

# Estimated resident population 2018: Data sources and methods





Crown copyright ©

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

#### **Citation**

Stats NZ (2020). *Estimated resident population 2018: Data sources and methods*. Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).

ISBN 978-1-99-003209-7 (online)

#### **Published in September 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>1 Purpose and summary</b> .....	<b>5</b>
Purpose .....	5
Summary .....	5
<b>2 Estimated resident population at 30 June 2018</b> .....	<b>7</b>
2.1 Population measures .....	7
2.2 About the ERP .....	8
2.3 Māori descent ERP .....	9
2.4 How the ERP is derived .....	9
2.5 Data sources and flows .....	11
2.6 The ERP and estimated uncertainty intervals.....	12
2.7 Historic comparison of the ERP and contributing components .....	15
<b>3 The adjusted resident population at census</b> .....	<b>16</b>
3.1 2018 Census counts .....	16
3.2 Net census undercount.....	17
3.3 Residents temporarily overseas.....	20
3.4 Calculating the resident population at census .....	21
<b>4 Population change from census date to 30 June 2018</b> .....	<b>23</b>
4.1 Births and deaths .....	25
4.2 Net migration.....	25
4.3 Imputation in component datasets .....	27
4.4 Calculating the ERP at 30 June 2018.....	28
4.5 Measuring uncertainty in the ERP .....	28
<b>5 Adjustments to the ERP</b> .....	<b>29</b>
5.1 Adjustment for implausible population group estimates.....	29
5.2 Demographic reconciliation adjustment .....	30
<b>6 References</b> .....	<b>32</b>
<b>Appendix 1</b> .....	<b>33</b>
<b>Appendix 2</b> .....	<b>37</b>
Geographic areas used in population estimation.....	37

## List of tables and figures

### List of tables

1 National summary of ERP at 30 June 2018 and contributing component estimates .....	10
2 ERP at 30 June 2018 and 95 percent uncertainty intervals for selected population groups.....	14
3 ERP and contributing components, at 30 June 1996–2018 .....	15
4 Net census undercount estimates by population group, 2018 .....	18
5 Estimates of components contributing to the ERP at 30 June 2018, for selected population groups .....	23
6 Level of item non-response observed in population variables by component dataset .....	27
7 Output variables for the ERP at 30 June 2018 .....	33
8 Overview of sources of uncertainty and methods to measure uncertainty in the ERP .....	33
9 Overview of sources, variables fully observed and variables with non-response in the ERP components .....	34

### List of figures

1 National population measures and their inclusions .....	8
2 Data sources and flows input to producing the ERP at 30 June 2018 .....	12
3 ERP at 30 June 2018 by age and sex for ethnic groups.....	13
4 Demographic reconciliation adjustment as percentage of the ERP at 30 June 2018, by age 0–14 years and sex.....	31

# 1 Purpose and summary

## Purpose

*Estimated resident population 2018: Data sources and methods* describes the data sources and methods used to produce the 2018-base estimated resident population (ERP) at 30 June 2018. It includes descriptions of the data sources and methods used to determine the contributing components at national and subnational levels, by age, sex, level 1 ethnic groups, and Māori descent. The report is aimed at technical users who want to gain an understanding of the magnitude and quality of the contributing components used to produce the estimated resident population at 30 June 2018.

## Summary

The 2018-base estimated resident population (ERP) of New Zealand is an estimate of all the people who usually live in New Zealand at 30 June 2018. The ERP is derived from the 2018 census usually resident population count adjusted for net census undercount, residents temporarily overseas, and estimates of population change due to births, deaths, and migration, between census night (6 March 2018) and the mid-year reference date (30 June 2018).

In this report, we describe the data sources and the methods used to estimate each of the components, the methodological steps for compiling the ERP, and the adjustments included in the ERP. While the methods to compile the ERP in 2018 have remained largely consistent with those followed in 2013, some changes have been introduced due to the extended use of administrative data.

All population measures, including census counts and population estimates, have some uncertainty around them. This report also describes the methods and the contributing data sources we used to measure the uncertainty around the ERP at 30 June 2018.

The process to compile the ERP included using administrative data sources and census coverage results. Administrative data has always been used in population estimation, and this has been extended to strengthen the accuracy of the 2018-base ERP. In summary, the ERP has been compiled by:

- using the 2018 Census dataset that had been created by combining administrative data and census forms ([2018 Census: How we combined the administrative data and census forms data to create the census dataset](#))
- using the results of net census under-coverage as measured by the [Post-enumeration survey 2018: Methods and results](#)
- accessing linked administrative data and 2013 Census data, compiled by Stats NZ's [Integrated Data Infrastructure](#) (IDI), to obtain information on resident location, ethnic group, and Māori descent in residents temporarily overseas on census night (RTO) and migration records
- accessing linked address notifications data in the IDI for estimating internal migration to and from geographic areas between census and mid-year
- deriving natural increases from birth and death registrations, which included information about resident location, ethnic group, and Māori descent
- accessing linked border movements for identifying migrant arrivals and departures between census and mid-year by the outcomes-based migration measure ([Outcomes versus intentions: Measuring migration based on travel histories](#))

- including an adjustment to the ERP at 30 June 2018 of children (aged 0–14 years) as a reconciliation of the differences in administrative estimates and ERP for these ages. The demographic reconciliation adjustment was based on administrative records of births, deaths, and international migration.

We acknowledge that Stats NZ has only partly engaged with customers on the extended use of administrative data in census and population estimates, and further work is needed on the social acceptability of these methodological improvements.

## 2 Estimated resident population at 30 June 2018

This chapter provides information on the different population measures. It provides an overview of the data and methods used to produce the ERP at 30 June 2018 and includes a national summary of the ERP broken down by components. We discuss the uncertainty estimates of the ERP and conclude with a tabulated summary for selected population groups with estimated uncertainty intervals. The chapter includes the following sections:

[2.1 Population measures](#)

[2.2 About the ERP](#)

[2.3 Māori descent ERP](#)

[2.4 How the ERP is derived](#)

[2.5 Data sources and flows](#)

[2.6 The ERP and estimated uncertainty intervals](#)

[2.7 Historic comparison of the ERP and contributing components.](#)

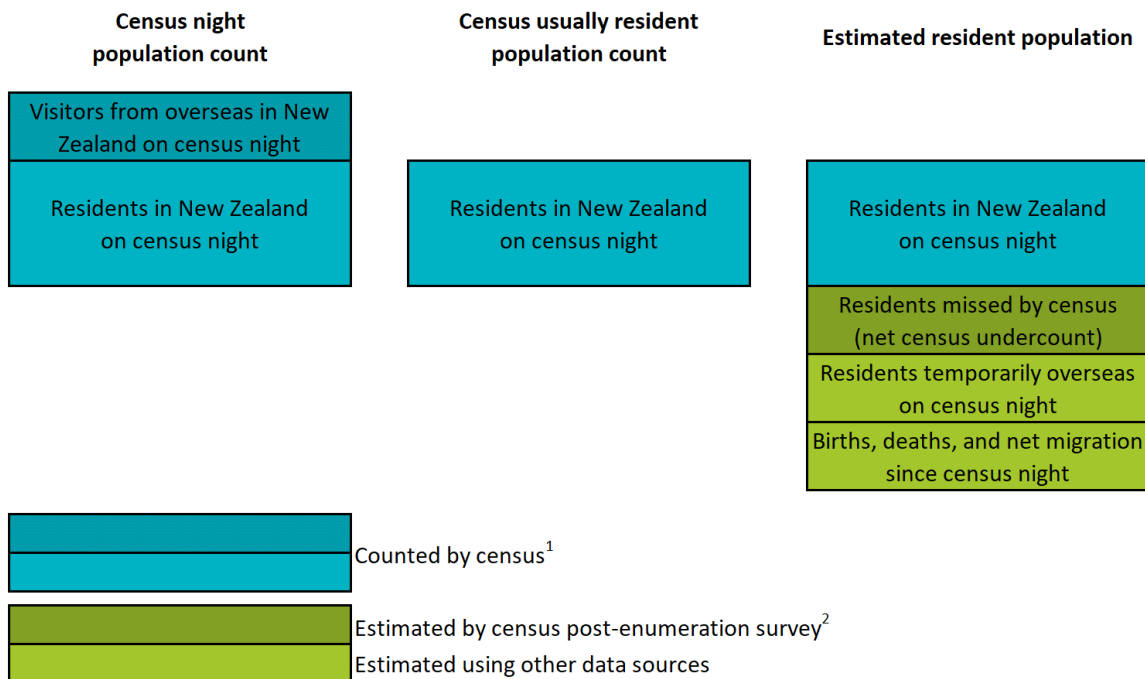
### 2.1 Population measures

The term 'estimated resident population' (ERP) clarifies that it is a demographic estimate and differentiates it from two other population measures: 'census usually resident population count' and 'census night population count'.

Figure 1 illustrates the sources and population groups used to define the three population measures at the national level.

Figure 1

### National population measures and their inclusions



1. The 2018 Census included admin enumerations which are described in the '2018 Census counts' section of this report.  
 2. Coverage of the 2018 Census was measured by the 2018 Post-enumeration Survey.

The ERP is produced following the availability of the census usually resident population count:

- to represent the best possible measure of the resident population for a wide variety of customers and uses (including use in demographic projections and indices)
- to meet the need for updated population estimates beyond census date as an indication of demographic change
- to satisfy United Nations 1952 recommendations to adjust for known deficiencies in census (United Nations, 1952).

## 2.2 About the ERP

The estimated resident population is as at midnight of the reference date (30 June 2018). A 'resident' is a person who usually lives in an area. This is a statistical, not legal, definition and is generally based on a person's self-identified usual address. The term 'resident' may be used differently in other contexts, such as in economic statistics or to describe immigration status.

In international travel and migration statistics, a resident is a person who has been living in New Zealand for 12 months or more as defined by travel histories and the '12/16-month rule'.

See [Defining migrants using travel histories and the '12/16-month rule'](#) and further information at [Outcomes versus intentions: Measuring migration based on travel histories](#).

We derived the 2018 ERP by age, sex, usual resident location at meshblock level, level 1 ethnic group (European or Other, Māori, Asian, Pacific, Middle Eastern/Latin American/African (MELAA)), and Māori descent. We derived estimates at meshblock level to provide flexibility in aggregating the ERP to all statistical output geographies needed by our diverse range of customers (see Appendix 2).



Population estimates are published for a variety of geographic and demographic variables, while some breakdowns are available on request. See table 7 in Appendix 1 for further details.

### Links to data

See [Estimated resident population \(2018-base\): at 30 June 2018](#) for data and commentary about the 2018 ERP. (Note: the 2018 ERP and 2019 national population estimates were released together.)

Tables available in [NZ.Stat](#) include:

- [Estimated resident population \(ERP\), adjustments to derive ERP at 30 June 2018 \(from census usually resident population count\)](#)
- [Estimated resident population \(ERP\), national population by ethnic group, age, and sex, 30 June 1996, 2001, 2006, 2013, and 2018](#)
- [Estimated resident population \(ERP\), subnational population by ethnic group, age, and sex, at 30 June 1996, 2001, 2006, 2013, and 2018](#)
- [Estimated Māori descent resident population \(ERP\), at 30 June 2001, 2006, 2013, and 2018](#)
- [Estimated components of subnational intercensal population change \(RC, TA\), 1981-2018.](#)

## 2.3 Māori descent ERP

The ERP is produced for the six level 1 ethnic groups and for the Māori descent population. Published tabulations of the ERP at 30 June 2018 include the Māori descent ERP by age, sex, and geographic area. Māori descent ERP at 30 June 2001, 2006, and 2013 are available but have not been included in published tabulations.

A requirement of the [Local Electoral Amendment Act 2002](#) (Schedule 1A) has been to provide updates of estimated Māori and general electoral population for wards, districts, and cities. This includes using the most recent information on the census Māori descent population (electoral) adjusted for net under-count, as measured by the post-enumeration survey (PES), and RTOs on census night, and estimated population change for the Māori descent electoral population between census and a given reference date. For this purpose and more general Māori population analyses, we have included the Māori descent indicator as a variable in all contributing data sources used to derive the ERP at 30 June 2018.

The Māori descent ERP is not used for general election purposes (for example, boundaries). Under the [Electoral Act 1993](#), census counts are used to calculate electoral populations and electoral boundaries. A Māori descent indicator is derived for all residents enumerated in the census for the purpose of determining New Zealand's electoral districts. Census collected individuals' Māori descent responses in terms of categories 'yes', 'no', 'don't know', or no response. Requirements of the electoral population calculations assume a Māori descent indicator associated with all census Māori responses that were not 'yes' or 'no'. The 2018 data sources and methodology for deriving the Māori descent indicator for electoral purposes is described in [Deriving the 2018 Māori descent electoral population](#).

## 2.4 How the ERP is derived

The ERP at 30 June 2018 is compiled from the 2018 census usually resident population count, with adjustments at census date, combined with components of population change between census date and 30 June 2018. The process includes the adjustments set out below.

The ERP at census date consists of:

- the 2018 census usually resident population count
- the adjustment for net census undercount, as measured by the 2018 PES
- the adjustment for the number of residents temporarily overseas on census night, based on short-term resident traveller data.

Population change between census night and 30 June 2018 consists of:

- natural increase (births less deaths), using birth and death registrations
- net international migration (migrant arrivals less migrant departures), based on the 12/16-month migration rule (outcomes-based method)
- net internal migration (in-migration less out-migration between areas of New Zealand), using internal movements identified by administrative address notifications.

The ERP at 30 June 2018:

- combines the resident population estimates at census date and the estimated population change between census date and 30 June 2018
- includes a demographic reconciliation adjustment of children aged 0–14 years, based on demographic estimates from administrative records of births, deaths, and international migration.

The data sources that have been used to produce the components are described in section 2.5, and table 9 in Appendix 1 provides an overview of their contribution to the data flows for compiling the ERP. The methods for compiling the ERP are described in sections 3.4 and 4.4, and the adjustments included when compiling the ERP are described in chapter 5.

### Summary of components contributing to the ERP at 30 June 2018

Table 1 summarises the size of the components producing the national total ERP.

**Table 1**

<b>National summary of ERP at 30 June 2018 and contributing component estimates</b>		
<b>Component</b>	<b>Estimates</b>	<b>Percentage of ERP</b>
<b>At census date</b>		
Census usually resident population count	4,699,755	95.9
Net census undercount	124,900	2.5
Residents temporarily overseas	77,200	1.6
<b>Population change between census date and 30 June 2018</b>		
Natural increase (births less deaths)	8,000	0.2
Net migration (migrant arrivals less migrant departures)	7,300	0.1
Total population change	15,400	0.3
Demographic reconciliation	-16,600	-0.3
<b>Estimated resident population at 30 June 2018</b>	<b>4,900,600</b>	<b>100.0</b>
<b>Source:</b> Stats NZ		

At the national level, the main contributing components added to the census usually resident population count for estimation of the resident population at 30 June 2018 were the net census undercount (2.5 percent of the ERP) and residents temporarily overseas at census date (1.6 percent

of the ERP). The contribution of population change between census date and mid-year was offset by the demographic reconciliation adjustment to the ERP at 30 June 2018.

Tabulated summaries of the components contributing to the ERP at 30 June 2018 for population groups are presented in table 5 (chapter 4) of this report.

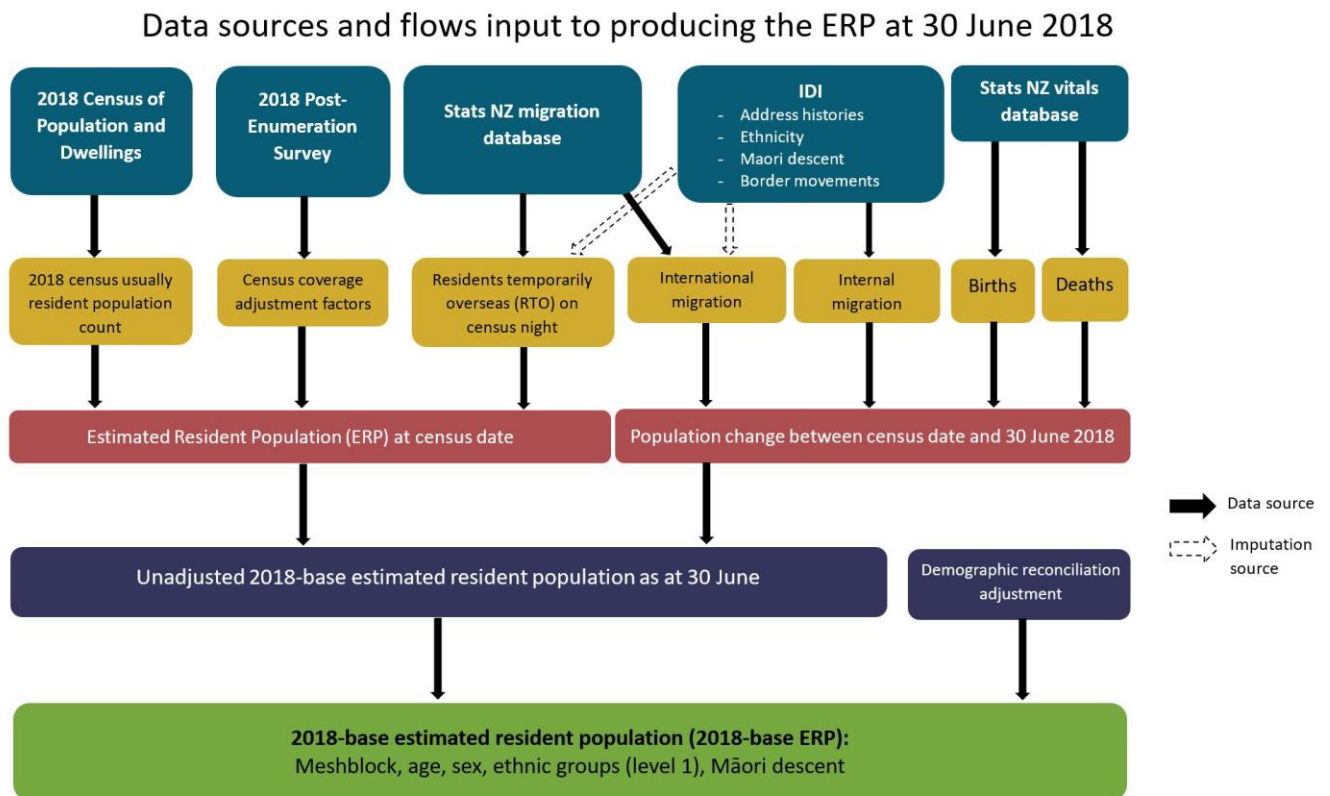
## 2.5 Data sources and flows

The individual components contributing to the ERP at June 2018 were produced using a wide range of sources. The 2018 Census of Population and Dwellings provided the counts of the usually resident population on census night, and results from the 2018 PES provided estimates of the under- and over-coverage of the census usually resident population count. The coverage-adjusted estimates of the usually resident population formed the basis for adding components entirely produced from birth and death registration data and other administrative sources.

Natural increase between the census date and mid-year used birth and death registration data collected by Department of Internal Affairs (DIA) and compiled by Stats NZ for statistical purposes. The ERP process accessed integrated border movements, collected by the New Zealand Customs Service and compiled by Stats NZ, to produce counts of RTO and migrant arrivals and departures between census date and 30 June 2018.

Figure 2 illustrates the data sources (the first row of text boxes) that were used to create the component datasets (the second row of text boxes) for compiling the ERP. We produced the resident population estimates at census and compiled the components of population change between census and mid-year (the third row of text boxes). These were combined before the demographic reconciliation adjustment was applied to produce the ERP at 30 June 2018.

Figure 2



Source: Stats NZ

Stats NZ’s IDI database was accessed to merge resident location, ethnic group, and Māori descent information with border movements data, and to define records of internal migration, their ethnic group, and Māori descent indicators. The IDI regularly updates tables of integrated address notifications and ethnic group information from a range of administrative sources. The IDI sources accessed included 2013 Census for information on resident location, ethnicity, and Māori descent. Chapters 3 and 4 in this report describe in more detail how the data was used to compile the ERP. Table 9 of Appendix 1 provides details on the variable information accessed by the individual sources.

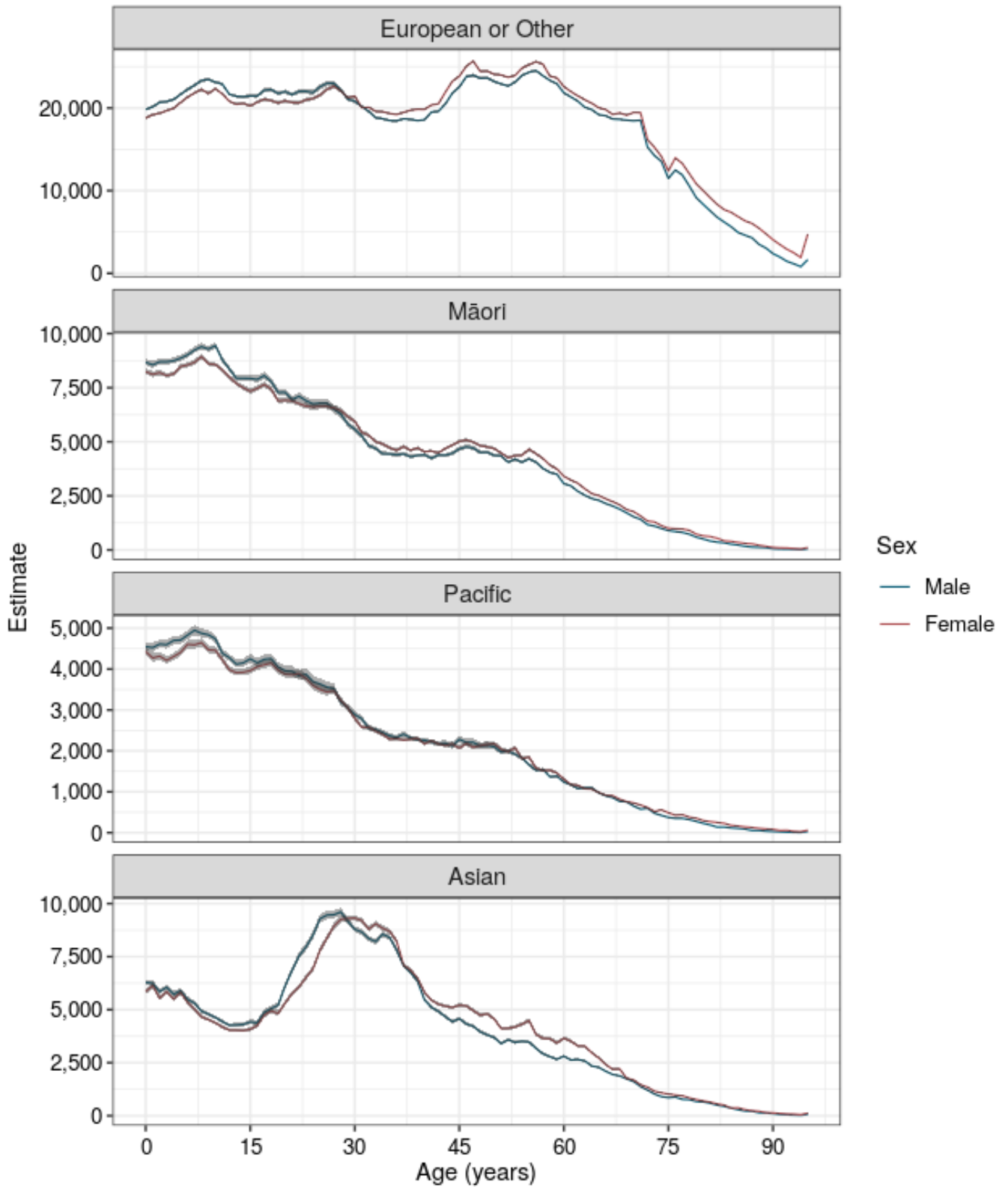
## 2.6 The ERP and estimated uncertainty intervals

The ERP at 30 June 2018 includes estimates of uncertainty intervals for cross-tabulations of the demographic variables. The ERP compilation incorporated uncertainty resulting from the contributing data and estimation processes where this was made possible. The process did not incorporate uncertainty around data quality in individual sources or from adjustments made to the ERP at 30 June 2018 (chapter 5). The main source of uncertainty in the ERP was the estimated net census undercount as measured by the PES.

Measures of uncertainty in the ERP, represented as 95 percent intervals given by the 0.025 and 0.975 percentile limits, show wider intervals for young adult males compared with young adult females, and relatively wider intervals for the Pacific and Māori ethnic groups compared with the European or Other and Asian ethnic groups. Figure 3 illustrates the uncertainty intervals (shaded area) for the ethnic group ERP by age and sex. The mean estimates at each age (0, 1, 2, ..., 95+) are represented as line graphs.

Figure 3

ERP at 30 June 2018 by age and sex for ethnic groups  
 Mean estimate and percentiles [0.025, 0.975]



Source: Stats NZ

Table 2 summarises the ERP for selected population groups and their percentages of the total ERP at 30 June 2018. Percentile estimates, that is, the 0.025 percentile (lower) and the 0.975 (upper), provide 95 percent uncertainty interval measures for the ERP (mean estimate) at 30 June 2018.

**Table 2**

<b>ERP at 30 June 2018 and 95 percent uncertainty intervals for selected population groups</b>				
<b>Population group</b>	<b>ERP at 30 June 2018 (mean estimate)</b>	<b>Lower percentile estimate (0.025 percentile)</b>	<b>Upper percentile estimate (0.975 percentile)</b>	<b>Percentage (mean estimate) of total ERP</b>
Total	4,900,600	4,886,700	4,915,400	100.0
<b>Sex</b>				
Male	2,430,200	2,421,500	2,440,400	49.6
Female	2,470,500	2,462,000	2,479,100	50.4
<b>Broad age groups</b>				
0–14 years	946,400	941,600	951,700	19.3
15–39 years	1,676,900	1,668,700	1,685,800	34.2
40–64 years	1,542,300	1,536,500	1,548,500	31.5
65 years and over	734,900	732,000	738,100	15.0
<b>Broad ethnic groups<sup>(1)</sup> and Māori descent</b>				
European or Other <sup>(2)</sup>	3,441,700	3,432,500	3,451,700	70.2
Māori	816,500	810,800	822,900	16.7
Pacific	407,700	403,200	412,900	8.3
Asian	770,600	763,800	778,200	15.7
MELAA <sup>(3)</sup>	77,000	76,500	77,400	1.6
Māori descent	941,200	935,100	947,500	19.2
<b>Subnational</b>				
Auckland region	1,654,800	1,645,500	1,664,900	33.8
Wellington region	525,900	522,500	529,300	10.7
Rest of North Island	1,572,900	1,566,200	1,579,800	32.1
Canterbury region	622,800	619,400	627,000	12.7
Rest of South Island	523,500	520,600	526,900	10.7
<p>1. People who reported more than one ethnic group are counted once in each group reported. This means that the total number of responses for all ethnic groups can be greater than the total number of people who stated their ethnicities.</p> <p>2. Includes New Zealander responses.</p> <p>3. Middle Eastern/Latin American/African.</p> <p><b>Source:</b> Stats NZ</p>				

Estimated relative uncertainty intervals were typically narrower for the larger population groups but also reflected different patterns of census coverage and their estimates of uncertainty in the ERP for population groups. Based on results of the 95 percent uncertainty intervals:

- for the total population, the estimated uncertainty of the ERP was +/- 0.3 percent of the mean estimate (This was higher for males (+/- 0.4 percent) compared with females (+/- 0.3 percent).)

- uncertainty estimates were higher for the younger population groups, that is, for the age group 15–39 years it was estimated at +/- 0.5 percent, and for the age group 40–64 years, it was +/- 0.4 percent of the ERP
- the ERP for the Pacific ethnic group had higher uncertainty than the other ethnic groups with the interval estimated as 1.1 percent lower and 1.3 percent higher than the ERP (This compared with intervals for the Asian ethnic group as 0.9 percent and 1.0 percent differences to the ERP and for Māori as 0.7 and 0.8 percent differences to the ERP.)
- the relative uncertainty interval of the ERP for Auckland was -0.6 percent; +0.6 percent, and for the Wellington region, it was -0.7 percent; +0.6 percent.

Not all sources of uncertainty in the ERP are measurable. Results from the census have not included uncertainty measures because there have been no attempts to measure uncertainty arising from census enumeration and processing errors (for example, misrecognition during scanning, uncertainty from age-sex imputation). Census coverage estimates are sourced from a sample survey (PES) and are therefore subject to sample errors. The modelled approach to coverage estimation included coherent measures of uncertainty in estimates of census under- and over-coverage. Section [4.5 Measuring uncertainty in the ERP](#) describes the sources and methods used to measure uncertainty in the ERP.

## 2.7 Historic comparison of the ERP and contributing components

Historic demographic trends have influenced the changes observed in the sizes of the components contributing to the ERP. Other changes have been the result of changes to methods and measurement rules for estimating the individual components. Table 3 presents a historic series of the estimated components produced for the ERP at 30 June 1996–2018.

**Table 3**

<b>ERP and contributing components</b>					
At 30 June 1996–2018					
Component	1996	2001	2006	2013	2018
<b>At census date<sup>(1)</sup></b>					
Census usually resident population count	3,618,300	3,737,280	4,027,947	4,242,048	4,699,755
Estimated net census undercount	58,900	80,900	80,100	104,200	124,900
Residents temporarily overseas <sup>(2)</sup>	39,800	54,500	64,500	81,700	77,200
<b>Population change between census date and 30 June</b>					
Natural increase (births less deaths)	8,600	8,600	10,000	9,100	8,000
Net migration (migrant arrivals less migrant departures)	2,500	-4,500	-1,300	200	7,300
Total population change	11,100	4,100	8,700	9,300	15,400

Demographic reconciliation <sup>(3)</sup>	3,900	3,700	3,400	4,800	-16,600
<b>Estimated resident population at 30 June</b>	3,732,000	3,880,500	4,184,600	4,442,100	4,900,600
<p>1. Census date was 5 March in 1996 and 2013; 6 March in 2001 and 2018; and 7 March in 2006.</p> <p>2. The apparent drop in RTOs between 2013 and 2018 reflects a change in estimation methodology, not a real drop in the number of NZ-residents travelling overseas.</p> <p>3. The demographic reconciliation was applied to ages 0–4 years in 2006; 0–9 years in 1996, 2001, and 2013; and 0–14 years in 2018.</p> <p><b>Note:</b> Due to independent rounding, figures may not sum to stated totals.</p> <p><b>Source:</b> Stats NZ</p>					

The net migration estimate contributing to the 2018 ERP was higher than the contributions made in previous census years. This reflects the higher net migration experienced in March–June 2018 than in other years.

There was an apparent drop in the number of residents temporarily overseas at the date of the 2018 Census. This was a reversal of the historic trend that had seen steady increases in numbers of short-term New Zealand-resident departures. The apparent reduction reflects a change in the method of estimation, from using the intended duration of absence of New Zealand-resident travellers reported on departure cards to using linked New Zealand-resident traveller departures and arrivals to ascertain how many were away on 6 March 2018.

Historically, the demographic reconciliation adjustment has contributed around 0.1 percent of the total ERP. The adjustments typically covered ages 0–9 years (ages 0–4 years in 2006). In 2018, the absolute value of the adjustment was higher, at 0.3 percent of the ERP, partly because the wider span of ages 0–14 years was included.

## 3 The adjusted resident population at census

This chapter describes the data sources used to produce the ERP at census date. We provide details of the 2018 Census counts and how they were used; and outline data and methods used in the adjustments to estimate the net census undercount and the residents temporarily overseas on census night (RTO). The sections included in this chapter are:

[3.1 2018 Census counts](#)

[3.2 Net census undercount](#)

[3.3 Residents temporarily overseas](#)

[3.4 Calculating the resident population at census](#)

### 3.1 2018 Census counts

For the first time in a New Zealand census, the 2018 Census used data from alternative sources to fill gaps when the characteristics of people or dwellings were not provided on census forms. These alternative sources were the 2013 Census and a range of administrative (admin) data sources, such as birth registrations and tax information. The use of statistical imputation for the remainder of the missing data was extended to a much greater range of variables than in previous censuses. See [Data](#)



[sources, editing, and imputation in the 2018 Census](#) and [2018 Census information by variable and quality](#) (refer to the Data Sources section of each variable).

Where people were listed as a member of a household but completed no individual form, and where people were counted through admin enumeration, alternative sources and imputation were the source of nearly all individual census characteristics.

The use of admin enumerations improved the census count over previous censuses for some subgroups of the population, and together with the alternative data sources and imputation, for many variables, helped maintain (or in some cases improve) the quality of information seen in previous censuses. However, for other variables, higher rates of, for example, missing data meant that quality was lower than in previous censuses, and those subgroups with lower response rates to the field collection, such as Māori, Pacific peoples, and young adults were more adversely affected.

The starting population for deriving the 2018 ERP is the 2018 ‘census usually resident population count’. The census usually resident population count of New Zealand is a count of all people who usually live, and are present, in New Zealand on census night. This count excludes visitors from overseas and residents who are temporarily overseas on census night.

Residents were counted at the meshblock of their usual residence. Those who were temporarily elsewhere in New Zealand on census night were included in the meshblock they usually live in. A meshblock is the smallest geographic area we use for statistical purposes – see ‘Geographic areas used in population estimates’ in [Appendix 2](#).

Other relevant information sources about the 2018 Census data are:

- [Overview of statistical methods for adding administrative records to the 2018 Census dataset](#) for detailed information on admin enumeration
- [Processing and evaluating the quality of 2018 Census data](#) for information on the 2018 Census data and how it was processed and evaluated
- The assessment of key census variables by the census External Data Quality Assurance panel published in the reports:
  - [Initial Report of the 2018 Census External Data Quality Panel](#)
  - [2018 Census External Data Quality Panel: Assessment of variables](#)
  - [Final report of the 2018 Census External Data Quality Panel.](#)

## 3.2 Net census undercount

It is usual that some people will be missed by the census and some people will be counted more than once. Census coverage patterns vary by age, sex, ethnic group, and geographic area of usual residence. Consequently, the census coverage adjustments for different population groups, which often are interrelated, reflect these coverage patterns.

Typically, more people are missed by the census than are counted twice, so the adjustment is referred to as net census undercount (NCU). NCU is the difference between census undercount (people missed by the census but who were meant to be counted) and census overcount (people erroneously counted or counted more than once by the census).

Table 4 provides a summary of the net census undercount adjustments for selected population groups as measured by the 2018 PES.

**Table 4**

<b>Net census undercount estimates by population group</b>		
<b>2018</b>		
<b>Population group</b>	<b>NCU (mean estimates)</b>	<b>NCU as percentage of ERP for population group at 30 June 2018</b>
Total	124,900	2.5
<b>Sex</b>		
Male	73,400	3.0
Female	51,400	2.1
<b>Broad age groups<sup>(1)</sup></b>		
0–14 years	29,900	3.2
15–39 years	63,500	3.8
40–64 years	29,000	1.9
65 years and over	2,600	0.3
<b>Broad ethnic groups<sup>(2)</sup> and Māori descent</b>		
European or Other (including New Zealander)	62,300	1.8
Māori	35,500	4.3
Pacific	19,700	4.8
Asian	24,500	3.2
MELAA <sup>(3)</sup>	2,000	2.6
Māori descent <sup>(4)</sup>	39,000	4.1
<b>Regions</b>		
Auckland	47,900	2.9
Wellington	12,000	2.3
Rest of North Island	38,800	2.5
Canterbury	14,200	2.3
Rest of South Island	12,000	2.3
<ol style="list-style-type: none"> <li>1. Age at 30 June 2018.</li> <li>2. People who reported more than one ethnic group are counted once in each group reported. This means that the total number of responses for all ethnic groups can be greater than the total number of people who stated their ethnicities.</li> <li>3. Middle Eastern/Latin American/African.</li> <li>4. Māori descent population for New Zealand's electoral population. This is different from the Māori descent indicator used in 2018 Census. See <a href="#">Deriving the 2018 Māori descent electoral population</a> for more details.</li> </ol>		
<b>Source:</b> Stats NZ		

## Demographic characteristics of the net census undercount

The estimated NCU varied by age, sex, ethnic group, and geographic region. When observing the NCU as percentages of the ERP for population groups (table 4 above) some key differences were noted:

- Males (3.0 percent of ERP) had a higher percentage undercount than females (2.1 percent of ERP).

- Younger adults aged 15–39 years (3.8 percent) had a higher percentage undercount than other age groups (for example, 1.9 percent of ERP for the 40–64 years age group).
- The NCU percentage for the Pacific ethnic group (4.8 percent of the ERP) was higher than the NCU percentages for other ethnic groups. This compared with 4.3 percent for Māori and 3.2 percent of the ERP for the Asian ethnic group.
- Overall, the NCU percentage of ERP for North Island regions was slightly higher than the NCU percentage for South Island regions. At the broad region level, Auckland had the highest NCU, estimated at 2.9 percent of the ERP. The NCU for combined North Island regions outside Auckland and Wellington was estimated at 2.5 percent of the ERP.

## Measuring the net census undercount

The PES measures levels of under- and over-coverage in the census. The 2018 PES was carried out in April–July 2018, following the census on 6 March 2018. It is an interviewer-administered survey with an electronic questionnaire administered at the face-to-face interview. All usual residents in a selected private dwelling were asked for basic demographic information and to provide any addresses where they might have filled in a census form, or had one filled in for them. The 2018 PES sample consisted of approximately 15,000 private dwellings across New Zealand, and PES questionnaires containing information from over 37,500 people were received. Individuals responded to the survey.

Results from data integration of 2018 Census and PES records provided indicators of PES respondents' inclusion in the census, that is, under-counted if the respondent should have been counted but was not found in the census, and over-counted if the respondent was erroneously included or counted more than once by the census. The pool of PES respondents that were correctly included in the census formed the basis for estimating probabilities of under- and over-coverage in the census.

The report [Post-enumeration survey 2018: Methods and results](#) describes the survey design and provides results from the integration of the PES to the 2018 Census. It also describes the survey estimation approach using this data.

The 2018 PES estimation model was generally consistent with the modelled approach implemented for estimating the census coverage adjustment in the 2013 ERP (Statistics New Zealand, 2016). It provided stabilised estimates of under- and over-coverage probabilities at the disaggregated level, simultaneously. Hierarchical models for estimating under- and over-coverage probabilities were specified at the individual level and their structures included the multi-level survey design, and household and geographic area effects. They included logistic regressions with individual level effects and important interactions of age, sex, and ethnic group.

The 2018 PES model has been refined to enable estimation of coverage probabilities to the territorial authority and Auckland local board area (TALB) level (the 2013 model estimated to regional council areas). The model allowed the age profile of census coverage to vary by sex and ethnic group within each TALB. However, geographic areas within a TALB were assumed to have a consistent coverage pattern.

The NCU is measured as the difference between the coverage-adjusted census estimates and the census usually resident population count. We combined estimates of under- and over-coverage probabilities to calculate a census coverage adjustment ratio by age, sex, ethnic group, Māori descent, and TALB. The coverage adjustment ratio is applied to the census usually resident

population file for calculation of the NCU. Section [3.4 Calculating the resident population at census](#) describes the calculations of the NCU as part of the ERP compilation process.

The modelled approach for estimating census coverage based on the PES sample provides direct measures of uncertainty for demographic cross-tabulations of the NCU. Section [4.5 Measuring uncertainty in the ERP](#) provides a technical description of the uncertainty measure associated with the NCU as well as other ERP components.

### 3.3 Residents temporarily overseas

Residents temporarily overseas on census night (RTO) are not in scope to be included in the census usually resident population count. Estimates of the resident population at census date include RTOs.

The number of RTOs (civilian and military combined) at the time of the 2018 Census was 77,200. This included 76,900 civilian and 300 military RTOs. Table 6 in chapter 4 includes the distribution of RTOs at national and broad subpopulation levels. Over one-third of RTOs (47 percent) were estimated to be usually resident in Auckland, with the next largest allocation (8 percent) being to Christchurch City. A historic comparison of the RTO count at census is available as table 3 in section 2.7.

The definition of 'resident' for assigning RTO or short-term resident traveller status is compatible with the new rule for identifying international migrants based on the outcomes-based measure ([Defining migrants using travel histories and the '12/16-month rule'](#)). To be classified as an RTO, a person must be a resident after their departure and on their arrival. Therefore, according to the outcomes-based measure, a person must satisfy these rules:

- a. They must be a resident after their departure.
- b. They must be a resident on the date of their arrival.
- c. They must not stay overseas for 12 months or more out of 16 months after their departure date.
- d. The 2018 Census date, 6 March 2018, must fall between the person's departure and arrival dates.

At the national level, the number of civilian RTOs by age and sex is derived by applying the above rules to linked international travel and migration data. The number of RTOs at census night was calculated by matching New Zealand-resident traveller arrivals to (earlier) departures, to determine those out of the country at midnight on 6 March 2018.

The ERP is derived at a more detailed level than what is available in short-term resident traveller data. RTO records include information on age, sex, and country of citizenship. The merging of border movements data with integrated administrative sources compiled by Stats NZ's IDI therefore assigned resident meshblock location, ethnic group, and Māori descent indicators to the civilian RTO records. This was followed by an imputation process to fill in missing location, ethnic group, and Māori descent information (see section [4.3 Imputation in component datasets](#) for more information about the imputation process).

Because international travel and migration data does not generally include movements of military personnel, the number and distribution of military RTOs (meshblock, age, sex) used information supplied by the New Zealand Defence Force. The ethnic group(s) of military RTOs were imputed using historic administrative data on ethnic group distributions of military RTOs supplied by the New Zealand Defence Force.

### 3.4 Calculating the resident population at census

Our objective is to estimate the resident population at 30 June 2018 ( $ERP_{Jun_{mase}}$ ). For flexibility in derivations to meet customer needs, it is calculated at a more detailed level than published: meshblock of usual residence ( $m$ ), year and quarter of birth (to enable generic age calculation) ( $a$ ), sex ( $s$ ), ethnic group indicator (for each of European, Māori, Pacific, Asian, MELAA, and Other) and Māori descent indicator ( $e$ ).

The process includes the combining of contributing components followed by calculation and adjustment steps to produce the ERP at 30 June 2018. The first step derives the estimated resident population on census night (given by equations 3–5 below). This is the resident population in New Zealand on census night as counted by the 2018 Census of Population and Dwellings and adjusted for net under-coverage ( $CenAdj_{mase}^i$ ). A small proportion of the usually resident population is temporarily overseas on census night ( $RTO_{mase}^i$ ), and they are included to obtain estimates of the resident population on census night (equation 1 below).

We simulate 1,000 iterations ( $i = 1, \dots, 1,000$ ) of each dataset, and the  $i$ 'th iteration of the census night ERP is then:

$$1. \text{erpCen}_{mase}^i = \text{CenAdj}_{mase}^i + RTO_{mase}^i$$

#### Coverage adjustment of the census

A census coverage adjustment ratio is calculated using estimated census under- and over-coverage of the usually resident population in New Zealand on census night as measured by the PES. Draws from posterior distributions of under- and over-coverage probabilities ( $\phi_{under}$ ,  $\phi_{over}$ ), respectively, at a broader demographic level than the ERP calculations (noted by  $d$  below and represented as a 'TALB by age by sex by ethnic group by Māori descent' cell) provided iterations of the adjustment factor. The census coverage adjustment ratio ( $R_d$ ) is calculated as:

$$2. R_d = (1 - \phi_d^{over}) / (1 - \phi_d^{under})$$

The expected value of the adjusted census count for cell  $d$  is denoted as  $CenAdj_{mase}^i$  (for iteration  $i$ ) and calculated by applying the census adjustment ratio to the census records in scope for coverage estimation ( $Census_{mase}^{in\_scope}$ ):

$$3. \text{CenAdj}_{mase}^i = R_d^i * \text{Census}_{mase}^{in\_scope} + \text{Census}_{mase}^{out\_of\_scope}$$

The census records in scope for coverage adjustment were records of the usually resident population identified as being in New Zealand on census night, having returned a census form before the PES field enumeration started, and not enumerated as populations living in prison and defence establishments.

#### Calculating the net census undercount

The net census undercount ( $NCU_{mase}^i$ ) is the difference between the coverage adjusted census counts and the census usually resident population count ( $Census_{mase}$ ), that is, it is calculated for cell  $mase$  and iteration  $i$  as:

$$4. NCU_{mase}^i = \text{CenAdj}_{mase}^i - \text{Census}_{mase}$$

In general, the cell values of the net census undercount are not integers. Results of the net census undercount may be produced as part of the ERP process and will for this purpose include random rounding to whole numbers.

## The resident population estimates at census

The estimated resident population at census ( $CenERP_{mase}^i$ ) combines estimates of the coverage adjusted census usually resident population and the residents temporarily overseas on census night. For the  $i$ 'th iteration, the ERP at census is calculated as:

$$5. \quad CenERP_{mase}^i = CenAdj_{mase}^i + RTO_{mase}^i$$

## 4 Population change from census date to 30 June 2018

The ERP at 30 June 2018 is compiled by using the adjusted resident population estimates at census date as a basis for combining with datasets representing the components of population change between the census date and mid-year. Datasets of population change refer to the natural increase (the difference between births and deaths) and the net migration (the difference between those leaving an area and those arriving).

The following sections describe the data processes used to estimate population change between the census date and mid-year and the calculation process for calculating the ERP at 30 June 2018. The sections included in this chapter are:

[4.1 Births and deaths](#) (including adjustment to account for late or delayed registrations)

[4.2 Net migration](#) (including migrant arrivals less migrant departures (international migration) and migration of residents between geographic areas of New Zealand (internal migration))

[4.3 Imputation in component datasets](#)

[4.4 Calculating the ERP at 30 June 2018](#)

[4.5 Measuring uncertainty in the ERP.](#)

Table 5 provides a summary of the sizes of the components contributing to the ERP at 30 June 2018 for selected population groups.

**Table 5**

<b>Estimates of components contributing to the ERP at 30 June 2018</b>							
For selected population groups							
	At 2018 Census			Component change (between 6 March and 30 June 2018)		At 30 June 2018	
Population group	Census usually resident population count	Net census under-count	Residents temporary overseas	Natural increase (births less deaths)	Net migration (migrant arrivals less migrant departures)	Demographic reconciliation	Estimated resident population
Total	4,699,755	124,900	77,200	8,000	7,300	-16,600	4,900,600
<b>Sex</b>							
Male	2,319,558	73,400	39,300	4,000	3,600	-9,800	2,430,200
Female	2,380,197	51,400	37,900	4,000	3,800	-6,800	2,470,500
<b>Broad age groups<sup>(1)</sup></b>							
0–14 years	923,403	..	..	..	..	..	946,400
15–39 years	1,576,119	..	..	..	..	..	1,676,900
40–64 years	1,485,069	..	..	..	..	..	1,542,300
65 years and over	715,167	..	..	..	..	..	734,900
<b>Broad ethnic groups<sup>(2)</sup> and Māori descent</b>							

<b>Estimates of components contributing to the ERP at 30 June 2018</b>							
For selected population groups							
	At 2018 Census			Component change (between 6 March and 30 June 2018)		At 30 June 2018	
Population group	Census usually resident population count	Net census under-count	Residents temporary overseas	Natural increase (births less deaths)	Net migration (migrant arrivals less migrant departures)	Demographic reconciliation	Estimated resident population
European or Other (including New Zealander)	3,349,213	62,300	40,500	3,300	-2,300	-11,200	3,441,700
Māori	775,836	35,500	5,800	4,200	-300	-4,500	816,500
Pacific	381,642	19,700	5,600	2,500	500	-2,300	407,700
Asian	707,598	24,500	29,800	3,400	8,000	-2,600	770,600
MELAA <sup>(3)</sup>	70,332	2,000	3,000	500	1,400	-300	77,000
Māori descent <sup>(4)</sup>	896,567	39,000	6,200	4,600	-200	-5,100	941,200
<b>Regions</b>							
Auckland	1,571,718	47,900	36,600	4,000	300	-5,700	1,654,800
Wellington	506,814	12,000	7,700	700	400	-1,700	525,900
Rest of North Island	1,516,020	38,800	17,600	2,300	3,900	-5,700	1,572,900
Canterbury	599,694	14,200	8,800	700	1,400	-2,000	622,800
Rest of South Island	504,842	12,000	6,500	400	1,400	-1,600	523,500
<p>1. When analysing population change for any given age group, cohort effects also have an impact. For example, those people aged 20-24 years at 30 June 2018 are not exactly the same people as those aged 20-24 years at 6 March 2018. For this reason, individual adjustments are not shown here.</p> <p>2. People who reported more than one ethnic group are counted once in each group reported. This means that the total number of responses for all ethnic groups can be greater than the total number of people who stated their ethnicities. This is the ethnic group or groups that people identify with or feel they belong to. Ethnicity is a measure of cultural affiliation, as opposed to race, ancestry, nationality, or citizenship. Ethnicity is self-perceived, and people can belong to more than one ethnic group.</p> <p>3. Middle Eastern/Latin American/African.</p> <p>4. Māori descent population for New Zealand's electoral population. This is different from the Māori descent indicator used in 2018 Census. See <a href="#">Deriving the 2018 Māori descent electoral population</a> for more details.</p> <p><b>Symbol:</b> .. not available  <b>Source:</b> Stats NZ</p>							

See the following table in [NZ.Stat](#) for a full breakdown by ethnic group, sex and, geography:

- Estimated resident population (ERP), adjustments to derive ERP at 30 June 2018 (from census usually resident population count).



## 4.1 Births and deaths

The number of births and deaths occurring in the period between census night and mid-year is recorded through the national registration system, administered by the Registrar-General under the auspices of the Department of Internal Affairs. As birth and death registrations are required by law, this data is of high quality.

By law, both parents of a child born in New Zealand must jointly notify Births, Deaths and Marriages as soon as is reasonably practicable after the birth (deemed by the Registrar-General as generally being within two months of the birth). Deaths are required to be registered within three working days after disposal of the body concerned.

However, not all births that occurred between census night (6 March 2018) and 30 June 2018 were registered at the time of the ERP compilation. Therefore, we included an allowance for the expected number of births that would have occurred during this period but had not yet been registered at the time of ERP compilation.

Historical data indicated there was likely to be less than 20 deaths occurring during the period 6 March to 30 June 2018 that had yet to be registered as of mid-2020 for inclusion in the 2018 ERP. This number was considered too small to include as an adjustment.

To ensure the births and deaths data was as complete as required for the compilation of the ERP, the data included the following adjustments:

- Around 300 births were added to the estimate of births between census night and 30 June 2018 to account for delayed and late registrations based on completed birth registrations for previous years. The ethnic group, Māori descent, and resident region information for these births were established by sampling from donors of previous late registrations as the group of late registrants has different characteristics to those who register early.
- Small proportions of birth and death registrations had missing ethnic group, Māori descent, and/or resident meshblock location information. An imputation process filled in missing responses for these variables in the datasets. (See section [4.3 Imputation in component datasets](#) for more information about the imputation process).

## 4.2 Net migration

The net migration for geographic areas is estimated as the difference between the number of people who arrive in an area and the number of people who depart from the same area. At the national level, this is equivalent to international migrant arrivals less international migrant departures. At the subnational level, it is the combined inflows to a geographic area of internal and international migrants less the combined outflows of internal and international migrants from a geographic area.

The international and internal migration components are separately estimated at the TALB level by age, sex, ethnic group, and Māori descent. We calculate the combined net migration at the TALB level, and this is distributed to the meshblocks for the TALB according to the distribution of the estimated resident population at census (see section [4.4 Calculating the ERP at 30 June 2018](#) for more information about the calculation of net migration).

Estimation of internal and international migration between census and 30 June 2018 are described as two different processes below.

## International migration

Migrant arrivals and departures between the census date and 30 June 2018 were defined by the outcomes-based international migration measure using linked travel histories and a rule for determining a change in resident status, which defines:

- a migrant arrival as an overseas resident who arrives in New Zealand and cumulatively spends 12 out of the next 16 months in New Zealand
- a migrant departure as a New Zealand resident who departs New Zealand and cumulatively spends 12 out of the next 16 months out of New Zealand.

This rule is independent of both the individual's legal residence status and the information stated on arrival and departure passenger cards.

The numbers of migrant arrivals and departures were classed as final by the outcomes-based migration rule and available by variables such as age, sex, TALB, and country of citizenship. The migration data are linked in Stats NZ's IDI for resident location, ethnic indicators, and Māori descent information. Migrants may have address notifications recorded before and after the migrant arrival or departure date. The address notifications closest to the reference dates are selected irrespective of the time sequence of address notification in relation to the movement.

Varying levels of missingness were observed by the linking processes to the resident location, ethnic group, and Māori descent information in the IDI. Table 6 in section 4.3 gives a summary of relative levels of missingness across these variables in the migration records. Migrant arrivals in general have more missing data than departures and, overall, the Māori descent variable has the highest level of missing data. Resident location information was imputed at the TALB level due to the large number of records with missing data and the limited input data available for the imputation process.

Section [4.3 Imputation in component datasets](#) explains the imputation method and details for the specific variables in the migration datasets.

## Internal migration

The internal migration flows between geographic areas have been estimated using integrated administrative address notifications data compiled by Stats NZ's IDI from key data sources (table 9 of Appendix 1).

Measurement of population change due to residents' internal migration between the census date and 30 June 2018 is undertaken by transforming individuals' linked address histories to statistical summaries of internal migration to and from areas over the period. The steps taken are summarised below:

- From residents' administrative address notification histories, we observed paired origin and destination locations at the meshblock level. From this data, internal movements in and out of meshblock areas were identified for the time period.
- For the short time period observed, it was assumed that net internal migration estimates for TALB areas derived by the count of movements to and from areas represented an accurate measure of net internal migration. Estimates of net internal migration for the same period of previous years have shown consistent results.
- Migrants arriving in and departing from New Zealand and infants born during the period were excluded from internal migration estimation.

- The migration flows to TALB areas, including universities or other key tertiary education providers, were adjusted to account for the time lag in address notifications data compared with time of movement. For the time period, there were large numbers of address changes by students who moved to certain TALBs for tertiary study. We assumed that most of these moves occurred before tertiary courses commenced, that is, by the end of February.
- Small adjustments were applied to the internal migration flows to ensure the net migration estimates balanced at the national level.

The internal migration dataset was created at the TALB level. From integrated administrative data in the IDI, information on date of birth, sex, ethnic group, and Māori descent were included at the unit record level. An imputation process followed to fill in missing ethnic group and Māori descent indicator responses (see section [4.3 Imputation in component datasets](#) for a description of the imputation methods).

The ERP process required estimates of net migration, that is, international and internal flows combined to calculate in-migration less out-migration at the meshblock level. The process for compiling the ERP is described in section [4.4 Calculating the ERP at 30 June 2018](#) and includes a step to distribute the net migration component derived at the TALB level to meshblocks within the TALB.

### 4.3 Imputation in component datasets

The processes for compiling the migration datasets (RTO, international and internal migration) included unit record data integration with address notifications histories and ethnic group information compiled by a range of administrative sources (see table 9 of Appendix 1 for more information about the sources used). The integrated administrative sources also included the 2013 Census file. Māori descent information was accessed by merging the 2013 Census data (2018 Census data was not available as an integrated source in the IDI at the time of compiling the ERP).

Records entering the ERP process and sourced by registration data and administrative collections had complete information on date of birth and sex. However, merging of resident location, ethnic group, and Māori descent indicator information encountered missingness in one or several of these variables (see table 6 below).

Levels of missingness were also highly dependent on whether the information was observed for RTOs, migrant arrivals, migrant departures, or for records of internal migration. There were nearly negligible proportions of missingness in birth and death records. By contrast, 27 percent of migrant arrivals and 17 percent of migrant departures had missing ethnic groups, and 12 percent of migrant arrivals and 10 percent of migrant departures had missing resident location information.

**Table 6**

<b>Level of item non-response observed in population variables by component dataset</b>				
Component	National total (number)	Non-response (percentage of total)		
		Location of usual residence	Ethnic group indicator	Māori descent indicator
RTOs	77,200	11	8	48
Births	18,300	0	0	3
Deaths	10,500	1	0	8
Migrant arrivals	36,700	12	27	91
Migrant departures	29,300	10	17	63
Internal migration	...	...	1	40

**Source:** Stats NZ

The compilation of the component datasets included a hot-deck imputation process for filling in missing values, that is, missing values were replaced with an observed value from a ‘similar’ record within the dataset. The process imputed the variables sequentially in order of least to most missingness. The hot-deck imputation method specified suitable donor pools, using auxiliary variables, such as age, sex, and country of citizenship in migration records. It applied a random sequential algorithm that included random sorting of variables before imputation. Hot-deck imputation was carried out 1,000 times to introduce a degree of randomness in the selection of resident location, ethnic group, and Māori descent information.

The hot-deck imputation approach, as described above, was used to maintain the distributions given by observed values and to minimise any bias introduced by the process. The approach ensured the donor pools were specified at suitable broad demographic levels in order to minimise the risk of repeated use of the same donor, as well as to avoid the use of outliers in the data.

## 4.4 Calculating the ERP at 30 June 2018

The steps to produce the unadjusted ERP at 30 June 2018 combines the resident population estimates at census date and the components of population change between the census date and 30 June. The natural increase (excess of births over deaths) is added at the required demographic level (meshblock, age, sex, ethnic group, and Māori descent). However, for computational efficiency the net migration contribution to population change has been calculated at the TALB level (see section [4.2 Net migration](#)). The unadjusted ERP at 30 June is calculated in the following steps:

1. adding births
2. subtracting deaths
3. adjusting for negative population cells (see section [5.1 Adjustment for implausible population group estimates](#))
4. aggregating the combined ERP dataset to TALB level
  - adding net international migration
  - adjusting for implausible population cells due to migrant departures (see section [5.1 Adjustment for implausible population group estimates](#))
  - adding net internal migration
  - adjusting for implausible population cells due to internal out-migration (see section [5.1 Adjustment for implausible population group estimates](#)) and adjusting to balance net internal migration at the required demographic level
5. distributing net migration to the meshblock level according to demographic distributions in the ERP dataset.

The final step combines a demographic reconciliation adjustment of ages 0–14 years with the unadjusted ERP at 30 June 2018 (refer to section [5.2 Demographic reconciliation adjustment](#) for the derivation on the demographic reconciliation adjustment). In summary, the final computational step for the  $i$ th iteration combines the components of population change births ( $Bth_{mase}^i$ ), deaths ( $Dth_{mase}^i$ ), and net migration ( $netMig_{mase}^i$ ) with the ERP at census date ( $CenERP_{mase}^i$ ) before including the final demographic reconciliation adjustment ( $DemAdj_{mase}^i$ ) in the calculations of the ERP at 30 June 2018.

$$ERPJun_{mase}^i = CenERP_{mase}^i + Bth_{mase}^i - Dth_{mase}^i + netMig_{mase}^i + DemAdj_{mase}^i .$$

## 4.5 Measuring uncertainty in the ERP

Uncertainty measures incorporated in the components sourced from registration and administrative data resulted from the imputation processes for missingness observed in resident geographic location, ethnic group, and Māori descent information (see table 8 in Appendix 1). However, these

processes have not incorporated uncertainty around the selection of records and related data quality measures in the auxiliary data sources.

As a result, the estimated uncertainty in the 2018-base ERP is a partial indication of uncertainty and builds on the work begun with the 2013-base ERP (see [Measuring uncertainty in the 2013-base estimated resident population](#)).

The estimated uncertainty intervals available with the ERP results at 30 June 2018 comprised contributions from census coverage estimation as measured by the PES and from imputation processes implemented at the demographic variable level when compiling the component datasets (mainly RTOs and net migration).

Census coverage estimation is using hierarchical models for estimating under- and over-coverage probabilities, and these models provide coherent and integrated measures of uncertainties of the coverage probabilities. By taking draws from their posterior distributions and incorporating these draws into the calculation of the coverage adjustment ratio (equation 2 in section 3.4), we have directly accounted for the uncertainty associated with the calculation of the net census under-coverage.

The 2018 PES modelled approach is described for the general audience in the report [Post-enumeration survey 2018: Methods and results](#) and is compatible with the approach used in 2013.

The process for compiling the ERP used a simulation approach to account for the uncertainties in the contributing data. The coverage-adjusted census population estimates, and the imputation processes used to compile the component datasets, generated 1,000 independent iterations. Each of the ERP calculation steps outlined in sections [3.4 Calculating the resident population at census](#) and [4.4 Calculating the ERP at 30 June 2018](#) above were carried out over all these iterations.

The ERP results in published tabulations refer to estimated means of the data. Estimates of uncertainty intervals of ERP results refer to the 95 percent credible intervals or as defined by the 0.025 and 0.975 percentile estimates.

## 5 Adjustments to the ERP

This chapter describes the methods used for the two different adjustments in the ERP process:

[5.1 Adjustment for implausible population group estimates](#)

[5.2 Demographic reconciliation adjustment.](#)

A set of adjustments were applied as an iterative process to adjust implausible estimates for population groups. The purpose of the adjustment to the ERP at 30 June 2018 was to demographically reconcile the population estimates for ages 0–14 years with the administrative-based estimates.

### 5.1 Adjustment for implausible population group estimates

The steps for calculating the ERP at mid-year included removing population in two stages, that is, by subtracting deaths and by subtracting out-migration. These steps observed implausible population estimates for small proportions of variable combinations at the detailed level. The components that are input to the ERP process have been compiled independently at the most detailed demographic level. Imputation processes and other data edits have resulted in the size of some population groups

in deaths and out-migration data not being plausible when compared with the size of the resident population estimates at census.

A deterministic adjustment approach re-distributed records with implausible population estimates across records with plausible estimates. It was an iterative approach that maintained the population group distributions with least missingness and resulted in adjustments being made to records with higher levels of imputation.

The integral processes to adjust for implausible population group estimates due to inconsistencies in deaths and out-migration data had negligible impact across all demographic distributions. Consequently, these adjustments have not been included as an explicit component of the ERP.

## 5.2 Demographic reconciliation adjustment

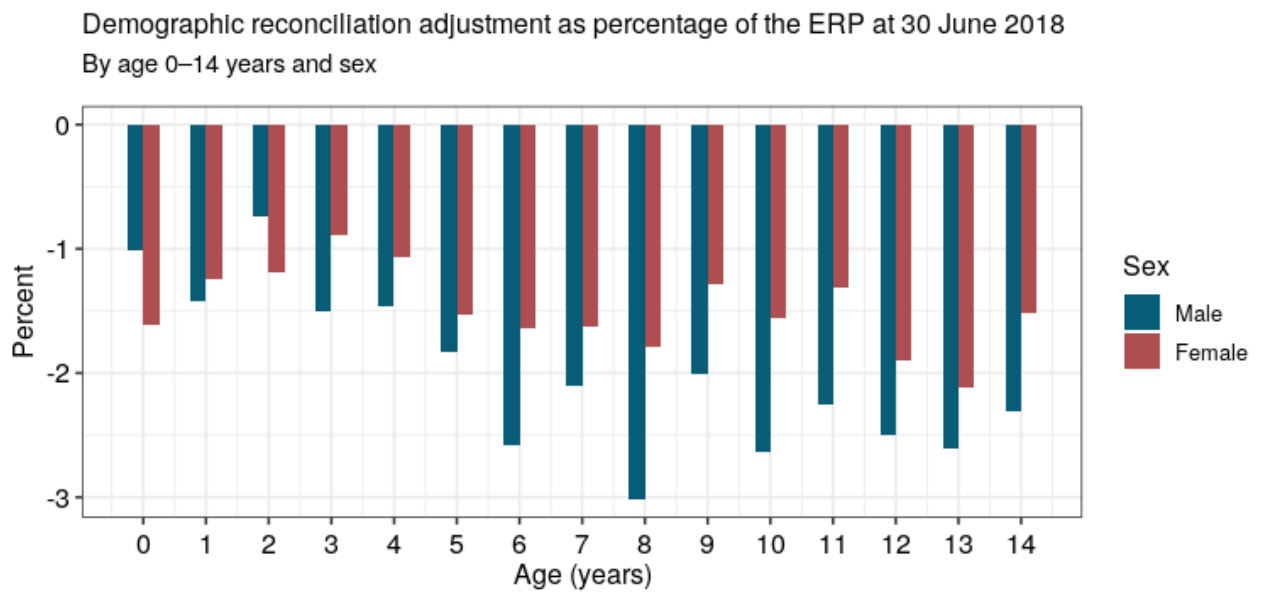
The adjustment to the ERP at 30 June 2018 for ages 0–14 years accounted for the discrepancy observed between the derived resident population estimates from administrative data and the unadjusted ERP for these ages. Administrative-based population estimates enabled by birth and death registrations, and migrant arrivals and departures for cohorts born in the September quarter of 2003 and onwards, were used as benchmark estimates for children. The discrepancy, based on the administrative benchmark, was a result of a too-high adjustment for net census under-coverage of children, by the PES. By comparison, the 2018 census usually resident population count for this age group was low.

The administrative-based estimates are regarded as high quality. This reflects the very high coverage of birth and death registrations in New Zealand and the outcomes-based migration measure (available from 2001) providing accurate accounts of changes in resident status due to international migration among the youngest ages.

The demographic reconciliation adjustment to the ERP at 30 June 2018 was applied as a numerical alignment with the administrative-based estimates by single years of age and sex for children. [Section 4.4 Calculating the ERP at 30 June 2018](#) describes how the ERP compilation process implemented the demographic reconciliation adjustment at the subpopulation level.

The demographic reconciliation adjustment decreased the estimated resident population, aged 0–14 years, by 16,600 (9,800 males and 6,800 females). Table 5 (section 4) includes estimates of the demographic reconciliation adjustment as an ERP component for selected population groups. Figure 4 summarises the adjustment, as a percentage of the ERP at 30 June 2018, by single year of age and sex. The adjustment was larger for males than females at most ages and aligned the sex ratio to that observed from administrative-based estimates.

**Figure 4**



Source: Stats NZ

The administrative-based resident population estimates aged 0–14 years at 30 June 2018 were derived independently from compiling birth and death registrations, and net migration by year, and quarter of birth and sex (from September quarter 2003). A small allowance was made for unregistered births that occurred before June 2018, by inflating birth registrations for years ending June 2015 to 2018 relative to birth notifications.

The unadjusted ERP at 30 June 2018 aggregated to year and quarter of birth and sex for ages 0–14 years formed the basis for comparison with the administrative-based estimates. The adjustment factor (by year and quarter of birth and sex) was calculated as the ratio of the mean of the unadjusted ERP (over 1,000 iterations) to the administrative-based estimates. The adjustment factors (by year and quarter of birth and sex) were applied consistently across subpopulation domains, such as geographic area, ethnic group, and Māori descent.

## 6 References

Statistics New Zealand (2016). [Measuring uncertainty in the 2013-base estimated resident population \(Stats NZ Working Paper No 16–04\)](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).

Statistics New Zealand (2020). [Post-enumeration survey 2018: Methods and results](#). Retrieved from [www.stats.govt.nz](http://www.stats.govt.nz).

United Nations (1952). [Manual I: Methods of Estimating Total Population for Current Dates](#) (United Nations Publications, Sales No. 52.XIII.5).



## Appendix 1

Table 7 summarises the ERP tabulations available by specified categories of each of the variables geographic area, age, sex, and ethnicity and Māori descent.

**Table 7**

<b>Output variables for the ERP at 30 June 2018</b>			
<b>Geography</b>	<b>Age</b>	<b>Sex</b>	<b>Ethnicity and Māori descent</b>
National	Single year of age (0, 1, ..., 100+), 5-year age groups, Broad age groups, Median age, Total all ages	Male, Female, Total	European or Other (including New Zealander), Māori, Pacific, Asian, Middle East/Latin America/African (MELAA), Māori descent
Regional council areas (RCs)	5-year age groups, Broad age groups, Median age, Total all ages		
Territorial authority and Auckland local board areas (TALBs)			
Statistical area 2 (SA2s)			
Urban rural areas			
District health board areas (DHBs)			
<b>Source:</b> Stats NZ			

Table 8 provides an overview of the sources of uncertainty that contributed to the measure and are incorporated into the ERP process.

**Table 8**

<b>Overview of sources of uncertainty and methods to measure uncertainty in the ERP</b>		
<b>ERP component</b>	<b>Source of uncertainty included in the ERP</b>	<b>Method for uncertainty measure</b>
Census usually resident population count	NA	<a href="#">Data sources, editing, and imputation in the 2018 Census</a>
Estimates of census coverage	Estimate of coverage adjustment ratio by age, sex, TALB, ethnic group, and Māori descent indicator variables (section 3.4)	Estimates of under- and over-coverage probabilities by Bayesian hierarchical models use multiple draws from their posterior distributions
Residents temporarily overseas on census night (RTO)	Non-response from integration of short-term resident traveller records with resident location, ethnic group, and Māori descent indicator information (section 4.3)	Hot-deck imputation approach

<b>Overview of sources of uncertainty and methods to measure uncertainty in the ERP</b>		
ERP component	Source of uncertainty included in the ERP	Method for uncertainty measure
Births and deaths	Non-response in ethnic group and Māori descent indicator information	Hot-deck imputation approach
International migration	Non-response from integration of short-term resident traveller records with resident location, ethnic group, and Māori descent indicator information (section 4.3)	Hot-deck imputation approach at the TALB level
Internal migration	Non-response from integration of internal migration records with ethnic group and Māori descent indicator information (section 4.3)	Hot-deck imputation approach at the TALB level
Demographic reconciliation	NA	NA
<b>Source:</b> Stats NZ		

Table 9 summarises the sources used to produce each component and the variables observed by the sources. This includes specification of the variables that required imputation for missing responses due to data integration processes for compiling the component datasets.

**Table 9**

<b>Overview of sources, variables fully observed and variables with non-response in the ERP components</b>			
ERP component	Sources used	Variables fully observed	Variables with non-response
Census usually resident population count	2018 Census of Population and Dwellings	Meshblock of usual residence, year and quarter of birth, sex, ethnic group (level 1 indicators), Māori descent indicator (electoral)	There are no variables with non-response due to the use of administrative data in the 2018 Census ( <a href="#">Data sources, editing, and imputation in the 2018 Census</a> ).
Census coverage adjustment ratio	<a href="#">Post-enumeration survey 2018: Methods and results</a>	TALB of usual residence, age (0, 1, 2, ..., 100+), sex, ethnic group (level 1 indicators), Māori descent indicator (electoral)	PES demographic variables used for coverage estimation have been imputed for missing values
RTOs	<ul style="list-style-type: none"> <li>Short-term resident traveller records (Stats NZ migration database)</li> <li>Compiled address notifications from administrative sources, including 2013 Census (Stats NZ IDI)</li> </ul>	Year and quarter of birth (age), sex, and country of citizenship (Country of citizenship has non-response but is not imputed.)	Meshblock, level 1 ethnic group indicators, Māori descent indicator

<b>Overview of sources, variables fully observed and variables with non-response in the ERP components</b>			
ERP component	Sources used	Variables fully observed	Variables with non-response
	<ul style="list-style-type: none"> <li>• Compiled ethnicity information from consented administrative sources including 2013 Census (Stats NZ IDI)</li> <li>• Māori descent information (electoral) from 2013 Census (Stats NZ IDI)</li> <li>• Military RTOs from Ministry of Defence</li> </ul>		
Births	Registrations (Stats NZ vitals database)	Year and quarter of birth, and sex (Included an adjustment for late registrations.)	Meshblock of usual residence, level 1 ethnic group indicators, Māori descent indicator, parental ethnic group indicators, parental Māori descent indicators
Deaths	Registrations (Stats NZ vitals database)	Year and quarter of birth, sex	Meshblock of usual residence, level 1 ethnic group indicators, Māori descent indicator.
International migration	<ul style="list-style-type: none"> <li>• Migrant records (Stats NZ migration database)</li> <li>• Compiled address notifications from administrative sources incl. 2013 Census (Stats NZ IDI)</li> <li>• Compiled ethnicity information from consented administrative sources including 2013 Census (Stats NZ IDI)</li> <li>• Māori descent information (electoral) from 2013 Census (Stats NZ IDI)</li> </ul>	Year and quarter of birth, sex, country of citizenship, country of birth (Country of citizenship and country of birth have non-response but are not imputed.)	TALB of usual residence, level 1 ethnic group indicators, Māori descent indicator
Internal migration	<ul style="list-style-type: none"> <li>• Migrant records (Stats NZ migration database)</li> <li>• Compiled address notifications from administrative sources incl. 2013 Census (Stats NZ IDI)</li> <li>• Compiled ethnicity information from consented administrative</li> </ul>	TALB, year and quarter of birth, sex	Level 1 ethnic group indicators, Māori descent indicator

<b>Overview of sources, variables fully observed and variables with non-response in the ERP components</b>			
ERP component	Sources used	Variables fully observed	Variables with non-response
	sources incl. 2013 Census (Stats NZ IDI) <ul style="list-style-type: none"> <li>• Māori descent (electoral) information from 2013 Census (Stats NZ IDI)</li> </ul>		
<b>Source:</b> Stats NZ			

## Appendix 2

### Geographic areas used in population estimation

The ERP at 30 June 2018 was compiled at the most detailed geographic level, a meshblock, which allows aggregation of the ERP to all the required statistical geographies. Each component of the ERP has been estimated at that level, enabled by the administrative sources used (see table 6 in chapter 4). Population estimates are prepared to the following statistical output geographies grouped by national or subnational levels.

#### National level

New Zealand. For statistical purposes, the term 'New Zealand' refers to 'geographic New Zealand'. This includes offshore islands, but does not include the Cook Islands, Niue, or the Tokelau Islands.

#### Subnational level

##### Regional council areas (regions)

The boundary of each regional council conforms as far as possible to one or more water catchment areas and takes into account factors such as natural resource management, land use planning, and environmental matters. Generally, regional council areas contain complete territorial authorities. Where territorial authorities straddle regional council boundaries, the affected area has been statistically defined in complete area units.

##### Territorial authority areas

A territorial authority is a city council or district council. There are 67 territorial authorities comprising 12 cities, 53 districts, Auckland, and Chatham Islands territory. 'Community of interest' was a key factor in determining the extent of territorial authority areas.

##### Auckland local board areas

Twenty-one local board areas were established as part of the new local government arrangements for Auckland in 2010. Statistics at the local board level are often presented alongside territorial authority level statistics to provide greater detail for Auckland.

##### Statistical area 2 (SA2)

The SA2 geography aims to reflect communities that interact together socially and economically. In populated areas, SA2s generally contain similar-sized populations. SA2s in city council areas generally have a population of 2,000–4,000 residents while SA2s in district council areas generally have a population of 1,000–3,000 residents. SA2 replaces the area unit geography from the Standard Areas Classification 1992. Based on boundaries at 30 June 2020, there are 2,255 SA2s covering all of New Zealand.

##### Urban rural

Urban rural is a statistical geography that classifies New Zealand into areas that share common urban or rural characteristics. It is used to disseminate a broad range of Stats NZ's social, demographic, and economic statistics. They are statistically defined areas with no administrative or legal basis. **Urban areas** are characterised by high population density with many built environment features where people and buildings are located close together for residential, cultural, productive,

trade, and social purposes. **Rural areas** represent land-based areas outside urban areas. A rural settlement is a cluster of residential dwellings about a place that usually contains at least one community or public building. **Other rural** areas are the mainland areas and islands located outside urban areas or rural settlements. Other rural areas are defined by territorial authority.

For more information about the urban rural classification and other statistical output geographies, see the [Statistical standard for geographic areas 2018](#).

## Meshblock

A meshblock is the smallest geographic unit for which we collect statistical data. Meshblocks vary in size, from part of a city block to large areas of rural land. Each meshblock borders on another to form a network covering all New Zealand, including coasts and inlets. The meshblock classification extends out to New Zealand's 200 nautical mile exclusive economic zone and is digitised to the 12-mile (19.3-km) limit. Meshblocks are aggregated to build larger geographic areas, such as SA2, territorial authority, and regional council areas. At the time of the 2018 Census, there were 53,589 meshblocks in New Zealand.