

# Applying 2018 Census data to electorate allocation





Crown copyright ©

[See Copyright and terms of use](#) for our copyright, attribution, and liability statements.

#### **Citation**

Stats NZ (2020). *Applying 2018 Census data to electorate allocation*. Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).

ISBN 978-1-98-858396-9 (online)

#### **Published in July 2020 by**

Stats NZ Tatauranga Aotearoa  
Wellington, New Zealand

#### **Contact**

Stats NZ Information Centre: [info@stats.govt.nz](mailto:info@stats.govt.nz)

Phone toll-free 0508 525 525

Phone international +64 4 931 4600

[www.stats.govt.nz](http://www.stats.govt.nz)

# Contents

<b>Purpose and summary .....</b>	<b>5</b>
<b>Executive summary .....</b>	<b>6</b>
<b>About this paper .....</b>	<b>7</b>
<b>The data used in electoral calculations and their sources .....</b>	<b>8</b>
Māori electoral option data .....	8
<b>Uncertainty in the census count about representation .....</b>	<b>10</b>
Under- and over-coverage in the population count .....	10
Attribute non-response .....	11
Mis-enumeration .....	11
<b>Census data: its qualities and its application to electoral calculations .....</b>	<b>13</b>
Non-response mitigations in the 2018 Census .....	13
Quality of the count: previous censuses .....	13
Quality of the count: 2018 Census .....	16
Effect of mitigations on electoral district calculations .....	20
<b>Discussion .....</b>	<b>35</b>
<b>Conclusion .....</b>	<b>36</b>
<b>References .....</b>	<b>37</b>

# List of tables and figures

## List of tables

1 Results of the MEO .....	9
2 Net national undercount.....	14
3 Net undercount of those with Māori ethnicity.....	14
4 Selected contributors to the census count in 2013 and 2018 .....	28

## List of figures

1 Relative difference of 2013 and 2018 Census from benchmark, by TALB area.....	19
2 Evolution of number of general electorates since the 2013 Census, showing mitigation record source.....	21
3 Evolution of Māori electorates since the 2013 Census, by mitigation record source .....	22
4 Evolution of Māori electorates since the 2013 Census, by Māori descent attribute source.....	23
5 2018 North Island electoral populations, by 2014 electorate boundaries and record source.....	26
6 2013 North Island electoral populations, by 2014 electorate boundaries and record source.....	29
7 2018 South Island electoral populations, by 2014 electorate boundaries and record source.....	30
8 2013 South Island electoral populations, by 2014 electorate boundaries and record source.....	31
9 2018 Māori electoral populations, by 2014 electorate boundaries and record source .....	32
10 2013 Māori electoral populations, by 2014 electorate boundaries and record source .....	33
11 2018 Māori electoral populations, by 2014 electorate boundaries and Māori descent source .....	33
12 2013 Māori electoral populations, by 2014 electorate boundaries and Māori descent source .....	34

# Purpose and summary

## Purpose

*Applying 2018 Census data to electorate allocation* explains how the 2018 Census data has been applied to electoral allocation and shows how the 2018 Census non-response mitigations contributed to the number of electoral districts and electoral populations. Novel methods of non-response mitigation were required for the 2018 Census due to the lower than expected response rate.

While mitigating for non-responses in the 2018 Census, Stats NZ also worked to understand how those mitigations might affect electoral calculations. The census population count helps determine the numbers and boundaries of the electorates, and a sufficiently low census response will reduce the number of electorates and increase the geographical size of the electorates. This paper provides a summary of Stats NZ's work, including some background information for context and to aid understanding.

This paper shows how each source of information for the census contributed to the electoral calculations. We demonstrate the effects both from the point of view of adding people to the census file by the source of the record and the sources of information used to derive Māori descent. These effects are illustrated for calculating the number of electorates and their population.

## Summary of key points

For the 2018 Census, Stats NZ used a range of administrative data sources as well as data from the census field collection to obtain the population count and the demographic characteristics and attributes of that population. This was necessary in order to mitigate the lower than expected response rate to the census. After each census, the Government Statistician uses data from the census and the Electoral Commission to provide a report to the Surveyor General. The report specifies the new electoral populations, the new number of electorates and the electoral population that is legally required for each electorate. This report allows the Representation Commission to review the electoral boundaries and make the appropriate changes to comply with the Electoral Act 1993. A lower than expected response rate can affect electoral representation by suggesting the need for fewer electorates than there should be or electorates with larger boundaries than appropriate.

This paper shows how the 2018 Census non-response mitigations contributed to determining the number of electorates and their populations. The number of electorates determined by using final 2018 Census data is the same as the number that would have been produced using the estimated resident population. The use of administrative data for the Māori descent attribute has likely improved the count of people of Māori descent, and therefore improved the representation of people of Māori descent. Evidence suggests that the electoral calculations for both the number of electoral districts and the drawing of the boundaries will be at least as robust as the calculations following the 2013 Census, if not more robust.

While data from the 2013 Census is included for context, the different error structure of the 2013 and 2018 files makes it difficult to make any useful comparisons. The judgements about inclusions in the 2013 file are not reproducible as they result from a series of operational decisions. By contrast, the addition of administrative data in the 2018 Census is done transparently. With this qualification, the effect of the non-response mitigations on the sizes of the electorates is reproduced for 2013 and 2018 Censuses.

## Executive summary

This paper is the final in a suite of three that analyses the use of 2018 Census data for electoral purposes. Work started soon after the 2018 Census to understand how the non-response mitigations applied to the Census 2018 data contributed to calculating both the number of electoral districts and the electoral populations. The first two papers were produced by Dot Loves Data under contract to Stats NZ, while this paper was produced by Stats NZ.

The first paper, [Electoral boundaries sensitivity analysis of 2018 Census data](#) (Stats NZ, 2019c), considered the sensitivity of electoral boundaries to the addition of records to meshblocks. The second paper, [Population counts sensitivity analysis of 2018 Census data](#) (Stats NZ, 2019e), investigated the sensitivity of varying both the number of people added to households and the number of people added to meshblocks from admin data. The sensitivity of meshblock sizes is important for setting electoral boundaries and determining electoral populations. Some of the conclusions Dot Loves Data reached from these investigations were:

- The 2018 Census is robust to the thresholds in the mitigations for the purpose of determining electoral boundaries and representation.
- The number of Māori electorates remains at seven, regardless of what threshold is chosen.
- The number of North Island general electorates is 49, and this would drop to 48 only if a further 35,000 or more admin records were to be eliminated from the final census file.
- Choices made in census processing affect fewer than 2.7 percent of the population.
- Admin enumerations have increased the representation of Māori and Pacific peoples in census data.
- Neighbourhood deprivation is the most important predictor of sensitivity in a meshblock.
- Although some groups of people remain difficult to count, such as young adults, children, and Māori, the use of administrative data has made the uncertainty in these counts practically negligible for many uses of census data.

This final paper shows the effects of the mitigations to the 2018 Census data on the calculations for the number of electoral districts and their electoral populations at each step of the process. The paper discusses how each mitigation has contributed to the final electoral calculations shown in [The mathematics of electorate allocation in New Zealand based on the outcome of the 2018 Census and Māori Electoral Option 2018](#) (Stats NZ, 2019g). We reached the following conclusions:

- The number of electorates using final 2018 Census data is the same as that which would have been produced using the 2018 estimated resident population (ERP).
- There is some evidence that the overall undercount in the 2018 Census is similar to, if not lower than, that found in the 2013 post-enumeration survey (PES). However, the 2018 PES will provide an official measure.
- There is also evidence that both the undercount and the variation in the undercount have been reduced at lower-level geographical areas, at least at the territorial authority local board area (TALB) level (figure 1). This improves representation through the calculation of the number of electoral districts and their electoral populations.
- The use of administrative data has improved the count of people of Māori descent (see: [Deriving the 2018 Māori descent electoral population](#), Stats NZ 2019b) and thus improved representation for those of Māori descent. This improvement applies both to the calculation of the number of electoral districts and their electoral populations.

As a result of these findings, this paper suggests that the electoral calculations for both the number of electoral districts and the drawing of the boundaries will be at least as, if not more, robust than after the 2013 Census.

## About this paper

For the 2018 Census, Stats NZ developed new methods to deal with unit non-response instead of using unit imputation, which is also known as the adding of substitute records. These new methods used administrative or admin data for enumeration. Stats NZ has done detailed analyses of the resulting census dataset and its applications to assess these new methods. This paper is concerned with applying the census data to the calculation of the number of electoral districts and electoral populations. It covers the following topics:

- It begins with a brief introduction to the data needed for the electoral calculations and their sources.
- This is followed by a brief review of how the electoral calculations rely on the census, including how specific weaknesses in a census will affect electoral calculations.
- After reviewing the non-response mitigations used in the 2018 Census, the paper considers the historical context of the count. It compares the traditional census process with the enumeration process of the 2018 Census.
- It goes on to illustrate the calculations for the number of electoral districts, showing the effect of the non-response mitigations. This is followed by the calculations of the electoral district populations. The effects of the 2013 Census unit substitution and Māori descent imputation are shown for comparison.
- The paper ends with a discussion of the final results in the context of representation.

The paper also discusses electoral calculations, but readers should refer to Stats NZ 2019g for the full details. Similarly, the paper briefly summarises the use of administrative data in the 2018 Census, but the paper [Overview of statistical methods for adding admin records to the 2018 Census dataset](#) (Stats NZ, 2019d) provides more information.

The results of the analysis inform our understanding of the effect of non-response mitigations in relation to electoral allocation. The results also show how the 2018 Census data relates to other sources of information on the population, such as the estimated resident population (ERP). These analyses are novel and are presented here for readers' interest.

## The data used in electoral calculations and their sources

The electoral calculations seek to determine:

1. the number of general and Māori electoral districts
2. the required electoral populations of the general and Māori electoral districts, known as the quota.

The results of these calculations are used by the Representation Commission to assess and, if required, set or reset the boundaries of the electoral districts. The calculations require the population counts of the general population and of those of Māori descent, which come from the census. When combined with Māori electoral option (MEO) data from the Electoral Commission, those counts enable us to calculate the number of Māori and general electorates and to determine the ideal number of people for each electorate – the quota.

The locations of the counts of usual residents affect the positions of electorate boundaries. These boundaries are drawn to ensure an electorate will have a certain electoral population, taking into account topographical and community information. The location of the usual residents by island also affects the number of electorates and the quotas for the different types of electorate. We provide a brief summary of the calculations below; for more comprehensive details on the electoral calculations, see Stats NZ 2019g.

The mathematics of the electoral calculations is quite straightforward, but the interactions can be complex. Part of the reason for this is that the process requires allocation of the total population to the Māori and general electoral populations. If one part of that allocation process contains an error, then the other part will also be affected.

It is first necessary to derive the Māori electoral population, using the count of people of Māori descent and the results of the MEO. The general electoral population is then simply the total population less the Māori electoral population. This calculation is needed both when determining the number of electoral districts and when calculating the population within an electoral district. The calculation for the number of Māori electoral districts and the general electoral districts in the North Island is obtained by dividing the Māori electoral population and the North Island general electoral population by the South Island quota. This ratio is rounded as specified in the Electoral Act 1993 to give the number of electoral districts in each case. The number of general electoral districts in the South Island is fixed at 16.

### Māori electoral option data

The MEO is where those of Māori descent who are eligible to vote choose to be on either the general or the Māori roll. The ratio of those of Māori descent who choose the Māori roll over the total enrolments of those of Māori descent on both the Māori and general rolls determines what proportion of the population of Māori descent is counted in the Māori electoral population or the general electoral population. From now on, this will be referred to as the option ratio. These calculations are discussed in more detail in Stats NZ 2019g.

The choice between rolls is exercised after each census and is combined with the census data to recalculate the number of electorates and their boundaries. In general, an increase in the option ratio results in a larger Māori electoral population. The option ratio fell after the 2018 option was exercised. The option ratios are shown in table 1.

**Table 1**

<b>Results of the MEO</b>					
<b>Option year</b>	<b>1997</b>	<b>2001</b>	<b>2006</b>	<b>2013</b>	<b>2018</b>
<b>Ratio</b>	53.6	55.4	57.8	55.7	52.8

The option ratio is now at its lowest point since the electoral reforms of 1993.

## Uncertainty in the census count about representation

In this section, we'll look at how the results of the electoral calculations rely on census data in a general sense.

### Under- and over-coverage in the population count

We think of under- and overcount as the difference between the true population and the population counted by census. Overcount is where people have been counted when they should not have been, while undercount is where people have not been counted when they should have been. Overcount can arise through, for example, duplicate forms or including out-of-scope records. Duplicate records can occur where a person responds more than once, but the duplicates are not identified and removed from the file. An out-of-scope record could include an overseas resident as a usual resident of New Zealand. Undercount typically means those who should have been counted but were not, whether from being missed, 'silent' refusals, or other reasons. Typically, a census will have elements of both types of count, with undercount the greater component, leading to net undercount.

Coverage refers to expressing this count in terms of a proportion or percentage. Coverage is the proportion of the population counted in a census and is typically net under-coverage.

The effect of under-coverage in the data used to calculate the number of electoral districts will depend on where and what population is undercounted. The relative differences in the undercount between populations is also relevant. For example, suppose that the South Island population of Māori descent is undercounted relative to the South Island general population. The South Island general population will then be larger than it should be, and the South Island quota will also be too large. Since the South Island quota is the divisor for determining the number of North Island and Māori electoral districts, this will bias these calculations downwards. Whether this under-coverage reduces the number of electoral districts for the North Island and Māori will depend on the degree of the under-coverage and how close the results of the calculations are to a change in the number of electoral districts. At the island level, it is important that the counts are as accurate as possible when calculating the number of electoral districts.

At an electoral district level under- or over-coverage affects the representation of the people within that electoral district. Note that net under-coverage at a national level does not preclude net over-coverage at lower levels of geography. A new quota is calculated for each type of electorate, South or North Island general, and Māori electoral districts. See Stats NZ 2019g for further details.

The boundaries are set so that the electoral populations are within a certain range of the quota. Suppose the boundary has been set but there is undercount within the area defined by the boundary. The Member of Parliament (MP) for that electorate will be considered to be representing the correct quota but will be in reality representing a greater number of people: the electoral population calculated from the count and those who were not counted. In this example, the people within that electorate are under-represented, that is, more people are represented by one MP than should be. Conversely, counting more people than are really in an electorate will lead to over-representation of the true population within that electorate. Since the electorates are constructed from meshblock counts, accurate counts at that level of geography are important. For clarity, a meshblock (MB) is the smallest geographic unit, typically containing approximately 30–60 dwellings and consisting of around 60–120 residents (Stats NZ, 2017b).

## Attribute non-response

Even when someone is counted in the census, we may not receive all the attribute information required for electoral purposes. Attributes relevant to electoral calculations, and therefore representation, are an indication of being of Māori descent or not and a usual residence address which can be placed in a meshblock.

If there is no response to the question of Māori descent or an inconclusive response, then a response will be derived. See Stats NZ, 2019b for an explanation of this process. Note that this derivation is for electoral purposes only. If this derivation was not done then the population of Māori descent, and therefore the Māori electoral population, would be underestimated. Such an underestimation would reduce the ratio for the number of Māori electoral districts but increase the ratio for the number of general electoral districts. The Māori population would still be represented but not in the correct electoral population. The proportions within the general and Māori electorates would not correctly represent the Māori population.

If an individual record lacks a good address it will generally be coded to the lowest level of geography possible, starting with meshblock and moving up to statistical area 1 (SA1), statistical area 2 (SA2), etc. Thus, the accuracy of the geographic location of the counts improves as we move up the geographical hierarchy. To be represented a person must be placed in an electoral population within an electoral district boundary. Since everyone must be placed in a meshblock for representation purposes, a meshblock will be imputed for a person if required. The imputation uses all the address information available for the record, by finding another record with similar characteristics and copying that record's meshblock. All records will have a usual residence meshblock for electoral purposes. See [Data sources, editing, and imputation in the 2018 Census](#) (Stats NZ, 2019a) for further information on this process, termed nearest neighbour imputation. The consequence of placing an individual record into the wrong meshblock is described in more detail below.

## Mis-enumeration

Mis-enumeration occurs when a person's usual residence is recorded in a different meshblock to their correct usual residence. This typically arises in the creation of the address list either in the field or before the field phase of the census.

The design of the electoral districts mitigates the effect of possible lower accuracy at smaller geographies due to field mis-enumeration or the imputation process. As the districts are relatively large compared to meshblock populations, the quota now being in the order of 65,000, a small undercount in one meshblock is likely to be offset by that person being included in a nearby meshblock.

If the effect is not mitigated by the placement of the boundaries of the electoral district, then the population will either be too high or too low, leading to under- or over-representation as described above. However, any effect is limited to the difference between the 'true' population and the 'counted' population. For example, if an enumeration error resulted in one meshblock having a count of 99 and its neighbouring meshblock 101, where the true count in each was 100 and the electoral boundary was between the meshblocks, then the electoral populations would be in error by one, leaving aside questions of Māori or general populations.

For each electoral district there is a tolerance for the size of its electoral population. When the electorate boundaries are set, the electoral population may vary by up to  $\pm 5$  percent from the quota (sections 36 and 45(7) of the Electoral Act 1993). However, the setting of those boundaries must

consider, among other things, the population projections of electoral populations (sections 35(3)(f)(v) and 45(6) of the Electoral Act 1993). The aim in considering the projections is to reduce the chance of under- or over-representation due to population changes prior to the next review, following the next census. This tolerance also mitigates the effect of any mis-enumeration.

## Census data: its qualities and its application to electoral calculations

In this section we review the non-response mitigations to the 2018 Census then provide some historical context around the qualities of census data. We consider how the mitigations in the 2018 Census may have addressed some of the historical features of census data. Finally, we combine the information on electoral calculations with the 2018 Census data to explore how the mitigations have led to the electoral results.

### Non-response mitigations in the 2018 Census

As a result of lower than expected response rates, the 2018 Census used admin data about usual residents to supplement the responses received through field collection. This approach was conservative, recognising the limitations of using admin data. The mitigations focused on adding people who were New Zealand residents, who were in New Zealand on census day, and who did not respond to the census. People were placed into their households where possible, or into their meshblock when this improved census information for small geographies (Stats NZ, 2019d).

People were added to the census population according to four criteria, worked through sequentially:

1. People who were listed on a census household form but did not submit an individual census form were included in the household.
2. Where there was information from a household that indicated others lived in the household who had not completed a form, we added them into the household if they were found through admin data.
3. Where there was no response from an address, we checked the address and, where appropriate, added people into the household via admin data using a model that predicts those people likely to have reliable data.
4. If we didn't have sufficient evidence to reliably place a person in a household, they may have been added to the meshblock where they live. People were included in a meshblock using a model that predicts those that are likely to have reliable admin meshblock data.

In this paper, we consider people added to households via admin data (points 2 and 3 above) as one non-response mitigation, resulting in four types of census records described in relation to electoral calculations:

- 2018 Census responses
- household listings (1 above)
- admin household enumerations (2 and 3 above)
- meshblock enumerations (4 above).

### Quality of the count: previous censuses

As described above, uncertainties in the census population count can affect both the calculations for the number of electoral districts and the locations of the boundaries. Historically, there has been net undercount in the census both at the national level and both under- and overcount at meshblock level with undercount dominating. Since 1996, a PES has been conducted after each census to assess how well the census counted the population. The overall net undercount for previous censuses is shown in table 2.

**Table 2**

<b>Net national undercount</b>		
<b>Census</b>	<b>Net undercount (%)</b>	<b>Sampling error (95% confidence)</b>
1996	1.6	0.2
2001	2.2	0.3
2006	2.0	0.4
2013	2.4	0.5
<b>Note:</b> 1996 was the first census to have a PES.		

Undercount affects representation by reducing the electoral populations, both general and Māori. At the national level net undercount has a lesser effect on representation, subject to the effect on the calculations for the number of electorates and the populations within electorates (the quota). Net undercount can have a much more noticeable effect on representation at lower levels of geography, especially if it is uneven across the country. These effects are discussed later in this section.

Another crucial aspect of the count for electoral calculations is the accuracy of the count of those of Māori descent. Undercount of those of Māori descent will reduce the Māori electoral population. Table 3 shows the net undercount, as measured by the PES, of those with Māori ethnicity. Māori descent and Māori ethnicity are different concepts, with the expectation that the number of people of Māori descent, having Māori ancestry, is larger than those identifying with the ethnicity, a cultural affiliation. The overall national net undercount is repeated for comparison in table 3.

**Table 3**

<b>Net undercount of those with Māori ethnicity</b>				
<b>Census</b>	<b>Net undercount of those with Māori ethnicity (%)</b>	<b>Sampling error (95% confidence)</b>	<b>Net national undercount (%)</b>	<b>Sampling error (95% confidence)</b>
1996	3.7	0.7	1.6	0.2
2001	4.4	1.0	2.2	0.3
2006	3.1	1.2	2.0	0.4
2013	6.1	1.3	2.4	0.5

It is clear from the PES data in tables 2 and 3 that there is an overall undercount of the population but the undercount of those with Māori ethnicity is greater.

Total non-response in the previous censuses includes not only the net undercount from the PES, but also the number of people for whom a substitute form was raised in the census and who are therefore included in the census count (Stats NZ, 2014). It seems reasonable to expect that, based on the consistently higher rate of undercount for Māori found in the PES, the people counted in the census through substitute records may also represent a higher proportion of Māori than those counted through a census response. This suggests the relative non-response of Māori could be even higher than when only considering PES results.

For the 2013 Census, we can estimate the Māori ethnic response as 89.2 percent from the 2013 PES. Of the 10.8 percent shortfall in population count, under-coverage is 6.1 percent and non-response is 4.7 percent. This non-response figure assumes equal proportions of non-response across ethnicities. However, we can use the under-coverage rates found in the PES and allow the substitute records raised to follow those proportions. Using simplifying assumptions and only looking at ethnicity, it is possible to estimate that the non-response of those with Māori ethnicity in 2013 may have been closer to 9.7 percent, giving a response rate of 84.2 percent.

The method used in previous censuses for imputing Māori descent used the distribution of the responses to inform the imputation. It is assumed that the missing population is missing at random, that is it follows the same distribution as the responding population. This assumption is weakened where there is evidence that the non-responding population is different in some way, such as there being a greater proportion of a characteristic in the non-responding population than in the responding population.

A higher level of non-response amongst people with Māori ethnicity, based on the census coverage results, suggests that the non-response of the population of Māori descent may not be compensated for adequately when missing at random is assumed, leading to attribute error as discussed above. This suggests the representation of people of Māori descent would be underestimated and therefore the Māori electoral population underestimated with consequential effects on representation. The exact level is difficult to estimate. While we expect the population of Māori descent to be greater than the population of Māori ethnicity, we have historically mitigated the non-response by imputing Māori descent for the units imputed or substitute records. For further information see [Imputation of Māori Descent for Electoral Calculations](#) (Stats NZ, 2000).

There is a difference from a representation perspective between undercount and non-response. For the censuses up to and including 2013 non-response could be addressed through unit imputation or substitute records. Such a record would have a usual resident meshblock and would have its Māori descent status imputed. With this information the record for that person could be included in the electoral population and thus the person represented for electoral purposes. If a person did not respond and did not have their record imputed then that person was not included in the count but rather was a part of the undercount. In this case the person is not a part of the electoral population and is not represented.

The 2013 PES also reports that undercount varies by geographic areas. For example, the northern North Island had a net undercount of 3 percent (sample error 0.7 percent) while the southern North Island had 1.3 percent (sample error 0.4 percent). These differing rates suggest there will be variation in how well people are counted in any given electoral district. Undercount at the national level reduces the electoral district ratios which may affect the numbers or the quotas of the electoral districts. If the undercount of electoral populations is uniform across the country then the representation will be relatively equally affected. Localised regional undercount is a greater problem because it is more likely to affect representation in a few electoral districts. For example, higher undercount for a given area means the true population is greater than the counted population, so the electoral district contains more people than allowed for.

At the small geographies, in 2013 and 2006, there was a target to obtain at least a 90 percent response rate in each subdistrict. This is the area given to each collector to deliver and collect forms from, made up of several meshblocks. In 2006, 3.7 percent of subdistricts did not achieve this level, and in 2013 it was about 11 percent.

## Quality of the count: 2018 Census

The 2018 Census has used novel methods, as described above, to obtain the final census data. How successful this has been can only be assessed by using data from the 2018 PES, which is not yet available. However, in the absence of those results, it is still possible to assess the relative quality of the two censuses by comparing and contrasting the 2013 and 2018 Census processes. The approach used in the 2013 Census is broadly comparable to earlier censuses.

In the past, censuses have relied heavily on observation by a collector and direct interaction with the census process by people within a limited time period. The 2018 process allowed other sources of information to contribute to the count in the absence of response or observation. This mitigates both under-coverage and non-response at a national and subnational level. The traditional method of dealing with non-response – unit imputation or the raising of substitute forms – relies on some form of observation or determination that people have been missed. However, the traditional method does not directly address under-coverage, there being no mechanism to do so.

A further advantage of the 2018 Census process is that where a record for a person is included from admin data, that person's Māori descent can also be included if it is available. This use of administrative data has likely improved the count of people of Māori descent in the census (Stats NZ, 2019b). Consequently, there is also an improvement in representation of those of Māori descent. As discussed, there was a shortfall in coverage in the 2013 Census and the 2018 Census responses that the mitigations for the 2018 Census have strived to address. How well this has been addressed will become clearer when the 2018 PES is available. In particular, this is the first PES to ask about Māori descent and not just ethnicity.

The 2018 Census also took advantage of other available estimates of the resident population of New Zealand. These afforded an understanding of the effects of the mitigations for the 2018 Census on the counts in relation to the estimates. As we will see in the next section the mitigations have enabled the census counts to approach these other estimates. Again, we would expect the PES to provide an estimate of the accuracy of the count in the census.

As discussed previously, the accuracy of the counts of population and those of Māori descent are important at both national level and smaller geographies. The counts at the smaller geographies ensure that the electoral boundaries are in the right place for good representation. Undercount, non-response and attribute error apply across all geographies, but at meshblock level mis-enumeration becomes relevant.

In 2013 and earlier censuses, one of the most common ways for mis-enumeration to occur was by a collector enumerating dwellings on both sides of a street when the middle of the street was the boundary between two different collector areas. The dwelling file would then have a set of addresses listed as being in the wrong meshblock. Unfortunately, there is a lack of information on the quantification of this problem. Typically, any mis-enumerations were fixed as they were found in the census process because people were aware this problem could occur, although it is possible a small number may remain in the final data.

The 2018 Census started with an address list that already had meshblock locations and was run in the context of the organisation having tools to deal with spatial data. This allowed the use of spatial data to improve the quality and management of address data. The address list was checked extensively, including in the field, and changes were tied to the geospatial data. These processes reduced the opportunity for mis-enumeration from the field and made identification of any errors and their correction much easier. Records from responses or the household listing will have address data that has been through this process.

However, the addition to the census file of records from administrative data does open another pathway for mis-enumeration, where the admin record has the wrong meshblock for a person. This can occur, for example, through the data quality of the address or through the usual residence of a person changing at census time before the admin data has had time to update. More detail is available on the quality of the address information in the admin data in [Experimental population estimates from linked administrative data: 2017 release](#) (Stats NZ, 2017a).

Mis-enumeration cannot be viewed in isolation. It can be argued that having a greater number of people counted, and thus represented, is a gain that can tolerate some mis-enumeration. This gain is obtained at both national level, in calculating the number of electoral districts, as well as at local level for the electoral district boundaries. The only time mis-enumeration will affect representation is if it occurs between meshblocks that straddle an electoral district boundary. The effect will be limited to the differences in counts of usual residents and usual residents of Māori descent. As discussed above, there are practical tolerances in the legislation for the allocation of electoral populations to the electoral districts.

We can compare the 2013 and 2018 Censuses in the context of representation. The 2013 Census had a greater undercount by certain ethnic groups and geographical areas, but of those who were counted their meshblock is likely to be accurate. The 2018 Census, by contrast, appears to have better counts of those ethnic groups and a lower undercount at the smaller geographic areas. However, there are more inaccuracies of location for individuals where they may be in the wrong meshblock when their record comes from admin data alone.

The 2013 Census data tended to provide boundaries which were more accurate for the counted population but at the expense of the count overall and at lower geographies. This reduced the result of the calculation for the number of electoral districts and made under-representation more likely. In this case, we thought there were a certain number of people in an electoral district when in fact there were more people. The Member of Parliament was therefore representing more people than the ideal.

In comparison, the 2018 Census will better represent the number of people overall and at electoral district level. The calculations of the number of electoral districts and the number of people who should be in each electoral district will be more accurate. However, the boundaries may not be as accurate because of the less accurate meshblock counts. In compensation, the number of people in each electoral district overall should be more accurate. More people will be represented but it is slightly more possible that there will be people not in their correct electorate. There may be both under- and over-representation at electoral district level.

We can look at the approximate magnitudes involved in the two censuses. The 2013 PES estimated the net undercount at 104,000 with the gross undercount at 135,000. We can estimate the number of people in the 'wrong' meshblock in the 2018 Census to be 110,000 from information in [Predicting the quality of admin location information for use in the 2018 Census](#) (Stats NZ, 2019f). The magnitudes are similar but the error in the meshblocks in the 2018 Census will only manifest itself if the electoral boundary is drawn between the meshblocks someone was counted in and where that person should have been counted. Even then the effect will be limited to the differences in counts. If the meshblocks sharing the mis-enumeration are in the same electoral district then representation will be unaffected.

## Comparing the quality of the Censuses

The 2018 Census should have better representation both at the overall level, for calculating the number of electorates, and at the lower level, for the electoral populations, than the 2013 Census.

This gain in representation outweighs the loss in accuracy due to less precise location information, and is preferred to the undercount and thus under-representation in the 2013 Census.

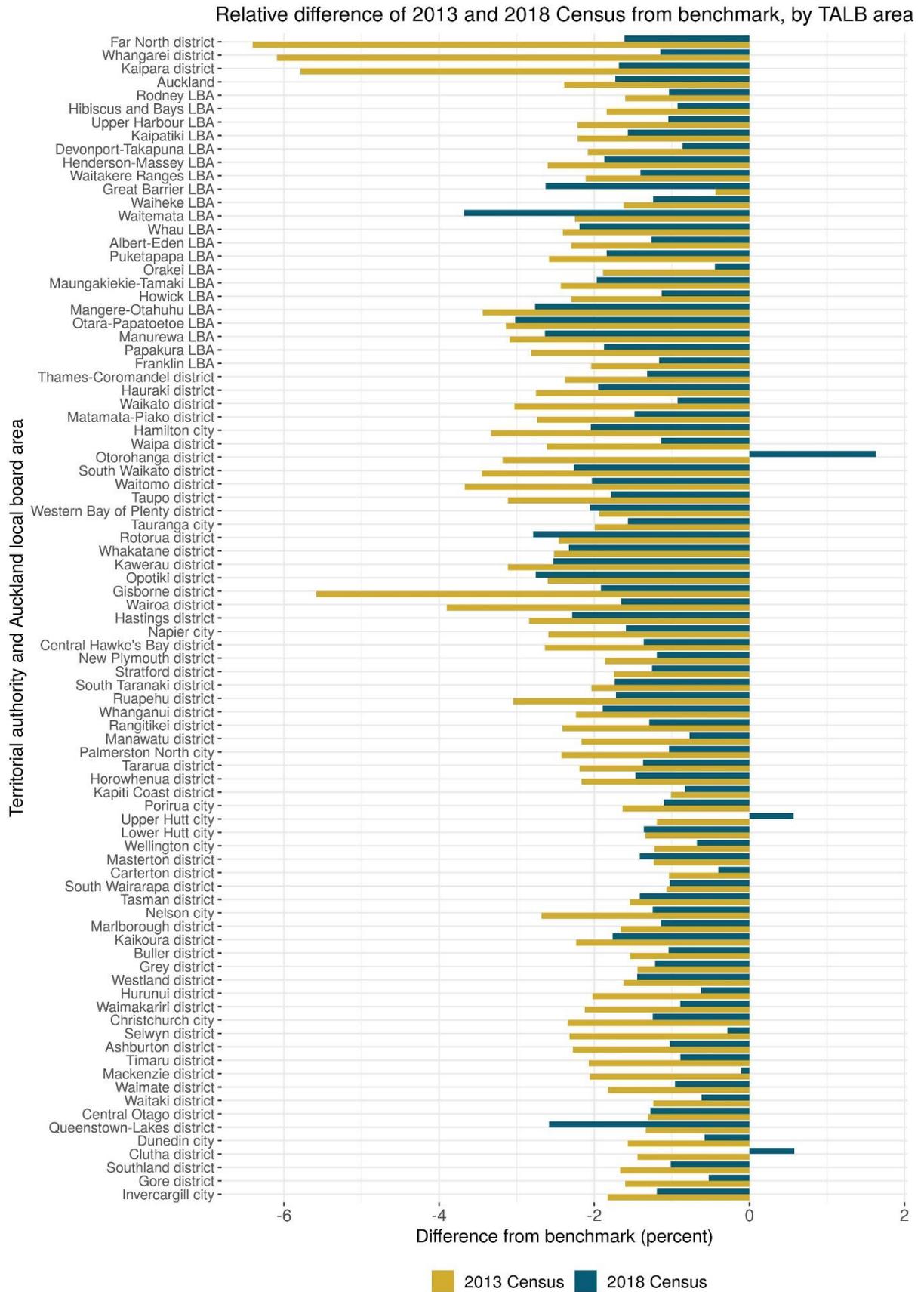
It is difficult to make a good assessment of the accuracy of the counts at meshblock level but an assessment has been made at Territorial Authority/Auckland Local Board level (TALB). In Stats NZ 2019d, a comparison is made at the TALB level for the 2013 and 2018 Censuses. A plot of the relative differences from the benchmarks is included in that paper as figure 13 and this is reproduced below as figure 1. The benchmarks for the 2013 Census use the estimated resident population (ERP) while the 2018 benchmark is from the dual system estimation model (DSE). The measures used employ the same concepts of population as the census. Further details are available in Stats NZ 2019d. The figure has TALB on the vertical axis and the difference from the benchmark in percent for each area on the horizontal axis.

The results in the figure show an improvement in the 2018 Census counts against the benchmarks. The majority of areas show undercount for both the 2013 and 2018 Censuses. The 2013 Census undercount levels tend to have a wide range and in some cases are quite high. In contrast, the 2018 Census areas have a narrower range of undercount and the range is at a lower level. This is notable for electoral purposes because it indicates an improvement in the count across the country.

For representation purposes, we need a good count and a count which is consistently good across all geographic areas. Wherever we have pockets of undercount or overcount it is possible for representation at an electoral district level to be less than the ideal. Such pockets can lead to imbalances in representation at electoral level as well as when calculating the number of electoral districts.

The figure of relative differences provides evidence that the mitigations for the 2018 Census have improved coverage across TALB areas and reduced the variation between TALB areas. These areas are much larger than meshblocks, the building blocks of electoral districts, but about two-thirds of the TALBs have smaller populations than any given electoral district. This provides some evidence that representation will be improved within electoral districts in 2018.

Figure 1



## Summary: the census data and its qualities

The counts of usual residents and usual residents of Māori descent at census night are fundamental to good representation. The undercount apparent in previous censuses would have had some effect on representation. It is likely that the greater undercount of those with Māori ethnicity also suggests a greater undercount of those of Māori descent. Errors of mis-enumeration need to be kept to a minimum but are likely to be of less practical importance. Evidence suggests the 2018 Census has improved population coverage at smaller geographic levels and therefore representation within electoral districts.

## Effect of mitigations on electoral district calculations

We now turn to the electoral calculations and illustrate how those calculations changed as each mitigation was carried out on the 2018 Census data. The first part looks at the calculations for the number of electoral districts for the general and Māori electorates. The population counts from the 2013 Census, 2013 ERP, and the 2018 ERP using the 2013 Census base are included for comparison.

The second part looks at the electoral populations within each electoral district and how these have evolved with each mitigation. The 2013 Census is included as a comparison.

## Calculating the number of electoral districts

There are two ways the various components of the mitigations employed for the 2018 Census contribute to the electoral populations. These are:

- from a record source point of view
- from an attribute source point of view.

The record source looks at where the record for a given person in the 2018 Census data came from. This lets us see how the calculations have changed with the addition of each type of record or enumeration. The attribute source considers the source of the information for the Māori descent attribute. We can see the effect on the calculations as we add the attribute data from each source. In this second case I have assumed the number of records is as given in the final data set, after all mitigations, and I have only allowed a change in a record's attribute.

The details of the calculations for the number of electoral districts are set out in the Electoral Act 1993 and are described in Stats NZ 2019g. The sources of admin data used in the 2018 Census are described in Stats NZ 2019d, while more details on the Māori descent attribute are in Stats NZ 2019b.

The cumulative effect of adding each source of data on the calculation for the number of general electoral districts or seats in Parliament is shown in figures 2–4. The calculations for the number of electoral districts are described in the first section of this paper and the outcome of the calculation is referenced hereafter as “the ratio”. The fraction part of the resulting ratio must be over one half for the number of electorates to be rounded up. The figures show the results of the ratio calculation with horizontal lines to separate the number of electorates on the y-axis. The x-axis is an indication of time.

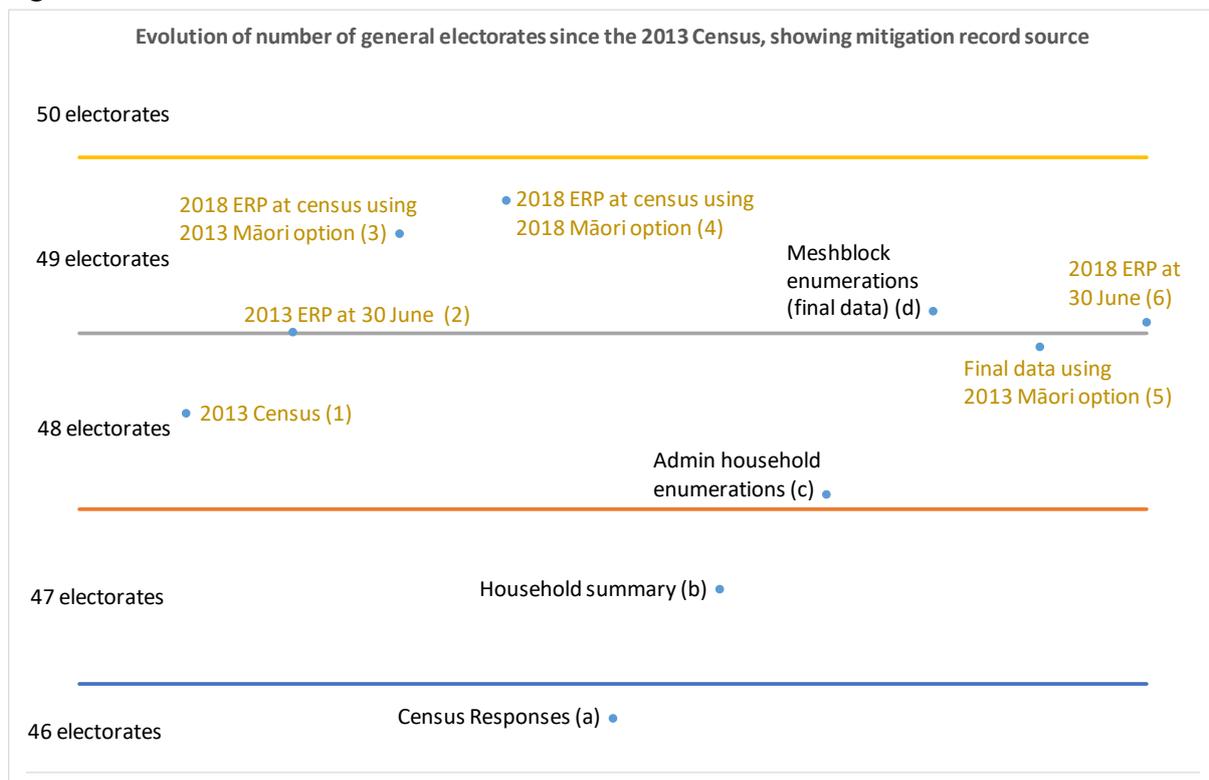
We also compare the calculation of electoral districts using the 2018 Census against calculations using the 2013 Census, the 2013 ERP and the 2018 ERP. Comparisons between the MEO for 2013 and 2018 are also shown due to the fall in the option ratio. The points directly related to the 2018 Census and its mitigations are identified by letters, while those with numbers are for context.

## General electoral districts

Figure 2 shows how the calculation of the number of general electorates has evolved during census processing. The points for context are in a lighter font and are followed by a number in parentheses, while those directly related to the 2018 Census are labelled in a darker font and are followed by a letter in parentheses. The points for context have been added to assist interpretation.

From left to right, after the 2013 Census, we reached 48 general electorates (1). The result of the 2013 MEO is used to derive this result. If we calculated the number of electorates using the estimated resident population (ERP) at the end of June 2013 (2), we would have crossed the threshold for 49 general electorates. The June 2013 ERP takes account of population growth since the 2013 Census day and the undercount in that census. Point 3 shows the number of general electorates using the ERP as at the 2018 Census but using the results of the 2013 MEO. This calculation of the ERP at census day was prepared for use in-house to assist the census process. See [The mathematics of electorate allocation in New Zealand based on the outcome of the 2013 Census and Māori Electoral Option 2013](#) (Stats NZ, 2013) for further details of the electoral calculations after the 2013 Census.

**Figure 2**



Point 4 uses the 2018 MEO result with the 2018 ERP as at the 2018 Census. As the option ratio fell after the 2018 MEO, fewer people were in the Māori electoral population and more were represented in the general electoral population. Keeping all other quantities equal, this resulted in an increase in the general electorate ratio, the number of South Island general electorates plus the calculation for the North Island general electorates.

The next four points illustrate the cumulative effect of the steps in creating the census data set as detailed previously. First (a) is the resulting general electorate ratio with census responses only, then those on the household summary or dwelling form only are added (b). Third is the admin household enumerations (c), and finally the admin enumerations within meshblocks (d). At this point, we have

49 general electorates, one more than after the last census. For each of these four points, a–d, we used the result of the 2018 MEO. The point from the final data set (d) conforms with the published electoral calculations in Stats NZ 2019g.

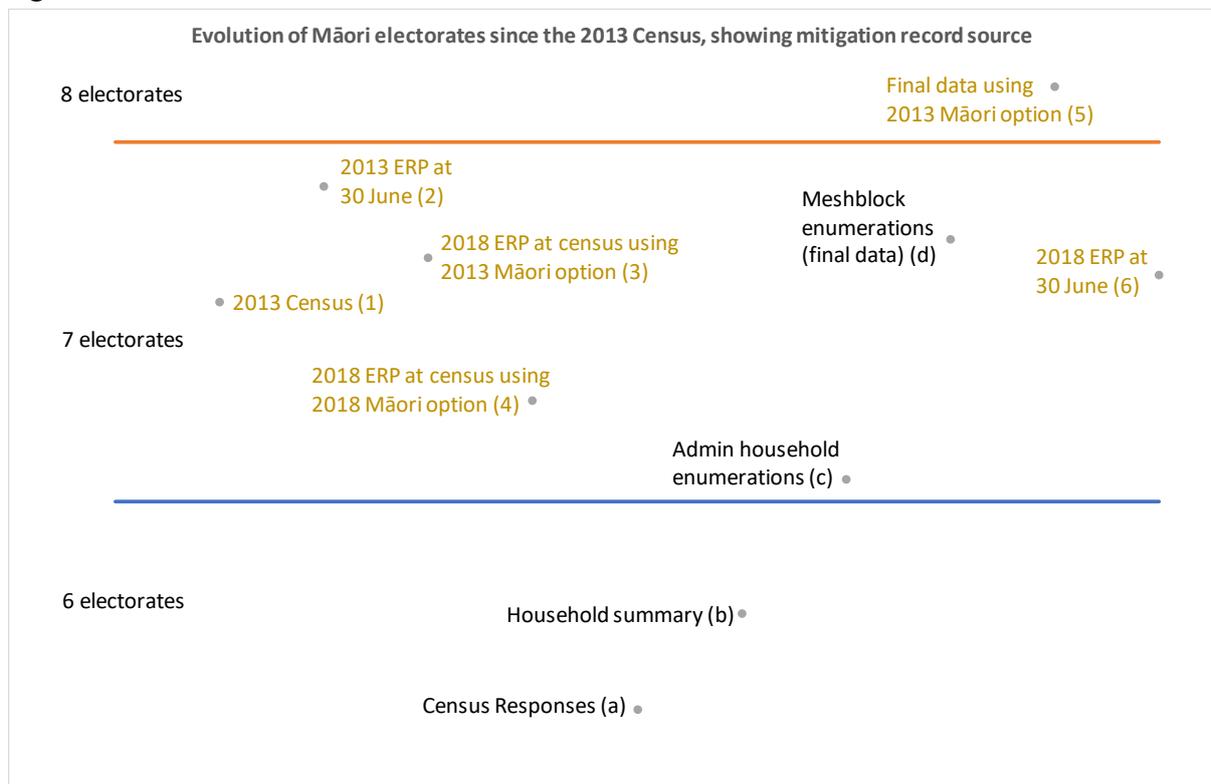
Two further points have been added for interest. Point 5 indicates that the number of general electorates would have remained the same as after the 2013 Census if the 2013 MEO ratio had been used with the 2018 Census data. Point 6 shows the resulting general electorate ratio, using the revised 2018 ERP at 30 June and the results of the 2018 MEO, which gives the same number of general electorates as the 2018 Census.

### Māori electoral districts

For the Māori electoral districts, we can view the evolution of the electoral district calculations by record source for non-response mitigation or by source of the Māori descent attribute.

Considering the record source first, figure 3 shows the ratio plotted and horizontal lines delineating numbers of electorates.

**Figure 3**



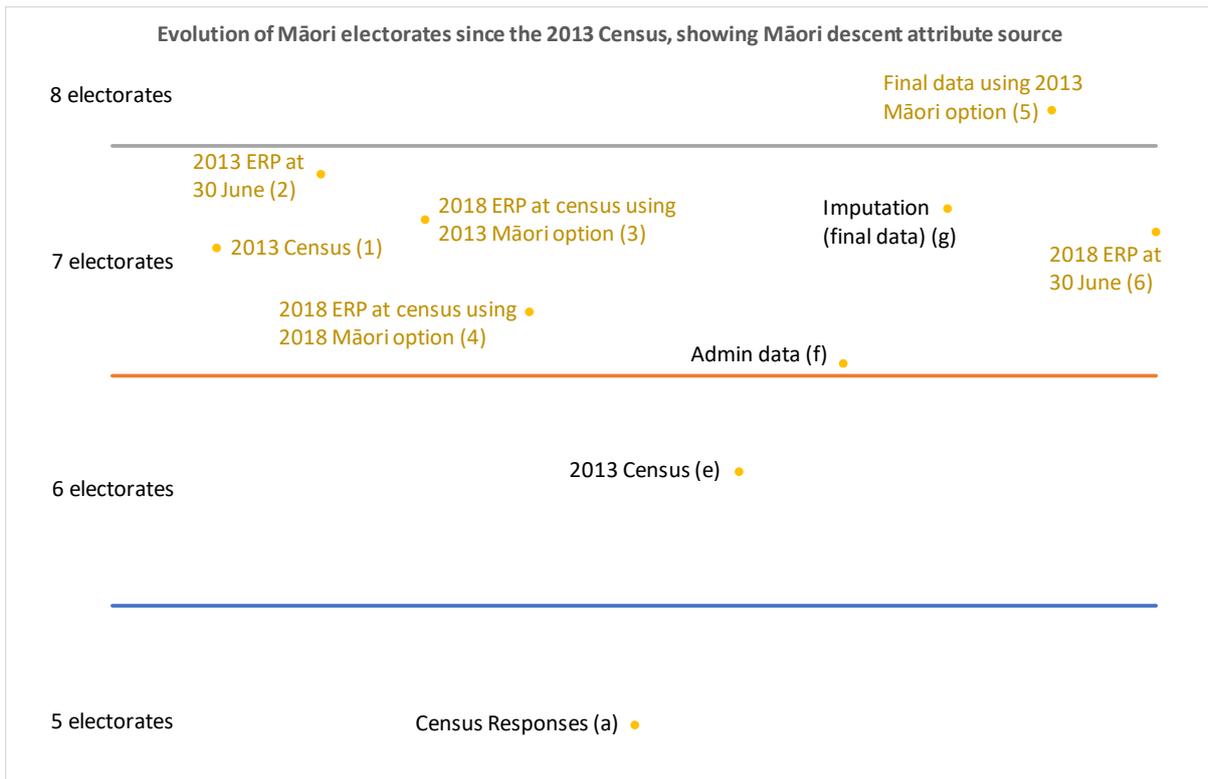
From left to right, the first two points illustrate the difference between the 30 June 2013 ERP (2) and the 2013 Census count (1), both using the outcome from the 2013 MEO. This difference is mainly due to undercount of the population of Māori descent in the census. Points 3 and 4 illustrate the difference due to the change in the MEO, using the same 2018 ERP as at the 2018 Census. The 2013 MEO is used for 3, and the 2018 MEO is used for 4.

The next four points similarly show the results for calculating the ratio for responses (a), adding those on the household summary or dwelling form (b), admin enumerations within households (c) and meshblock enumerations (d). Point d represents the final set of 2018 Census data. Note that the number of Māori electorates remains at seven as it was before the 2018 Census.

Point 5 illustrates the calculations for the 2018 Census data using the results from the 2013 MEO. Finally, point 6 shows the revised ERP at the end of June 2018 with the 2018 Māori option result. Note that had the ratio of people of Māori descent and eligible to vote choosing to be enrolled in the Māori roll remained the same as after the 2013 option, the number of Māori electoral districts would have increased to eight after the 2018 Census.

The second way of looking at the results is to recalculate the number of Māori electorates by adding the sources of the attribute of Māori descent, shown in figure 4. This analysis assumes the population is fixed at the total number of records in the final data, and the count of people of Māori descent is increased as additional sources are used to derive descent.

**Figure 4**



The first four points are the same as in figure 3. The next four points look at the effect of the attribute source data. Had the Māori electoral districts been calculated using census responses only, it would have resulted in five electorates (a). The next three points, show the effects of determining Māori descent by adding 2013 Census attribute data (e), admin data (f), and finally imputation (g). Point g represents the final set of 2018 Census data and aligns to point d in figures 2 and 3. Note that this is imputation of Māori descent for electoral purposes only. The majority of records has either a response from the 2018 or 2013 Census or an admin source for Māori descent. The final two points illustrate again the result using the 2013 Māori option ratio with the final 2018 Census data (5), and the revised estimate of the usually resident population at the end of June 2018, with the 2018 MEO results (6). Further details on the Māori descent electoral variable can be found in Stats NZ 2019b.

It is clear from figures 2 and 3 that the meshblock enumerations have made a substantial difference to the electoral district calculations, improving representation over census responses alone, as has including Māori descent information from 2013 Census and admin data.

The two figures also illustrate how the mitigations have moved the number of electorates progressively closer to the ERP value. The ERP is not the same concept as the census count, however, the two should be relatively near each other. While this approach does show the relative contribution of the census components to the electoral calculations, the final census count would not be based solely on responses received. The final census count has always included some form of non-response mitigation since 1996.

## Summary

Both census and electoral enrolment data are used to calculate the number of general and Māori electorates. The enrolment data, and in particular the results of the MEO, affect how the census counts are applied to calculating electorate populations and hence the number of electorates.

The number of general electoral districts has increased to 49, which is consistent with the 2018 ERP. The number of Māori electoral districts has remained at seven, partly due to the fall in the ratio of people of Māori descent who are eligible to vote and have chosen the Māori roll.

The figures show how the calculations have changed with each mitigation to the census data as well as their relative contribution. In the case of the Māori electoral districts, this is shown by both record source and Māori descent information source. It is apparent that the meshblock enumerations have made a substantial difference to the electoral district calculations, as has including Māori descent information from the 2013 Census and admin data. Each of these effects has improved representation.

## Composition of electoral districts

We now turn to the electoral district populations calculated using the 2018 Census data on the 2014 boundaries and how these have been affected by the non-response mitigations. This will help us understand how each non-response mitigation has contributed to each electoral district's population. It will also show if the mitigations contributed more or less to particular electoral districts.

The required electoral population of an electoral district is driven by the South Island quota, which is the general electoral population of the South Island divided by 16. The general electoral population is made up of the non-Māori population plus a proportion of the population of Māori descent, determined by the ratio of Māori on the Māori roll to total Māori registrations. This ratio has reduced from the level in the 2013 MEO. In effect, this reduction in ratio increases the general electoral population and is one of several factors leading to the increase in the South Island general electoral population.

The other contributing factors are the population growth generally in the South Island and the increase in the population of Māori descent. The increased South Island general electoral population has led to an increase in the South Island quota from 59,679 to 65,458.

At least three factors affect the number of general electoral districts in the North Island. These are:

- the increased proportion of people of Māori descent, which reduces the North Island general electoral population
- the reduced ratio of Māori on the Māori roll to total Māori registrations, which increases the general electoral population
- the South Island quota.

There are two ways we can look at the populations in the current 2014 electoral districts. These are:

- electoral populations: the general populations for the general electoral districts and the Māori electoral population for the Māori electoral districts
- census night usual resident population.

The first directly relates to the electoral purpose. The population counts are mediated by the enrolment information used to calculate the electoral populations. The disadvantage of this approach is that it obscures the exact counts of the populations within geographic areas of the electoral districts.

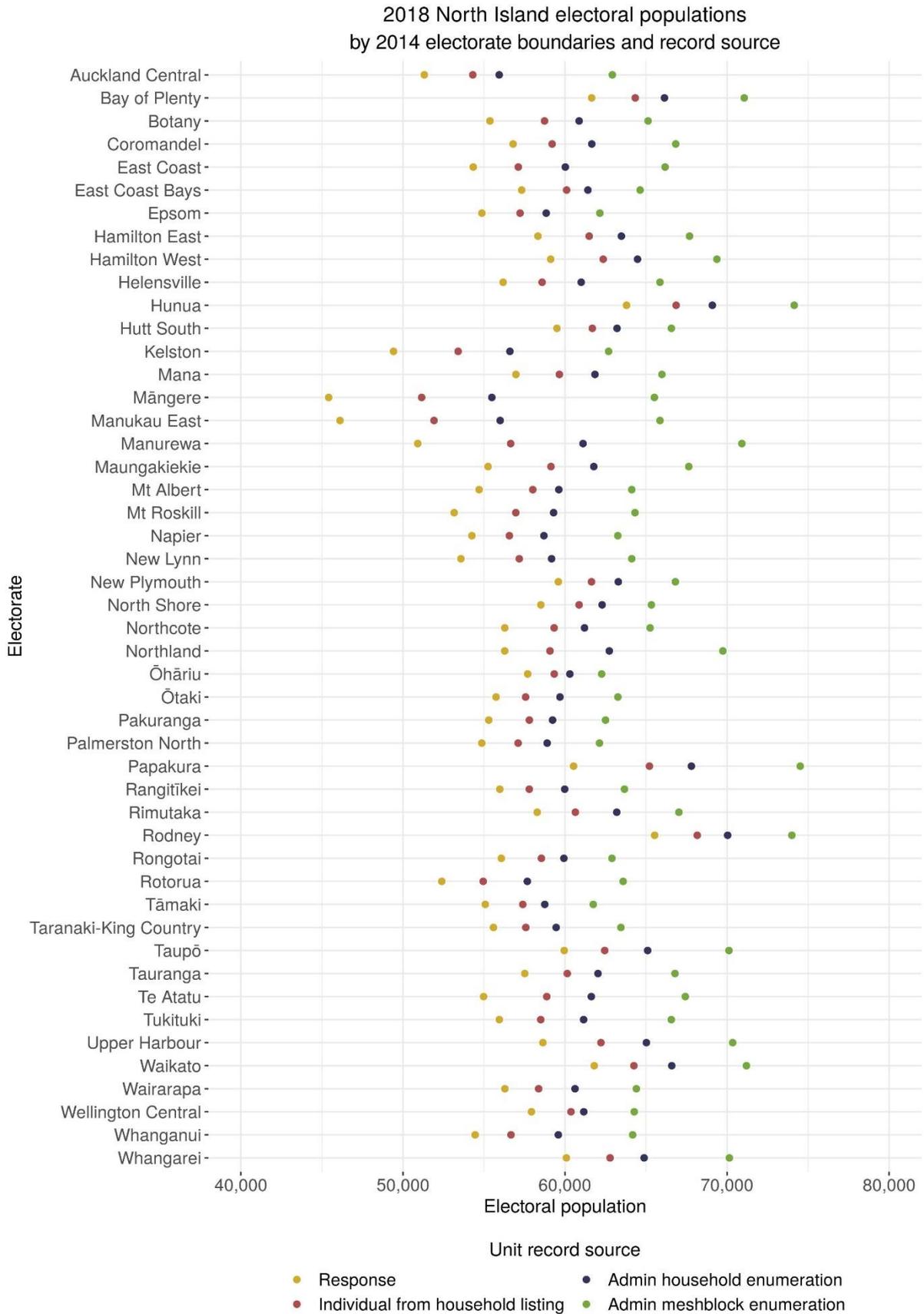
The second approach considers only the population counts within the districts as measured by the census. It is difficult to interpret these results for electoral purposes as they are heavily influenced by proportions of people within that district on each roll. Since our primary focus is the electoral purpose, we only consider electoral populations, the first perspective.

In the previous section, the electoral district calculations considered both record source and source of the Māori descent information. The same approach works for the electoral populations of the districts, looking in turn at the North and South Islands' general electoral populations and the Māori electoral population for both 2018 and 2013 data.

The plots in figures 5–12 show the current electorates down the side and the populations from left to right. The dots represent the electoral populations as calculated from 2018 Census responses, adding individuals on the household listing, adding admin household enumerations and finally meshblock enumerations or, for the 2013 Census, responses and imputations. All calculations are based on using the outcomes from the 2018 MEO.

## North Island general electoral populations: comparison between 2013 and 2018

Figure 5



The plot of the North Island general electorate populations in figure 5 shows a consistent pattern of additions from the various record sources. A number have larger additions than others, which tends to reflect a lower response from some ethnic groups and their subsequent inclusion via other sources. The electorate populations with the most additions of records, whether from information on the household listing, admin households or admin meshblock enumerations, are: Māngere, Manukau East, and Manurewa. This suggests that these electoral districts had greater non-response and that their representation has benefitted most from the mitigations. Ōhāriu has the least additions. Wellington Central has the least additions from admin households, reflecting the lower confidence with which we have been able to derive households from administrative data there.

The plots of the 2018 electoral populations tend to show some variability in the total populations and at a higher level than the previous electoral populations in some cases. This is simply because these plots reflect the populations of the current boundaries that were set based on the 2013 Census populations and enrolment data. Further, the electoral district quotas based on the 2018 Census data have all increased by between five and seven thousand on those from the 2013 calculations based on the 2013 Census data. The Representation Commission will assess how the current boundaries need to change in the light of the 2018 Census results, while meeting the requirements of the Electoral Act 1993.

For 2013, we had two sources of records and two sources of Māori descent information. The two sources of records are respondents and unit imputation (or substitute forms). The two sources of Māori descent are respondents and attribute imputation. The methods used in the 2013 Census for Māori descent imputation for electoral purposes are the same as those used in the 2001 and 2006 Censuses (Stats NZ, 2000). The methods for 2018 are detailed in Stats NZ 2019b.

The plot in figure 6 for 2013 shows a similar pattern to 2018 in terms of response and unit imputations but at a reduced magnitude due to the differences in response rates for the two censuses. As there was only one non-response mitigation, raising substitute records, there are only two dots to show for each electoral district. While Māngere, Manukau East, and Manurewa still had high numbers of unit imputations, the greatest number was in the Northland electoral district. The lowest number in 2013 was in the Ōhāriu electoral district.

Both the 2013 and 2018 data have elements of underlying variability that may not be apparent from the plots. For 2018, decisions have been made around the records to be included from admin data, in particular the meshblock enumerations. The cut-off point for inclusion of records and the accuracy of the meshblock data for that record will contribute to the uncertainty of the final estimate. There is a trade-off between these two quantities – we want to add records with good quality addresses but not be so strict that the count and representation suffer. Stats NZ 2019d includes more detail on the setting of this cut-off point.

The 2013 Census also has uncertainty in its population estimates. For example, the 2013 PES estimated the population to be undercounted in the 2013 Census by 2.4 percent (95 percent confidence interval of 1.9 percent to 2.9 percent). These people will be allocated across the country, with greater concentrations in regions of greater under-coverage and in some ethnicities over others. We have looked at coverage and its variability above. The imputation for Māori descent also introduces some uncertainty. While difficult to quantify, there was nevertheless uncertainty in the 2013 Census results. However, those results were not subject to the same analysis as the 2018 Census. The model-based nature of the 2018 mitigations make the 2018 Census more amenable to exploring the effects of uncertainty. The 2013 Census, because it uses a more traditional approach, is much more difficult to perturb to explore the effects of uncertainty. This is particularly the case for the unit imputation or the raising of substitute forms. While it is easy enough to perturb the number

of units imputed in a given case, it is more difficult to perturb the process of detection leading to the imputation.

We can compare the two censuses by which features contributed to the count.

**Table 4**

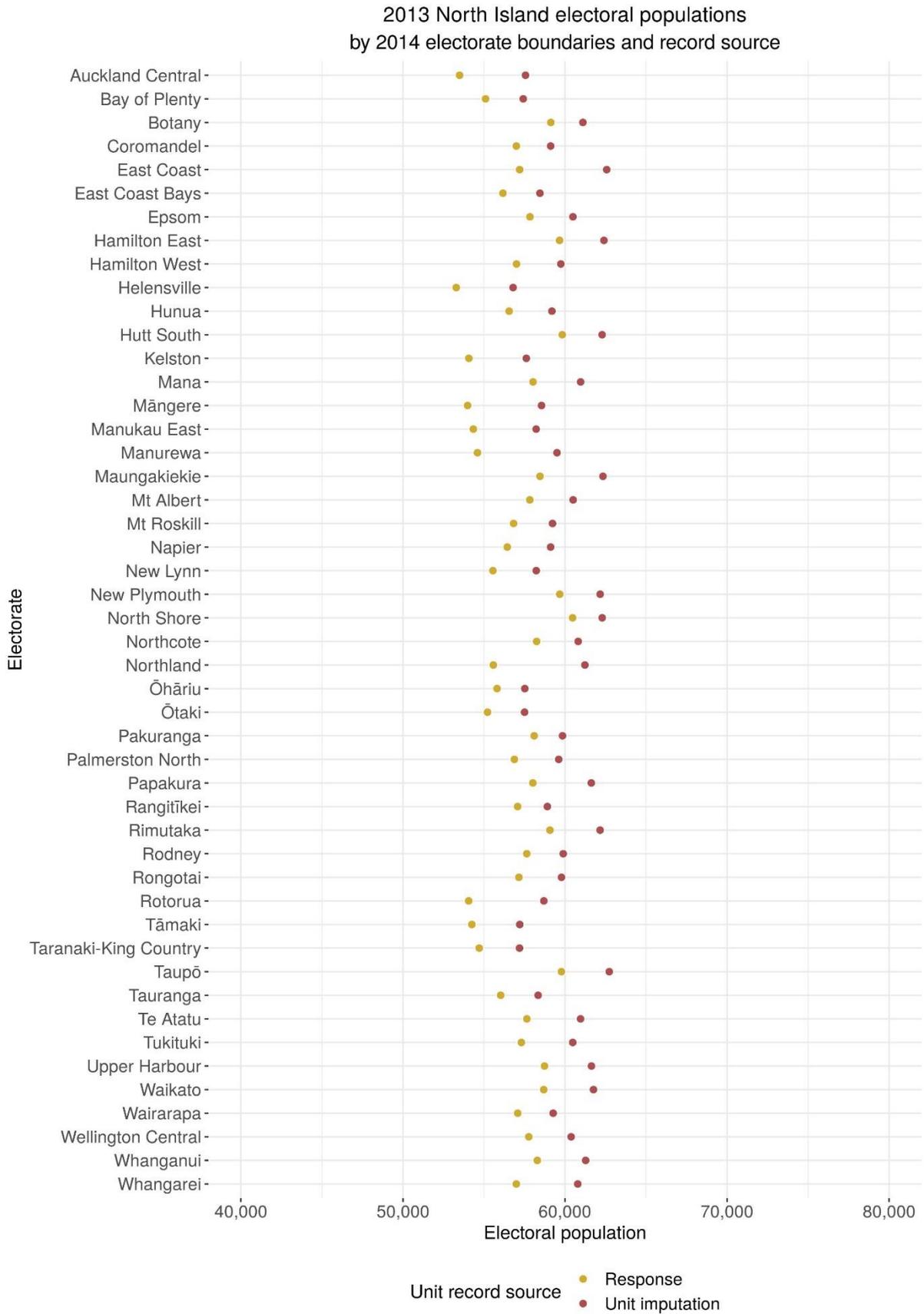
<b>Selected contributors to the census count in 2013 and 2018</b>			
<b>Contributor</b>	<b>2013 Census</b>	<b>2018 Census</b>	
Publicity	✓	✓	Responses
Place (e.g. urban, rural, apartment)	✓	✓	
Propensity of people in a location to respond	✓	✓	
Collectors' influence (delivery and collection)	✓		
Influence of mailout or list leave process		✓	
Reminder letters		✓	
Non-response follow-up		✓	
			Non-response mitigations
Substitute records or unit imputation	✓		
Listings on household form		✓	
Admin household enumeration		✓	
MB enumeration		✓	

There were elements of collector influence in the 2018 Census and non-response follow-up in the 2013 Census, but these were to a much-reduced extent.

For the 2013 Census, the decisions on raising substitute records were largely operational. In the 2018 Census, the difference between the final count and the responses is greater, and the processes used for the non-response mitigations are model based rather than operational based. This allows us to examine the effects on the count of varying thresholds in the selection models. For the 2013 Census, we can still vary the levels of the substitute records, but it is much more difficult to assess the effect for a change in process.

It is worth bearing this in mind when reading the plots showing the evolution of the electoral populations by record and attribute source. They show the contributions of the various sources but not the uncertainty associated with each quantity. The uncertainty may vary by location. The finer the geography, the more difficult a meaningful estimation of uncertainty, with meshblocks being the most difficult.

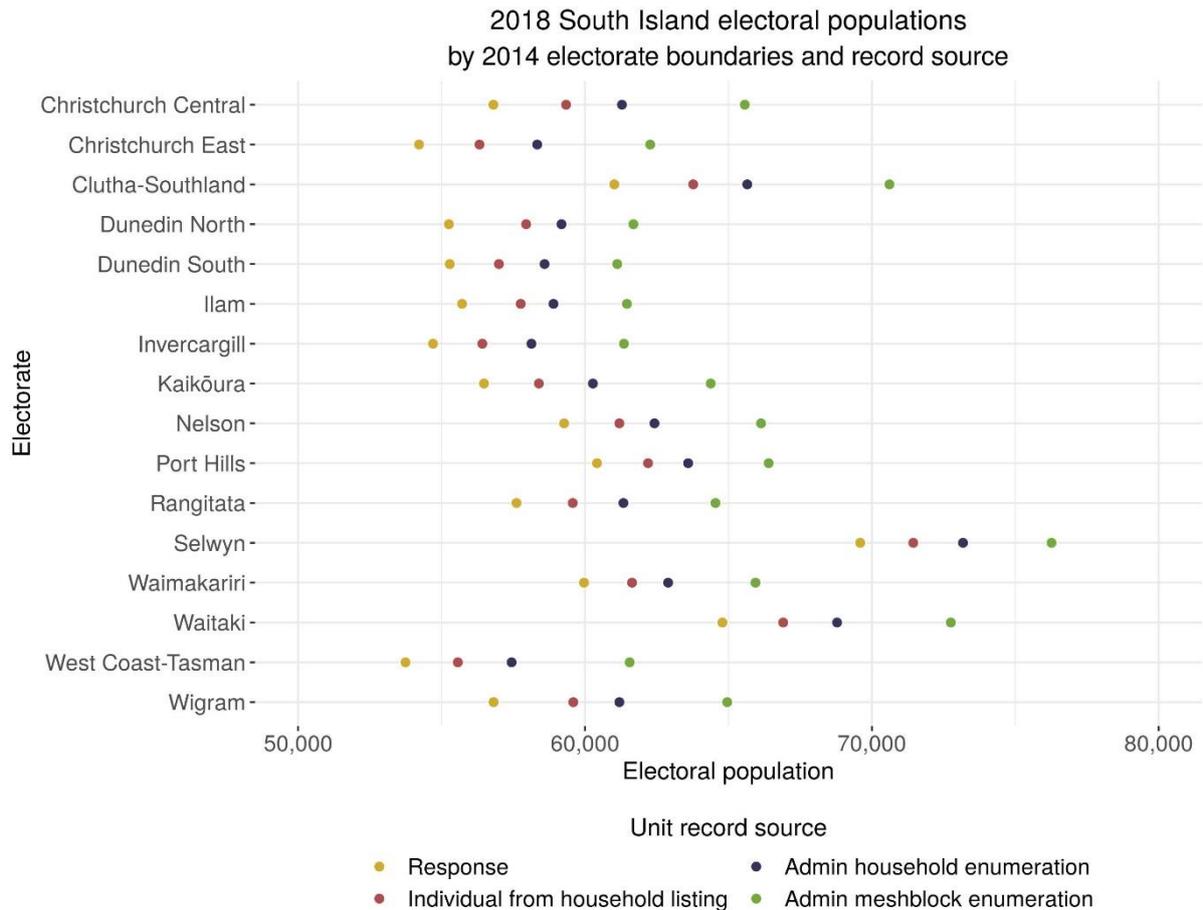
Figure 6



### South Island general electoral populations: comparison between 2013 and 2018

The South Island has the same number of general electorates for 2013 and 2018 because the number is fixed at 16 in the Electoral Act 1993. As noted above, the quota for the size of the electoral population has increased over the results from the 2013 Census. Figure 7 shows the effect in the electoral populations of the mitigations in the 2018 Census for the South Island general electoral districts. Note that the scales are different for the 2018 North and South Island electoral population figures.

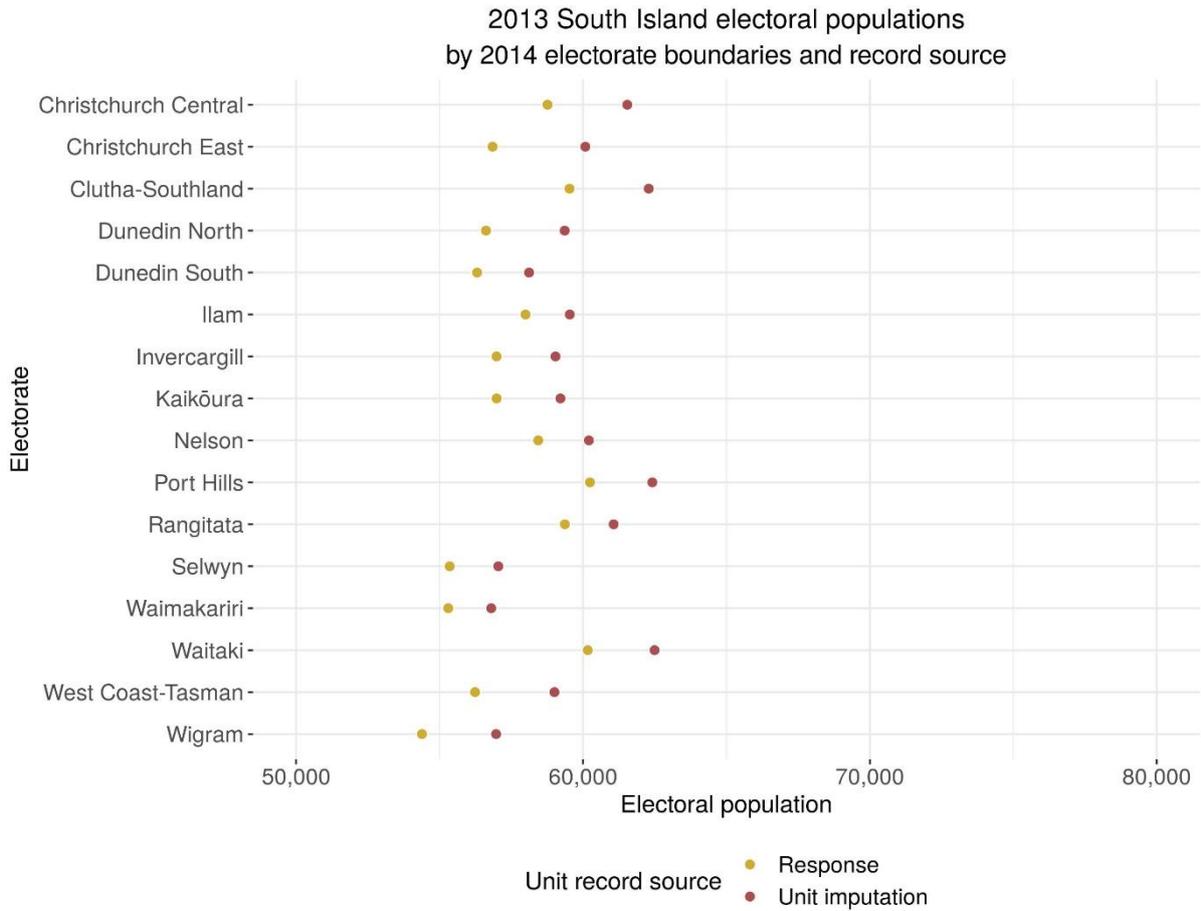
**Figure 7**



The patterns of additions are similar to those for the North Island. However, the magnitude of the records added by the non-response mitigations tend to be less than those for the North Island. This reflects the pattern of greater non-response in the North Island. For the 2018 Census, the general electoral districts with the most admin records added were Christchurch Central and Clutha-Southland, while the least were added in Ilam. The large growth in the Selwyn electorate is clearly shown when compared with 2013 in figure 8. In comparison with the North Island general electorates, the South Island general electorates have additions from the mitigations at the lower to middle range of the North Island distribution.

Across the whole country, in the 2013 Census, the greatest number of unit imputations were made in the Christchurch East general electoral district, and the least were made in Waimakariri. Comparing the proportions of records added, Christchurch East is the fourteenth highest. Clutha-Southland, Christchurch Central, Wigram, Dunedin North, and West Coast-Tasman are grouped together with five North Island general electorates just above the median, and the remaining 10 South Island general electorates are in a group of the lowest 18 electorates.

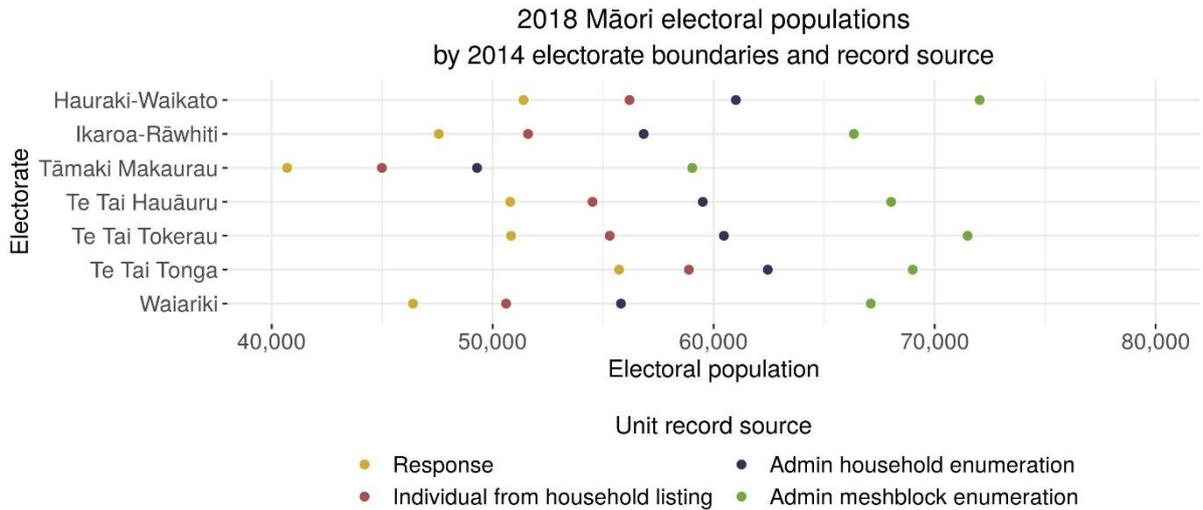
**Figure 8**



**Māori electoral populations: comparison between 2013 and 2018**

The plots in figures 9–12 show the electoral populations for the Māori electorates both by record source and source of Māori descent information. There is a similar pattern of records added for the Tāmaki Makaurau and Te Tai Tokerau electorates between 2018 and 2013. In 2018 Waiariki and Te Tai Tokerau had the largest contributions from the non-response mitigations. The fewest records were added in Te Tai Tonga. While the fewest unit imputations were made in Te Tai Tonga in the 2013 Census, the greatest number was in Tāmaki Makaurau.

**Figure 9**

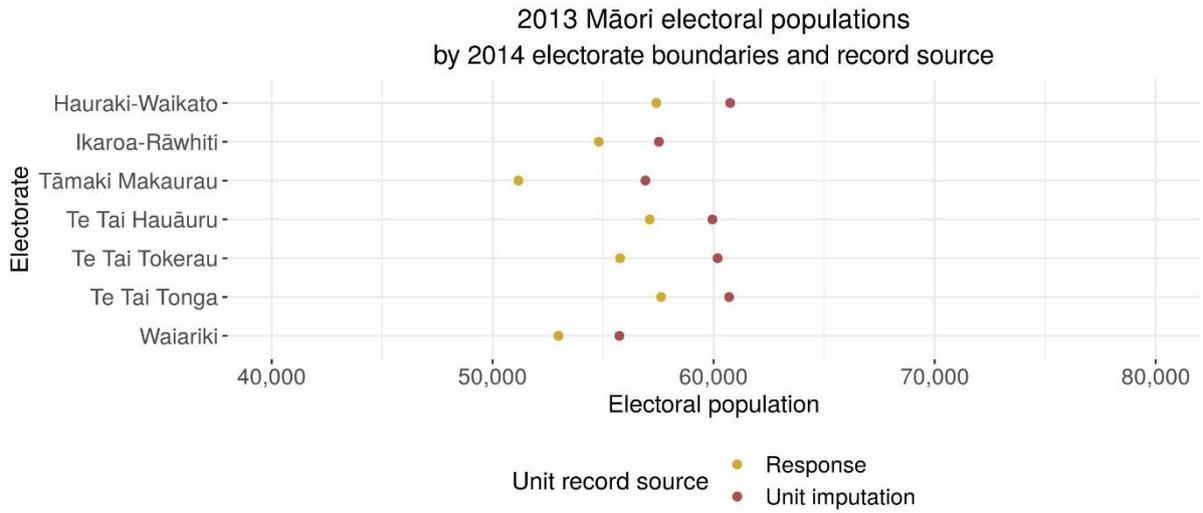


We can compare the proportion of records added across both the general and Māori electoral districts. For 2018, of the nine electoral districts where over 25 percent of the population has been added through the mitigations, six are Māori electoral districts and three are North Island general electoral districts. The six Māori electoral districts are in the North Island. The three North Island general electorates are Māngere, Manukau East, and Manurewa. All the South Island electoral districts have additions less than the median, with the dozen electoral districts with the lowest proportion of additions comprising eight South Island electoral districts.

The degree of mitigation required has been highest in the Māori electoral districts and in the general electoral districts of the North Island. These are the electorates that have had the greatest improvement in their representation for electoral purposes.

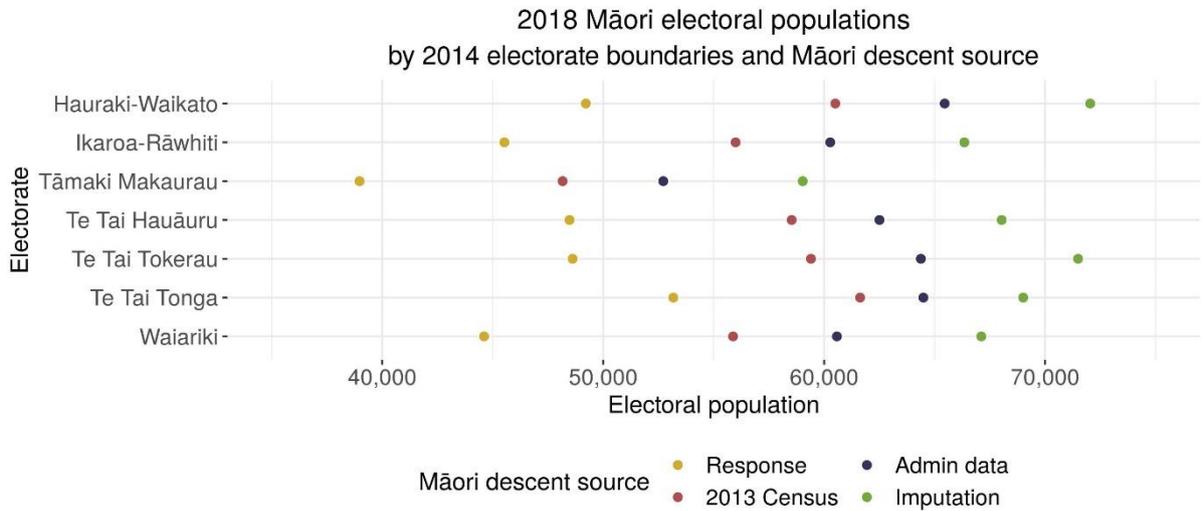
In 2013, of the eight electoral districts where the proportion of unit imputations exceeds 7 percent, two are Māori electoral districts and the remainder North Island general electoral districts. The Māori electoral districts are Tāmaki Makaurau and Te Tai Tokerau. The six North Island electoral districts are Northland, East Coast, Manurewa, Rotorua, Māngere, and Auckland Central. The South Island electoral districts are more spread out, with three above the median proportion of additions, although of the lowest 13 electoral districts, eight are South Island electoral districts.

**Figure 10**

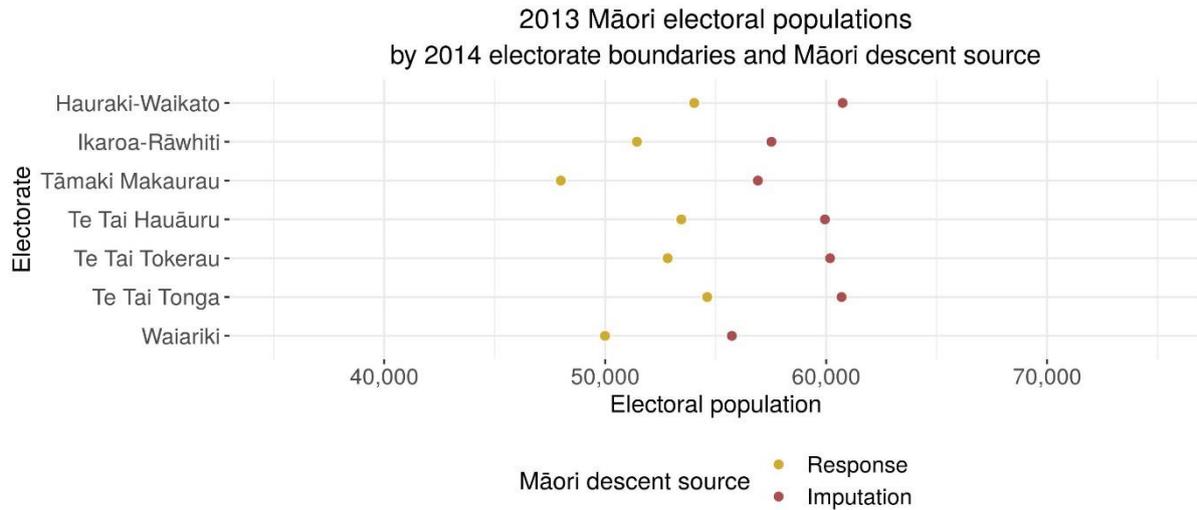


These final two plots in figures 11–12 show the Māori electoral population for 2018 and 2013 by source of the Māori descent information. Of interest in the 2018 plot in figure 11 is the large contribution from 2013 Census information and admin data. The largest receivers of Māori descent information were Hauraki-Waikato and Te Tai Tokerau by number but Tāmaki Makaurau by proportion. Te Tai Tonga consistently required the least information from other sources due to the higher response rate. In 2013, the greatest number and proportion of records given Māori descent by imputation was in Tāmaki Makaurau, and the least was in Te Tai Tonga.

**Figure 11**



**Figure 12**



**Summary**

The effect of the non-response mitigations on the electoral district populations have been graphed to show the effects of adding records from various sources, and sources of information for the attribute of Māori descent. There are a few electoral districts that have benefited most from the non-response mitigations, although all electoral districts have gained from the mitigations, which has improved representation.

We have provided the results for the 2013 Census alongside those for the 2018 Census for comparison, although care must be taken with any comparison. Both censuses are subject to uncertainties but in different ways. The earlier censuses were influenced by operational decisions, which are difficult to derive uncertainty measures for and assess the effects of varying those decisions on the final level of unit imputations. For the 2018 Census, a model-based approach was used for the non-response mitigations. For analysis of the model uncertainty and its possible effects on electoral results see Stats NZ 2019c.

The ability to obtain attribute data from admin sources, especially the 2013 Census, has been valuable for determining Māori descent. Representation of those of Māori descent has been improved over the response levels through the improved count of the Māori electoral population.

## Discussion

The response to the 2018 Census and the subsequent mitigations have precipitated a need to further explore our understanding of the census and how it contributes to the electoral calculations. In doing so we have found a number of limits in what can be evaluated quantitatively. However, in the interests of understanding the quality of the 2018 Census data for the electoral calculations the attempt had to be made.

Our analyses show how the non-response mitigations have contributed to the number of electoral districts and electoral populations. While this approach does illustrate the relative contribution of each source of census records, the final census count is not based solely on responses received but always includes some form of non-response mitigation, as has been the case in censuses since 1996.

Representation requires good counts of the general population and the population of Māori descent. These counts are required both at the island level and at the lower level, currently meshblocks. The island level counts are used to derive the number of electoral districts and the appropriate electoral populations required for each electorate. The meshblock counts are used to create the boundaries of the electoral districts to obtain the appropriate electoral populations.

The under coverage of the population has been increasing at the national level in the census since 1996, with the greatest undercount being 2.4 percent ( $\pm 0.5$  percent with 95 percent confidence) according to the 2013 PES. Even though the undercount tends to be higher in the north of the North Island, overall the counts at higher geographic levels, such as island level, tend to have an undercount similar to the national level. This robustness at national level lends strength to the calculations of the number of electoral districts. Indications so far for the 2018 Census suggest a lower undercount at the national level than previously, although this will be measured by the 2018 PES.

However, while the counts of the populations themselves are robust there is evidence of a reduction in the coverage of the population of Māori descent in the 2013 Census. We know from the 2013 PES that people with Māori ethnicity were undercounted at a rate greater than the general population and that this undercount has been increasing since 1996. It seems reasonable that the population of Māori descent was similarly undercounted, even if not at the same rate.

The 2018 Census mitigations have been able to access data where people have already indicated Māori ancestry, giving greater reliance on provided information rather than imputation. The records added through mitigations tend to have a higher representation of people of Māori descent, which would be expected if the census undercounts this population. Further, the increase in the population of Māori descent recorded in the 2018 Census suggests that the process for imputing Māori descent for electoral purposes in the 2013 Census did not impute enough people of Māori descent (Stats NZ, 2019b). While the data used in the 2018 Census mitigations did exist in 2013, it was not available or sufficiently understood to be used in that census. It is now better understood and able to be used for statistical purposes, such as mitigating the lower than expected response rate in the 2018 Census.

The implication of the apparent improvement in coverage of the population of Māori descent in the 2018 Census is a better calculation of the Māori electoral population. This should improve the accuracy of the allocation of the population between the Māori electoral population and the general electoral population. The first step for representation is to be counted; obtaining more accurate attribute information, whether of Māori descent or not, enables more accurate allocation between the Māori or general electoral populations. These populations then inform the number of electoral districts for the Māori or general populations. If the attribute of Māori descent is not as accurate as the count of people, then people will still be represented in the electoral populations but not

necessarily in the correct population. This may have implications for the number of electorates allocated to the two electoral populations as those populations are the numerators in the calculation.

Representation will also be affected by differences in coverage across regions. In this case, under-coverage can affect both the calculations of the number of electoral districts and their boundaries. In the latter case, an electorate could have more people in it than allowed for. Mis-enumeration may also affect representation if the mis-enumeration has occurred between an electoral district boundary, although this effect is much less than under-coverage in general. That is, the gain from representation through being counted probably outweighs small errors of precision in the location of that person.

We have illustrated the effect of the mitigations, both at the level of calculating the number of electoral districts and in calculating electoral populations within electoral districts. These results are also affected by enrolment registrations. While the census provides the counts of the usual residents and the populations of Māori descent, the electoral populations themselves are the result of enrolments and, in particular, the result of the MEO. It is possible to illustrate the effect of a change provided all the other variables are held constant, but the real world is seldom that simple.

We also compared a range of exploratory calculations about how using different population data would affect the electoral calculations and found the final 2018 Census data to result in similar calculations to those using the 2013-base ERP at June 2018. We have illustrated the results of these calculations, and have shown how the addition of population records with each of the mitigations has contributed to the number of electoral districts.

Our exploration of electoral populations by electoral district again shows the contribution of each non-response mitigation and highlights the variation across geographies and electorates in terms of the relative contribution of each mitigation. While 2013 data is included as a point of comparison, the different error and non-response structures of the 2013 and 2018 files make any useful comparison difficult. What is highlighted is that the judgements about inclusion in the 2013 file are not reproducible in a 2018 context as they result from a series of operational decisions. The process of unit imputation is statistically viable only if high enough response rates are achieved. By contrast, the addition of admin data is done transparently, allowing for detailed analysis of each step of the process.

## Conclusion

There is some evidence that the overall undercount in the 2018 Census is similar to, if not lower than, that found in the 2013 PES, however the 2018 PES will provide an official measure. There is also evidence that both the undercount and the variation in the undercount have been reduced at lower geographical areas, at least at the TALB level (figure 1). It is also apparent that the count of people of Māori descent has improved (Stats NZ, 2019b). This evidence thus suggests that the electoral calculations for both the number of electoral districts and the drawing of the boundaries will be at least as, if not more, robust than after the 2013 Census.

## References

- Stats NZ (2000). [Imputation of Māori Descent for Electoral Calculations](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2013). [The mathematics of electorate allocation in New Zealand based on the outcome of the 2013 Census and Māori Electoral Option 2013](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2014). [Coverage in the 2013 Census based on the New Zealand 2013 post-enumeration survey](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2017a). [Experimental population estimates from linked administrative data: 2017 release](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2017b). [Statistical standard for geographic areas 2018](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2019a). [Data sources, editing, and imputation in the 2018 Census](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2019b). [Deriving the 2018 Māori descent electoral population](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2019c). [Electoral boundaries sensitivity analysis of 2018 Census data](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2019d). [Overview of statistical methods for adding admin records to the 2018 Census dataset](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2019e). [Population counts sensitivity analysis of 2018 Census data](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2019f). [Predicting the quality of admin location information for use in the 2018 Census](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).
- Stats NZ (2019g). [The mathematics of electorate allocation in New Zealand based on the outcome of the 2018 Census and Māori Electoral Option 2018](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).