi with more than 6,000 members.
66. Accurate small area and small population data is highly valued by many census information users. For example, the Ministry of Health contracts the University of Otago to produce the NZ Deprivation Index using census attribute data at meshblock level. Local government utilises small area data for planning and service delivery. Government agencies consulted on this paper are also concerned to ensure retention of the valuable small area and small population level data, currently used to guide service provision and resource allocation, especially in targeting vulnerable groups. Further work is necessary to confirm customer needs for small area and small population data.

#### **ACTIVELY WORK TOWARDS AN ADMINISTRATIVE DATA CENSUS**

67. Statistics NZ considers an administrative data census supported by a redeveloped household survey programme to be a promising future direction for the New Zealand census.
68. However, investigations show that moving to an administrative data census supported by surveys is not possible now, as there remains significant uncertainty in the range and accuracy of data that can be provided. Continued development and testing is required to enable Cabinet to consider an administrative data and sample survey census as a proven viable option alongside other census models.

69. Therefore, Statistics NZ seeks direction from Cabinet to actively work towards an administrative data census. This will entail progressing the issues outlined throughout this paper, including:

- Improving the quality of key administrative data across government;
- Resolving legislative issues in the Electoral Act 1993 and the Statistics Act 1975;
- Further development of innovative statistical modelling, publishing an experimental series and testing results through and against the 2018 Census;
- Exploring new data sources to meet critical customer information needs, including scoping the redevelopment of Statistics NZ's household survey programme;
- Confirming key customer needs for small area and small population data; and
- Further understanding the risks and costs of census transformation.

### **Managing risk and uncertainty**

70. Moving away from a full enumeration census would represent the greatest transformation in official statistics since the New Zealand census was first established in 1851. The census is seen internationally as one of the foundations of open democracy and underpins trust in government.

71. There would be considerable risk involved in moving to a new system before it is soundly proven. Inaccurate census information would impact business investment, undermine New Zealand's international financial position, and could cost the government and local communities considerably in poor decisions. [A territorial authority<sup>9</sup>] estimated that a greater than 5 percent error in population estimates alone could cost the Council upwards of \$17 million over 10 years.

72. Other countries which have moved to administrative data census models have done so over long periods of time. Denmark took more than 15 years to move to an administrative data census, despite beginning with a well-established population register system.

73. Considerable further work and an extensive testing programme are required before recommendations could be made to confidently move away from the current modernised census. Despite some promising signs, the proposed programme of work and positive intentions to actively work towards an administrative data future, it may prove not to be viable in the foreseeable future.

### **Timing of next steps**

74. Improving the quality of administrative data, developing and testing the administrative data series and designing the redevelopment of the household survey programme are needed to reduce the risks and uncertainties of changes to census. The most critical of these is the time necessary to adequately test the new administrative series through repeated benchmarking against full enumeration census data, with a substantial "benchmark test" against one census followed by a full "dress rehearsal" of an administrative data census series against a further full enumeration census.

75. Therefore Statistics NZ currently considers one further full enumeration census after 2018 is likely to be required, in order to ensure Cabinet can confidently take major decisions on the future of census.

76. The timing of further reporting to Cabinet is expected as follows:

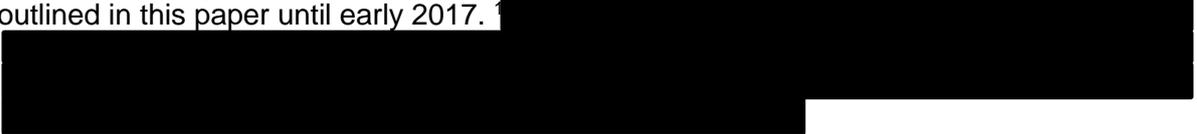
- 31 October 2016 – a progress report on the identified areas of further work, including the first release of the experimental series and agency plans to improve administrative data quality, and identifying the investment required to progress work beyond the 2016/17 year, which may take the form of an indicative business case.

---

<sup>9</sup> Territorial authority redacted under Section 9 (2) (ba) (i)

- 2018/19 – a report to seek a decision on the nature and timing of the post-2018 census, for Budget 2019.
- 2020 – a major census transformation report back, following the first benchmark test, enabling development of a more detailed transformation plan.
- Mid-2020's – a detailed business case is expected following full dress rehearsal testing of an administrative data census against a further full enumeration census.

### **Financial Implications**

77. There are no immediate financial implications. Current funding for census transformation, referred in paragraph 9 d) above, is sufficient to complete the Statistics NZ work programme outlined in this paper until early 2017. 

78. The key administrative data agencies have agreed to take all reasonable steps to implement the administrative data system changes either within existing baselines or as part of budgeted data system changes.

79. Statistics NZ will bring the census transformation programme of work into the business case process at an appropriate time, in consultation with relevant agencies, when risks and costs are clearer and significant investment is required.

### **Human Rights, Treaty of Waitangi and Privacy Considerations**

80. There are no change proposals in this paper.

### **Legislative and regulatory implications**

81. There are no legislative and regulatory implications at this time.

### **Consultation**

82. The following agencies have been members of the Census Transformation Interagency Working Group and have been extensively consulted on this paper: Department of Internal Affairs (including Local Government Commission and the Government Chief Information Office), Inland Revenue, Land Information NZ, Ministry of Business, Innovation and Employment, Ministry of Education, Ministry of Health, Ministry of Justice, Ministry of Social Development, Social Policy Evaluation and Research Unit (Superu), and Te Puni Kokiri.

83. The Electoral Commission, NZ Customs Service, Accident Compensation Corporation and Ministry of Transport were also consulted. Consultation on this paper was also undertaken with central agencies: the Treasury and State Services Commission. Department of Prime Minister and Cabinet has been informed.

84. The Office of the Privacy Commissioner has been consulted and is generally supportive of the integration of government administrative data for census purposes, provided it is conducted by Statistics NZ under the strict privacy and confidentiality rules that currently apply to integration of data in the IDI.

85. Discussions on census transformation have been held with Auckland Council and Christchurch City Council, as key users of census information; and with Wellington City Council, regarding the potential administrative data available from local government.

### **RECOMMENDATIONS**

86. It is recommended that the Cabinet Economic Growth and Infrastructure (EGI) Committee:

---

<sup>10</sup> Redacted under Section 9 (2) (f) (iv)

1. **note** that in February 2014 Cabinet considered progress on census transformation investigations and options for the next census, agreed to conduct a modernised 2018 Census, using the full enumeration model, and:
  - 1.1. directed Statistics NZ, in consultation with other government agencies, to undertake a programme of work, including technical, user engagement and legislative and policy work streams, on census transformation by October 2015;
  - 1.2. directed Statistics NZ and the Ministry of Justice to lead work on identifying options to address the electoral implications of changes to the census model by October 2015;
  - 1.3. directed Statistics NZ to report back to Cabinet by 31 October 2015 with a final report on census transformation investigations, drawing on the work in recommendations 1.1 and 1.2 above, including the potential for census models fundamentally based on administrative data, sufficient to enable Cabinet to agree a preferred future direction for the New Zealand Census; and
  - 1.4. noted that \$2.5 million over 18 months, to complete the further Census Transformation work programme (outlined above), was included in the Budget 2014 Business Case for The 2018 Census and in the related appropriations.[EGI Min (14) 2/11]
2. **note** that the “digital first” strategy and other features of the modernised census model will reduce census costs for the 2018 Census and are expected to further reduce costs in future modernised censuses;
3. **note** that Statistics NZ considers an administrative data census, supported by a redeveloped household survey programme, to be a promising, but not yet possible, future direction for the New Zealand census;
4. **note** that, while the investigations referred to in Recommendation 1.1 above show that an administrative data census has potential for the future:
  - 4.1 administrative data is not currently of sufficient quality or coverage to meet critical customer needs for census information,
  - 4.2 there are legislative barriers to be resolved,
  - 4.3 further work and testing is necessary to demonstrate the viability of an administrative data census model, and
  - 4.4 an administrative data census may yet prove not to be a technically feasible or cost effective option;
5. **note** that an administrative data and sample survey census is not expected to meet all information needs currently met by census, in particular some attribute information for small areas and small population groups;
6. **note** that any future census must, at a minimum, be able to produce general population counts and Māori descent population counts at a meshblock level for the setting of electorate boundaries, and must be able to meet other critical information needs;

#### *Improving administrative sources*

7. **note** that key administrative data sources for census purposes are currently identified as: Health, Education, Benefits, Inland Revenue, Immigration, Border Movements, and Births, Deaths and Marriages;
8. **direct** agencies responsible for the key administrative data sources identified in recommendation 7 above to, in conjunction with Statistics NZ, the Government Chief Information Office (GCIO) & Land Information NZ (LINZ), with a target date of 31 March 2018:
  - 8.1 develop and implement a concerted cross agency strategy, co-ordinated by SNZ, focused on improving the consistency of data collection and data entry;

- 8.2 move to standardised collection of name, sex, date of birth, address, ethnicity, Māori descent and iwi (where these are collected), in accordance with relevant standards developed by SNZ, GCIO & LINZ; and
  - 8.3 ensure data is recorded with sufficient data description and time reference to enable efficient and effective data re-use, led by SNZ in conjunction with GCIO;
9. **note** that Statistics NZ will, in conjunction with GCIO and LINZ, work with the agencies referred to in recommendation 8 to develop plans for these actions and will report back to Cabinet by 31 October 2016 on any implications arising from those plans, including financial implications;

*Resolving legislative barriers*

- 10. **note** any decision to move away from a five yearly full enumeration census will require changes to the Electoral Act 1993 and the Statistics Act 1975;
- 11. **note** that Statistics NZ is undertaking a review of the Statistics Act 1975 which will consider opportunities for greater flexibility in the sources and collection modes for census data and the levers necessary to ensure supply and assure quality of administrative data for statistical purposes;
- 12. **note** that Statistics NZ, Ministry of Justice, and the Electoral Commission consider there are viable options for further investigation to address the electoral implications of a move away from a five yearly full enumeration census;
- 13. **direct** Statistics NZ and Ministry of Justice to identify opportunities to, and further investigate options in more detail for, decoupling the electoral system from the census and report back to Cabinet on these investigations by 31 March 2019;

*Further work and testing*

- 14. **note** that Statistics NZ intends to release, from 2016, an experimental series of population estimates, compiled from available administrative data and expanding over time, in order to test methodologies and monitor the effects of improvements in data quality;
- 15. **direct** Statistics NZ to, in conjunction with other government agencies, actively work towards an administrative data census by:
  - 15.1 Improving the quality of key administrative data across government;
  - 15.2 Further developing the required statistical modelling, resulting in an experimental series to be tested through and against the 2018 Census;
  - 15.3 Exploring new data sources to meet critical customer information needs, including scoping the redevelopment of Statistics NZ's household survey programme;
  - 15.4 Confirming key customer needs for small area and small population data; and
  - 15.5 Further understanding the risks and costs of census transformation.
- 16. **direct** Statistics NZ to report back to Cabinet by 31 October 2016 on progress on the work outlined in recommendation 15, identifying the investment required to progress work beyond the 2016/17 year;
- 17. **note** that a further report on progress will be made in 2018/19, in association with a decision on the nature and timing of the following census (i.e. the post-2018 Census), followed by a major census transformation report in 2020, following the first benchmark test, enabling development of a more detailed transformation plan; and
- 18. **note** that, given the risks and uncertainties inherent in making major changes to the New Zealand census, Statistics NZ currently considers a further full enumeration census after 2018 is likely to be required, in order to ensure Cabinet can confidently take major decisions on the future of census

Hon Craig Foss  
 Minister of Statistics  
 Date:

# Models for future Census Transformation

|                                      |   | Range of information   | Timeliness  | Public burden | Sustainability | Cost  |      |        |
|--------------------------------------|---|--|---|---------------|----------------|-------|------|--------|
| Increasing administrative data usage | <b>1 5 yearly modernised (long form)</b><br>A modernised survey of the entire population at a single point in time every 5 years.   | ✓ Meets customer requirements<br>✓ Good level of accuracy<br>✓ No legislative change required<br>✓ Increasing use of admin data  | ✗ Costs increase with population<br>✗ Declining response rates  | Full          | Medium         | High  | Less | High   |
|                                      | <b>2 10 yearly (long form)</b><br>A survey of the entire population at a single point in time every 10 years.   | ✓ Cheaper than 5 yearly model<br>✓ Provides small area statistics<br>✓ Good level of accuracy<br>✓ Increasing use of admin data  | ✗ Declining response rates<br>✗ Similar cost to model 4 but data quickly becomes outdated<br>✗ Population only rebased every 10 years         | Full          | Low            | Lower | Less | Medium |
|                                      | <b>3 10 yearly (short form)</b><br>+ redeveloped household survey programme<br>A short survey to count the entire population at a single point in time supported by surveys to provide attribute information. | ✓ Increasing use of admin data<br>✓ Annual attribute information   | ✗ Limited small area statistics<br>✗ Declining response rates<br>✗ Similar cost to 5 yearly model<br>✗ Population only rebased every 10 years | Reduced       | High           | Lower | Less | High   |
|                                      | <b>4 Admin data</b><br>+ redeveloped household survey programme<br>Linked admin data for population statistics and supported by surveys to provide attribute information.                                     | ✓ Cheaper than 5 yearly model<br>✓ Reduces respondent burden<br>✓ High re-use of admin data<br>✓ Full range of attributes<br>✓ Flow on improvements to government data | ✗ Limited small area statistics<br>✗ Dependent on quality and supply of admin sources   | Reduced       | High           | Lower | More | Medium |
|                                      | <b>5 Admin data only</b><br>Linked admin data for population statistics and some attributes.  | ✓ Cheaper than 5 yearly model<br>✓ Reduces respondent burden<br>✓ High re-use of admin data<br>✓ Flow on improvements to government data                               | ✗ Limited range of attributes<br>✗ Limited small area statistics<br>✗ Dependent on quality and supply of admin sources                        | Limited       | High           | Low   | More | Low    |
|                                      | <b>6 Population register</b><br>Compulsory register of people residing in NZ, and requiring mandatory updating, holding core demographic information. Linked admin for attributes.                            | ✓ Cheaper than 5 yearly model<br>✓ Reduces respondent burden<br>✓ High re-use of admin data  | ✗ Low public acceptability<br>✗ Limited range of attributes<br>✗ Dependent on quality and supply of admin sources                             | Limited       | High           | High  | More | Low    |

*Actively working towards an administrative data census...*

# What Census information can we get from administrative data?

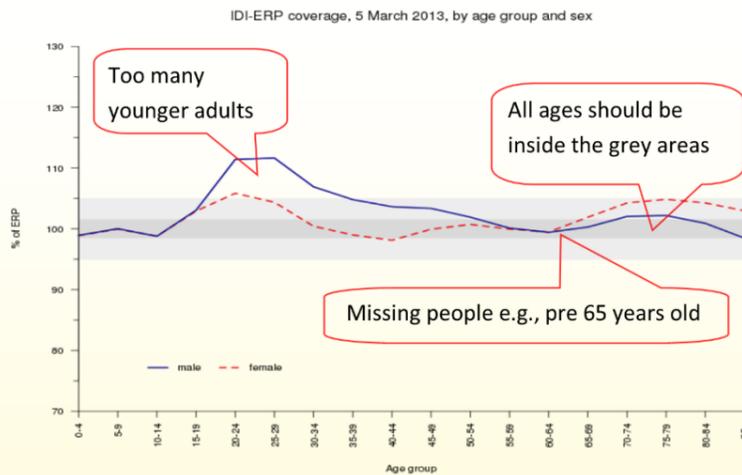
*Some, but it is not good enough yet, and some is missing*

## Census Information creates a rich picture of the nation and for local communities

Whatever method we use to produce census information, we need to:

- **Count** everyone once, only once, and in the right place
- Provide **attributes**: migration, education & training, work & income, ethnicity & culture, health & disability, households & families, transport, and housing.

## POPULATION COUNTS from linked administrative sources are improving



Statistics NZ has developed prototype methods for getting population counts. *Statistics NZ selects people who are 'active' in tax, benefits, health, or education and remove people who have died or left the country.*

Results are not too bad at national level, **BUT** are not good enough for some breakdowns by age and sex groups.

**Still a lot of work to do. Statistics NZ needs:** better ways of selecting and removing people; better linking; a coverage survey and statistical model to fix errors in administrative data.

- ### Administrative data advantages
- ✓ Re-uses information
  - ✓ Reduces respondent burden
  - ✓ More frequent
  - ✓ Longitudinal
  - ✓ Cost reduction in the long term

- ### Administrative data disadvantages
- X Depends on supply and quality of admin sources
  - X Distributed government collection makes consistency hard (e.g. address, ethnicity)
  - X A lot more complex
  - X Misses people and includes wrong people
  - X Limited information about attributes

## Admin data results for ETHNIC GROUPS are not good enough yet

Statistics NZ has combined ethnicity data from six different government sources to get an ethnic group for nearly everyone (98%).

**BUT** accuracy is not good enough: best for European only, but counts are too high for Maori and Pacific; and not good for people with multiple ethnic groups.

**Government needs** improved data collection, consistent with standard, across distributed systems.

- Examples:
- Best results from DIA Birth Registrations
  - Improving results from Ministry of Health

## Admin data results for GEOGRAPHIC AREAS are not good enough yet

Statistics NZ has combined address data from five different government sources to get a meshblock for nearly everyone (99%).

**BUT** accuracy is not good enough yet for small areas (79% have the same meshblock as census).

**Government needs:**

- Government agencies to use an address validation system.

## Admin data for ATTRIBUTES gives only some parts of the picture

- ### Admin data is good for
- Study participation
  - Personal income
  - Country of birth
  - Years in NZ
  - Industry

- ### Admin data is partly useful for
- Highest qualification
  - Legal marriages
  - Children born
  - Work & labour force status
  - Maori descent
  - Tenure of household
  - Weekly rent

- ### Admin data is poor for
- Iwi
  - Households
  - Families
  - Language spoken
  - Occupation
  - Hours worked
  - Religion
  - Means of travel to work
  - Fuel types to heat dwelling

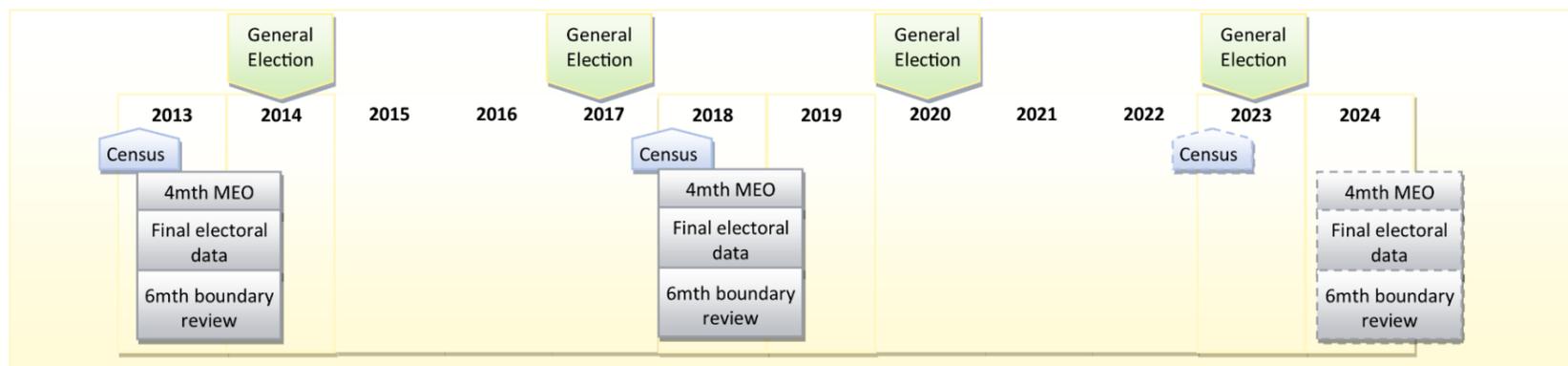
# Electoral options for managing a transformed census

The most practical options for further investigation to manage the electoral implications of a transformed census involve decoupling the electoral elements from the census.

**Status quo** Under the Electoral Act, the number of electorates and electorate boundaries can only be revised on the basis of census data. The completion of a census is the trigger point for the start of both the Māori Electoral Option (MEO) delivered by the Electoral Commission, and the boundary review process conducted by the Representation Commission.

The **MEO** is a four month period when electors of Māori descent decide whether to be enrolled on the general or the Māori roll. The interval between each MEO is either 4, 5, or 6 years, depending on the timing on the census and general election and roll choice generally applies for two general elections. The **Boundary Review** is triggered by the census. The electoral populations and the number of electorates are calculated using data from the census and MEO.

## Inter-connections between the current 5-yearly census and the electoral system



## Practical options for managing the electoral implications of a transformed census

A change in census model will require Government to make decisions on the preferred timing and triggers for the MEO and boundary review, and a preferred data source for calculating electoral populations and setting electoral boundaries.

This table summarises the possible solutions investigated for managing the triggers for MEO and boundary review, and the data source used to calculate electorate populations and set electorate boundaries. It shows under which census transformation models a solution is either feasible for further investigation, or possible but not preferable. The feasibility of options that involve a decoupled trigger depends on both the data source used and the exact timing of the MEO boundary review.

Key: ✓ - feasible for further investigation  
 \* - possible, but not preferable  
 x - not feasible/applicable

| Triggers and data sources                                   | Census transformation model |                        |                            |
|---|-----------------------------|------------------------|----------------------------|
|   | 5-yearly survey based       | 10-yearly survey based | Administrative -data based |
| <b>Solutions for triggering the MEO</b>                     |                             |                        |                            |
| Status quo – triggered by a 5-yearly census                 | ✓                           | x                      | x                          |
| 10-yearly – triggered by a 10-yearly census                 | x                           | *                      | x                          |
| 5-yearly – triggered by Electoral Act (decoupled)           | ✓                           | *                      | ✓                          |
| 6-yearly – triggered by Electoral Act (decoupled)           | ✓                           | ✓                      | ✓                          |
| 'Continuous' MEO  | ✓                           | ✓                      | ✓                          |
| <b>Solutions for triggering the boundary reviews</b>        |                             |                        |                            |
| Status quo – triggered by the 5-yearly census (and MEO)     | ✓                           | x                      | x                          |
| 10-yearly – triggered by the census (and MEO)               | x                           | ✓                      | x                          |
| 5-yearly – triggered by the Electoral Act (decoupled)       | ✓                           | *                      | ✓                          |
| 6-yearly – triggered by the Electoral Act (decoupled)       | ✓                           | ✓                      | ✓                          |
| 9-yearly – triggered by the Electoral Act (decoupled)       | ✓                           | ✓                      | ✓                          |
| <b>Solutions for population data source</b>                 |                             |                        |                            |
| Ordinary resident population – traditional census model     | ✓                           | *                      | x                          |
| An alternative population data source (e.g. ERP)            | ✓                           | ✓                      | ✓                          |
| Government Statistician certificate of electoral population | ✓                           | ✓                      | ✓                          |
| Enrolled electors   | *                           | *                      | *                          |

| Solutions for triggering the MEO   | Solutions for triggering the boundary review  | Data source for the boundary review  |
|--|---|--|
| <ul style="list-style-type: none"> <li>Triggered by the census (5 or 10 yearly)</li> <li>Triggered by the Electoral Act (decoupled)</li> <li>'Continuous' MEO</li> </ul>   | <ul style="list-style-type: none"> <li>Triggered by the census (5 or 10 yearly)</li> <li>Triggered by the Electoral Act (decoupled)</li> </ul>  | <ul style="list-style-type: none"> <li>Estimated Resident Population (ERP)</li> <li>Certificate of population</li> <li>Enrolled electors</li> </ul>  |
| <ul style="list-style-type: none"> <li>Replacing the MEO with a 'continuous' system would allow electors of Māori descent to change rolls once every electoral cycle.</li> <li>By choosing a trigger based on the timing of the electoral cycle restrictions to changing electoral and statistical systems independently are minimised.</li> <li>A continuous or 6 yearly decoupled MEO would enable the MEO to be timed to reduce the chances of it occurring in the same year as an election.</li> </ul> | <ul style="list-style-type: none"> <li>By choosing a trigger based on the electoral cycle, rather than the census, restrictions to changing electoral and statistical systems independently are minimized.</li> <li>A trigger based on the electoral cycle could occur at 6 or 9 yearly intervals. A 3-yearly review is considered too short as frequent changes could cause confusion among voters and would be administratively costly.</li> <li>A boundary review triggered by the census could occur on a 5 or 10 yearly survey cycle.</li> </ul> | <ul style="list-style-type: none"> <li>Using a Certificate of Population is consistent with the Local Electoral Act and would allow the Government Statistician to decide the best data source to be used at the point in time it was required.</li> <li>The ERP could provide a data source if Statistics NZ developed it to the minimum requirements for using it for electoral processes.</li> <li>Using enrolled electors is not considered practical as it does not provide accurate population data due to significant differences in enrollment rates.</li> </ul> |