2023 Post-enumeration Survey: Standard design for coverage estimation

Citation
Stats NZ (2023). *2023 Post-enumeration Survey: Standard design for coverage estimation*. Retrieved from www.stats.govt.nz.

ISBN 978-1-99-104929-2

Published in March 2023 by
Stats NZ Tatauranga Aotearoa
Wellington, New Zealand

Contact
Stats NZ Information Centre: info@stats.govt.nz
Phone toll-free 0508 525 525
Phone international +64 4 931 4600
www.stats.govt.nz

Contents

List of tables and figures

List of tables

List of figures

Purpose and summary

Purpose

This paper outlines the planned standard design for coverage estimation of the 2023 Census using the 2023 Post-enumeration Survey (PES). The standard design includes all statistical design components involved in coverage estimation, as well as any dependencies between the statistical components and the sample survey operation. The standard design excludes detailed information about the sample survey operation itself, which will be covered in other operational planning documents.

Stats NZ, like other national statistical organisations, makes concerted efforts to ensure as high a coverage as possible in its five-yearly census of population and dwellings. However, censuses everywhere tend to miss some people or count people more than once.

The PES is a household survey undertaken shortly after the census to evaluate the completeness of census coverage, and to report on census response rates. It involves an independent re-enumeration of a statistically designed sample of private dwellings and the people within them covered by the national census. The PES responses are linked to the census records (both responses and admin enumerations) to find out who was missed by census, counted more than once, or counted in error (for example, people who were overseas on census night, babies born after census night). The key outputs produced include the net census undercount, gross undercount, and gross overcount. These are produced at a national level and by a range of subpopulations of interest.

At Stats NZ the project tasked with design, implementation, and output of coverage measures for 2023 Census is the 2023 Census Coverage project. A core deliverable of this project is design and operation of the 2023 PES. Throughout this document we use the 2023 Census Coverage project to refer to the broader context of the work programme or the team involved, and we use 2023 PES to refer to the survey and related statistical design.

It is helpful for readers to consider and note that the PES aims to measure the coverage of the census, not the quality of the census. In fact, PES exists to assess the degree of bias in the census. As the PES is a household survey, it is also subject to sample and non-sample error. We use a range of statistical design techniques to minimise systematic non-sample errors as much as possible. There is always a likelihood of some residual bias in the PES population estimates, but we can be confident they are significantly smaller than the bias in the census.

In addition to outlining the key design components, this report also summarises design work undertaken in response to the change in the census model, and lessons learnt from the 2018 PES. This report also introduces the next phase of design work – the development of an adaptive design.

Summary of contents

This report is organised into four sections, with the first three each containing several chapters.

Section 1 provides necessary context to support this report including:

- Introduction to the PES

- Introduction to coverage methodology
- Overview of PES approach to working for and with Māori
- Overview of the 2023 Census design
- Privacy, security, and confidentiality.

Section 2 focuses on transitional work between the 2018 and 2023 PES cycles, and includes the following chapters:

- Improving the quality of the PES dwelling frame
- Recommendations from 2018 PES independent review
- Coverage measurement of a combined census model by design
- 2018 PES lessons incorporated into the 2023 PES design.

Section 3 provides detailed descriptions of all key design components in the planned standard design for coverage estimation of the 2023 Census via a 2023 PES. Chapters included in this section are:

- Producing high-quality population estimates for Māori
- Scope and population definitions
- Survey objectives
- Sample size and quality targets
- Questionnaire content
- Dwelling enumeration approach
- Sample design
- Linking
- Treatment of missing data
- Estimation.

Section 4 outlines the work included in the next phase of the design: developing an adaptive design that can respond to potential issues placing pressures on the statistical assumptions of the standard design.

Key changes from 2018 PES design

The proposed design of the 2023 PES is largely consistent with the design implemented for the 2018 PES (Stats NZ, 2020a), with a few key changes made in response to lessons learnt in 2018, change in census taking model, and changes in the wider statistical system.

Key changes from the 2018 PES design include:

- a shift to collect gender instead of sex
- redesign of dwelling enumeration approach
- update to sample design stratification variables and an increase in sample size
- linking PES to an admin ever-resident population (referred to throughout as the Integrated Data Infrastructure (IDI) Spine, see Appendix 1) to enable triangular linking to improve PES linking with census
- extension of imputation methods to include imputation from admin data
- population estimates produced using usual residence location (instead of PES dwelling location) and aiming to produce estimates by three gender categories.

Section 1: Context

1.1 Introduction to the PES

The main purpose of the Post-enumeration Survey (PES) is to measure the level of coverage (undercount and overcount) in the census. Coverage measures are key performance indicators for census, constituting a major part of the evaluation component for the census. This evaluation component is a government requirement when requesting funding for large programmes, such as the census. Each PES captures information and lessons that can be used to improve the following census.

In New Zealand, census counts are not adjusted directly to incorporate errors in counting identified by a PES. Coverage measures are used within Stats NZ to adjust the base population (Stats NZ, 2020a). This base population is the estimated resident population (ERP), which is used to derive post-censal population estimates and projections. Census coverage measures improve the accuracy of the base population, which in turn leads to more accurate post-censal estimates and projections.

Census coverage is not uniformly distributed. Younger people, males, students, new migrants, and members of smaller or disadvantaged groups are all more likely than other groups to be missed (and therefore undercounted) in the census (O'Hare, 2019). All these groups are of high interest for policy making. As the ethnic and geographic diversity of the New Zealand population increases, the importance of measuring and accounting for the differences in undercount also increases. These differences are essential elements in producing robust ethnic and subnational estimates and projections, and vitally important for ensuring equality and equity of treatment of people across society. Thus, production of robust coverage measures is a critical component in Stats NZ meeting our responsibilities to Māori as Te Tiriti partners (Stats NZ, 2023a).

Population estimates and projections are widely used by central and local government to assess current and future needs for facilities and services in health, education, housing, and social welfare, and in the allocation of public funds to provide these services. Incorporating adjustments for census coverage into the base estimated resident population enables identified population groups to be as accurately represented in population statistics as possible, thereby facilitating better appropriation of funds. Businesses and community organisations also use estimates and projections data, for planning, research, and marketing purposes. Population estimates are used to improve the quality of numerous other surveys, by providing benchmarks for calibration of sample survey estimates.

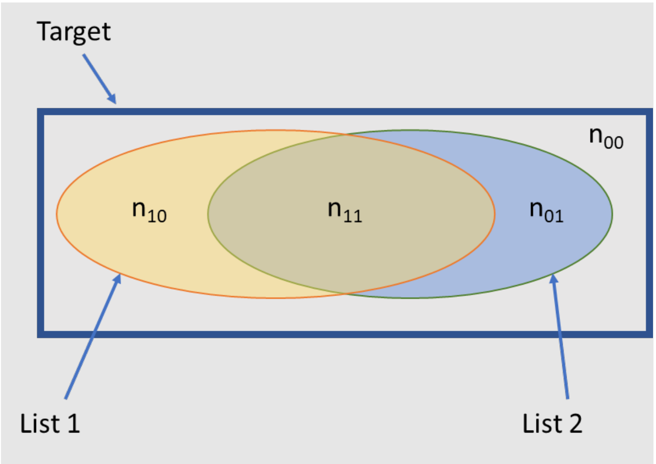
Earlier PESs (1996, 2001, and 2006) produced coverage measures of private dwellings in addition to the coverage measures produced for the New Zealand population. In 2013 and 2018, we encountered challenges in producing robust coverage estimates of private dwellings, in part due to the modernisation of the census model. Following the challenges experienced in 2013 and 2018, dwelling coverage measurement is out of scope for the 2023 PES.

1.2 Introduction to coverage methodology

The methodology used at Stats NZ to measure and estimate census coverage is based on dual system estimation (DSE), also known as capture / re-capture. In a typical DSE situation, two partial captures of individuals are collected from the target population. The two partial captures are linked, and the union is used to identify individuals on both lists, and those only found on one of the two lists.

In figure 1, the solid box outline represents the target population, with yellow and blue ovals representing sample captures for list 1 and list 2. Records captured in both lists are noted as n_{11}, those only in list 1 as n_{10} and those only in list 2 as n_{01}. Records that are not captured in either of the lists are noted as n_{00}.

Figure 1: Dual system estimation



Alternatively, the DSE problem can be depicted in a contingency table as shown in table 1. Binary indicators indicate a record being captured by a list (L_1 and L_2).

Table 1: DSE contingency table

		List 2 (L_2)		
		1	0	
List 1 (L_1)	1	n_{11}	n_{10}	n_1
	0	n_{01}	n_{00}	
		n_2		Target (N)

The target population is the sum of the inner cells. However, only three of those cells have observed counts, and the fourth is unknown: n_{00}. Therefore, an estimator equation is used to calculate the estimated total population, with the Lincoln Petersen estimator (Petersen, 1896; Lincoln, 1930) the most common.

$$\widehat{N} = \frac{n_1 \times n_2}{n_{11}}$$

From the estimated target population, it is straightforward to derive the under-coverage in a given list (for example, list 1) as $\widehat{N}$ - n_1.

In the PES application of DSE, the two lists are sourced from capture of individuals in the PES sample (list 1, L_1), and the large capture of individuals in the census dataset (list 2, L_2). We use L_1 and L_2 as inclusion indicators for lists 1 and 2, respectively. Because the PES list is a sample, the estimation framework gets adapted to operate primarily on the top line of the contingency table (where capture in list 1, $L_1 = 1$). Records captured by the PES are used to estimate the probability of being in list 2 (census) given capture in list 1 (PES): $\hat{P}(L_2 = 1 \mid L_1 = 1) = n_{11}/n_1$

We also assume that PES and census are independent, therefore, the probability of being in census is equal for PES and non-PES individuals:

$$\hat{P}(L_2 = 1 \mid L_1 = 1) = \hat{P}(L_2 = 1) = n_{11}/n_1$$

This probability is then used to derive an adjustment ratio that can be applied to the observed census count to estimate the total population. This rearrangement is referred to as a 'one-way' DSE.

$$\hat{N} = n_2 \times \frac{1}{\hat{P}(L_2 = 1)} = n_2 \times \frac{1}{n_{11}/n_1} = \frac{n_1 \times n_2}{n_{11}}$$

The DSE approach is theoretically quite simple but the estimation assumptions that underpin the method can be challenging to uphold in a real-life application like the PES (Stats NZ, 2021a).

- **Causal independence** – the likelihood of being recorded on one list (that is, census) having no relationship with the likelihood of being recorded on the other (that is, PES). This is a difficult assumption to achieve and validate since it is not directly observable. We focus efforts on minimising the risk of causal dependence through ensuring operational independence (separate staff, different address information, no overlap of PES field operation with census field operations, strict embargo of PES areas from census). There is always a risk of some residual causal dependence due to a small section of the population who actively avoid both the census and the PES and take measures to remain invisible.
- **Homogeneity of capture** – all records having the same likelihood of being captured in a given list. For example, everyone has the same likelihood of being counted by the census. This assumption is not achievable at the total population level so instead we achieve this by aiming to collect information about individuals that may have a meaningful relationship with their likelihood of under- or over-coverage. We then use this information to group people together with others that have similar coverage likelihoods. This includes demographic information such as age, sex/gender, ethnicity, and geographic location of residence.
- **Perfect linking** – error-free linking between the two lists (census and PES records) including no missed links, and no records incorrectly linked. We achieve this through a combination of automatic and clerical (manual) linking. We take a conservative approach to automatic linking to reduce the risk of false positive links (two records linking when they don't refer to the same person), with clerical linking to resolve any false negative links (two records not linking when they do refer to the same person) remaining after automated linking. In 2018 we also introduced linkage acceptance criteria to filter both PES and census records to only include those with high-quality information to support accurate linking. In 2023 we plan to also link PES responses to the IDI spine (see Appendix 1). We will use the three different links (PES to census, PES to IDI spine, census to IDI spine) to identify inconsistencies and to perform triangular linking.
- **No erroneous inclusions** – records that are not part of the target population for capture (such as babies born after census night, deaths prior to census night, residents temporarily overseas on census night). We achieve this through detecting erroneous inclusions in the census file, excluding them from the estimation of gross under-coverage and performing a separate estimation of gross over-coverage. The two estimates are then combined to

produce net under-coverage. In PES we extend this group to also include people who were counted more than once.
- **Closed population** – no entries or exits to or from the target population in between the different capture points. Realistically, this is not achievable or controllable; however, we use data about the population (such as births, deaths, and border crossings) to approximate this assumption at a national level. This assumption is harder to achieve at a subnational level due to lack of specific information about internal migration. Historically, and for 2023, we have assumed this is negligible with the small amount of internal movement occurring between census night and the PES survey period.

1.3 Overview of PES approach to working for and with Māori

The 2023 Census Coverage project is committed to meeting our responsibilities to Māori as Te Tiriti partners (Stats NZ, 2023a). Application of these responsibilities ensure the fair, equitable, impartial, positive, and accurate measurement and description of response and coverage by census. The 2023 Census Coverage project has set the following outcome objectives to uphold these responsibilities:

- contribute to the broad objective of rebuilding trust and confidence in the census by providing robust and transparent measures of coverage achieved by census
- raise awareness of the purpose of the 2023 Post-enumeration Survey
- ensure that robust estimation of coverage of Māori populations is enabled
- communicate outcomes to Māori as an assurance of a commitment to the principle of awhi atu awhi mai (reciprocity).

These outcomes will be supported by:

- applying a te ao Māori perspective to design and implementation
- representation in governance decisions
    - the 2023 Census Coverage Board will include at least one member who will be a strong advocate and bring a te ao Māori perspective
- representation in quality assurance processes
    - the 2023 Census Coverage Quality Assurance Panel will include at least one member who will be a strong advocate and should ideally be a recognised expert in indigenous statistics and indigenous data sovereignty
- building cultural capability of staff working on the project
- use and implementation of relevant policy and guidance such as the Ngā Tikanga Paihere framework (Stats NZ, 2020b), and Mana Ōrite Relationship Agreement.

1.4 Overview of 2023 Census design

In October 2021 Stats NZ published the *2023 Census: High Level Design* which outlines the priorities, scope, and key design features for the 2023 Census (Stats NZ, 2021b). The 2023 Census mission is to “re-build trust and confidence and deliver quality data to our customers” (Stats NZ, 2021b, p. 8). This is supported by four strategic initiatives:

1. drive high response rates
2. deliver quality data to meet customer needs
3. build trust to gain commitment
4. build trust and value for Māori through data equity.

Key features of the 2023 Census design include (Stats NZ, 2021b):

- more, earlier, and ongoing community engagement activity, including supporting communities to drive their own response to the census through Community Counts initiatives
- more census collectors at census time and census collections will start earlier
- people will have the option of doing the census online or on paper forms, with up to 50 percent of households receiving paper forms
- increased assistance to complete the census, including help on the doorstep in some areas and engagement centres in areas where they are needed most
- questions about sexual identity and gender included for the first time in the New Zealand census
- increased focus on accessibility, with different formats supporting a wider range of people in Aotearoa New Zealand
- increased focus on building data capability with communities, Māori, and iwi, because we know that understanding the value of census data and how to use it is a strong motivator to participation
- using a combined methodology, which involves collecting responses using census forms and using administrative data (other data sources) when census responses are missing – while used for the 2018 Census to address data quality issues, a combined method model is an intentional part of the design of the 2023 Census to maximise the quality of census data
- planning for limited change to the census forms and the products and services we will provide post-census (we will validate what was intended for the 2018 Census with partners, customers, and stakeholders before finalising for the 2023 Census).

In addition, Stats NZ received funding "to run an iwi-led census data collection pilot to build iwi-Māori data capability and capacity alongside the 2023 Census" (Stats NZ, 2021b, p. 5).

Table 2 summarises the key design features and how they interact with coverage measurement. Note, a lot of the features are designed to improve response rates, particularly for those groups typically undercounted. Higher and more evenly distributed response rates across population groups are very beneficial for the PES estimation framework and reduces the risk of estimation bias.

Table 2: Relationship between 2023 Census design features and 2023 PES

2023 Census design feature	Implications for 2023 PES
earlier and sustained engagement with communities	operational challenges to ensure PES sample areas remain embargoed from census when engagement activities can overlap with PES dwelling enumeration operation
earlier census collection	increased risk of over-coverage and differences in key location information (from people completing the census in advance)
more census collectors	reduced resource pool for PES to recruit from (to support PES independence from census)

paper or online response (with increased paper available)	increase in paper responses creates an increased risk of the PES estimation assumption of perfect linking being valid (due to potential reduced quality of key linking information such as name and date of birth)
more assistance to complete census forms	no specific implications for PES
the inclusion of gender and sexual identity questions	PES will align with the Stats NZ data standard and collect gender (Stats NZ, 2021f)
increased focus on accessibility, different formats, and language support	PES also increasing focus on accessibility, different formats, and language support
building data capability with communities, Māori, and iwi	PES to ensure the field workforce is diverse and representative of communities
a combined census model by design	new and different drivers of coverage error which create challenges for the PES estimation methods and assumptions (Stats NZ, 2021a) • increased risk of duplicate over-coverage impacting the PES perfect linking assumption • increased risk of erroneous inclusion over-coverage from admin enumerations
limited change to the census forms, and products and services	PES questionnaire change to collect gender rather than sex
iwi-led collection	potential to alter expected coverage patterns – areas need to be considered as part of PES sample design

1.5 Privacy, security, and confidentiality overview

Stats NZ is bound by the Data and Statistics Act 2022 and the Privacy Act 2020 to keep the information we collect confidential (Stats NZ, 2021c). Stats NZ's ability to ensure statistically sound, safe, responsible, and culturally appropriate use of both response and administrative data is essential in building and maintaining the trust of our respondents, customers, and stakeholders.

The 2023 Census Coverage project takes a 'privacy by design' approach to Information privacy, security, and confidentiality. This approach ensures we effectively manage personal and other confidential information of respondents, staff, customers, and other people and organisations we deal with through the PES cycle.

Public willingness to provide data and information is central to achieving our goals and is enabled by a high level of trust and confidence in the way we manage and secure this information. We are mindful of public expectations about privacy, security, and confidentiality to maintain and develop our social licence to collect PES responses, and to combine them with administrative data to provide meaningful outputs.

Privacy

Information privacy is the ability of an individual or group to keep their personal information to themselves. The Privacy Act 2020 promotes and protects individual privacy, establishing principles for the collection, use, and disclosure of personal information.

A comprehensive privacy impact assessment (Stats NZ, 2023c) is a key part of the 2023 Coverage project's focus on privacy for understanding and building social licence for PES activities. This assessment is required to inform decisions on data management, which in turn inform detailed design of processes and systems. As well as examining how response and administrative data is used in coverage measurement of the 2023 Census combined model, this assessment also prescribes the controls necessary to minimise privacy risk and to ensure we can build the social licence needed.

When researching and considering the use of administrative data in the 2023 Census Coverage project, we have been, and will continue to be guided by, the Ngā Tikanga Paihere framework (Stats NZ, 2020b) and the Five Safes framework (Stats NZ, 2020c) to ensure safe, responsible, and culturally appropriate use of data.

Security

Security is the mechanism we use to ensure the confidentiality, integrity, and availability of the systems used by the 2023 Census Coverage project, and the information those systems hold:

- confidentiality – reflects our capability to limit access to information and enables us to protect respondents' personal information, our administrative data, and other confidential information obtained and held by the project.
- integrity – ensures consistency, accuracy, and trustworthiness of data over its entire lifecycle, and as well as enabling privacy; it also enables our goal of delivering quality data
- availability – census coverage systems can be accessed and used whenever they are needed by those with approved access, meaning they can be used to conduct operations and enable our goal of timely release of quality data.

Confidentiality

Stats NZ's confidentiality and disclosure control best practice will be followed. Statistical confidentiality and disclosure control ensure that respondents' personal or household information are not divulged nor can be deduced from published statistics or from microdata. The purpose of this is to ensure that respondent and household information remain confidential, and the risk of disclosure is minimised.

Generally, this is achieved through a combination of:

- aggregating data to an extent that individual personal or household records cannot be identified or inferred
- adding random noise (perturbing) in certain data points in microdata and aggregate data

- suppressing certain data points in aggregate data
- de-identifying personal data and removing correlation identifiers
- controlling access to microdata for specific approved statistical and research processes.

Section 2: Lessons from 2018 PES

This section focuses on transitional work between the 2018 and 2023 PES cycles, and includes the following chapters:

- Improving the quality of the PES dwelling frame
- Recommendation from the 2018 PES independent review
- Coverage measurement of a combined model by design
- 2018 PES lessons incorporated into the 2023 Census design.

2.1 Improving the quality of the PES dwelling frame

A key focus area for improvement on 2018 PES in 2023 is the quality of the PES dwelling frame (Stats NZ, 2023a). The PES requires a frame of permanent private dwellings from which to select a proportion of dwellings for interview. To support the estimation assumption of independence, PES aims to create the sample frame in such a way that the chance of an individual or dwelling being found by PES is unrelated to the chance of that same individual or dwelling being found by census.

2018 PES experienced significant challenges in achieving a high-quality frame. Traditionally (1996–2013), PES used a combination of administrative address list data (obtained from New Zealand Post in some cycles, and Quotable Value Limited in others), desktop canvassing, and field canvassing to create its list of private dwellings, while census completed a scratch listing of all private dwellings – collated at the same time as provisioning everyone with their census forms. The benefit of this approach was that the two resulting lists of dwellings were inherently independent from each other. However, 2018 Census moved to a modernised model using administrative address list data (obtained from New Zealand Post and Land Information New Zealand), historic census addresses, and field canvassing to create the 2018 Census dwelling frame. This shift in census process presented 2018 PES with a dilemma – how to maintain independence from census without compromising on quality or cost.

2018 PES ended up using an externally provided address list (built from cadastral data provided by territorial authorities), followed by a dwelling enumeration operation. The key lesson from the 2018 PES experience was the unexpectedly high level of net under-coverage in the sourced address list (~20 percent). This proved particularly challenging for the field staff to completely resolve through dwelling enumeration, with the PES sampling frame assessed as having 8 percent net under-coverage of permanent private dwellings (based on comparison with census). This was significantly higher than that achieved in 2013 PES (less than 1 percent net under-coverage). This poor-quality frame led to dwelling coverage measurement for 2018 Census being de-scoped from 2018 PES and placed pressure on the estimation assumptions for person coverage estimation. While dwelling coverage measurement remains out of scope for 2023 PES, the quality of the dwelling frame is a key priority for supporting the estimation assumptions for person coverage estimation.

Following the experiences and lessons of 2018 PES, two key pieces of work were undertaken to refine the design of the dwelling frame to ensure improved quality for 2023 PES:

1. evaluation of available address list sources to support a dwelling enumeration process
2. a quasi-experimental field test to evaluate frame quality resulting from different dwelling enumeration processes and resourcing strategies.

Address list sources to support a dwelling enumeration process

Conceptually there are limited options available for creating a list of private dwellings in New Zealand for any Stats NZ survey (including PES). These options are:

1. create a complete list of private dwellings from scratch (without a starting address list) – **scratch frame**
2. use an externally provided list of addresses and refine through dwelling enumeration processes – **external list**
3. use an internally available list of addresses and refine through dwelling enumeration processes – **internal list**

Each of the options described above were evaluated using the following criteria:

- degree of independence from sources being used by 2023 Census
- quality of data and fitness for use
    - information available to select addresses associated with permanent private dwellings
    - coverage of addresses associated with permanent private dwellings
- cost to implement for both the field test (in year 2022) and PES proper (in year 2023).

This evaluation determined the internal list (address sourced from the Stats NZ Statistical Location Register (SLR)) as the preferred option as it provides the best opportunity for PES to achieve a high-quality dwelling frame. However, the 2023 Census address list will also be sourced from the same place so there will be some degree of dependence with census. This dependence risk can be mitigated by ensuring there is variation embedded in the PES address list selection (where possible, noting census will continue to refine their address selection processes), and through PES conducting an independent dwelling enumeration to clean the address list into a sample frame of private dwellings. The PES dwelling enumeration operations will be conducted by different staff to those working on 2023 Census – this is to prevent the same person working in the same area for both operations and thus finding the same dwellings for both operations.

For more details on the option comparison and evaluation, see Appendix 2.

Experimental test to evaluate dwelling enumeration approaches

In April 2021, the 2023 Census Coverage project undertook a small-scale dwelling enumeration field test to explore drivers of quality and cost to inform an optimal dwelling enumeration approach to support a high-quality dwelling frame for 2023 PES. The test aimed to explore a range of different dwelling listing processes, and the notion that dwellings exist on a spectrum of 'hard-to-find' to 'easy-to-find' (for more information see Appendix 3). The test also explored the relationship between staff experience and quality. The test was a quasi-experimental design as a full experimental design was not achievable within cost and time constraints.

From previous experience and international research, four dwelling enumeration approaches were identified for testing on cost and quality:

1. list-based field enumeration (baseline)
2. double field pass followed by conflict resolution
3. list-based desktop canvassing
4. scratch field enumeration.

The goal outcome from each of the dwelling enumeration approaches is the same: create a comprehensive list of all permanent dwellings (private and non-private) in a defined geographic area. For each dwelling, record address information, classification information about type of dwelling, along with any notes that would help field staff locate the dwelling for an interview.

List-based field enumeration achieves this goal through field activity: they walk the geographic area with a provided list of addresses. Field staff are encouraged to explore and make contact if they are unable to confidently classify a provided address and are trained to seek out and record missing dwellings.

The double field pass followed by conflict resolution is like the list-based approach, but more intensive to ensure higher quality. In this approach, two field staff members separately and independently complete the list-based enumeration for an area. The results of these two field activities are compared and any differences between the two lists are investigated and resolved to an agreed truth. This resolution process involves both staff who completed the initial activities, as well as a team leader.

List-based desktop canvassing works from a provided list again, but it is cleaned through desk-based investigation instead of field activity. Desktop staff are instructed to use a range of desk-based resources (such as satellite imagery, rates databases, internet searching) to classify addresses and record relevant information. They are also trained to search for and record any additional dwellings they discover in their sample area.

Scratch field enumeration is the most different of the approaches because it does not start with a list of addresses or dwelling records. Field staff start with only a list of all known streets in the area and build the list of dwellings from ‘scratch’. Scratch enumeration can be more time intensive because staff also need to record the address information for all records.

Table 3 presents the design of the field test. Key elements of the design include:

- independent dwelling enumeration using four different enumeration processes
- an embedded comparison of staff experience
- a sample of geographic areas selected to represent a range of situations and difficulty in finding dwellings
- consistent geographic areas used for all dwelling enumeration passes
- a consistent list of starting addresses used for all list-based dwelling enumeration approaches.

To analyse and evaluate the quality of the different enumeration approaches, a gold standard list for dwellings for the test areas was created by combining the detected dwellings from all collected lists and investigating any inconsistencies. Each of the enumeration approaches was then compared against the gold standard list to measure under-coverage and over-coverage of dwellings achieved by that approach. This data, along with information related to difficulty of finding dwellings in each of the areas, were used to model the relationships between dwelling difficulty, enumeration approach, and under- and over-coverage errors.

Table 3: 2023 PES dwelling enumeration test design

Difficulty to find dwellings	Starting address list	Enumeration process	Staff experience
Consistent set of geographic areas representing a broad range of hard-to-find and easy-to-find dwelling areas	Consistent list of addresses selected from the SLR	List-assisted field enumeration (baseline)	Stats NZ field interviewer
			Temporary contractor
		Double field pass & resolve	Both Stats NZ field interviewer and temporary contractor
		Desktop enumeration	Temporary contractor
	Base street records only	Scratch enumeration	Temporary contractor

The key outcomes of the test were:

- the double field pass and resolution approach achieved the highest quality with lowest number of under- and over-coverage errors, but it came at significant cost
- experienced staff achieved higher quality than temporary contract staff
    - experienced staff were particularly valuable in the more difficult areas with complex situations such as multiple dwellings at one address, or in non-residential areas where dwellings are sparce
- list-based desktop enumeration and scratch field enumeration had the highest error rates
    - desktop has value for simple residential areas, and some rural areas (where there is recent satellite imagery and / or Street View).

The outcomes above, along with detailed cost analysis, were used to inform the dwelling enumeration approach for 2023 PES (documented in Section 3: Components of the 2023 PES standard design).

2.2 Recommendations from the 2018 PES independent review

Prior to the publication of the 2018 PES results and methodology, Stats NZ engaged two external and independent reviewers to quality assure the estimation methodology, Bayesian coverage models, and the soundness of the outputs. James Brown (University of Technology Sydney) focused his review on the estimation methodology, while John Bryant (Bayesian Demography Limited) focused his review on the construction and implementation of the Bayesian coverage models. Both reviewers joined key Stats NZ staff for a discussion of the methodology. From the review and discussion, two key areas recommended for exploration ahead of 2023 PES were:

1. extending the estimation framework to account for local area (or sub-population) coverage error (Brown, 2020)

2. developing an uncertainty framework to better reflect and communicate the subjective elements of the estimation framework and the possibility of having multiple plausible sets of assumptions that produce different results (surfaced during discussion).

Inconsistent reporting of population characteristics

Brown (2020) noted the 2018 PES estimation is based around a PES respondent's usual residence location, but the key variable in determining coverage (was a PES person counted in census or not) is determined at a national level. Brown asked, "what if an individual's identified usual residence private dwelling location in census is different" (Brown 2020, p5) to that recorded by PES? Brown noted the potential for over- and under-coverage at the local geographic area if the distribution of inconsistent PES and census locations do not balance out. Brown recommended Stats NZ consider extending the 2018 PES framework to allow for a localised definition of coverage.

Following this recommendation, the 2023 Census Coverage project investigated differences in reporting of usual residence location. The project extended this work to also assess differences in reporting of ethnicity. This was driven by the importance of accurate ethnicity counts in the New Zealand population system, a commitment to honouring Te Tiriti o Waitangi, and the increased likelihood of differences in reporting resulting from the extended use of admin data in the combined census model. In addition to assessing the magnitude of reporting differences evident in data, the project team also investigated a range of methodological solutions for incorporating an adjustment into the broader coverage estimation framework (see Appendix 4).

The preferred methodological approach was to develop a Bayesian model to model differences in reporting separately from the under- and over-coverage models.

2018 PES treatment of inconsistent reporting

In the 2018 PES estimation methodology, we use both the characteristics reported in the PES and the characteristics reported in the census when producing coverage-adjusted population estimates for small geographies and subpopulations of interest. The characteristics reported in the PES are used in estimating the under- and over-coverage probabilities for various combinations of characteristics. Those coverage probabilities are then combined to form a coverage adjustment ratio (for each combination of characteristics). The coverage adjustment ratio is then applied to the relevant count.

For example, when producing a coverage-adjusted population estimate for people usually resident in the Canterbury region, we use all PES records who report residing in Canterbury and their coverage indicators (were they counted in the census file – regardless of where they were counted, and how many times) to determine the coverage probabilities for Canterbury. The coverage adjustment ratio is then applied to the count of all census records usually living in Canterbury.

This approach makes four key assumptions:

1. the census correctly records the characteristics for individuals it captures correctly (without coverage error)
2. the PES correctly records the characteristics for individuals missed by the census, or overcounted by the census

3. any inconsistencies in reporting of characteristics between the PES and the census are independent of coverage error
4. any inconsistencies in reporting of characteristics represent natural variation and not systematic biases.

The first three of the assumptions above are challenging to explicitly assess so we focus our analysis on the fourth. We investigate natural variation and systematic bias by analysing the magnitude and direction of gross differences between reporting in PES and census, along with the variation in gross and net differences across subgroups.

Differences in reporting of usual residence location

Differences in reporting of usual residence location was investigated at two geographic levels for which coverage estimates are produced: regional council (16 areas across New Zealand excluding 'area outside region') and territorial authorities and Auckland local boards ((TALB) 87 areas across New Zealand excluding 'area outside territorial authority', 85 of which were present in the 2018 PES sample). Analysis was conducted using 2018 PES data with the total PES dataset subset to those records that had linked only once to a 2018 Census record.[1] The analysis dataset contained 30,360 PES records.[2]

The analysis approach included:

- magnitude and direction of gross reporting differences at a PES record level for each combination of inconsistently reported geographies, for instance:
  - count of PES records with census geography recorded as Canterbury and PES geography recorded as Wellington – PES geography recorded as Canterbury and census geography recorded as Wellington
- range of net difference in counts of records for each geography
  - count of records using census geography compared with using PES geography
- range of net difference as a percentage of total PES records for each geography.

Analysis results for regional councils:

- 333 of the 30,360 records (1.1 percent) report a different PES regional council than reported in census
- the gross difference ranged from -9 to 21, with a median and lower quartile of 0, and upper quartile of 0.25
- the net count difference ranged from -36 up to 30, with a median of 1
- the net difference as a percentage of PES count ranged from -1.9 percent up to 2.2 percent, with a median of 0.1 percent.

Analysis results for territorial authorities and Auckland local boards:

- 840 of the 30,360 records (2.8 percent) report a different PES TALB than reported in census
- the gross difference ranged from -9 to 12, with a median, lower, and upper quartiles all 0
- the aggregate count difference ranged from -24 up to 22, with a median of 0

---

[1] Records with multiple links were excluded to ensure analysis results reflect only differences in reporting and are not clouded by over-coverage.

[2] All counts have had random rounding to base 3 (RR3) applied for confidentiality.

- the net difference as a percentage of PES count ranged from -6.3 percent up to 14.9 percent, with a median of 0.1 percent.

At both levels of geography, there is evidence of differences in reporting. For most areas the differences suggest natural variation in reporting rather than systematic patterns. This is supported by the median of gross differences being 0 for both regional councils and TALBs. There is no evidence to suggest that PES has any systematic bias in the capturing of usual residence geography. This is further supported by the aggregate count differences being small, and the net differences for each geographic area even smaller. The investigation conclusion was to not proceed with development of a Bayesian model to adjust for differences in reporting of location. This was driven by the small magnitude of observed differences, as well as the limited data available to support the model.

Differences in reporting of ethnicity

Differences in reporting of ethnicity was investigated at level 1 of the Stats NZ Ethnicity classification (Stats NZ, 2005). The analytical approach was modified to reflect the multiple response nature of the ethnicity concept. Of the two standard approaches to presenting ethnicity information in New Zealand (total response, and combined categories), this analysis used total response output in which individuals are counted in all their reported ethnicity groups.

The Statistical Standard for Ethnicity defines ethnicity as:

"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." (Stats NZ, 2020d)

Two different base sets of PES records were used in analysis: all PES records with no coverage error (containing 30,360 records [3]), and a narrower set of PES records where records with any ethnicity imputation (either by PES or census) were excluded (26,670 records). Analysis results for level 1 ethnicity[4] are summarised below.

Observed differences:

- when records with any ethnicity imputation are excluded, 2,766 of 26,670 records (10.4 percent) report at least one difference ethnicity in PES than reported in census
- when records with ethnicity imputation are included, the observed difference in reporting increases to 11.4 percent (3,447 of 30,360 records)
- overall, more ethnicities are reported by census than PES, and this is seen in the four major ethnicity groups (Māori, Pacific, Asian, and European)
- PES reported higher rates of Middle Eastern, Latin American, and African (MELAA), and Other ethnicity.

Gross differences and directions (based on only those PES record with inconsistent ethnic group indicators):

[3] All counts have had RR3 applied for confidentiality.

[4] Level 1 ethnicity consists of six categories (excluding residual): European, Māori, Pacific Peoples, Asian, Middle Eastern / Latin American / African, Other ethnicity.

- for every level 1 ethnicity group there were bi-directional differences in reporting suggesting natural variation in reporting rather than systematic biases
- for all level 1 ethnicity groups except MELAA, PES reporting was lower than census reporting:
    - for example, for the 813 PES records with inconsistent reporting of the Māori ethnicity indicator, 43 percent of those were records who reported Māori ethnicity in the PES and not in census, compared with 57 percent with the reverse reporting pattern
    - this is consistent with overall more ethnicities being reported in census than PES.

Net count and percent differences:

- the net difference in count of PES records with Māori ethnicity as recorded by PES compared with census was -90 (-2.2 percent)
- the net difference in count of PES records with Pacific ethnic group as recorded by PES compared with census was -15 (-0.7 percent)
- the net difference in count of PES records with Asian ethnic group as recorded by PES compared with census was -54 (-1.4 percent)
- the net difference in count of PES records with MELAA ethnic group as recorded by PES compared with census was 21 (5.5 percent)
- there were much larger net differences for European (-711, -3.9 percent) and Other ethnicity (579, 63.6 percent), however, we suspect these differences may reflect inconsistencies in the coding of people reporting ‘New Zealander’ for ethnicity
- the net difference in count of PES records with the broader grouping of European, MELAA, or Other ethnicity (used in PES estimation) was -111 (-0.6 percent).

Scenarios in table 4 are defined according to reporting ethnicity in PES. For instance, if a record in PES has more than one ethnicity and has the same ethnicities reported in census, it is categorised as ‘multiple ethnicities, consistent’. If a record in PES has more than one ethnicity and has completely different ethnicities reported in census, it is categorised as ‘multiple ethnicities inconsistent’. If a record in PES has more than one ethnicity and at least one ethnicity reported in census is the same, it is categorised as ‘multiple ethnicities mixed consistent’.

Table 4 shows for PES 2018:

- almost 81 percent of records reported same single ethnicity in PES and census
- almost 7 percent of records reported same multiple ethnicities in PES and census.

Table 4: Differences in reporting of ethnicity between PES and census

Scenarios	Percentage
Single ethnicity, consistent	80.82
Multiple ethnicities, consistent	6.92
Single ethnicity, inconsistent	3.78
Multiple ethnicities, inconsistent	0.07

Single ethnicity, mixed consistency	3.78
Multiple ethnicities, mixed consistency	4.64

While there is evidence of PES having lower net counts for each of the ethnicity groupings used in estimation modelling, there is no evidence to suggest this is driven by systematic biases in PES collection rather than natural variation in reporting. We are confident that our estimation assumptions hold and that we can continue to use the existing estimation approach for 2023 PES.

Measurement and communication of uncertainty due to estimation assumptions

The second key recommendation to come out of the 2018 PES independent review process was for 2023 PES to develop a more comprehensive uncertainty framework than the existing credible intervals which primarily communicate uncertainty due to sampling error.

Challenges encountered during the 2018 PES estimation phase highlighted the level of subjectivity involved in defining the estimation methodology. As outlined in section 1, there are five critical assumptions that underpin the DSE approach to coverage estimation. Several of these assumptions can come under pressure at various stages of the census and PES taking process. Designing a robust estimation process involves evaluating the validity of each of the assumptions, and the level and direction of bias that can be introduced if an assumption is weak or compromised. The art of estimation design is balancing these assumptions and biases in such a way that the results produced are fit for purpose. The 2018 PES reviewers acknowledged this element of PES estimation design, suggesting that in the process of refining 2018 PES, we defined multiple plausible approaches – each with subtle variations in balancing the assumptions and biases. The reviewers recommended developing a framework for capturing and communicating the uncertainty and subjectivity contained within those multiple plausible approaches.

This development is not within scope for the present standard design but is a priority for the adaptive design (see Section 4: Adaptive design).

2.3 Coverage measurement of a combined census model by design

2018 Census collection response rates were lower than expected, resulting in a significant shift in New Zealand’s approach to census taking. For the first time, New Zealand shifted from a traditional full field enumeration census (supplemented with ‘substitute’ records) to a combined model census in which records from an admin population were added to the 2018 Census dataset to mitigate non-response (Stats NZ, 2019a). Admin enumerations were selected for inclusion in the 2018 Census dataset in three ways:

- as usual residents in a private dwelling when sufficient household membership information is available in the admin data
- as people present in a non-private dwelling on census night (prisons and defence establishments)
- as usual residents in a small geographic area (meshblock) when there was sufficient evidence to support this would improve the quality of count information for that area.

The shift in census model poses several challenges for PES and its estimation assumptions. The 2018 PES methodology included some modifications in response to the combined census model such as special treatment of admin enumeration in prisons and defence establishments. Part of the focus for the 2023 PES is to give dedicated time and research into the implications of the combined model, allowing PES to respond by design instead of reactively as was the case in 2018. This section outlines research undertaken in response to the combined census model, specifically:

- coverage estimation of people usually resident in non-private dwellings
- increased potential for over-coverage due to false negative links between census and the admin population
- minimising erroneous inclusions in the admin population (work leveraged from the Census Transformation programme).

Coverage estimation of people usually resident in non-private dwellings

The population of interest for census coverage measurement are all people who usually reside in New Zealand and were present in New Zealand on census night. These people may reside in temporary or permanent accommodations, in private or non-private dwellings (NPDs). As a household survey, the PES samples permanent private dwellings and the individuals within. To achieve coverage measurement of the full population of interest, PES uses private dwelling responses to predict census-response patterns of the total population. Therefore, PES methodology assumes the census coverage probability does not differ by dwelling type and assumes the coverage patterns for people residing in NPDs and temporary accommodation are the same as those measured for people residing in private dwellings. Internally, it has long been understood that this assumption is unlikely to strongly hold, but short of PES sampling from all dwelling types (deemed operationally impractical), it is necessary for producing fit for purpose population estimates.

The shift to a combined census model in the 2018 Census led to refinements of the assumption outlined above (let us refer to it as the NPD assumption). The introduction of admin enumerations into particular types of NPDs (prisons and defence establishments) systematically shifted the assumed coverage patterns for those NPDs. Both prisons and defence establishments had close to 100 percent coverage in the 2018 Census dataset, which clearly invalidates the NPD assumption in those cases (Stats NZ, 2020e). The 2018 PES adjusted the estimation methodology to account for this – people residing in those NPDs were not eligible for a coverage adjustment.

As part of preparing for a 2023 Census, combined census model by design, the 2023 Census coverage project embarked on dedicated research to reassess the NPD assumption and identify any further areas for improvement. The key components of the research were:

- data exploration of population distributions by NPD type in 2006, 2013, and 2018 Census datasets
- canvass and assess admin data sources related to NPDs and their residents
- use combined census and admin data to assess coverage patterns in 2018 Census.

The key findings from the research were:

- residents of NPDs comprise a small but highly variable part of the New Zealand population

- the number of usual residents in NPDs varies considerably between NPD types
- educational institutions and aged care facilities represent almost 60 percent of people usually residing in NPDs
- there are currently no admin data sources directly related to residents of NPDs available in the Stats NZ IDI:
    - there are several sources with census night occupancy information (for example, hospitals, prisons, residential care facilities, and some night shelters)
    - potential for this data to be used for usual residence with development of residency rules based on duration of stay
- identification of people residing in NPDs can be achieved through the IDI address information (and linking to NPD addresses):
    - there is currently a lot of noise in the data under this approach suggesting not all the NPD -> address -> person relationships are of high quality
    - development work is needed to develop a prediction model to evaluate and identify high quality address information.

For 2023 PES, the key outcome of this research was that there is not currently any information readily available that can be used to refine and improve the PES estimation approach and assumptions. However, there were several promising avenues that are being progressed by the Census Transformation programme for use in the future.

Detection and measurement of coverage error caused by false negative links between census responses and the admin population

The move to a combined census model has resulted in the possibility of linkage error in the census to admin population linking. False negative linkage error can result in duplicate records being included in the census dataset, creating over-coverage. PES aims to detect this over-coverage through linking a PES respondent to all matching records in the census dataset. If PES successfully links to both records of the duplicate pair, then they will go into the over-coverage model and be accounted for. If PES only links to one of the duplicate census records, then they will appear as having no coverage error and the duplicate record will not be accounted for in coverage estimation. Ahead of 2023 PES and coverage measurement of a combined census model by design, the 2023 Census Coverage project undertook research to confirm if the planned PES processes are sufficient to detect and account for duplicates resulting from false negative linkage error.

Historically, PES tried to link all PES and census records regardless of the quality of linking information. From the linking information we would determine if a PES record was or was not counted in the census and how many times they were counted (a person 'was' could be 0, 1, 2, or more). If a PES record did not have sufficient linking information for us to be confident if they were or were not counted in the census, we would impute their 'was' variable. During the design phase for 2018 PES, we proposed a change for how we would treat records with insufficient information to link with confidence: rather than attempting linkage and then imputing an outcome, we would exclude records with poor linking entirely and rely on estimation processes to account for them. It was anticipated that this would reduce bias in the final estimation of under-coverage in the census response file. This improvement to the measurement of under-coverage came with the trade-off of decreased ability to detect duplicates in the census file where one or more of the census responses had low quality linking information. This was a reasonable trade-off for the planned census design which typically resulted in low amounts of over-coverage and duplicates. However, the change in the census design from a full-field enumeration to a combined model may have had unexpected consequences when combined with the linkage acceptance criteria.

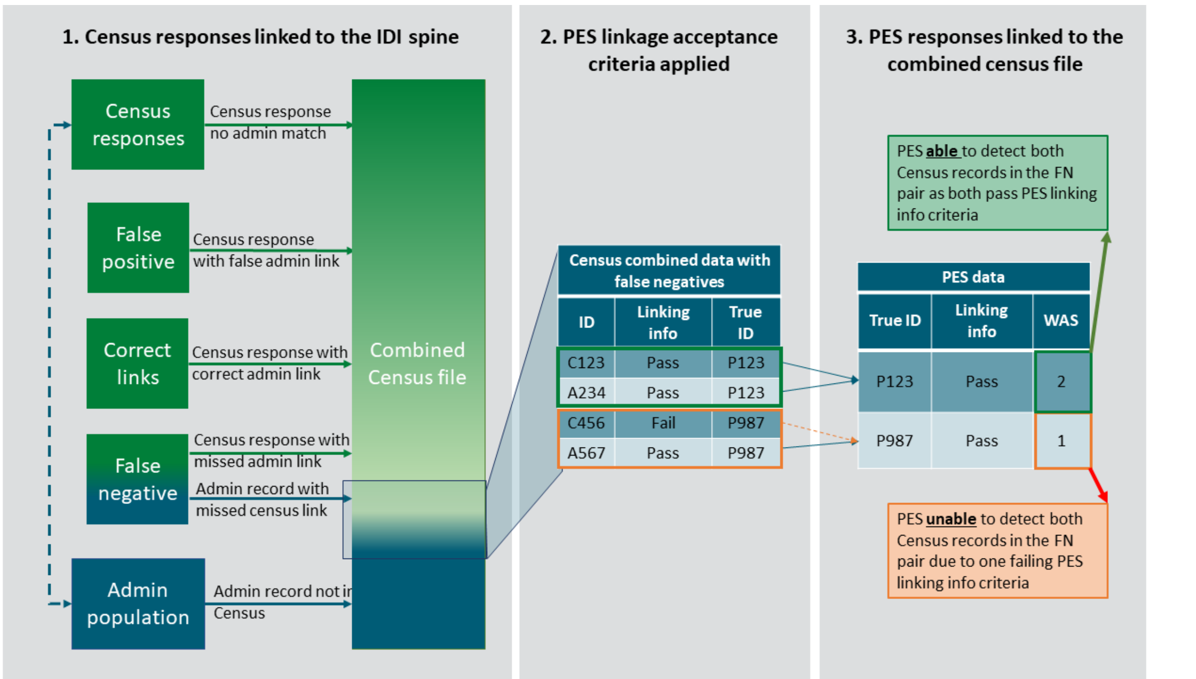
The primary focus of this research was to:

- investigate the relationship between false negative links in the census dataset and the PES linkage acceptance criteria
- evaluate and estimate the potential for bias in the PES results from reduced ability to detect duplicate pairs.

Figure 2 demonstrates how unidentified linkage error within the combined census file may interact with the linkage acceptance criteria. The magnified blue box is a snapshot of some of the duplicates caused by false negative linkage error between census responses and admin enumerations. Records C123 and A234 both refer to the same individual record: P123. Both records have enough linking information to pass the linkage acceptance criteria. This means that both advance to the step where they can be linked with PES. Consequently, both duplicate records are identified and accounted for in the final estimates.

Meanwhile, records C456 and A567 both refer to the same individual record: P987. Record C456 does not have enough linking information to pass the linkage acceptance criteria. Therefore, it is removed from the linking file and is not available to link with PES. Record A567 has enough linking information to pass the linkage acceptance criteria and is linked to the spine. Because C456 was not linked with the PES file, we are unable to recognise that there were duplicate records in this case.

Figure 2: Flowchart of interaction between false negative linkage error and linkage acceptance criteria



To support analysis of the relationship between false negative duplicates and the linkage acceptance criteria – we needed a known set of duplicates caused by false negative linkage error in the census responses to admin population link. This group of duplicates was derived using a method described in (Stats NZ, 2019b) which uses an indicator, M*, to show that a census record

should be in the IDI spine. An M* value of 1 indicates that a census response provides information that demonstrates that it should have a true match in the IDI spine. The proportion of the census records who were in the M* subset and did not link to the IDI spine were used to estimate a false negative linkage rate of 1.2 percent in the 2018 Census (Stats NZ, 2019b).

This M* indicator can be cross tabulated with a linkage acceptance indicator to assess any relationship between the two.

Table 5 presents the summary counts by the two indicators. A Fisher exact test (Fisher, 1922) was used to evaluate if there was a statistically significant relationship between the two indicators. The Fisher test resulted in an odds ratio of 359 (p-value < 0.001) indicating a significant relationship.

Table 5: Relationship between whether a record is in M* and if it passes the linkage acceptance criteria

	LAC Pass	LAC Fail
M*	3,311,901	27
Non-M*	860,454	2,520

While there is a significant relationship, the very low number of records that failed the linkage acceptance criteria (2,547 records failed compared with 4,172,355 passing) suggests that this relationship is only a small contributor of bias in the coverage estimates for 2018 PES. It may not be valid to assume the same low number of records to fail the linkage acceptance criteria in 2023 as observed in 2018, particularly with planned increased use of paper forms in 2023 which are known to increase risk of false negative links. Sensitivity testing was conducted exploring the proportion of census responses received from paper forms and rate of undetected false negatives due to the linkage acceptance criteria, and the resulting bias in PES estimation (see Appendix 5 for detailed information). The testing results indicated negligible bias resulting from undetected false negative duplicates even if 100 percent of census responses are received via paper (assuming similar false negative rates as observed in 2018).

We also investigated whether false negative linkage error within the census dataset disproportionately affected certain ethnicity groups more than others. We found that records with European ethnicity, Māori ethnicity, and Pacific ethnic group were each under-represented in the false negative cohort in comparison with their representation in all census responses. Meanwhile, records with Asian ethnicity, MELAA ethnicity, and Other ethnicity were over-represented. However, as the linkage acceptance criteria failure rate is so low, this false negative error is expected to be detected and accounted for in the current methodology. Improvements to linkage design and processes are expected to reduce this impact further.

Minimising erroneous inclusions in the admin population

The admin resident population used in 2018 Census is constructed using the IDI-ERP (an estimated resident population (ERP) derived from the IDI). The required IDI-ERP was derived for census night, 6 March 2018, and subset to residents present in New Zealand on census night (Stats NZ, 2019a). The IDI-ERP is known to have some coverage errors – both over- and under-coverage. The 2018 Census used the IDI-ERP as the base-population for admin enumerations for inclusion in dwellings assuming negligible over-coverage in these records.

However, the same assumption is not made for admin enumerations into meshblocks – the IDI-ERP is adjusted to remove over-coverage caused by people selected in the IDI-ERP who are not

New Zealand residents at the time of the census (Stats NZ, 2019a). The resulting admin population is called the 'IDI-ERP_Sure' to reflect that it is designed to only contain people we are 'sure' belong to the New Zealand resident population. The IDI-ERP_Sure is created using stricter activity rules than those used to create the IDI-ERP – people are required to have at least two activity sources compared with one, with at least one of these sources being either tax or health. This approach aims to remove the over-coverage while minimising the number of valid records also being removed.

The IDI-ERP_Sure approach is an approximate approach to identifying and removing over-coverage since it is not possible to directly find and exclude the over-coverage records. This has the potential to be problematic for PES as it places pressure on the assumption of no erroneous inclusions. PES would aim to detect erroneous inclusions (such as residents temporarily overseas, or babies born after census night) through interview, and then account for that type of over-coverage in estimation. The difficulty with erroneous inclusions in an admin population (in this context) is that most of the coverage uncertainty relates to the migration status of individuals on census day (Stats NZ, 2019a).

Uncertainty in migration status can create over-coverage in an admin resident population due to time lags in admin data and differences between the official and operational approaches to classifying border crossings as short-term or long-term. The official measure of international migration requires 16 months of data following the date of border crossing into or out of New Zealand. This measure (using the 12/16 rule) defines a migrant arrival as an overseas resident who has spent at least 12 of the 16 months following the initial date of arrival in New Zealand (Stats NZ, 2018a). The IDI-ERP used by 2018 Census approximates the 12/16 rule using a 6/12-month rule which defines a resident as someone who has spent at least 6 of the 12 months spanning the relevant reference date in New Zealand (Stats NZ, 2019a). The 6/12-month rule is applied solely to the 12-month window and is not conditional on their resident status at time of arrival or departure. Over-coverage can occur if a person was in New Zealand on census night, they pass the operational measure, but once more admin data is received, they do not pass the official measure indicating they should not have been classified as a resident. This type of over-coverage is not able to be measured directly by the PES since it requires a full 16 months of admin data from the reference date (census night).

The Census Transformation programme has continued to refine the methods used to derive an admin resident population for New Zealand, with a series of improvements included in the 2021 release of an experimental administrative population census (APC) (Stats NZ, 2021d). In particular, these relate to moving closer to the 12/16 month used in the official measure where this information is available, and continuing to use the 6/12 rule when more timely data is needed. Two of the other improvements relate to over-coverage:

- "Removal of ACC (Accident Compensation Corporation) as an activity source. An investigation of the quality of activity source indicators found that ACC was less reliable than other sources. Most ACC activity is associated with activity within the health system. Around 10,000 individuals who were found to be active solely due to ACC claims were removed from the admin resident population." (Stats NZ, 2021d, p. 15)
- "Duplicate records are removed. This results in around 20,000 records removed from the admin resident population because there were clear indications of two records for the same person." (Stats NZ, 2021d, p. 16)

2.4 2018 PES lessons incorporated into the 2023 Census design

Part of the value of a PES is that it can be used to inform and guide improvements for future censuses (Stats NZ, 2021a). Measures of coverage and response rates can inform field collection design, prioritisation algorithms, and non-response follow up activities. Analysis of over-coverage patterns can inform processing methods and non-response mitigation methods. Following the introduction of the combined census model in 2018 Census, 2018 PES identified three areas for potential improvements to 2018 Census processing methods and non-response mitigation methods:

1. census definition of 'what is a response'
2. census treatment of residents temporarily overseas on census night
3. census approach to linking census responses with an admin population.

Definition of 'what is a response'

The 2018 Census introduced a set of minimum information requirements for a census form return to be considered a response and included in the census dataset (Stats NZ, 2019a). The information requirements were designed around having enough evidence that a unique person is being counted. The information requirements used in 2018 were "valid values for two out of three of the fields: name, date of birth, and meshblock" (Stats NZ, 2019a, p. 9).

This response definition is conceptually and operationally similar to the PES concept of linkage acceptance criteria which is designed around having enough quality information to be confident in achieving a correct link in another data source (should a matching record exist). The 2018 PES linkage acceptance criteria required all records to have a valid name (with at least two characters for each first and last name) along with either high quality location information or date of birth, and at least one other piece of information such as country of birth. These criteria were initially set slightly broader (name or date of birth required when high quality location information was available) but were tightened during 2018 PES linking to reduce the number of uncertain link outcomes occurring.

Comparing the PES rules with those used by census, there are two key components making PES stricter:

- PES requires all records to have valid name information
- PES requires an additional piece of information beyond name and date of birth when location is not known.

The 2023 Census Coverage project evaluated the 2018 Census response definition to see if there would be any value to PES and census in tightening the response definition.

Here are the key findings from this evaluation.

- A number of facetious and non-valid responses were found in the 2018 Census dataset. If these had been detected and facetious information removed, the records may not have met the response definition.
    - 2023 Census have agreed an action to refine and improve processes for identifying and resolving records with invalid name and date of birth information.

- There is evidence that the response definition fields are not sufficient to prevent false negative links occurring in the census to admin link which can impact on coverage estimation.
    - Consideration was given to how the response definition might be tightened without being detrimental to the census data.
    - Most options for tightening the definition (such as requiring date of birth, or country of birth if born outside of New Zealand) risk excluding respondents who have not completed an individual form (individuals on the household listing only).
    - Respondents who have not completed an individual form are not a representative subset of the total population, and tend to be similar types of people as those that get missed by census altogether.
    - They are also critical contributors to the creation of families and households.
    - Excluding these individuals would have significant negative impact on the quality and equity of census data, which outweighs the potential benefit to PES.

Removal of residents temporarily overseas from the census count

Residents temporarily overseas (RTOs) on census night who are mistakenly counted as present in New Zealand on census night are a known source of over-coverage (Stats NZ, 2020e). Since the first PES in 1996, the PES questionnaire has aimed to detect and measure this source of coverage error. For the first time, 2018 PES used administrative border crossing data (dates of departure from New Zealand and dates of arrival into New Zealand) to identify RTOs (Stats NZ, 2020e). This admin data combined with the census file enabled PES to directly observe RTO over-coverage in the census count.

The 2018 Census also used border crossing data as part of the combined census model, specifically using border crossings to identify people not in New Zealand on census night and exclude them from being eligible for inclusion in the census file as admin enumerations (Stats NZ, 2019a). The 2023 Census will extend this use of border crossing data to identify census responses from RTOs and to exclude them from the census usually resident and census night population counts, removing them as a source of over-coverage. Note, these responses will remain available for use in the creation of families and households.

Approach to linking census responses with the admin population

The 2018 Census responses and admin enumerations were combined into the overall census file using a series of automated linking passes (Stats NZ, 2019c). The conservative automated linking process was designed to minimise false positive linkage resulting in a rate of 0.6 percent +/- 0.3 percent in 2018. To ensure unidentified duplicates did not result in inflated census counts, a statistical adjustment was included in the admin enumeration process to reduce the number of records eligible for inclusion in the census file.

It was estimated that the false negative rate in the 2018 census file was 1.21 percent (41,500 records). This was determined through use of the M* indicator (Stats NZ, 2019b). M* is a binary indicator applied to a census record that provides information that enables us to be confident that it will appear in the IDI spine. A record with an M* indicator of 1 that did not link to the spine in the automated linking process is assumed to be a false negative. This estimation hinges on the assumptions that records that meet M* conditions are in the spine, and that records that do not appear in M* will have the same rate of missed matches as those in M* (Stats NZ, 2019b).

The 2018 Census automated linkage process used the Fellegi-Sunter method to compare census responses with the administrative data spine (Stats NZ, 2019c). In this linking method all records from one dataset are compared to records from another dataset on each linking variable. A total weight that summarises the comparison results of the linking variables is given to each pair of records. This weight reflects the likelihood that two records refer to the same person. Thresholds were chosen to determine what linking weights were considered to indicate a high-quality match or non-match.

These linkage weight thresholds were designed to be conservative to minimise the number of false positives between census responses and administrative data enumerations. This reduces the amount of under-coverage within the combined census file. However, the PES is better designed to be able to account for under-coverage from false positives than duplicates in the file caused by false negatives. Adjusting the linkage thresholds within the linkage process to allow for a more even balance between the false positive and false negative error rates would improve the capability of the 2023 Census Coverage project to provide accurate coverage estimates.

The 2023 Census Coverage project investigation into the 2018 Census false negatives found that certain demographics were over-represented. Children under the age of five appeared in the false negatives cohort at thrice the rate that they do in census responses, while paper forms appeared at twice the rate. In contrast, census respondents born in New Zealand, of European ethnicity, or those who filled out online forms were under-represented. These results have been shared with the 2023 Census team resulting in an agreement taken by census to evaluate the balance between false positive and false negative errors in the 2023 Census linking design, and if possible, to take steps to reduce the false negative rate while maintaining an acceptable false positive rate.

Further investigation into the missing links for young children indicated by the M* methodology indicated the M* assumption (records that meet M* conditions are in the spine) does not hold for this group. There can be a considerable lag for some young child migrants being added to the spine. For 2023 the rules for creating the M* subset will need to take this into account.

Section 3: Components of the 2023 PES standard design

This section provides detailed descriptions of all key design components in the planned standard design for coverage estimation of the 2023 Census via a 2023 PES. Chapters included in this section are:

- Producing high-quality population estimates for Māori
- 2023 PES scope and population definitions
- 2023 PES survey objectives and quality targets
- 2023 PES questionnaire content
- 2023 PES dwelling enumeration approach
- 2023 PES sample design
- 2023 PES linking
- 2023 PES treatment of missing data
- 2023 PES estimation framework.

3.01 Producing high-quality population estimates for Māori

As outlined in earlier sections of this document, a core purpose of measuring census coverage is to improve the accuracy of the base population estimates, which in turn leads to more accurate post-censal estimates and projections. These improvements are particularly important because census coverage is not uniformly distributed. As the ethnic and geographic diversity of the New Zealand population increases, the importance of measuring and accounting for the differences in undercount also increases.

Also core to the 2023 Census Coverage project is a commitment to meeting our responsibilities, as Te Tiriti partners, to Māori. As described in Overview of PES approach to working for and with Māori, the 2023 Census Coverage project has set several objectives to uphold this commitment. The objective of interest here is the statistical steps we take to ensure robust estimation of the Māori population (both Māori ethnicity and/or descent) and the achieved coverage and response rates by census. The following paragraphs provide high-level information about how each statistical design component contributes to the robust estimation of Māori populations. Detailed information about each of the statistical components is covered in subsequent sections.

In addition to the above, this section also includes a specific discussion on the production of coverage-adjusted population estimates for Māori descent. This discussion is based in the tension of having two Māori descent concepts and what that means for PES processes and outputs.

Scope, survey objectives, and quality targets

The PES scope and population definitions aim to balance practical and operational considerations with robust statistical design. In the 2023 PES design process, there is a strong focus on producing robust estimates for Māori. For the most part, the scope and population definitions are carried over from 2018 PES. Two areas of specific review and consideration were the relationship between the PES dwelling definitions and emergency and transitional housing, and the statistical assumptions relating people living in private dwellings with those living in non-private dwellings.

The 2023 PES has an intensified focus given to improving detection and sample representation of dwelling establishments that provide emergency and transitional housing to the public. Information from the Ministry of Housing and Urban Development (Ministry of Housing and Urban Development, 2020) shows households of Māori ethnicity make up 56 percent of the total number of households accepted into the Housing First programme as at 30 June 2020 (since the programme was first launched in March 2017). Improving PES detection and representation of emergency and transitional housing establishments will improve coverage and population estimates for Māori.

The PES population definitions rely on an implicit assumption that the dwellings and individuals included in scope for the PES are representative of the total population and sufficient to reflect the variation in coverage patterns achieved by census. We know that this assumption does not hold for all types of non-private dwellings and have investigated ways in which we might alter or strengthen this assumption (see Coverage estimation of people usually resident in non-private dwellings). The outcome of this research showed there is currently no information readily available that can be used to refine and improve the PES estimation approach and assumptions.

The 2023 PES survey objectives are set in a manner so that they support the requirements and uses of PES data by both the 2023 Census, and the 2023 base estimated resident population (ERP). Both the 2023 Census and 2023 base ERP require estimates of net undercount of the population identifying with Māori ethnicity, and with Māori descent.

The 2023 PES quality targets include a subset of key performance indicators (KPIs) for the precision of coverage estimates by ethnicity and by Māori descent (table 6). Precision is measured using Bayesian credible intervals (CI). The inclusion of these KPIs signify the importance and commitment of the project to aim for robust, precise, and unbiased measures of the Māori population.

Table 6: 2023 PES precision targets directly related to Māori

Category	**2023 proposed target CI** (max percent +/-)	**2018 achieved CI** (max percent)
Māori		
• Māori ethnicity (Level 1)	**1.0**	-0.71 – 0.81
• Māori descent (electoral)	**1.0**	-0.63 – 0.67

Questionnaire content

The PES is a short questionnaire with content focused on collecting key demographics, information to support linking PES respondents with census records, and statistical scoping questions (used to determine eligibility for inclusion in estimation). Questionnaire topics and format are designed to align with best practice and published statistical standards. The statistical standards for concepts such as ethnicity and Māori descent undergo considerable consultation and regular review (Stats NZ, 2020f).

Key changes to 2023 PES questionnaire content that are relevant for Māori and iwi include:

- a change in the format of the Māori descent question to better align with the statistical standard and census concept
- the inclusion of iwi affiliation.

Iwi affiliation has been added to the PES questionnaire to align with a Stats NZ in-principle decision to collect iwi affiliation in all Stats NZ's household surveys. The (high-level) rationale is twofold: ensuring Māori/iwi better see themselves in our questionnaires and with a view to building a pool of iwi affiliation data over time and survey vehicles, to enable greater insight for Māori/iwi (especially in the absence of a reliable administrative data source for this variable). Iwi affiliation data collected by the PES will not be used in coverage estimation or for output.

Dwelling enumeration

The design of the dwelling enumeration approach for 2023 PES has focused on applying lessons learnt from 2018 PES and making changes to ensure significant quality improvements. This focus has been broad and aimed at resolving a global quality issue impacting all of PES. The improvements made to the dwelling enumeration approach are designed to produce a high-quality sample frame, which in turn supports PES producing robust statistics for all population groups.

Sample design

The overarching goal of the 2023 PES sample design is to balance the statistical outcomes (such as accuracy measures, and relationship with statistical assumptions) with operational and other considerations (such as cost, sample size, and respondent burden). Producing accurate and robust population estimates for Māori is a key consideration for the sample design.

Sample design process

The sample design process incorporates delivery for Māori in the following ways:

- basing the data used in simulations on historic coverage patterns which are known to vary by demographics such as Māori descent and ethnicity
- stratifying the population implicitly with variables that are correlated with coverage patterns and priority population groups to enable implicit over-sampling
- including accuracy measures of Māori ethnicity and descent populations in the design evaluation criteria and processes.

By using stratification variables such as the census enablement strategy and socio-economic characteristics captured in the NZ Deprivation index (NZDep) that are highly correlated with coverage and response patterns, we ensure we implicitly over-sample subgroups with poorer census coverage. We know from 2018 PES and prior that coverage patterns vary by ethnicity and by Māori descent. By designing our sample design to target groups with poor coverage, we implicitly over-sample Māori and Pacific populations. By implicitly over-sampling rather than explicitly over-sampling we can balance the statistical outcomes with respondent burden.

Accuracy measures for Māori populations are a key factor in our evaluation processes for comparing the merits of the different design options considered. In addition to Māori populations, accuracy for Pacific, youth age groups, and geographic grouping also inform design evaluations.

The specific metrics used to evaluate designs are:

- sampling rates of iwi-led areas
- sampling rates of areas containing emergency and transitional housing
- sampling rates at regional council and TALB geographies

- comparison of statistical measures such as precision, bias, and their interaction.

2023 PES sample design

The final sample design being used for the 2023 PES delivers to Māori through:

- implicit over-sampling via strata design, particularly the NZDep and 2023 Census enablement strategy stratification variables
- high sampling rates of iwi-led areas and areas containing emergency and transitional housing
- the selected design showed the best overall accuracy for Māori (note, we place more importance on the bias / precision interaction than we do on precision itself)
    - mid-range precision scores for both Māori descent and Māori ethnicity compared with other designs
    - smallest bias scores for both Māori descent and Māori ethnicity compared with other designs
    - highest scores for the bias / precision interaction – with 99 percent and 98 percnt of simulated runs containing the ‘true’ population value within the credible interval bounds for Māori descent and Māori ethnicity, respectively.

The national over-sampling rates, in the final primary sampling unit (PSU) selection, for key groups of interest are:

- Māori descent 3.3 percent
- Māori ethnicity 2.7 percent
- Pacific ethnic group 0.4 percent
- Asian ethnic group 0.6 percent.

Table 7 below provides regional council and TALB summaries of the implicit over-sampling of Māori descent and ethnicity subgroups achieved through the stratified sample design.

Table 7: Summary of implicit over-sampling rates Māori descent, Māori, Pacific, and Asian ethnic groups

Summary measure	Subgroup	Regional council	TALB
% geographic areas with evidence of over-sampling	Māori descent	88%	57%
	Māori ethnicity	81%	56%
	Pacific ethnic group	44%	45%
	Asian ethnic group	50%	45%
Average degree of over-sampling across the geographic areas	Māori descent	3.6%	2.3%
	Māori ethnicity	3.3%	1.6%
	Pacific ethnic group	-0.2%	-0.8%
	Asian ethnic group	0.6%	1.7%

Linking

The 2023 PES linking design utilises a range of linking methods to achieve the critical estimation assumption of perfect linking. These methods include automated linking, clerical pairwise linking, and clerical triangular linking.

Analysis of 2018 PES linking data found a relationship between ethnicity and linking methods:

- records with Māori, Pacific, and Asian ethnic groups all had a notably low proportion of near-exact links (defined in Appendix 12: Near-exact link indicators) compared with other ethnicities
- they were also more likely to have links created in later passes of automated linking, suggesting that the quality of linking data was lower than those records that linked in earlier, more conservative, passes (more information on automated linking and pass design can be seen in Automated linking)
- they were over-represented in the unlinked records at the end of automated linking.

These results support the need and value of having a multi-staged automated approach, supported by clerical linking.

Treatment of missing data

The PES estimation framework requires all key demographic variables to be complete. The variables included are ethnicity, Māori descent, age, gender, country of birth, and usual residence location (TALB level). The 2023 PES approach for treating missing data uses the 2018 design as the base, with extensions to consider the use of alternative data sources. This builds on the extensive use of alternative data sources in the 2018 Census and improves the 2023 PES imputation design as it will reduce the amount of statistical imputation required.

The 2023 PES imputation design outlines three main approaches to treating missing information:

1. deterministic imputation where the missing value is set to be the same as the value observed in a linked dataset:
    a. 2023 Census values including responses to the census, alternative data sources that would be used to fill the gaps in the census (responses, alternative admin sources, and imputed values in the census) or
    b. alternative data sources (historic census and admin data)
2. statistical donor imputation where a value is selected from a pool of similar donor records.

When considering a te ao Māori perspective, we have been cognisant of the proposed imputation of ethnicity and Māori descent information, as well as the use of administrative data sources. During the design process we have consulted with internal advisors within Stats NZ to ensure the methods and data sources proposed are appropriate and culturally acceptable.

Estimation

PES estimation aims to produce high-quality estimates of gross under-and over-coverage and net coverage for all populations and cross-tabulations of interest. We include a range of demographic groupings as model covariates including binary ethnicity indicators for four ethnicity groups (Māori, Pacific, Asian, and the fourth group is European, MELAA and Other ethnicity combined), and Māori descent. We specifically test and improve the fit of our models to accurately capture the coverage patterns for Māori. In 2018 PES these included adding additional covariate interaction terms such as Māori ethnicity with sex / gender and age, and Māori ethnicity with Pacific ethnic group.

Producing coverage-adjusted populations for Māori descent

Census produces Māori descent outputs using two classifications. The output released publicly aligns with the Census Māori Descent classification (Stats NZ, 2018b) with four top level categories released.[5] Census also produces an output supplied directly to the Electoral Commission under the Electoral Act, which is then used to update electoral boundaries. This output aligns with the Census Māori Descent Electoral classification (Stats NZ, 2022) which has two categories.

Census Māori Descent categories

01. Māori descent
02. No Māori descent
04. Don't know
99. Not elsewhere included

Census Māori Descent Electoral categories

01. Māori descent
02. No Māori descent

The 2018 PES produced coverage-adjusted population estimates for the Census Māori Descent Electoral categories. This was primarily because of customer need via the ERP. However, following lessons from 2018 Census and 2018 PES, it is important to consider the options available such as the possibility of producing outputs for both classifications. In doing so we consider the customer needs, te ao Māori perspective, and technical challenges.

Customer need considerations:

- 2023 Census requires a coverage adjusted population estimate of the Māori descent population at a national level based on the Census Māori Descent classification (the first category only)
- the ERP requires coverage-adjusted population estimates of both the Māori descent and No Māori descent populations based on the Census Māori Descent Electoral classification (the first two categories)
- the ERP requires these population estimates for a range of cross-tabulations with other demographics
- note, the ERP requirements are based on known historic and current data need.

Te ao Māori perspective:

- is supported by the statistical standard as the agreed and recommended way to collect Māori Descent

[5] The 2018 Census use of alternative sources and statistical imputation resulted in all records having valid values and no records in the residual 'Not elsewhere included' category. This category continues to be output to support time series analysis.

- ongoing discussions and engagement with Māori support the importance and validity of the don't know category
- the Census Māori Descent classification pre-dates the Census Māori Descent Electoral classification
- the Census Māori Descent Electoral classification was developed specifically for electoral purposes following changes to the Electoral Act 1993.

Technical challenges and limitations:

- Producing population estimates for both classifications within the existing modelling framework is challenging due to the high degree of overlap in these two variables. Having two or more highly correlated independent variables in the models creates a multi-collinearity problem. This problem decreases the stability of the model results and impacts model covariate selection, causes bias in the model coefficients, and leads to model overfitting. Producing estimates from such a model can result in reduced quality for both variables and their relationship with coverage.
- Producing population estimates using the Census Māori Descent classification only also comes with quality concerns due to the variation in size across the categories. The two large categories (Māori descent and No Māori descent) have sufficient data that we can be confident in the quality of the estimates produced. The small Don't know category has very sparse amounts of data from which to model the relationship with coverage patterns. Further research is needed to determine how to best manage the coverage estimation assumptions for this group (out-of-scope for the standard design). Without this research, we cannot be confident in the quality of the estimates produced.

All things considered, there is no clear pathway that is without trade-offs. Continuing with the 2018 PES approach to only produce population estimates for the electoral concept does not align with a te ao Māori perspective or enable measurement of the 2023 Census KPIs. Similarly, producing population estimates for the census concept alone does not meet the requirements for the ERP and beyond, and requires additional research to ensure quality outputs for all categories. Producing estimates of both classifications supports the te ao Māori perspective, and the full range of customer need but is not currently technically feasible. The proposed standard design aims to ensure both the 2023 Census and ERP requirements are met but is unable to produce a full suite of outputs for the census classification.

The proposed design for 2023 PES:

- Māori descent information will be collected in line with the statistical standard, treating 'don't know' as a valid response category
  - this will ensure high quality data capture for 2023 PES outputs and to support future research
- 2023 PES processing of Māori descent information will derive the Māori Descent Electoral classification for use in estimation and outputs
  - this derivation requires imputation for all records with 'don't know' category responses
  - this will enable the granular cross-tabulations required by the ERP
- produce PES outputs (and KPI measures) using electoral classification
- derive a coverage-adjusted population estimate for the Census Māori Descent category 01 Māori descent but applying the coverage adjustment result from the electoral 01 Māori Descent group to the census count

    - this will produce the population estimate needed for measuring the 2023 Census KPIs for Māori descent
- 2023 PES will not produce coverage adjusted population estimates for all four categories of the Census Māori Descent classification
    - research ahead of the next PES will investigate methods for producing robust outputs for all categories.

3.02 Scope and population definitions

PES population definitions are unique compared with standard social surveys in that they need to reflect both the census target populations (all people and dwellings in New Zealand) and the census key performance indicators (KPIs). The scope and population definitions for 2023 PES remain the same as those used in 2018.

We adopt the Organisation for Economic Co-operation and Development (OECD) definitions for target and survey populations:

> "A target population is the population outlined in the survey objects about which information is to be sought and a survey population is the population from which information can be obtained." (OECD, 2008)

To measure both under- and over-coverage in the census, the PES target definitions need to be broader than those of census to include those people that should have been counted but were not, as well as those that were counted but should not have been.

Dwelling population definitions

The target population for dwellings in the 2023 PES is defined as all private dwellings that should have been (whether counted or not) or were (including those which should not have been) counted by the 2023 Census of Population and Dwellings. The population for inference is all private dwellings in New Zealand.

The 2023 PES survey population for dwellings is all private dwellings on the North Island, South Island, or Waiheke Island of New Zealand, overlaid with some inclusions and exclusions for practicality.

For ease of survey data collection, the 2023 PES survey population for dwellings **includes** permanent private dwellings and permanent non-private dwellings.

Permanent private dwellings:

- permanent unoccupied dwellings and dwellings in which people do not usually live are included both for enumeration and estimation purposes.
- private dwellings that are under construction are in scope for enumeration and estimation purposes.
- mobile dwellings that are fixed to one location and have permanent private residents are included for both enumeration and estimation purposes.
- dwellings in motor camps (such as caravans, campervans, cabins etc) that are fixed to one location and have permanent private residents are in scope for both enumeration and estimation purposes.

Permanent non-private dwellings (including those under construction):

- these are included in scope for the PES field enumeration of dwellings but are excluded from interview and estimation
- they are included to support a complete enumeration of private dwellings, including those associated with a non-private dwelling

- by excluding non-private dwellings (NPDs) from interview, we assume people residing in private dwellings can represent the coverage patterns of those residing in NPDs.

A note on emergency and transitional housing structures:

Transitional and emergency housing are two types of government-supported short-term housing available to individuals and families in New Zealand who are in urgent need of housing (Ministry of Housing and Urban Development, 2021). Transitional housing is operated by the Ministry of Housing and Urban Development, and emergency housing is operated by the Ministry of Social Development. Although designed for short-term stays, people and families can often end up staying in these types of housing for extended periods of time. There is a broad range in the types of housing provided including new build, residential homes, and non-private dwelling complexes re-purposed for private dwelling living.

For 2023 PES:

- any housing provided to an individual or family is self-contained and allows them to live independently from their neighbours, will be classified as a private dwelling
- in the case of motel units re-purposed for emergency housing, these will be classified as private dwellings if they are continuously used for private housing for six months or longer
- any housing provided where there are shared / communal facilities will be classified as a non-private dwelling welfare institution. This includes hostels and backpackers that have been re-purposed for emergency housing.

For ease of survey data collection, the 2023 PES survey population for dwellings **excludes** 'Other' private dwellings:

- mobile dwellings (for example, caravans, houseboats, house trucks, tents etc) that are intended to be transportable, and movable are excluded unless they have permanent private residents and are fixed to one location. These are excluded because of their movable nature and the likelihood that they will have changed location throughout the various phases of census and PES (PES enumeration, census enumeration, census night, and PES collection)
- improvised dwellings, vehicles lived in, vessels lived in, and places of habitation with no dwelling are also out of scope due to their temporary and movable nature
- dwellings in some very remote areas are excluded because of the poor cost-benefit trade-off of running field enumeration and interviews in remote areas.

Person population definitions

The 2023 PES target population for people is all usually resident people (whether in private dwellings or not) who either should have been counted or were counted by the census. This is the same as 2018 PES. The population for inference is all usually resident people in New Zealand on census night.

The 2023 PES population for people will consist of individuals who are usually resident in a New Zealand private dwelling or staying at one during the survey period (including both New Zealand residents and overseas visitors). Noting the 2023 PES population for dwellings is all private dwellings on the North Island, South Island, or Waiheke Island of New Zealand. The 2023 PES population for people has the following exclusions:

- people living in non-private dwellings (for example, prisons, hospitals, hotels)
- people living in 'other' private dwellings (for example, tents, caravans, yachts)
- people who died after census night
- overseas diplomats, their families, and people living with them
- people on islands outside of the North Island, South Island, and Waiheke (for example, Stewart Island, Chatham Islands).

Overseas visitors are included in the survey population as they may have been enumerated in census as usual residents and be contributing to the over-coverage. Similarly, babies born after census night are also included as they may have been enumerated in the census, thereby contributing to over-coverage.

3.03 Survey objectives

The survey objectives of the 2023 PES are to measure the response and coverage of New Zealand residents in the 2023 Census to enable the assessment of related 2023 Census KPIs, and for adjusting the 2023 Census counts for coverage error in the creation of the 2023 base ERP.

Relevant 2023 Census KPIs

2023 Census have determined a set of key performance indicators to measure the success of the 2023 Census (Stats NZ, 2021e). These are designed to align with the 2023 Census strategy to rebuild trust and confidence in census and deliver quality data to customers. The seven KPIs are presented in table 8, with those that are specifically dependent on the PES indicated in **bold type**.

Table 8: Key performance indicators for 2023 Census

KPI 1	**National response rate total population ≥ 90%**
KPI 2	**National response rate for people of Māori descent ≥ 90%[1]**
KPI 3	**National net coverage rate total population ≥ 98%**
KPI 4	**National net coverage rate for people of Māori descent ≥ 98%[1]**
KPI 5	A preliminary release of census data occurs within the year the census is held[2]
KPI 6	Improve the quality of iwi affiliation data • priority level for iwi affiliation increases from 2 to 1 • iwi affiliation quality metric 1 score improves from 0.71 (very poor) to ≥ 0.90 (moderate or above)
KPI 7	Percentage of Māori who have medium to very high trust and confidence in a 2023 Census by census day ≥ 90%
KPI 8	Improved trust and confidence of Tagata Moana in 2023 Census:

	a) Percentage of Pacific Peoples who have medium to very high trust and confidence in a 2023 Census by census day ≥ 90% b) National response rate for Pacific ≥ 90%

1. This measure uses the census concept of Māori descent which has three valid response categories: yes, don't know, and no.
2. The assumption is that census day is Tuesday, 7 March 2023.

To support these 2023 Census KPI measures, 2023 PES needs to produce the following.

- A coverage-adjusted population estimate for the total population. This provides the denominator for KPI 1 (response rate) and KPI 3 (net coverage rate).
- A coverage-adjusted population estimate for people of Māori descent (those people with Māori descent 'yes' when there are three valid response categories: yes, don't know, and no). This provides the denominator for KPI 2 (response rate) and KPI 4 (net coverage rate).

In addition to the KPIs, 2023 Census is also defining key result areas (KRAs) and programme performance measures (PPMs) to support, guide, and monitor programme performance at different stages. Table 9 outlines the KRAs and related PPMs that are relevant to 2023 PES.

Table 9: Key result areas (KRAs) and programme performance measures (PPMs) for 2023 Census – relevant to 2023 PES

KRA 1 Count of Units and Coverage	PPM 1.1 National population gross under-coverage rate ≤ 3% PPM 1.2 National population gross over-coverage rate ≤ 1%
KRA 2 Quality of attributes	PPM 2.1 Percentage of population counted from individual forms ≥ 90%
KRA 4 Data equity	PPM 4.1 Response rate for ethnicity sub-groups Māori, Pacific, Asian ≥ 90% PPM 4.2 Net coverage rate for ethnicity sub-groups Māori, Pacific, Asian ≥ 98% PPM 4.3 Response rate for 15–29-year-olds ≥ 90% PPM 4.4 Net coverage rate for 15–29-year-olds ≥ 98% PPM 4.5 Response rate by regional council area ≥ 90% PPM 4.6 Response rate by territorial authority and Auckland local boards ≥ 90% PPM 4.7 Net coverage rate by regional council area ≥ 98% PPM 4.8 Net coverage rate by territorial authority and Auckland local boards ≥ 98%

Estimated resident population requirements

The estimated resident population (ERP) requires net undercount estimates for the population by the following demographic breakdowns:

- single year of age (0-99, 100+)
- three-category gender (male, female, other)
- level 1 ethnicity (European, Māori, Pacific, Asian, MELAA and Other)
- level 2 ethnicity – select groups (Samoan, Chinese, Indian)
- Māori Descent Electoral (yes / no)
- country of birth as New Zealand born / Overseas born
- territorial authority and Auckland local board areas.

2023 PES survey objectives

The PES survey sample is designed to produce estimates of gross undercount, gross overcount, and net coverage for the following key population groups:

- total usually resident population in New Zealand on census night
- gender (male, female)[6]
- age (0–14 years, 15–29 years, 30–44 years, 45–64 years, 65–74, 75+ years)
- Māori ethnicity and descent (electoral) populations
- ethnicity (level 1 groups)
- birthplace (New Zealand born, overseas born)[7]
- geographical area
    - regional council
    - territorial authorities and Auckland local board areas.

The PES is not specifically designed to:

- locate people who intend to remain invisible and take sufficient steps not to be found
- look for missing people via other methods such as through administrative records (which some international statistical agencies use to improve their coverage estimates)
- address the classification difficulties with dwellings (such as what is a private dwelling) where the counting or not of a dwelling could have an impact on the PES results and could result in people being systematically missed, or double-counted, by either census or PES
- address the effects of differences in enumeration practices between the census and the PES.

3.04 Sample size and quality targets

The 2023 PES aims to measure the coverage of New Zealand residents in the 2023 Census. Dwelling coverage rates achieved by the 2023 Census are out of scope for 2023 PES.

The target sample size for the 2023 PES is approximately 16,500 private dwellings, which is a 10 percent increase on 2018 PES.

[6] A third gender category is included in the estimation design but is not specifically considered in the sample design or precision targets due to a lack of auxiliary information available to inform strata design.

[7] The 'unknown' birthplace category is included in the estimation design but is not specifically considered in the sample design or precision targets due to a lack of auxiliary information available to inform strata design.

2023 PES key performance indicators

Key performance indicators (KPIs) for the 2023 PES are determined by the 2023 Census Coverage project in consultation with 2023 Census and the team responsible for production of the ERP. The proposed KPIs for 2023 are a development of the targets set in 2018 PES. These KPIs cover survey response, completeness of the sample-frame, and precision of survey outputs, and they will be used to inform the sample design, field operational planning, and modelling of coverage data.

Dwelling frame targets

The dwelling frame used to select the PES sample must be of high quality, to ensure every dwelling in PES areas has a chance of being selected, and to enable dwellings that are missed in the census to be included in the PES frame. A high-quality dwelling frame requires a good starting address list, augmented by sound check-enumeration practices to update the list, and identify and correct errors.

The dwelling-frame target is new for 2023 (see table 10), and this follows significant issues with frame coverage identified in the 2018 PES. Two measures (calculated following all 2023 PES operations) will be used to assess the quality of the frame and the KPI:

- sample loss due to incorrect frame classification as a private dwelling (as proxy for over-coverage in frame)
- dwellings found during interview (as proxy for under-coverage in frame).

Table 10: 2023 PES dwelling frame and interview collection KPIs

KPI	2023 PES target	Status in 2018 PES
Achieved sample rate at national level	≥82%	82% target (81% achieved)
Achieved sample rate at subnational (regional council areas)[1] level	≥80%	No target set 80% ASR achieved for 10 of 16 regions, 4 of the remaining 6 were above 75%
Survey response rate at national level	≥90%	90% target (91% achieved)
dwelling-frame accuracy (enumeration)	≤2% dwelling under-coverage	No target set (≈8% net under-coverage achieved)

1. Area outside of region is not included.

Interview collection targets

In line with previous cycles, two key measures will measure and report on the quality of the interview collection phase:

- survey response rate calculated as the percentage of eligible households that respond to the survey (see Appendix 6)
- achieved sample rate (ASR) calculated as the number of eligible households responding as a proportion of all dwellings sampled.

ASRs are used as a proxy measure for survey response rates, as they provide measures of response in real time while the survey is still in the field. This enables action to be taken during field collection if rates are below acceptable levels. Due to the different calculation methods, the ASR is typically significantly lower than the response rate; in 2018 the ASR was 10 percentage points lower than the response rate at the national level.

High survey response is required to meet the stringent data-quality standards required for a coverage survey. The main reporting measure will continue to be the national response rate, which is calculated in the latter stages of data processing. National targets for 2023 are proposed at the same level as 2018 PES, which were successfully met.

An ASR target for regional council areas is proposed for 2023, to reduce geographic differentials in response to the PES, which proved problematic in 2018. This is a change from 2018 PES, when no regional targets were set, to ensure regional rates are monitored closely during 2023 collection as well as the national rate. Regional ASR targets will be factored into the 2023 collection approach, to maximise our ability to meet both national and regional targets. A regional response-rate target was considered but has not been included in the KPI suite, with the regional ASR target considered to be more important.

Precision targets

Historically PES has set sample error precision targets to align with the direct weighted estimation methods used for 2013 PES and prior (and were planned for use in 2018 PES). Following the change in estimation methodology used for 2018 PES (a shift to using Bayesian models), the sample error precision targets are being replaced with credible interval targets. Due to the shape of the posterior distributions, which are output by the model, the credible intervals are not necessarily symmetrical. Note, we specify precision targets as a mechanism for managing sample error. We do not specify bias targets or other targets relating to non-sample errors – this is because we use a range of statistical design techniques and decisions to actively minimise and mitigate non-sample errors.

The 2023 PES precision targets have primarily been informed by what was achieved by 2018 PES.

Table 11 presents the 95 percent Bayesian credible interval (CI) targets for 2023 PES. We show the difference between the posterior distribution's median and 2.5th percentile as a negative value, and the difference between the median and 97.5th percentile as a positive value. Key changes from 2018 PES (aside from the shift away from sample error) are the more granular geographic categories and the introduction of Māori Descent Electoral.

Table 11: 2023 PES precision targets

Category	**2023 proposed target CI** (max percent)[1]	**2018 achieved CI** (max percent)[2]
National total population	**-0.4–0.4**	-0.29–0.32
Gender (male, female)	**-0.4–0.4**	-0.37–0.42
Broad age group • 0–14, 15–29, 30–44,	**-1.0–1.0**	-0.65–0.73

45–64, 65–74, 75+		
Māori • Māori ethnicity (level 1) • Māori descent (electoral)	**-1.0–1.0** **-1.0–1.0**	-0.71–0.81 -0.64– 0.72
Ethnicity (level 1, excl Māori) • Asian, European, Pacific, MELAA, Other	**-1.5–1.5**	-1.12–1.32
Geographic area[3] • Regional council • TALB[4]	**-1.5–1.5** **-1.8–1.8**	-1.22–1.73 -2.30–3.45

1. The 2023 proposed credible intervals targets are in reference to the median population estimate for a given group.
2. The 2018 achieved credible intervals are in reference to the median population estimate for a given group. The achieved credible intervals for all categories containing multiple subgroups represent the maximum observed boundary for each lower and upper across the subgroups.
3. Residual categories for area outside are not included.
4. For 85 percent of TALB areas.

3.05 Questionnaire content

The approach to the 2023 PES questionnaire is to retain most of the content from the 2018 PES, with changes made to questionnaire variables only where there is a strong case to do so. Drivers for change include societal changes over the five years between surveys, and methodological developments in both census and PES. Stats NZ provides a full summary of the topics evaluated and subsequent recommendations.

New question for 2023

Iwi affiliation is also proposed for inclusion to align with a Stats NZ in-principle decision (taken by Rachael Milicich – Deputy Government Statistician, Insights and Statistics and Rhonda Paku – Kaihautū) to collect iwi affiliation in all Stats NZ's household surveys. The rationale is twofold: ensuring Māori/iwi better see themselves in our questionnaires and with a view to building a pool of iwi affiliation data over time and survey vehicles, to enable greater insight for Māori/iwi (especially in the absence of a reliable administrative data source for this variable). Iwi affiliation data collected by the PES will not be used in coverage estimation or for output.

Modified questions

Gender

In April 2021, Stats NZ released a new data standard for gender, sex, and variations of sex characteristics. The standard provides "definitions and guidance on the collection of gender, sex, and variations of sex characteristics, as well as guidance for deriving cisgender and transgender population data" (Stats NZ, 2021f, p. 6). A fundamental principle of the standard is that of 'gender by default' – where collection and output of gender data is the default as opposed to sex. The principle specifies that "users should have a clearly established information need for collection and outputting sex data" (Stats NZ, 2021f, p. 8).

The PES has historically produced a 'sex' output. However, the question is stated as "Are you..." with the binary options of male or female and does not specify if it is asking for the respondent's 'sex at birth', 'current sex', or 'gender'. As a result of the question being open to interpretation, it is assumed to be more likely that PES respondents have answered with a concept closer to 'gender' than 'sex at birth', but currently there is no option for 'another gender.'

The 2023 PES will align with the statistical standard, shifting from the previously ambiguous sex question to gender for both collection and output (although additional research is needed to confirm if coverage estimates can be produced for all three categories). The 2023 PES will not collect sex information due to the lack of an established information need and operational challenges for respectfully collecting accurate information.

2023 PES will ask for gender information using the following question format:

What is your gender?

- *Male*
- *Female*
- *Another gender*
    - *please state (non-proxied responses only)*

[Male, Female, Another gender, Don't know, Refuse]

Māori descent

A question on Māori descent was added in 2018 PES, and this is retained for 2023 – albeit changed slightly to align it with the collection of the Māori descent variable in 2023 Census. In 2018 PES, the question treated the don't know response category as a residual category rather than a valid response category. Analysis of 2018 Census and PES data indicated the difference in question formats resulted in a different response distribution and far fewer don't know responses reported in PES than in census. For 2023 PES, the question format will be aligned with the census format as shown below:

Are you descended from Māori?
(If necessary) For example, do you have a Māori birth parent, grandparent, or great-grandparent?

- *Yes*
- *Don't know*
- *No*

[Yes, Don't know, No, Refuse]

This alignment will provide greater consistency with the Māori descent variable and concept as collected in census and will assist with reporting of the response rate for people of Māori descent. In addition to the change in question format, interviewers will also be trained and prompted to read out the response options to ensure respondents are informed that don't know is a valid response.

Full set of questionnaire topics

Table 12 below provides a summary of the topics and questions included in the 2022 PES field test. Questionnaire content for 2023 will be finalised following the conclusion of the field test.

Table 12: 2023 PES questionnaire content

Topic	Question set[1]
Demographics	• Full name / other names • Date of birth / age • Gender • Ethnicity • Māori descent, iwi affiliation • Country of birth / date of arrival to live in New Zealand
Location	• Usual residence address • Census night address • Other addresses where a census form may have been completed
Census	• Census form completion
Dwelling	• Dwelling address • Dwelling type • Joined dwelling • Number of storeys • Dwelling ownership / family trust • Number and names of usual residents and visitors at PES address • Additional dwellings at PES address
Scope	• Born after census night • Recent overseas travel / return to New Zealand • Visitor from overseas • Absence from usual residence • Stays at other private dwellings

1. Note the questions each respondent gets asked are determined by the routing from previous responses, and the specific wording depends on if it is a proxy or non-proxy response.

3.06 Dwelling enumeration approach

The 2023 PES approach to dwelling enumeration builds on the lessons learnt from 2018, and key research pieces outlined earlier in this document. The key lessons and decisions incorporated into the approach are as follows.

- Start with the highest quality list available, even if it is not fully independent from that being used by census.
- Prioritise experience and resources towards more complex areas with highest rates of dwellings that are hard to find, or hard to classify.
- Completing more than one enumeration pass of an area improves the quality of the resulting dwelling frame.
- Completing desktop canvassing ahead of field enumeration improves the quality of the resulting dwelling frame.

The resulting dwelling enumeration approach is as follows.

- Produce a fit-for-purpose, quasi-independent dwelling list for enumeration, derived from the SLR. The development of this list builds on work undertaken in 2021 by the coverage project team and is dependent on work planned by the SLR and Admin Data team to improve the quality of the SLR.
- Implement desktop canvassing using a modified version of the internally available SLR Editing Tool, accounting for the lessons learnt from the dwelling enumeration field test.
- Undertake desktop canvassing on 75 percent of the sample to improve data quality, targeting the desktop canvassing workload to the hardest-to-find areas. Note that adding more desktop canvassing beyond 75 percent provides diminishing returns.
- All dwellings are to be enumerated after desktop canvassing by field staff (either experienced or temp), subject to limitations imposed by COVID-19. Experienced field staff will be targeted to the hardest-to-find areas, where possible.

3.07 Sample design

The 2023 PES sampling design includes two stages: The first stage involves the selection of a sample of primary sampling units (PSUs) from the Household Survey Frame (HSF).[8] The second stage of the sample selection consists of selecting eligible dwellings within the selected PSUs.

The process of defining a suitable sampling design for selecting sample PSUs was completed in five stages:

1. preparation of a pretend population and stratification variable selection
2. evaluation of sample size and sample allocation to strata strata design
3. detailed evaluation of precision and KPIs
4. refinement of allocation and sampling methods
5. sample PSU selection.

In a typical sample design process, steps 2 and 3 can be completed together. However, the 2023 PES sample design was the first time Stats NZ has embarked on developing a sample design for a Bayesian estimation framework. Due to computational and time constraints, we were unable to run the full range of designs evaluated through the full Bayesian estimation framework. To reduce the computational load, we devised a two-step approach for evaluating sample designs. The first step was to evaluate the full range of considered designs (12 in total) using a standard weighted estimation simulation framework in which 1,000 sample draws produced 1,000 estimates, from

[8] The HSF is the standard sampling frame Stats NZ uses to select samples and manage overlap control between variety of household-based surveys which run either with Stats NZ or other government departments.

which we summarised various evaluation measures. The second step was to select the four most promising designs and run them through the full Bayesian estimation framework and do a detailed evaluation of each design against the full range of precision KPIs set out.

Throughout the sample design process, we were guided by the project commitment to produce robust estimates for Māori descent and ethnicity populations. We were also guided by the needs and priorities of the ERP (particularly estimates at TALB geography), and emerging areas of interest such as iwi-led collection areas, and areas with emergency housing establishments. As many of these enquiries are unable to be explicitly addressed in the strata design itself, we include presence of these in our evaluation measures.

Construction of the pretend population used in design simulations

To support comparisons between different sample design options, we construct a fixed 'pretend population' from which we can define the known 'true' population and select both the census capture of that true population and the PES sample as well. We use the term 'pretend population' instead of 'synthetic population' or 'simulated population' because our main dataset is based on the real 2018 Census records and doesn't have artificially created records.

To create this pretend population, we select all permanent private dwellings from the 2018 Census dataset using the same criteria used for creating HSF:

- dwelling type code "1" – private dwellings
- dwelling status code "11"– occupied dwellings
- dwelling description code not "06", "07", "08", "09" - exclude mobile dwelling not in a motor camp, improvised dwelling, or shelter, roofless or rough sleeper, non-private dwelling.

We then populate the dwellings with the usual residents from 2018 Census dataset. The final population of dwellings and individuals had the following limitations and simplifications.

- We did not include non-private dwellings, and all included dwellings were in scope for PES, which will not be the case for the real survey.
- We did not include empty dwellings, since they would not be able to be used for simulated census, PES, and coverage estimation processes.
- We did not include admin enumerated records from the 2018 Census dataset that were not associated with a dwelling. Allocating these records into a 'synthetic dwelling' may have distorted the real population structure and would come with technical complexity.

From this known population, we then create a fixed census capture, set up fixed PES survey response indicators, and a framework for selecting the PES sample dwellings.

To create the census capture, each individual record in the pretend population was assigned binary under-coverage and over-coverage indicators. Firstly, we assigned over-coverage indicators, and then, to all non-over-coverage records, we assigned under-coverage. Probabilities to be under- or over-coverage were derived from the PES 2018 outputs and varied based on TALB, four binary ethnicity indicators (Māori, Pacific, Asian, and a grouping of European, MELAA, and Other), a binary Māori descent indicator, binary sex as collected in 2018 census, and broad age groups. This was to ensure our census capture included realistic and varied coverage patterns. In creating the census capture in this way, we make the following assumptions:

- over-coverage and under-coverage are mutually exclusive and independent
- the coverage patterns observed in 2018 Census are representative of what can be anticipated for 2023 Census.

To enable the PES capture to reflect different draws of PSUs, we first create fixed survey response indicators on the full pretend population. Using the PES 2018 response patterns, we assign a response indicator to each dwelling record. The binary indicator of response, 1 – 'responded' and 0 – 'did not respond', is drawn as a Bernoulli random variable with a TALB-specific probability of a dwelling response. By basing this step on the 2018 PES response patterns, we assume these patterns are representative of what can be anticipated for 2023 PES.

For every draw of PSUs used in the sample design investigations, we created a PES sample capture. For each PSU selected in each draw, we selected eleven dwellings (based on the optimal cluster size design for 2018 PES). The response indicator created earlier is then used to remove non-responding dwellings. The remaining dwellings and the people in them form the sample used to produce estimation results.

This is a simplified approximation of the actual sampling process. Elements not accounted for include:

- dwelling enumeration errors / sample loss
- individual level non-response
- data reporting and collection problems and PES to census linking errors.

Despite the noted limitations of our approach, we are confident the pretend population and simulated data captures for census and PES are fit for purpose. We primarily use this data approach to ensure we control as many variables in the design process as possible. This allows us to isolate differences resulting directly from the design variations that are of interest.

Selection of stratification variables

The process of stratification tries to divide the population into homogenous subgroups and then select sample from each stratum separately. We use stratification to reduce sampling errors, manage sample size, and to target subgroups correlated with our outcome of interest (coverage patterns achieved by census).

2018 PES included a full sample design review and considered a broader range of stratification variables before proceeding with the following:

- 12 broad regions (Gisborne and Hawke's Bay are combined, as are the West Coast, Marlborough, Nelson, and Tasman)
- a mix of urban and rural areas
- the planned mode of delivery of census forms (enablement strategy)
- socio-economic characteristics captured in the NZ Deprivation index (NZDep).

For 2023 PES, we focused design investigations on possible updates to two of the stratification variables from the 2018 PES design: region classification, and planned census enablement strategy. The region classification was investigated in response to the 2023 Census setting programme performance measures for response and net coverage rates for regional councils, for which 2023 PES will provide the denominators for these calculations. Two approaches for grouping the census enablement strategy were investigated because of the relationship between enablement strategy and census response and coverage patterns.

By using stratification variables such as the census enablement strategy and NZDep that are highly correlated with coverage and response patterns, we ensure we implicitly over-sample subgroups with poorer census coverage. We know from previous 2018 PES and earlier that coverage patterns vary by ethnicity and by Māori descent. By designing our sample design to target groups with poor coverage, we implicitly over-sample Māori and Pacific populations. By implicitly over-sampling rather than explicitly over-sampling we can balance the statistical outcomes with respondent burden.

Evaluation of sample size and strata design

A range of sample design options were evaluated for 2023 PES. Key elements investigated were:

- Sample size
    - 15,000 dwellings (consistent with the sample size used in 2013 and 2018)
    - 16,500 dwellings
    - 18,000 dwellings
- Stratification variables
    - We held two stratification variables consistent with the 2018 PES design: urban/rural; and a 10-category grouping of the New Zealand Deprivation Index (NZDep)
    - With the new KPIs set for regional council, we explored two different region groupings
        - broad region with 12 region categories (used previously by PES, and common in most household surveys)
        - regional council with 16 region categories (directly aligned with the KPI)
    - two groupings of the 2023 Census enablement strategy variable – to protect against any structural variations in coverage patterns resulting from the 2023 Census field design, and to target the PES sample towards priority response groups.
        - 2-category census design:
            - areas that are exclusively mail-out
            - all other areas (any PSU with a mix of Mail-out with delivery or assist, and full delivery and or assist areas)
        - 3-category census design:
            - areas that are exclusively census mail-out
            - areas that are a combination of census mail-out and census deliver or exclusively deliver
            - areas that contain any census assist approach.

This resulted in a total of 12 sample designs considered (see Table 13) and run through a simulation approach in which 1,000 draws of PSUs were selected for each design and estimates produced.

Table 13: Summary of strata designs, and sample sizes evaluated

	Design details			
Design	Base	Region group	Census design group	Sample size
1a	Urban / Rural NZDep (10 groups)	Broad (12)	Mail-out Deliver + Assist	15,000
1b				16,500
1c				18,000
2a			Mail-out Deliver Assist	15,000
2b				16,500
2c				18,000
3a		Regional council (16)	Mail-out Deliver + Assist	15,000
3b				16,500
3c				18,000
4a			Mail-out Deliver Assist	15,000
4b				16,500
4c				18,000

The following evaluation measures were produced from the simulations for each design.

- Number of strata (target range of 100–120).
- Average number of TALB represented in the simulation sample (target of 85 out of 88 as the remaining three are offshore areas not in scope for household surveys).
- Representation of each of the three iwi-led areas being considered by 2023 Census. Representation is defined as selection of 5 percent or greater of the PSUs in the iwi-led areas. Iwi-led areas make up less than 2 percent of the national PSU framework and PES typically samples 6 percent of national PSUs. A target of 5 percent was set to reflect the need to over-sample these areas and was demonstrated to be realistically achievable.
- Representation of areas known to contain emergency housing establishments. Representation is defined as selection of 5 percent or greater of the PSUs known to contain at least one emergency or transitional housing establishment. The 5 percent threshold is set in the same manner as the iwi-led area threshold.
- Precision measures for all PES KPI groups (calculated as sampling error).
- Bias measures (as distance between estimate and truth) for all PES KPI groups (calculated as root mean square error).[9]

All designs produced comparable results for the first four measures listed above (full evaluation results for all designs can be found in Appendix 7).

A summary of the results of the simulation studies is listed below (full evaluation results for all designs can be found in Appendix 7).

- All designs produced number of strata within the target range.
- All designs produced selected PSUs from an average of 84 TALB. This was very close to the target of 85 and is considered acceptable since all designs had the same result.

[9] Some bias is controlled through the use of the pretend population and simulated sampling framework. However, the DSE assumption of homogeneity was not fully controlled so we would expect some residual bias in the estimator.

- All designs achieved good results for representation of two of the three iwi-led areas being considered by census.
    - One of the iwi-led areas proved challenging for many designs
        - None of the smaller sample size designs met the representation target for this iwi-led area
        - None of the group 2 (2a, 2b, 2c) designs met the representation target for this iwi-led area
    - Designs 3a and 4a (the smaller sample size) came close but failed to meet representation for another of the iwi-led areas
- All designs achieved the representation target for areas known to contain at least one emergency or transitional housing establishment.

The precision and bias measures for the different designs were evaluated via a ranking method. At each sample size, the four different strata designs were ranked best to worst on both bias and precision measures for each of the population groups for which PES has set KPIs. These rankings were then collapsed to summarise the rankings across all KPIs groups, Māori population groups, and geographic groupings. A weighted score was calculated for each strata design at each sample size for each of the collapsed groupings. The weighted calculated applied a weight of 3 for the design ranked first for each bias and precision, a weight of 2 for the design ranked second, and a weight of 1 for the design ranked third. These weight scores are then summed over the results groups of interest for each design.

Table 14. Summary of bias and precision weighted ranking scores at three sample sizes

a. Ranking summary for sample size of 15,000 dwellings

Design	All groups	Māori	Region	TALB
1	368	12	52	262
2	391	0	61	282
3	359	8	53	258
4	274	4	26	218

b. Ranking summary for sample size of 16,500 dwellings

Design	All groups	Māori	Region	TALB
1	357	8	56	269
2	415	4	60	293
3	322	0	40	240
4	298	12	36	218

c. Ranking summary for sample size of 18,000 dwellings

Design	All groups	Māori	Region	TALB
1	386	8	61	273
2	396	12	64	297
3	300	4	29	224
4	310	0	38	226

Table 14 presents the results of the bias and precision rankings. At the smallest sample size (15,000 dwellings) design 2 performs best overall but worst for Māori, and design 1 performed best for Māori and second overall. At the middle sample size (16,500 dwellings) design 2 is again best overall, with design 1 second. Design 4 performs best for Māori but worst overall. At the largest sample size (18,000 dwellings) design 2 performs best overall and for Māori, followed closely by design 1.

Designs 3 and 4, which use the 16-region stratification variable, did not improve the quality of designs on all sample sizes. Having more small-sized strata complicated the process of PSU selection. Therefore, we excluded designs 3 and 4 from the further analysis.

Based on the range of evaluation measures, the most promising designs were identified as:

1. design 2 with sample size of 18,000 dwellings
2. design 1 with sample size of 18,000 dwellings
3. design 1 with sample size of 16,500 dwellings
4. design 2 with sample size of 16,500 dwellings.

Detailed evaluation of precision and KPIs

The second stage of the sample design process was to run 100 of the 1,000 initial samples for each of the three preferred designs through the PES Bayesian estimation framework. The detailed results from the Bayesian framework are then used to complete a detailed evaluation of each design and ability to achieve the range of 2023 PES KPIs.

Here are the key components in this process.

- Select a sub-sample of 100 draws from the initial 1,000 samples used in the previous step. The initial 1,000 draws are randomly generated, so the sub-sample is achieved by selecting the first 100 draws.
- Running the 100 samples through Bayesian under- and over-coverage models. Note, the under-coverage model is a simplified version of what was used to produce 2018 PES results but is sufficient to correctly estimate the data generated by the simulated census of the pretend population. We also acknowledge that the model modifications will be required for the real estimation in response to the 2023 Census outcomes.
- The model results are then collapsed, and key results retained.
- Evaluation measures are produced for all PES KPI groups:
    - bias (distance between the estimate and truth) calculated as normalised root mean square error from the model output midpoints
    - precision calculated as the maximum credible interval range produced from the model outputs
    - the interaction between bias and precision calculated as the proportion of the 100 draws for which the ‘true’ value falls within the credible interval.

The key results from analysis of the evaluation measures are listed below.

- All designs performed well on the bias measure. There was very little variation across the designs and no obvious pattern to the variation.
- All designs performed well on the precision measure when evaluated against the PES KPIs for demographic groups. All designs meet the KPI targets for total population, gender, Māori descent, four ethnic groupings, and age groups.
- There were minimal precision gains across the PES KPI populations of interest in the two designs with the larger sample size (18,000 dwellings compared with 16,500 dwellings).

- All designs failed to achieve the PES KPI target for regional council geographic areas. One design met the PES KPI target for TALB,[10] with the others significantly off.
- The two designs with the smaller sample size (16,500 dwellings) were more stable and performed better on the bias / precision interaction measure compared with the two designs with the larger sample size (18,000 dwellings). This showed that the increasing sample size makes the results more sensitive to PSU selections, which can lead to the overfitting of the model to the observed data.
- Advice from Stats NZ household survey experts was to focus resource into achieving a high response rate for the smaller sample size (16,500 dwellings) rather than to increase the sample size.

The evaluation results for the designs were collapsed by ranking the designs (1 for the best performing and 4 the lowest performing) on each of the evaluation measures, performance for Māori, respondent burden, and cost. The key considerations when determining the preferred design were the interaction between bias and precision, precision, precision performance for Māori population groups, and respondent burden. Table 15 presents a summary of the results.

Table 15: Summary of sample design evaluation measures and considerations. *The numbers indicate the quality rank assigned to a design with 1 being the best-performing and 4 – lowest.*

	Design 1: 16,500	Design 2: 16,500	Design 1: 18,000	Design 2: 18,000
Bias	1	1	1	1
Precision	3	4	1	2
Interaction	1	2	4	3
Māori	1	3	2	4
Burden	1	1	3	3
Cost	1	1	2	2

The two best-performing designs across all measures are designs 1 and 2 with a sample size of 16,500 dwellings. Design 1 (which has two-category 2023 Census enablement strategy stratification variable) performs better than design 2 (which has a three-category 2023 Census enablement strategy stratification variable) on three of the four key considerations: the bias / precision interaction, precision performance for Māori, and precision across all KPI groups.

Allocating PSUs to strata

In this section, you can find how 1,500 PSUs were allocated to each sampling stratum.

Region

Any disproportionate sampling by stratum increases the sampling variance of national and other cross-strata estimations (for example, age by gender cross-tabulations). Therefore, the sample fraction in each region is designed to be proportional to that of the total population. That is, the number of PSUs selected from stratum h was:

[10] The TALB precision KPI used for sample evaluation was a target of 80 percent of TALBs with credible intervals of +/- 1.5 percent. This KPI was revised following the sample design to better reflect what was realistically achievable. The final KPI for TALB is 85 percent of TALBs with credible intervals of +/- 1.8 percent.

$$n_h = n \frac{N_h}{\sum N_h}$$

Where n is the total number of PSUs to select, and N_h is the number of PSUs in total in stratum h. Table shows the population proportion per region and the number of PSUs selected per region.

Table 16: Total number and selected number of PSUs by region

Broad region	Total number of PSUs	Percentage of total PSUs	Number of PSUs selected	Percentage of sample PSUs selected
Northland	876	3.8%	58	3.9%
Auckland	6,960	30.0%	456	30.4%
Waikato	2,253	9.7%	144	9.6%
Bay of Plenty	1,520	6.6%	97	6.5%
Gisborne / Hawke's Bay	1,089	4.7%	71	4.7%
Taranaki	647	2.8%	41	2.7%
Manawatū-Whanganui	1,296	5.6%	84	5.6%
Wellington	2,580	11.1%	165	11.0%
Tasman / Marlborough / Nelson / West Coast	995	4.3%	64	4.3%
Canterbury	3,179	13.7%	205	13.7%
Otago	1,213	5.2%	79	5.3%
Southland	566	2.4%	36	2.4%
NATIONAL	23,174	100	1,500	100

Urban and rural PSUs

To manage the operational costs of fieldwork, we classify areas into urban and rural areas (noting urban areas are less expensive to collect sample from than rural areas). We over-sample urban areas relative to their actual proportion in the population. For all regions excluding Auckland, they are over-sampled by a factor of 1.4 (consistent with standard household survey practice):

$$\frac{n_{urban}}{N_{urban}} = 1.4 \frac{n_{rural}}{N_{rural}}$$

Where 'urban' refers to PSUs that are classified as being in main urban areas. We use a different factor in Auckland to ensure better representation and sampling from the Auckland local boards (some of which are wholly rural or wholly urban). In Auckland, urban areas are over-sampled by a factor of 6. Table shows the total population of PSUs by the urban / rural classification, and the resulting sample sizes in each.

Table 17: Total number and selected PSUs by urban / rural classification

		Total number of PSUs	Percentage of total PSUs	Number of PSUs selected	Percentage of sample PSUs selected
Urban		14,790	63.8%	1,013	67.5%
Rural		8384	36.2%	487	32.5%

2023 Census enablement strategy

Like the regional allocation, the sample fraction in each of the two 2023 Census enablement strategy categories are designed to be proportional to that of the total population.

Table shows the total population of PSUs by the urban / rural classification, and the resulting sample sizes in each.

Table 18: Total number and selected PSUs by 2023 Census enablement strategy

	Total number of PSUs	Percentage of total PSUs	Number of PSUs selected	Percentage of sample PSUs selected
Exclusively mail-out	15,544	67.1%	1,021	68.1%
Other	7,630	32.9%	479	31.9%

New Zealand Deprivation Index

We align with the allocation methods used by the Household Labour Force Survey (Stats NZ, 2017), using Neyman allocation to the NZDep strata based on an indicator variable that shows the employment status of each person . This method allocates the sample to strata based on the strata variances and similar sampling costs in the strata. A Neyman allocation scheme provides the most precision for estimating a population mean given a fixed total sample size. Neyman allocation assigns sampling units (PSUs) within each stratum, proportional to the product of the population stratum size (N_h) and the within-stratum standard deviation (S_h), so we can achieve minimum variance for the population mean estimator. The sample size in any stratum is worked out according to the following equation:

$$n_h = n \frac{N_h S_h}{\sum N_h S_h}$$

where n_h is the sample size for stratum h , n is the total sample size, N_h is the population size for stratum h, and S_h is the standard deviation of stratum h of the variable for which you are trying to maximise survey precision (in this case employment status of people in stratum h).

Using employment status results in an over-sampling of more highly deprived areas and areas with poorer census coverage results. This leads to a higher proportion of the achieved sample being Māori, improving the quality of coverage adjusted population estimates for Māori.

Sample selection and overlap control

Territorial authority and Auckland local board areas (TALB) are unable to be explicitly included in the strata due to complexity and sample size. However, the PSU selection process is designed to implicitly account for TALBs and maximise the probability of selecting at least one PSU in each TALB.

As will all other Stats NZ household surveys, 2023 PES PSUs selection is conducted through an overlap control system. Below is an excerpt from the Household Labour Force Survey documentation that outlines the Stats NZ overlap control procedures (Stats NZ, 2017, p. 20).

> "At Stats NZ, we select PSUs for any household survey through the overlap control system. This is because it is desirable to minimise the sample overlap, as measured by the proportion of selected PSUs in common across surveys. That is, we try to reduce respondent burden as much as possible by minimising the number of respondents selected to take part in more than one survey.
>
> In any overlap control scheme, stratum size and stratum sampling fractions may conspire to make sample overlap unavoidable. In such a case it is desirable to have a scheme that minimises the extent of overlap as part of its normal operation. The overlap control we use implements the conditional sampling methods of Chowdhury et al. (2000), and Bell (2011) fusing the main ideas from these two methods, then extending the system to deal with frame reformation (changing the PSU boundaries), as in Lu (2012). This methodology allows us to share the household frame with other government departments and minimise overlap across organisations."

Summary of final sample design

A summary of the final 2023 PES sample design and sample selection is described as follows.

The Stats NZ national framework of PSUs were grouped in 112 strata based on:

- broad regions
- a mix of urban and rural
- the 2023 Census enablement strategy
- socio-economic characteristics captured in the NZ Deprivation Index.

Sample selection across TALBs is implicitly achieved through the selection process which aims to maximise the probability of selecting at least one PSU in each TALB.

The 2023 PES sample design does not include any explicit over-sampling of any demographic group. Instead, we achieve the necessary over-sampling implicitly through the strata design. The national over-sampling rates, in the final PSU selection, for key groups of interest are:

- Māori descent 3.3 percent
- Māori ethnicity 2.7 percent
- Pacific ethnic group 0.4 percent
- Asian ethnic group 0.6 percent.

Table 19 provides regional council and TALB summaries of the implicit over-sampling rates in the final PSU selection for groups of interest.

Table 19: Summary of implicit over-sampling rates Māori descent, Māori, Pacific, and Asian ethnic groups, based on final PSU selection

Summary measure	Subgroup	Regional council	TALB
% geographic areas with evidence of over-sampling	Māori descent	88%	57%
	Māori ethnicity	81%	56%
	Pacific ethnic group	44%	45%
	Asian ethnic group	50%	45%

Average degree of over-sampling across the geographic areas	Māori descent	3.6%	2.3%
	Māori ethnicity	3.3%	1.6%
	Pacific ethnic group	-0.2%	-0.8%
	Asian ethnic group	0.6%	1.7%

The target sample size for the 2023 PES is 16,500 permanent private dwellings. This will be achieved through selection of 11 dwellings from each of the 1,500 PSUs.

3.08 Linking

Linking PES records to the census dataset is a critical step in coverage estimation. It is through linking that we can detect which PES records were missed by census, which should have been and were counted, and any records counted more than once or counted in error. As outlined in section 1.2 Introduction to coverage methodology, perfect linking between PES and census is a core estimation assumption. To ensure this assumption is met, each PES record needs to link to every census record that is a true match for it. This section of the document will outline the methods and processes we use to ensure that assumption is upheld, including automated linking methods, clerical linking methods, and quality assurance processes. More detailed information on the 2023 PES linking methodology is published on the Stats NZ website (Stats NZ, 2023b).

High-level linking design

The over-arching goal of the 2023 PES linking design is to achieve perfect linking between the PES and census records. We define perfect linking as negligible false positive and negligible false negative links. We define the following quality targets which are measured via the quality assurance sampling process:

- less than 0.5 percent of PES records with 1 link are assessed as a false positive link
- Less than 0.5 percent of PES records with more than 1 link are assessed as having at least 1 false positive link
- Less than 0.5 percent of PES records with no link are assessed as a false negative link
- Less than 0.5 percent of PES records with 1 link are assessed as having a false negative link to an additional record.

The above targets are set at levels that are practically achievable, and by setting four different types with the same targets we are confident that net error will be close to zero percent for the full population as well as for smaller subpopulations.

To achieve these outcomes, we use a range of methods and processes including both automated and clerical linking (Stats NZ, 2023b). The 2023 Census responses will be linked to the IDI spine as part of census processing and creation of the complete census dataset. The census to IDI spine links, along with the complete census dataset are key inputs into the PES linking design. The key steps in our linking design are:

- rule-based methods to only accept records with sufficient information to support being able to link / not link each record with confidence (link acceptance criteria)
- automated pairwise linking of PES to census records, using a conservative multi-pass design to minimise false positive links

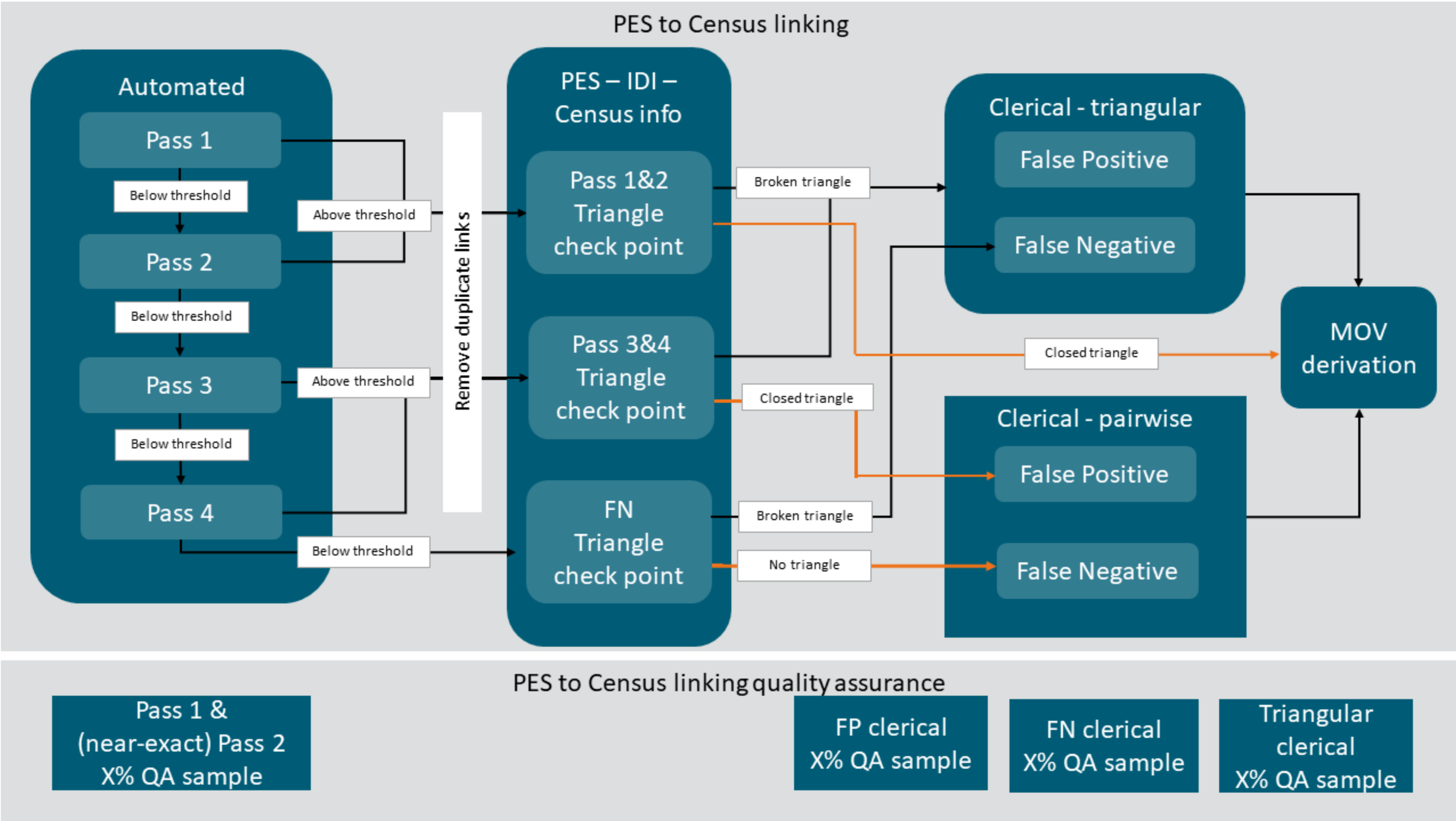
- automated pairwise linking of PES to IDI spine records, using a linking strategy designed to balance both false positive and false negative links
- removal of duplicate links created during automated linking
- identification of broken triangle scenarios between PES, census, and IDI spine records
- clerical resolution of broken triangles
- clerical review of a subset of automatic links to detect false positive links
- clerical review of all records with no PES – census link to resolve any false negative links
- derivation of a 'matching outcome variable' (MOV) for use in estimation processes
- quality assurance processes to confirm perfect linking assumption is met.

Figure 5 provides a visual representation of the role of automated linking, clerical triangular linking, clerical pairwise linking, and quality assurance in the high-level linking design. All datasets used in linking (records collected in PES, in the census dataset, and in the IDI spine) are first subset to only retain records that pass the link acceptance criteria. The datasets are constructed to include unique person and address combinations, to aid PES in exhausting all possible locations at which a person may have been counted in census. This can result in the creation of duplicate links (PES person linking to same census record more than once) during automatic linking.

As shown in Figure 3, we begin by completing multi-pass automated linking. Links created in the first two passes are accepted as links and do not go through any clerical linking unless they have a broken triangle link. Links created in the third and fourth passes have been shown to have greater risk of false positive error, so these links are directed to clerical review.

Following automated linking, we check and clean any duplicate links created so that only unique PES–census record pairs are retained. All PES records (and any associated links to census and/or IDI spine records) then get checked and divided into broken and closed or no triangles (more detail is provided in subsequent sections). Broken triangles proceed to clerical triangular linking for resolution. Closed triangles either proceed to clerical pairwise linking or to the MOV derivation depending on the type of link they have and how that link was created. Records with automatic links created in pass 1 or near-exact links in pass 2 proceed to the MOV derivation, records with automatic links created in passes 3 or 4 (or non-exact links in pass 2) proceed to clerical linking to check for false positive links, and records that have no PES to census link from automated linking proceed to clerical linking to check for false negative links. Once both clerical triangular and pairwise linking processes are complete, the MOV can be derived, and quality assurance processes started. The quality assurance process involves completing clerical review samples for automatic links created in pass 1 and near-exact links from pass 2, samples of records from clerical pairwise linking for detection of false positive and false negative links, and a sample of records that have been through clerical triangular linking.

Figure 3: High-level linking design. Note that pass 2 links are a special case, explained in Resolving false positive links.



Link acceptance criteria and input data

To determine whether records have enough linking information (name, age, date of birth) to not cause coverage error, all records used in PES linking (PES responses, the combined census dataset, and the IDI spine) are first evaluated on the quality of linking information available. Only records with high-quality linking information are retained and eligible for linking. The reasoning for this is to support PES in linking with confidence. Records that do not meet the link acceptance criteria are excluded from linking and estimation and are instead treated as under-coverage by the estimation models. The benefit of this approach is that it reduces the risk and magnitude of bias in the estimation due to the failure of the DSE assumption for perfect linking.

The link acceptance criteria for 2023 PES are consistent with those used in 2018 PES (Stats NZ, 2023b).. The criteria for a record to pass is dependent on whether it has high-quality location information (able to be defined at the meshblock level).

Acceptance criteria when location information is high quality:

1. Must have either
    a. Name (either from individual form, or household table) **OR**
        i. *Both first and last name fields must contain at least 2 text characters*
        ii. *We assume any detectable facetious or erroneous name information has been cleaned out*
    b. Date of birth (full date of birth including day, month, and year)
2. And at least one other item (different from the item above):
    a. Name (either from individual form, or household table)
        i. *Both first and last name fields must contain at least 2 text characters*
        ii. *We assume any detectable facetious or erroneous name information has been cleaned out*
    b. Date of birth or age
    c. Sex
    d. Country of birth.

Acceptance criteria when location information is not high quality*:

1. Must have both
    a. Name (either from individual form, or household table) **AND**
        i. *Both first and last name fields must contain at least 2 text characters*
        ii. *We assume any detectable facetious or erroneous name information has been cleaned out*
    b. Date of birth (full date of birth including day, month, and year)
2. And at least one other item:
    a. Sex
    b. Country of birth.

These link acceptance criteria get applied to all datasets used in PES linking to ensure consistent information is available for all pair-wise links and triangular links:

- PES responses
- census individuals (responses and admin enumerations)
- IDI spine records.

In the case of the pairwise link between the census responses and the IDI spine, the link acceptance criteria are applied to the already created census – IDI spine pairwise links. In this case, any record identified as failing the criteria is stripped from the census – IDI spine-linking results (only for PES purposes).

After the link acceptance criteria have been applied, the various datasets can be restructured ready to go through linking. For automated linking, we use a long-table format for the input data. In this format the 'unit' is the unique combination between person and address. The unit is set this way to enable us to efficiently search and link people using multiple different addresses and locations. Once automated linking is completed, we collapse the links created into a dataset where the 'unit' is the unique pair created between the PES person and the other list (either census or IDI spine).

Automated linking

2023 PES will complete two pairwise automated linking processes. The first to link PES records with the census dataset, and the second to link PES records with the IDI spine.

PES – census automated linking

The 2023 PES to census automated linking design is currently based on the design used in 2018 PES but will be refined during 2023 PES processing to best fit the 2023 PES data collected. Prior to linking, census data and PES data were first standardised, so that matching variables were coded compatibly.

The 2018 PES automated linking was a multi-stage probabilistic linked approach based in Fellegi and Sunter methodology (Fellegi & Sunter, 1969). In the Fellegi and Sunter method, all records from one dataset are compared to records from another dataset on each linking variable. A total weight that summarises the comparison results of the linking variables is given to each pair of records. This weight reflects the likelihood that two records refer to the same person.

The 2018 PES linking approach is designed to be a conservative linking strategy to minimise false positive links. The linking approach included four linking passes (shown in Table 20) executed using QualityStage (QS) software. In the first pass, records must share the same date of birth, sex, and address ID (blocking variables). First and last names are then compared (linking variables), and records with enough similarity are accepted as links. Given the high level of information included in the blocking variables, this is a high-quality and conservative pass.

Subsequent passes use additional information such as meshblock (small geographic area) and country of birth, while allowing for some differences in these variables. However, conservative cut-off weights have been chosen to minimise the number of false positive links.

Table 20: 2018 PES multi-pass linking design for PES – census automated linking

Pass	Blocking variables	Linking variables	Linking type
1	Date of birth Sex Address ID	First name Family name	Deterministic
2	Date of birth Sex	First name Family name Country of birth	Probabilistic

		Meshblock	
3	Meshblock	First name Family name Day of birth Month of birth Year of birth Sex Age Country of birth	Probabilistic
4	Sex Age Soundex first name Soundex family name	First name Family name Day of birth Month of birth Meshblock	Probabilistic

Two key changes to 2023 PES content and sample design are expected to alter the PES–census pass design. The first is the shift from collecting sex to collecting gender. In 2018 PES, sex is used as a blocking variable in three of the four passes. For 2023 PES linking, we may prefer to use gender as a linking variable instead of a blocking variable depending on data quality in both PES and census. The second change that will impact on the pass design is the increase in sample size for 2023 PES. We expect the sample size increase will impact the Fellegi-Sunter weights and acceptance thresholds for each pass.

PES–IDI spine automated linking

The 2023 PES to IDI spine automated linking design is based on the design used by 2018 Census to link census responses with the IDI spine (Stats NZ, 2019c). The pass structure and cut-offs have been loosely refined to fit 2018 PES data but will be further refined during 2023 PES processing to best fit the 2023 PES data collected. Prior to linking, census data and IDI spine data are first standardised, so that matching variables are coded compatibly.

Like the PES to census linking design, the PES to IDI spine linking design is a multi-stage probabilistic linked approach based in Fellegi and Sunter methodology. Unlike the PES to census linking design, the pass strategy is designed to balance both false positive and false negative errors, rather than focus on minimising false positive errors alone. This difference is because the PES to IDI spine linking does not have a clerical component to resolve any false negatives remaining after automated linking.

Table 21 presents the linking pass design. In general, the linkage procedures start with the most discriminating and high-quality blocking variables. Passes 1 and 2 require exact agreement on date of birth meaning and use name information to support. Passes 3 to 5 use the range of address information available in the IDI spine for blocking, supported by name and date of birth information. Pass 6 blocks on demographic and last name information, supported by date of birth and location.

Table 21: 2023 PES multi-pass linking design for PES–IDI spine automated linking

Pass	Blocking variables	Linking variables
1	Date of birth IDI spine meshblock	PES full name compared with IDI spine first names
2	Date of birth	First names Last names First initials of first names Gender Country of birth Most recent meshblock (IDI)
3	IDI spine meshblock	First names Last names Country of birth Gender Age Day of birth Month of birth Year of birth
4	Meshblock (IDI, using full address list)	First names Last names Country of birth Gender Age Day of birth Month of birth Year of birth
5	Most recent address ID (IDI)	First names Last names First initials of first names Age Day of birth Month of birth
6	Age Last names	First names Day of birth Month of birth Gender Most recent meshblock (IDI)

Clerical linking

Clerical linking plays a critical role in the high-level linking design used by PES (prior to 2013, it was the sole approach for linking PES records with census). The 2023 PES use of clerical linking has three roles:

1. to identify and resolve any false positive links created in passes 2–4 of the automated linking
2. to identify and resolve any false negative links following the automated linking
3. to investigate any broken triangles and resolve any connected false negative and positive links.

Resolving false positive links

The use of clerical linking to identify and resolve false positives links in 2023 PES is based on 2018 PES clerical linking and additional investigation to refine and prioritise clerical efforts to where they produce the greatest value.

The 2018 PES approach classified all links resulting from automated linking into two categories based on quality of the link: near-exact and non-exact. A near-exact link is defined using a conservative definition of a match with small tolerance for errors in names and date of birth. The full-definition is shown in Appendix 12: Near-exact link indicators. A non-exact link is any link that does not meet the definition for near-exact. This information in conjunction with which pass created a link can be used to guide clerical linking towards records with higher likelihood of false positive error. Table 22 presents this information for 2018 PES linking (Stats NZ, 2023b). There is a notable reduction in the percentage of near-exact links created in passes 3 and 4 compared with passes 1 and 2.

In 2018 PES, all near-exact links were automatically accepted as true matches and clerical linking was performed on all non-exact links created in automated linking (16 percent of all created links), regardless of which pass created the link. However, investigation into the clerical outcomes for non-exact links created in pass 1 along with the quality assurance sample of the near-exact links created in pass 1 found no evidence of false positive error (Stats NZ, 2023b). Based on this evidence, the approach for 2023 PES is to automatically accept all pass 1 links as true matches. Investigation into the clerical outcomes for non-exact links created in pass 2 did find a small amount of evidence of false positive error. There was no evidence of false positive error in the near-exact links created in pass 2. Based on this evidence, the approach for 2023 PES is to automatically accept all near-exact pass 2 links as true matches, with all non-exact links from pass 2 routed to clerical linking.

Table 22: 2018 PES automatic links by linking pass and quality

Pass	Links created	Near-exact links
1	24,510	21,969 (90%)
2	2,796	2,469 (88%)
3	3,012	1,317 (44%)
4	288	96 (33%)
Total	30,606	25,851 (84%)

Investigation into the clerical outcomes for links created in passes 3 and 4 found more evidence of false positive error in both the non-exact links than in passes 1 and 2. There is a higher proportion of non-exact links in these later passes, and a higher number of false positive links. As pass 3 is the first pass where date of birth is not a blocking variable, the final two passes are likely to be more probabilistic than the first two and therefore the near-exact links may also be less accurate. To ensure that perfect linking is upheld, the approach for 2023 PES is to route all links created in passes 3 and 4 to clerical linking.

Resolving false negative links

Clerical linking to resolve false negative links is based on 2018 PES and prior. Every PES record that fails to link to a census record in automated linking will be routed to clerical linking. Each of those records will go through an extensive searching exercise to either surface missed links, or confidently confirm that record as under-coverage in the census dataset. The key change in 2023 compared with 2018 regarding false negative linking is to introduce a quality assurance sample to evaluate and report on any residual false negative links.

Triangular linking

Triangular linking is the process of using multiple pairwise links to link three datasets together. In the 2023 PES context, the three datasets of interest are:

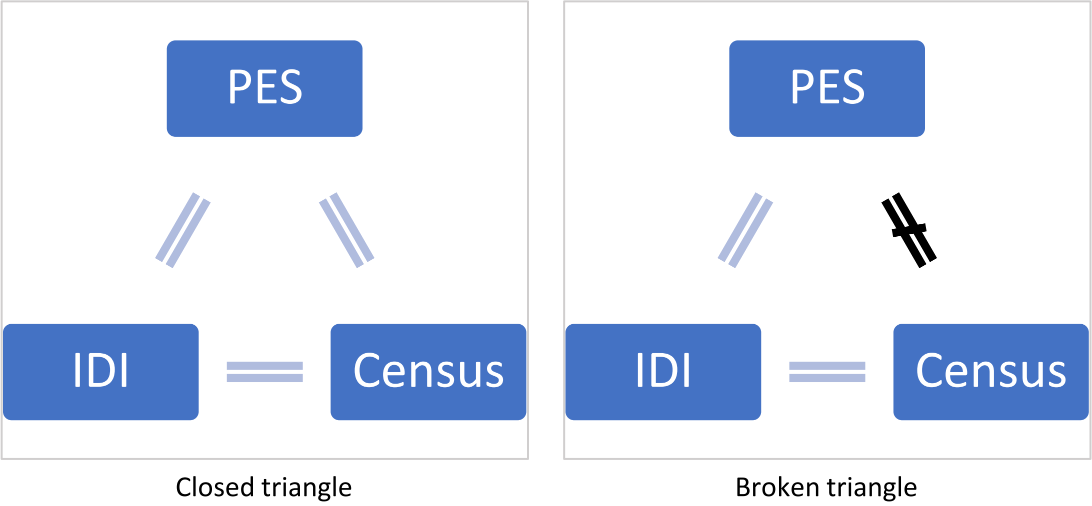
- 2023 PES responses
- 2023 Census individual dataset (responses + admin enumerations)
- IDI spine.

Throughout 2023 Census and 2023 PES processes, three pairwise links are established:

- PES–census (established by PES through automated and clerical linking)
- PES–IDI spine (established by PES through automated linking)
- Census–IDI spine (established by census through automated linking).

These probabilistic pairwise links are intransitive: a record in A links to a record in B; and the B record links to another record in C; this does NOT imply that the records in A and C will link. We can check for consistency in the links – looking for closed triangles and broken triangles. Figure demonstrates a closed triangle scenario (left) in which the pairwise links between all three datasets are consistent. The PES record links to a census record which is linked with the same IDI record as the PES is linked to Figure 4 also demonstrates a broken triangle scenario (right) in which not all the pairwise links are consistent. In the depicted scenario, the PES record has linked to an IDI record, but has not linked to the related census record – one of the pairwise links is broken.

Figure 4. Closed and broken triangles in PES linking



Both closed and broken triangles are highly informative for PES linking. A closed triangle increases confidence in the PES to census link and reduces concerns of false positive error. A broken triangle provides information and can surface both false negative and false positive linking errors. Figure 5 presents two scenarios where false negative error in the PES to census linking can be detected from broken triangles. In scenario 1 (left) the PES record does not link to a census record and is considered missed by the census. However, through automated linking, it is found to link to an IDI spine record; this IDI spine record is linked to a census record. In scenario 2 (right) PES record 1 was linked to census record 1 through automated linking. However, it is also linked to IDI spine record 1 which, in turn, is linked to census record 2. Triangular linking can evaluate whether each of these links is valid and adjust our coverage estimations to account for this initially undetected duplicate census record.

Figure 5: Examples of false negative scenarios detectable through broken triangles

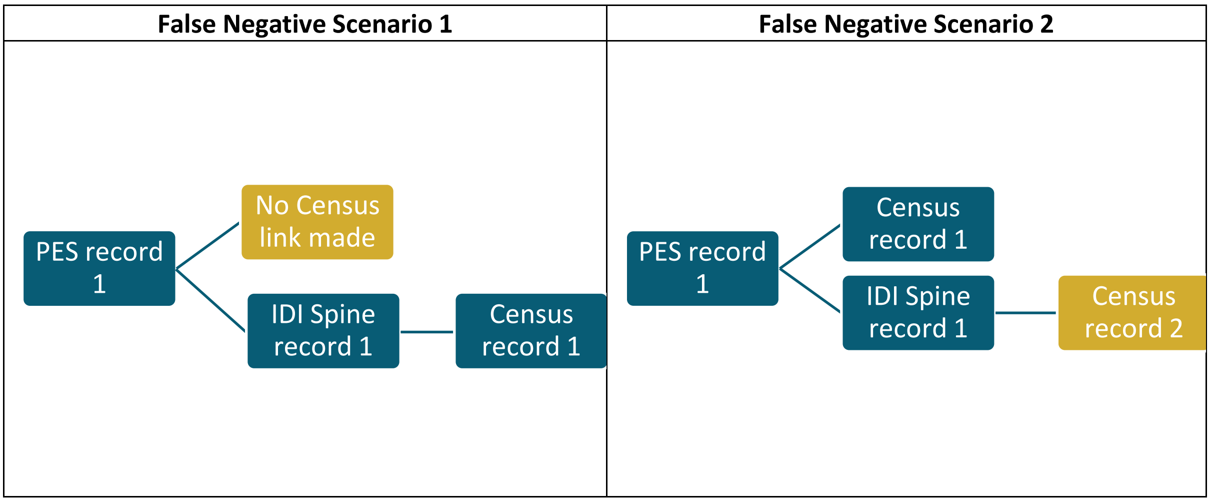
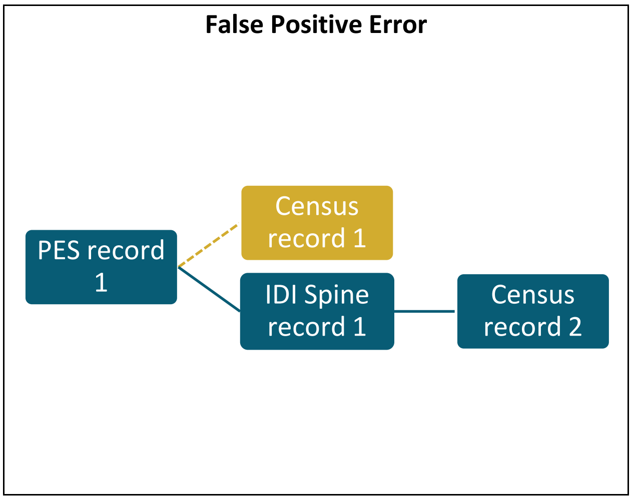


Figure 6 depicts what false positive error would look like in triangular linking. In this case, the PES focus would be detecting whether the link between PES record 1 and census record 1 was valid. However, having the additional information about the PES record 1–IDI spine record 1, and PES record 1 and census record 2 links, makes the assessment of the validity of the PES record 1–census record 1 link more robust.

Figure 6: Example of false positive error detectable through broken triangles



2023 PES linking will include clerical review of all broken triangles detectable following automated linking. Triangles will be detected prior to pairwise clerical processes beginning to reduce double-work and ensure records don't go through both pairwise and triangular clerical processes.

Linkage error and demographics

Both false positive and false negative linkage error typically affect certain demographics more than others. Analysis of 2018 PES data using the near-exact classification, automated pass information, and overall link outcomes found two relationships of interest: ethnicity, and census response mode.

Relationship between ethnicity and linking processes:

- records with Māori, Pacific, and Asian ethnic groups all had a notably low proportion of near-exact links compared with other ethnicities
- they were also more likely to have links created in later passes of automated linking (passes 3 and 4)
- they were over-represented in the unlinked records at the end of automated linking
- note – these groups are also those with higher rates of under-coverage in census.

Relationship between census response mode, response type, and linking processes:

- census records that completed a paper individual form tended to be linked in later passes or left unlinked at the end of automated linking
- fewer than 50 percent of the linked paper individual forms were near-exact links compared with 90 percent for linked online individual forms
- census records that did not complete an individual form but were listed on the household summary / dwelling form also tended to be linked in later passes or left unlinked at the end of automated linking.

These results support the need and value of having a multi-staged automated approach, supported by clerical linking.

Quality assurance

The 2023 PES approach to quality assuring the linking processes and assessing the validity of the perfect linking assumption is based in clerical review of links created / not created. The outcome of the quality assurance process is to be able to determine if the below quality targets have been met, and if each of the designed linking processes have operated as expected.

Quality targets:

- less than 0.5 percent of PES records with one link are assessed as a false positive link
- less than 0.5 percent of PES records with more than one link are assessed as having at least one false positive link
- less than 0.5 percent of PES records with no link are assessed as a false negative link
- less than 0.5 percent of PES records with one link are assessed as having a false negative link to an additional record.

Design linking processes to be assured:

- automated linking with no subsequent clerical review
- clerical pairwise linking to check for false positives
- clerical pairwise linking to check for false negatives
- clerical triangular linking that can surface both false positive and false negatives.

Table presents the relationship between the quality targets, linking processes, and the quality assurance sampling strategy. The quality assurance samples will be stratified by the link outcome scenario outlined in the quality targets, and by linking process. All automatic linking strata will have a 5 percent sample taken for quality assurance, and all other strata will have a 15 percent sample taken. The 15 percent sample of all clerical strata aligns with the sampling rate used in 2018 PES and ensures balance between representation and cost. The 5 percent sample of automatic linking strata is reduced from the 15 percent used in 2018 PES based on the results of 2018 which found no false positive links detected (Stats NZ, 2023b).

Table 23: Detailed summary of sampling strategy to assess linking quality targets

Quality target	Related processes	QA sample selection
False positive rate for records with link count =1	Automated linking Clerical pairwise – FP Clerical triangular - FP	5% sample of automated links 15% sample of clerical pairwise 15% sample of clerical triangular
False positive rate for records with link count >1		5% sample of automated links 15% sample of clerical pairwise 15% sample of clerical triangular
False negative rate for records with link count =0	Clerical pairwise – FN Clerical triangular - FN	15% sample of clerical pairwise 15% sample of clerical triangular
False negative rate for records with link count =1		15% sample of clerical pairwise 15% sample of clerical triangular 1% sample of records linked once during automated linking

		100% sample of records linked more than once during automated linking

3.09 Treatment of missing data

Following linking and prior to estimation, PES completes processes to treat any missing data fields for the PES records. The PES estimation models require all values in the input data to be complete and valid (unless otherwise accounted, for example, through an additional missingness or error models). Therefore, any records with missing values are removed from the estimation input dataset. Removing near complete records is not desirable as it decreases the size of the input sample and increases the uncertainty of our estimates.

Having completeness in our estimation variables is especially important if the missingness is not at random. This is especially problematic if records from subpopulations that are more likely to be under-covered are also more prone to missing variables. Removing the records that are missing in this example would lead to under-representation of under-covered people, and lead to a bias in our coverage estimates. Treating missingness in these variables is a crucial part of the PES estimation process and vital to producing high-quality coverage estimates of the population.

2023 PES imputation design

Throughout this section we use the term imputation as a general term referring to treatment of missing data. We use the term deterministic imputation when the missing value is set to be the same as the value observed in a linked dataset. The term statistical imputation is used when the missing value is imputed using statistical methods.

The 2023 PES approach for treating missing data uses the 2018 design as the base. Table summarises the variables that were eligible for imputation in 2018 PES, and those proposed for 2023 PES. The key updates to the variables to go through imputation are:

- shift from sex to gender
- Māori descent using the census classification is added, the electoral classification remains
- geographic location (at TALB grouping) has previously been based off the PES dwelling location, but in 2023 we aim to shift to using the respondent's usual residence location as it is better aligned with the customer needs and data uses.

Table 24: Variables eligible for imputation in 2018 and 2023 PES

Variable	2018	2023
Date of birth/age	Y	Y
Sex/gender	Y (sex)	Y (gender)
Ethnicity	Y	Y
Māori descent	Y (electoral concept)	Y (electoral and census concept)
Country of birth/NZ born indicator	Y	Y

Territorial authority and local board (TALB)	Y (PES interview address)	Y (preferred option: usual residence address)

The 2018 PES imputation methods have been extended in 2023 PES to consider the use of alternative sources. This builds on the extensive use of alternative sources in the 2018 Census and improves the 2023 PES imputation design as it will reduce the amount of statistical imputation required.

The 2023 PES imputation design uses a small number of methods, in a consistent hierarchy for most variables eligible for imputation. The exception is gender, which has a unique design.

The two methods used for 2023 PES (and previously in 2018 PES) are:

1. deterministic imputation where the missing value is set to be the same as the value observed in a linked dataset
    a. 2023 Census values (responses, alternative sources, and donor) or
    b. alternative data sources (historic census and admin data)
2. statistical donor imputation where a value is selected from a pool of similar donor records.

There are multiple sub-categories of deterministic imputation. PES records are linked to the census dataset, and to the IDI spine. These two links give us two different linked records from which we could impute missing PES data. Each of these records also has multiple sources for a given piece of data. A census record can have data collected through response, or deterministically imputed from a linked IDI spine record, or statistically imputed from a donor pool. Variable information relating to an IDI spine record can be from a range of administrative data sources depending on the variable of interest. Appendix 8 contains detailed information on the administrative sources used for each of the PES variables.

Table 25 presents a summary of the 2023 PES design for treating missing data. All variables, except for gender, have a consistent design and hierarchy of methods and sources. Deterministic imputation from census is the preferred method and all sources of census data are accepted. This is followed by deterministic imputation from admin data via the IDI. Any remaining missingness will be treated using statistical donor imputation. The imputation design for missing gender data in PES is currently proposed as deterministic imputation from census response, followed by statistical donor imputation. Alternative data sources (via the census dataset or the IDI) are not considered as there has not yet been sufficient research completed to ensure high-quality representation of all three gender categories. This research is currently being progressed with potential to inform the 2023 Census imputation design. Depending on the research outcomes, we may update the 2023 PES imputation design for gender.

Table 25: Summary of 2023 PES imputation methods and sources

Method	Linked record type	Available sources	Gender	All variables excl. gender & UR TALB
Deterministic	Census	Census response	Y	Y
		Census alternative source	N*	Y
		Census statistical imputation	N*	Y
Deterministic	IDI	Differs by variable	N*	Y

Statistical donor imputation	na	PES donor pool	Y	Y
* PES will assess the broader use of census imputed responses and administrative sources for treatment of missing gender data in the PES once the 2023 Census have finalised their design for treating missing gender data in the census.				

Change to geographic location

A key change in the 2023 PES imputation design compared with that used in the 2018 PES is geographic location used in PES estimation. In 2018, the geographic location variable represented the location of the PES interview dwelling. Consideration was given to the possibility of using the PES respondent's usual residence location in 2018 but was not progressed due to an inability to impute missing data without systematically biasing against the census under-covered population (who do not link to a census record so were not eligible for deterministic imputation).

The 2023 PES imputation design has revisited the possibility of imputing missing usual residence location – driven by the decision to link PES to the IDI spine, which provides an alternative source for deterministic imputation for any records that have not linked to census. Investigation into 2018 PES records that had their usual residence address missing showed that over half had a census match and could have their usual residence imputed from census. Of the records without a census link, 80 percent did link to the IDI spine and could have their usual residence imputed from administrative sources. For the remaining 20 percent that did not link to either census or the IDI spine, we will complete using statistical donor imputation. Given the small number of records in this group, we do not expect to introduce much bias.

3.10 Estimation

Following both linking and treatment of missing data, the set of complete PES records then go through a series of derivations (to create all information needed for estimation processes) and then on to estimation. The key outputs from estimation are produced for all population groups of interest:

- PES estimated population
- estimated gross overcount in the census dataset
- estimated gross undercount in the census dataset
- estimated net undercount achieved by census.

All four outputs are expressed as counts, with the net undercount also expressed as a percentage of the PES estimated population. Gross overcount is the number of people counted more than once in the census, along with people who were counted in the census, but should not have been counted. Gross undercount is the number of people who were supposed to be counted by the census but were not counted. The net undercount is the difference between gross undercount and gross overcount. Because the difference nearly always results in an undercount, the net undercount is a measure of how much smaller the counted census population is than the PES estimated population. Any case where the PES estimated population is smaller (that is, a net overcount) is expressed as a negative undercount. The level of coverage is the census resident total expressed as a percentage of the PES estimate of the total population. The net coverage and the net under-coverage add to 100.

This section of this document covers:

- Estimation concepts
- Derivations
- Output calculations
- Bayesian modelling framework, including model checking
- Credible intervals.

Estimation concepts

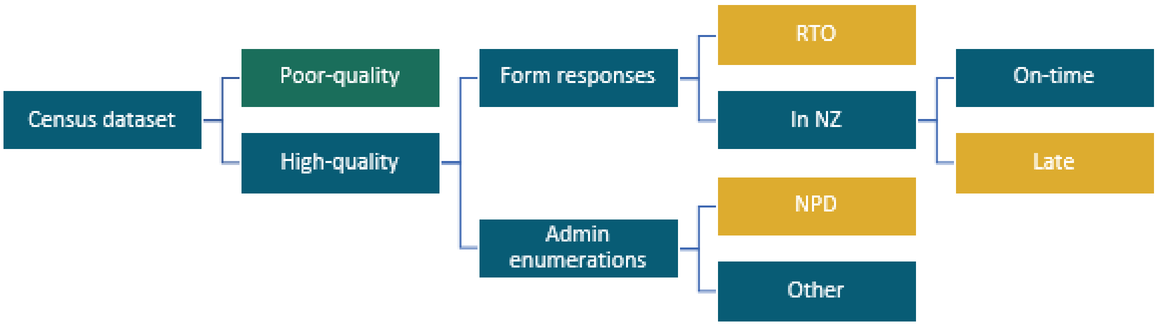
The core focus of the PES estimation design is to ensure the estimation assumptions (see section 1.2 Introduction to coverage methodology) are valid, and that any potential biases are controlled and mitigated. This relates in part to how the census file is created, and how different groups of records are treated differently throughout the estimation process. In this section we outline how and why we segment the census dataset.

The approach for 2023 PES coverage estimation aligns with that developed for 2018 PES in response to the 2018 Census dataset. Assuming no significant change to how the census dataset will be constructed, we segment the census dataset based on the following factors:

- quality of linking information – records in the census file are split into those with high-quality linking information and those without (see Link acceptance criteria and input data)
- type of record – high-quality records in the census file are split into those that come from a census form response, and those that are admin enumerations
- presence in New Zealand on census night – high-quality census form responses are split into those that were in New Zealand on census night, and those that can be identified from admin data as having been overseas on census night (residents temporarily overseas (RTO))
- time of response – high-quality census form responses for people present in New Zealand on census night are split into those that were received prior to the start of PES interviewing and those received after (late)
- type of admin enumeration – admin enumerations into prison and defence establishments are separated from the remaining admin enumerations (mostly in private dwellings or at meshblock level).

Figure 7 shows this segmentation in visual form. The blue boxes reflect the census records that are eligible for inclusion in coverage estimation. The ochre boxes show the census records that were excluded from coverage estimation for the most part, and the records in the green box were accounted for using an estimation adjustment.

Figure 7. Segmentation of the census dataset



Census eligible population

The census eligible population consists of all census records that are eligible for coverage estimation. This group is the combination of high-quality admin enumerations (excluding those in prison and defence establishments), and on-time census form responses (received prior to the start of PES field activity) for residents present in New Zealand on census night. On-time census form responses returned from non-private dwellings (NPDs) are included in this population.

Quality of linking information

Records are distinguished by the quality of their linking information in efforts to reduce the amount of linkage error between the PES and census records. Linkage error has been shown to introduce bias into coverage estimation. Records with poor quality information are treated as under-coverage throughout the estimation process since they do not have an opportunity to link. However, they are accounted for in the gross undercount calculation using an adjustment since they are not true under-coverage.

Admin enumerations in prisons and defence establishments

As part of the 2018 Census process, admin enumerations were added into prison and defence non-private dwellings, based on files of unnamed unit records provided by the Department of Corrections and New Zealand Defence Force Department of Defence,, respectively. This approach is expected to be repeated for 2023 Census. These individuals were highly unlikely to be part of the PES survey population, which excludes people living in non-private dwellings. In addition, the use of agency-supplied records means we were confident the census counts were reliable for the prison and defence populations and did not expect the same coverage patterns as for the rest of the population.

Therefore, these admin enumerations in non-private dwellings were excluded from the PES estimation process. They are included in the total population but are assumed to have no coverage error.

Residents temporarily overseas

A resident temporarily overseas (RTO) in this context is a person who usually lives in New Zealand but was overseas on census night. RTOs who responded to the census and are contained in the census dataset are identified using admin data on border movements (dates of departure from New Zealand, and dates of arrival into New Zealand). The same approach is used for identifying RTOs surveyed by PES. As we can observe over-coverage from RTOs directly through this admin data approach, we excluded all records identified as RTOs (in both the census and the PES) from coverage estimation.

The 2023 Census plans to use border-crossing data to identify census responses from RTOs and to exclude them from the census usually resident and census night population counts, removing them as a source of over-coverage. The 2023 PES retains the 2018 PES design as a mechanism for confirming the 2023 Census processes to exclude RTOs from the census dataset have been effective.

Late census returns

A late return is a census individual form that is completed and received by census on or after the date PES field activities begin. It is important to identify late returns and remove these from coverage estimation because the PES and census are required to be independent. Census forms returned after the PES went into the field may have been influenced by awareness of the PES, regardless of whether the individuals or households were in the PES sample. This would lead to a biased estimate of census coverage as the PES respondents would have a higher rate of completed census forms than the rest of the population. The problem is that we cannot distinguish between census forms returned late because of PES prompting and those returned late because of other reasons.

In general, we exclude late returns from coverage estimation and add them back into the population estimate at the end. There are two exceptions to this: late returns that contribute to over-coverage, and late returns that do not meet the requirement for high-quality linking information. Late returns that can be reliably identified as duplicate-driven over-coverage are included for over-coverage estimation as they indicate a systematic problem in census regarding resolution of multiple responses. This can occur in situations in which a census respondent provides both an on-time and a late return, or multiple late returns. Late returns with poor-quality linking information are treated the same as on-time returns with poor-quality linking information, that is they are treated as under-coverage throughout.

Derivations

In addition to preparing census data for coverage estimation, we also determine which PES records are to be included in estimation. Inclusion is based on eligibility, scope, quality of linking information, presence in New Zealand on census night, and relationship with census late returns.

To be included for estimation a PES record must meet the following:

- belong to a PES household that has participated in the PES survey (either complete response, part complete / part refusal response, or part response / unable to re-contact to complete)
- completed or partially completed an individual PES questionnaire
- have high-quality linking information.

Additional criteria for inclusion for under-coverage estimation:

- scope code of 1: usual resident of New Zealand in New Zealand on census night and during the PES period
- not have any evidence of being outside of New Zealand on census night (either via PES questionnaire or via linking to admin data).

Additional criteria for inclusion for over-coverage estimation:

- scope code of 2 or 5:
  - 2: babies born after census night
  - 5: visitor to New Zealand, usually resident overseas but in New Zealand on census night
- not have any evidence of being outside of New Zealand on census night (either via PES questionnaire or via linking to admin data).

See Appendix 9 for more details on person eligibility and scope codes.

The key variables used in estimation are given the names 'should' and 'was'.

Conceptually these are defined as:

'Should' = 1 if a PES respondent SHOULD have been counted in the census
= 0 if a PES respondent SHOULD NOT have been counted in the census

'Was' = the number of times a person WAS counted in the census
= 1 if a PES respondent WAS counted ONCE in the census
= 0 if a PES respondent WAS NOT counted in the census.

In practice, these derivations are further refined to account for PES scope and eligibility concepts, the estimation approach, and treatment of different types of census returns (such as RTOs, late returns, and admin enumerations). See Appendix 9 for detailed derivations.

From the 'should' and 'was' concepts, we can derive binary indicators (under01 and over01) for each record to indicate their inclusion in the under- and over-coverage estimation models.

Conceptually these are defined as:

Under01 = 1 if a PES respondent SHOULD have been counted in the census but WAS NOT

= 0 if a PES respondent SHOULD have been counted in the census and WAS

Over01 = 1 if a PES respondent SHOULD have been counted in the census ONCE and WAS counted more than ONCE

OR

= 1 if a PES respondent SHOULD NOT have been counted in the census and WAS

= 0 if a PES response SHOULD have been counted in the census ONCE and WAS counted only ONCE.

Output calculations

Estimated population for New Zealand

The PES estimated population for a subgroup with demographic attributes x and geographic attributes t is the sum of the coverage-adjusted census eligible population, the late census returns, and admin enumerations in prisons and defence establishments (non-private dwellings, NPDs).

$$PES\ estimated\ population\ (\boldsymbol{x},t) = census_{elig}(\boldsymbol{x},t) \times R(\boldsymbol{x},t) + census_{late}(\boldsymbol{x},t) + census_{NPD}(\boldsymbol{x},t)$$

National total is calculated as a sum of all demo-geographic subgroups:

$$Total\ PES\ estimated\ population = \sum_{\boldsymbol{x}\subset \boldsymbol{X}, t\subset T} PES\ estimated\ population\ (\boldsymbol{x},t)$$

where:

$census_{elig}$ = census count of high-quality admin enumerations (excluding those in prison and defence establishments), and on-time census form responses for residents present in New Zealand on census night (irrespective of responding in private dwelling or non-private dwelling)

R = coverage adjustment ratio produced from the combined under- and over-coverage models

$census_{late}$ = census count of late returns with high-quality linking information

$census_{NPD}$ = census count of prison and defence admin enumerations.

From the PES estimated population, we can then produce the net undercount as the difference between the PES estimation population and the census usually resident population count.

$$Net\ undercount(\boldsymbol{x},t) = PES\ estimated\ population(\boldsymbol{x},t) - census(\boldsymbol{x},t)$$

$$Net\ undercount(\boldsymbol{x},t)\ (\%) = \frac{net\ undercount(\boldsymbol{x},t)}{PES\ estimated\ population(\boldsymbol{x},t)} \times 100$$

$$\text{The PES estimate of the census level of coverage}(\boldsymbol{x},t)\ (\%) = \frac{census(\boldsymbol{x},t)}{PES\ estimated\ population(\boldsymbol{x},t)} \times 100$$

where:

$census$ = census usually resident population count.

Estimated gross undercount of the census eligible population

Only people who should have been counted by the census can potentially contribute to the gross undercount.

In terms of PES variables, anyone who has WAS = 0 and should have been counted will contribute to the undercount. Each PES record for anyone who should have been counted is assigned a binary under-coverage indicator:

Under01 = 1 if a PES respondent SHOULD have been counted in the census but WAS NOT

= 0 if a PES respondent SHOULD have been counted in the census and WAS.

The gross undercount is produced from the under-coverage model operating on the census eligible records, with an adjustment to account for records excluded from the census eligible population due to insufficient quality of linking information. This adjustment is included to account for the fact that the model estimates under-coverage for the eligible census records rather than all census records. The census records excluded from the estimation due to poor quality linking variables were, in fact, counted by census so are not considered part of the under-coverage of the full census file.

$$Gross\ undercount(\boldsymbol{x},t) = census_{elig}(\boldsymbol{x},t) \times (1 - p_{ocov}(\boldsymbol{x},t)) \times \frac{p_{ucov}(\boldsymbol{x},t)}{1 - p_{ucov}(\boldsymbol{x},t)} - census_{linkrej}(\boldsymbol{x},t)$$

$$Gross\ undercount(\boldsymbol{x},t)\ (\%) = \frac{gross\ undercount(\boldsymbol{x},t)}{PES\ estimated\ population(\boldsymbol{x},t)} \times 100$$

where:

$census_{elig}$ = census count of high-quality admin enumerations (excluding those in prison and defence establishments), and on-time census form responses for residents present in New Zealand on census night (irrespective of responding in private dwelling or non-private dwelling)

p_{ucov} = the estimated under-coverage probabilities for the $census_{elig}$ records

p_{ocov} = the estimated over-coverage probabilities for the $census_{elig}$ records

$census_{linkrej}$ = census count of records with poor-quality linking information (includes on-time returns, late returns, and non-prison or non-defence admin enumerations).

Estimated gross overcount of the census eligible population

For gross overcount, we needed to consider both the people counted more than once, and people counted when they should not have been. People may have been counted in the census more than once because they completed census forms at more than one address, someone completed an additional form on their behalf, or census processing included an admin enumeration for someone who had also completed a form.

Conversely, people may have been counted in the census when they should not have been. This can occur when people complete forms for babies born after census night, or forms are completed for people who were overseas on census night (either completed on their behalf, or by themselves before leaving or after returning).

In terms of PES variables:

- anyone counted more than once is any person who should have been counted and who was counted more than once (WAS > 1)
- anyone counted once (WAS > 0) when they SHOULD NOT have been.

Each PES record that was counted is assigned a binary over-coverage indicator:

Over01 = 1 if a PES respondent SHOULD have been counted in the census ONCE and WAS counted more than ONCE

OR

= 1 if a PES respondent SHOULD NOT have been counted in the census and WAS

= 0 if a PES response SHOULD have been counted in the census ONCE and WAS counted only ONCE.

Gross overcount is produced as the combination of the over-coverage probabilities for the $census_{elig}$ records and the over-coverage directly observed from RTOs:

$$Gross\ overcount(\boldsymbol{x}, t) = census_{elig}(\boldsymbol{x}, t) \times p_{ocov}(\boldsymbol{x}, t) + census_{RTO}(\boldsymbol{x}, t)$$

where:

$census_{elig}$ = census count of high-quality admin enumerations (excluding those in prison and defence establishments), and on-time census form responses for residents present in New Zealand on census night (irrespective of responding in private dwelling or non-private dwelling)

p_{ocov} = the estimated over-coverage probabilities for the $census_{elig}$ records

$census_{RTO}$ = census count of records identified as RTOs using admin data on border movements.

Modelling over- and under-coverage

After being linked to the census file and filtered to eligible records, the PES is used to infer census coverage. The PES provides a sample of the population, in which the presence, absence, and number of records in the census are recorded alongside geographic and demographic attributes. We modify the popular dual-system estimation (DSE) approach to estimate and correct census counts for under- and over-coverage. In a classical DSE framework where over-coverage is ignored, we would define a system shown in the matrix in table 25, where individuals are added to cells based on their presence or absence in each list.

To account for and quantify heterogeneity of inclusion in the census in New Zealand, the population is divided into categories by geographical areas, noted T, and demographic attributes (for example, age, ethnicity, and sex), noted X.

Table 26 cell counts are therefore expressed as a function of $\boldsymbol{x}$ and t, which represents any subset of $\boldsymbol{X}$ and T formed by a combination of demographic and geographic attributes. Counts $n_{00}(\boldsymbol{x}, t)$ and therefore $N(\boldsymbol{x}, t)$ are unobserved.

Table 26: Cross-tabulation of census and target population counts

		Census (L_2)		
		1	0	
PES (L_1)	1	$n_{11}(\boldsymbol{x}, t)$	$n_{10}(\boldsymbol{x}, t)$	$n_1(\boldsymbol{x}, t)$
	0	$n_{01}(\boldsymbol{x}, t)$	$n_{00}(\boldsymbol{x}, t)$	
		$n_2(\boldsymbol{x}, t)$		$N(\boldsymbol{x}, t)$

Complications arise when it is not relevant to ignore over-coverage. Given the PES is designed to detect such cases of over-coverage, we can use PES to estimate the probability of a census record to be over-coverage, $\widehat{p_{ocov}}(\boldsymbol{x}, t)$. Let $\widehat{p_{ucov}}(\boldsymbol{x}, t) = n_{10}(\boldsymbol{x}, t)/n_1(\boldsymbol{x}, t)$ be the probability of being under-coverage for an individual with characteristics $\boldsymbol{X} = \boldsymbol{x}$ and $T = t$ in the population.

An adjusted population count for such characteristics is defined as $N(\boldsymbol{x}, t) = R(\boldsymbol{x}, t) \times n_C(\boldsymbol{x}, t)$, where

$$R(\boldsymbol{x}, t) = \frac{1-\widehat{p_{ocov}}(x,t)}{1-\widehat{p_{ucov}}(x,t)}$$

is the coverage adjustment ratio applied to census counts in each demographic category $\boldsymbol{x}$ from $\boldsymbol{X}$ and each geographical area t from T.

Model development

Coverage-adjusted population estimates are required for small geographic and demographic levels, with specific outputs required for each cross-tabulation cell of single year of age, sex, ethnicity, New-Zealand born, Māori descent, and territorial authority and Auckland local board areas (TALB). To perform estimation at this granular level, it is not appropriate to use under- and over-coverage proportions observed in the PES sample as direct estimates. Instead, two Bayesian multilevel generalised linear models are used to estimate under-coverage and over-coverage probabilities for each cross-tabulation cell of interest.

In each model, binary under- and over-coverage indicators are regressed against demographic and geographic covariates. The hierarchical element of each of the models incorporates parameters relating to the PES sample design as well as geographic variables such as household, PSU, stratum, and TALB effects. Even though we are planning to use the full 2018 PES model as a starting point of our modelling framework, we will keep some flexibility in our estimation process, allowing to remove or introduce new model parameters in response to the 2023 Census performance and the 2023 PES outcomes. From each model, 1,000 probabilities for each under- and over-coverage are predicted for each of the demo-geographic combinations.

The estimated under- and over-coverage probabilities are then combined into a coverage adjustment ratio $R(\boldsymbol{x}, t)$ applied to census records for each unique combination of TALB and demographic attributes. After combining eligible census counts and the coverage adjustment ratio, the output has the form of 1,000 vectors of coverage-adjusted census counts for each demo-geographic combination.

For full technical specification of the models used in 2018 PES, see Appendix 10. Furthermore, Appendix 11 explains how the model accounts for the complex PES sample design.

Model checking framework

The model checking framework will include the following.

- Checking Stan (Stan Development Team, 2023)(modelling software) run traceplots, which show the parameter values over the iterations, to control Markov chain Monte Carlo (MCMC) run quality. If needed, the length of the MCMC chain was increased to ensure the convergence of the sampling chains.
- Calculation of Pareto Importance sampling Leave-one-out cross-validation (PSIS-LOOCV) and Pareto *k* estimates to control the quality of fits and compare resulting models (Vehtari,

Gelman, & Gabry, 2017). These measures facilitate the selection between models which use different combinations of the individual and geographical variables.

- Posterior Predictive Checks (PPC) to assess the fit of the model. Inferred models are used to generate distributions of under- and over-coverage counts using the PES dataset as an input and obtained distributions were compared with observed counts. Comparisons are conducted for the full PES dataset, as well as for the subsets of PES records representing groups defined by combinations of TALB, ethnicity, sex, and age.

Credible intervals

In Bayesian statistics, credible intervals are used to summarise the uncertainty around an estimated quantity of interest. While the goal of credible intervals is similar to the frequentist confidence intervals, the statistical definition and meaning is different and more straightforward. Credible intervals describe an interval within which the unobserved quantity of interest falls with a particular probability. For instance, a 90 percent credible interval defines a range of values within which the 'true' value has a 90 percent probability of belonging.

In Bayesian statistics, the estimation of a quantity of interest takes the form of a probability distribution (the posterior distribution) instead of a point value. Therefore, a credible interval is a summary of the posterior distribution. Following our previous example, if we want a 90 percent credible interval for a quantity we are trying to estimate, we need to find the central portion that corresponds to 90 percent of the posterior distribution. To do so, we select the 5 percent quantile (the value below which lies 5 percent of the distribution), and the 95 percent quantile (the value below which lies 95 percent of the distribution).

When producing credible intervals from our Bayesian models, we infer parameters of the under- and over-coverage models using STAN and its R interface rstan. The program implements a hierarchical Monte Carlo sampler (HMC) with the No-U-turn sampler algorithm (NUTS). We use three chains of length 8,000 and 6,000 for the under-coverage and over-coverage models. Chain length is set experimentally by increasing it until the effective number of chain samples N_{eff} and the Gelman-Rubin convergence diagnostic $\hat{R}$ reached acceptable values across all parameters, indicating appropriate mixing and convergence, respectively. The first half of each HMC chain is used for warm-up and the other half used for posterior sampling. For each of the two models, we further draw 1,000 posterior samples equally spaced in the posterior sampling part of the three chains with equal representation of each chain to obtain our final joint posterior sample of model parameters.

The joint posterior for model parameters is carried through to the calculation of $\widehat{p_{ucov}}(\boldsymbol{x}, t)$ and $\widehat{p_{ocov}}(\boldsymbol{x}, t)$ in the under- and over-coverage model, respectively. The coverage adjustment ratio $R(\boldsymbol{x}, t)$ is calculated using elementwise calculation from $\widehat{p_{ucov}}(\boldsymbol{x}, t)$ and $\widehat{p_{ocov}}(\boldsymbol{x}, t)$ posterior samples. Finally, the eligible census counts for each category is multiplied by the posterior vector $R(\boldsymbol{x}, t)$ to obtain a posterior sample of corrected census counts. After the addition of non-eligible counts (NPD and late returns) to the corrected counts, it is straightforward to extract credible intervals for the quantities of interest by computing the chosen quantiles of the 1000-long vector of posterior values.

Section 4: Adaptive design

The purpose of the adaptive design is to proactively consider the risk associated with the various statistical assumptions made in the standard, how to detect when a risk becomes an issue, and to have prepared statistical responses.

The scope of the adaptive design has been refined following lessons surfaced during the standard design work. Table outlines the work pieces considered for inclusion in the adaptive design, whether they remain in the final scope, and why they are included / excluded.

The highest priority piece of work relates to the statistical risks and assumptions embedded in the standard design. This includes the core assumptions that underpin the dual system estimation methodology, as well as additional assumptions associated with sampling.

Here are the three key deliverables included in this work piece.

1. A stocktake and risk assessment of the assumptions embedded in the standard design. This will be used to prioritise resource and focus for subsequent deliverables.
2. A set of methods and diagnostic tests that can be used to assess assumption validity. This work will focus on the assumptions with the highest risk profile. Assessment methods and diagnostic tests will include those which can be actioned against input datasets, as well as those that operate on PES estimation outputs.
3. A set of methodology options for repairing assumption validity where possible.
4. A framework for determining if / when the assumption violations are too extreme to repair using PES alone and a more significant change to the estimation framework might be required.

Table 27: Summary of work considered for scope of adaptive design

Work piece	Description	Included in scope	Rationale
Risks and assumptions	This work focuses on the statistical risks and assumptions associated with the standard design. Two key components of the work are to develop mechanisms for assessing the risks and assumptions (identifying when it shifts from a risk to an issue) and developing contingency methods for a subset of the more likely scenarios.	Yes – priority 1	This work is informed by lessons learnt in previous PES cycles and aims to shift work statistical risk management from being reactive (as in 2018 and 2013 PES) to being proactive.
Treatment of late census returns	This work investigates two different approaches for treating late census returns. 2018 PES experiences suggest there is a relationship between these records that the estimation assumptions so there is value in understanding that relationship (related to above).	TBC	This work may provide a helpful case study for the risks and assumptions work above. On its own is not a priority as the standard design approach to treating late returns is considered stable and not a source of major risk.
Gender research	This work explores the options available for producing population estimates for all three gender categories should the PES not have sufficient sample of all three categories to produce high quality estimates.	Yes – priority 2	Three-category gender is new to both 2023 Census and 2023 PES, so this is a cycle of considerable change, which comes with risk for data quality. We think it would be sensible for us to prepare alternatives to the standard design given gender is such a core demographic variable.
Uncertainty measures	This work responds to a recommendation from a 2018 PES external reviewer to explore ways of better communicating the subjective nature of balancing bias in PES estimation (see Measurement and communication of uncertainty due to estimation assumptions)	Yes – priority 3	This work is important for improving how we communicate statistical error and uncertainty in our estimates – particularly the role of bias in estimates.
Differences in reporting of demographics	This work builds on research completed as part of the standard design (see Differences in reporting of ethnicity). Completion of this research requires collaboration and partnership with external communities.	No	It is not feasible to progress this research to a conclusion in time for the outcomes to contribute to the 2023 PES methodology and outputs. This research has some overlap with the Stats NZ Census Transformation programme, so we are working through options to have it progress outside of the 2023 Census coverage project.

Investigation into census combined model components	This work involves understanding the different components of the census combined model, in particular admin enumerations into dwellings, and admin enumerations into meshblocks.	No	This work was initially planned to be completed prior to the standard design. The work was started but it quickly became apparent that it is very complex and may not yield any meaningful results. Given the low confidence in producing outcomes, this work is lower priority compared with everything else.
Investigation into usual residence concept	This work involves investigating the various conceptual and operational definitions of usual residence used by census, PES, and the ERP. The goal of the research was to understand how the different definitions relate to and impact PES.	No	This work is of interest but is not critical for standard or adaptive design for 2023 PES.

Glossary

admin enumeration – use of administrative data to add people to the usually resident census population when a census response has not been received.

causal independence – the likelihood of being recorded on one list has no relationship with the likelihood of being recorded on the other. For example, there is no structural or operational reason PES would be more or less likely to capture those records already captured by census.

census coverage (rate) – compares the number of people who were counted in the census with the number who should have been counted (as estimated by the PES).

census response (rate) – compares the number of people who responded to the census with the number who should have responded (as estimated by the PES). Historically a 'response' has been defined as the completion and return of an individual form. However, 2018 PES also produced a census response rate calculation using a minimum information definition of response as well.

clerical linking – the process of manually reviewing records not able to be linked by automatic processes and searching for missed links. In clerical linking, an analyst decides which record pairs are links and which are non-links. This process can allow for more nuance to be considered in the linking outcome than can be accounted for in automatic processes.

closed population – the population being captured in list 1 is the same as that for list 2 – no members enter or leave the population in between the two captures.

combined census model – a census in which some information on the numbers and characteristics of the population are derived from information taken from administrative data sources held for non-statistical purposes, but where other information that is not available from such sources is collected directly from individual persons and households by means of full or partial field enumeration or from other sample surveys.

Community Counts – a community-led engagement initiative being explored for use in 2023 Census. Community Counts involves Stats NZ taking direction from local representatives, influencers, and decision-makers in co-designing engagement and promotion initiatives for the census test (Stats NZ, 2021g).

confidentiality – the protection of data from, and about, individuals and organisations; and how we ensure that data is not made available or disclosed without authorisation.

desktop canvassing – the process of cleaning and updating an address list through desk-based investigation instead of field activity. Desktop staff are instructed to use a range of desk-based resources (such as satellite imagery, rates databases, internet searching) to classify addresses and record relevant information.

deterministic imputation – missing response information is obtained directly from linked information relating to the same individual.

emergency and transitional housing – forms of temporary accommodation provided to people who are in urgent need of shelter. Although designed to be short-term stays, the housing shortage in some regions has resulted in many occupants staying long-term.

erroneous inclusion – records that are not part of the target population for capture.

false negative link – two records that should have been linked because they correspond to the same unit (that is, they are a true match) but were not linked.

false positive link – two records that were linked in error and do not correspond to the same unit (that is, they are a not a true match).

homogeneity of capture – all records have the same likelihood of being captured in a given list. This is not achievable at a population level, so we use characteristics (such as age, sex) to group records until they have equal likelihoods of capture.

iwi-led collection – this initiative provides funding for Stats NZ to support its Treaty partners in building their data collection and analysis capability by assisting iwi to collect responses to the 2023 Census in two geographic areas. This initiative also aims to support the existing 2023 Census programme by improving response rates in priority groups (Government of New Zealand, 2021).

near-exact link – a conservative match requiring high degree of agreement in both name and date of birth information. For a first name to agree, we allow a single insertion, deletion, replacement, double, single, swap, two truncations, or two appendments. The conditions for last names are almost the same, except we allow only one truncation or one appendment. For date of birth to agree, we allow a single instance of a replacement, a swap, or a transposition of the day and month (if the year is the same). A near-exact link satisfies the conditions required for each variable – that is, the first name agrees, last name agrees, and date of birth agrees.

non-exact link – the links that do not satisfy the near-exact conditions are non-exact links.

overcount – people who should not have been counted by the census but were counted in error, and people who should have been counted once but were counted more than once.

over-sampling – the goal or action of sampling a particular group of interest at a higher rate than they contribute to the wider population. At Stats NZ, we use techniques such as stratification, allocation, and sampling selection methods to implicitly over-sample groups of interest. Less commonly, we may also use additional techniques such as booster panels to explicitly over-sample groups of interest.

perfect linking – error-free linking between the two lists including no missed links, and no records incorrectly linked.

PES sampling frame – a list of all private dwellings within PES geographic areas that are available for sample selection. This is the key mechanism through which PES accesses people in the PES target population.

privacy – a person’s ability to control the availability of data about themselves.

security – how an organisation protects its people, information, and assets from harm and builds resilience.

statistical donor imputation – missing response information is filled using response information from another unit, the donor. Typically, the donor is chosen in such a way that it resembles the imputed unit as much as possible on one or more background characteristics.

triangular linking – the process linking three datasets together by using three sets of pairwise links. The pairwise links can be used to create closed and broken triangles. Broken triangles can help detect linkage error in one or more of the pairwise links.

undercount – people who should have been counted by the census but were missed.

References

Brown, J. J. (2020). Discussion of 2018 PES Estimation Methodology. Unpublished document. Contact Stats NZ for access.

Chipperfield, J, Brown, J, & Bell, P (2017). Estimating the Count Error in the Australian Census. *Journal of Official Statistics*, 43-59.

Elleouet, J S, Graham, P, Kondratev, N, Morgan, A K, & Green, R M (2022). Small domain estimation of census coverage: A case study in Bayesian analysis of complex survey data. *Journal of Official Statistics*, 767-792.

Fellegi, I, & Sunter, A (1969). A theory for record linkage. *Journal of the American Statistical Association, 64*, 1183-1210.

Fisher, R A (1922). On the Interpretation of $\chi 2$ from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society, 85*(1), 87-94.

Government of New Zealand. (2021). *Iwi-led Data Collection for 2023 Census.* In Wellbeing Budget 2021: Securing Our Recovery. Retrieved from http://www.treasury.govt.nz.

Lincoln, F C (1930). Calculating Waterfowl Abundance on the Basis of Banding Returns. United States Department of Agriculture Circular(118), 1-4.

Ministry of Housing and Urban Development. (2020). *OIA response DOIA2021200488 Housing support in New Zealand.*

Ministry of Housing and Urban Development (2021). Transitional housing. Retrieved from www.hud.govt.nz

OECD (2008). Glossary of Statistical Terms. Retrieved from www.stats.oecd.org

O'Hare, W P (2019). Differential Undercounts in the U.S. Census: Who is Missed? Springer,

Petersen, C. C G (1896). The Yearly Immigration of Young Plaice Into the Limfjord From the German Sea. Report of the Danish Biological Station, 6, 5-84.

Račinskij, V, & Hammond, C (2019). Overcoverage estimation strategy for the 2021 Census of England & Wales (DOCX, 50 KB). Retrieved from https://uksa.statisticsauthority.gov.uk/

Stan Development Team (2023). RStan: the R interface to Stan. Retrieved from https://mc-stan.org/

Stats NZ (2005). Ethnicity New Zealand Standard Classification 2005 V2.1.0. Retrieved from www.aria.stats.govt.nz

Stats NZ (2017). Household Labour Force Survey sources and methods: 2016. Retrieved from www.stats.govt.nz

Stats NZ (2018a). Migration Data Transformation. Retrieved from www.stats.govt.nz

Stats NZ (2018b). Census Māori Descent v2.0.0. Retrieved from www.aria.stats.govt.nz

Stats NZ (2019a). Overview of statistical methods for adding admin records to the 2018 Census dataset. Retrieved from www.stats.govt.nz

Stats NZ (2019b). Dual system estimation combining census responses and an admin population. Retrieved from www.stats.govt.nz

Stats NZ (2019c). Linking 2018 Census respondents to the Integrated Data Infrastructure. Retrieved from www.stats.govt.nz

Stats NZ (2020a). Estimated resident population 2018: Data sources and methods. Retrieved from www.stats.govt.nz

Stats NZ (2020b). Ngā Tikanga Paihere: a framework guiding ethical and culturally appropriate data use (PDF, 1.6 MB). Retrieved from www.data.govt.nz

Stats NZ (2020c). The Five Safes framework. Retrieved from www.stats.govt.nz

Stats NZ (2020d). Statistical Standard for Ethnicity V1.0.0. Retrieved from www.aria.stats.govt.nz

Stats NZ (2020e). Post-enumeration survey 2018: Methods and results. Retrieved from www.stats.govt.nz

Stats NZ (2020f). Reviews of standards and classifications. Retrieved from www.stats.govt.nz

Stats NZ (2021a). Why do we do the Post-enumeration Survey (PES), and why is it hard? Retrieved from www.stats.govt.nz

Stats NZ (2021b). 2023 Census: High Level Design. Retrieved from www.stats.govt.nz

Stats NZ (2021c). Information, privacy, security, and confidentiality policy. Retrieved from www.stats.govt.nz

Stats NZ (2021d). Experimental administrative population census: Data sources and methods. Retrieved from www.stats.govt.nz

Stats NZ (2021e). Measuring the success of the 2023 Census. Retrieved from www.stats.govt.nz

Stats NZ (2021f). Data standard for gender, sex, and variations of sex characteristics. Retrieved from www.stats.govt.nz

Stats NZ (2021g). 2021 Census community engagement test. Retrieved from www.stats.govt.nz

Stats NZ (2022). Census Māori Descent Electoral V.1.0.0. Retrieved from www.aria.stats.govt.nz

Stats NZ (2023a). 2023 Post-enumeration Survey: High-level design. Retrieved from www.stats.govt.nz

Stats NZ (2023b). 2023 Post-enumeration Survey: Linking design. Retrieved from www.stats.govt.nz

Stats NZ (2023c). 2023 Post-enumeration Survey: Privacy impact assessment. Retrieved from www.stats.govt.nz

Vehtari, A, Gelman, A, & Gabry, J (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing, 27*(5), 1413-1432

Appendix 1: Stats NZ admin ever-resident population

Stats NZ (2019a) provides an excellent summary of the Stats NZ Integrated Data Infrastructure which is the administrative data used by both PES and census. The following excerpt is from page 9:

> **Integrated Data Infrastructure (IDI)**
>
> The IDI is a large research database which holds microdata about people and households. Data are gathered from a range of government agencies, Stats NZ surveys and the 2013 Census, and non-government organisations. The data are linked together, or integrated, to form the IDI.
>
> The basic structure of the IDI consists of a central 'spine' to which the other data collections are linked at the individual level (Black, 2016; Gibb et al, 2016). Broadly, the target population for the spine is all individuals who have **ever** been residents of New Zealand. Thus the spine comprises a reference list that is designed to include nearly all of those who interact with the range of data sources that are included in the IDI, with the exception of those who only enter New Zealand as short-term visitors.
>
> The spine is made up of the union of people in three data sources:
>
> - all births registered in New Zealand since 1920
> - all visas granted to migrants since 1997 (excluding visitor and transit visas)
> - all individuals issued with an IRD (tax) number.
>
> The IDI spine in 2018 included around 10 million individuals found in one or more of these sources. Migrants from Australia, the Cook Islands, Tokelau, and Niue do not require a visa, and are only included in the spine once they are issued a tax number. There is a low level of duplication in the IDI spine where links between these datasets have been incorrectly missed, meaning the same individual is included twice.

Appendix 2: Address list evaluation

Conceptually there are limited options available for creating a list of private dwellings in New Zealand for any Stats NZ survey (including PES). These options are:

1. create a complete list of private dwellings from scratch (without a starting address list) – **scratch frame**
2. use an externally provided list of addresses and refine through dwelling enumeration processes – **external list**
3. use an internally available list of addresses and refine through dwelling enumeration processes – **internal list.**

Table evaluates each of these options on the following criteria:

- degree of independence from sources being used by 2023 Census
- quality of data and fitness for use
    - information available to select addresses associated with permanent private dwellings
    - coverage of addresses associated with permanent private dwellings

- cost to implement for both the field test (2022) and PES proper (2023).

Table 28: Evaluation of address list options

	Scratch frame	External list	Internal list
Independence	High • any residual dependence could be due to poor quality field process	Medium • dwelling enumeration required to mitigate	Medium • dwelling enumeration required to mitigate • design work required to embed variation into the PES selection of addresses
Quality – dwellings identifiable	NA • no address list supplied	Medium • limited information available, design work required to target Stats NZ concept	High • extensive information available, design work required
Quality – coverage	Low • evidence of significant under-coverage – would require a contact-based field process to resolve • reliant on high quality field process	Medium • evidence of significant under-coverage in starting list (~20%) • reliant on high quality field process • some risk of confirmation bias	High • reliant on high quality field process • some risk of confirmation bias
Cost	High • greater field time and resource required • cannot support desktop canvassing • tool development required to support different field process	Medium • data source fee • field operation required • potential desktop canvassing required	Medium • field operation required • potential desktop canvassing required

Table summarises the key options and the implications of each.

Table 29: Advantages and disadvantages of address list options

Option description	Advantages	Disadvantages
Scratch frame	• High degree of independence if field operation results in high-quality list	• Relationship between quality and cost • High quality output requires high cost, high contact field process • Additional funding required • Tool development required • Cannot support desktop canvassing enumeration method
External list	• More perceived independence compared with internal list • Supports a range of dwelling enumeration designs (eg combination of field and desktop) • Uses existing tool functionality	• Additional source cost • Risk of unknown quality issues (such as high under-coverage) • Lack of information to support targeting private dwelling concept • Some risk of confirmation bias impacting on quality
Internal list	• Highest quality approach – uses fullest range of information available to target dwellings • Balances independence and quality • Supports a range of dwelling enumeration designs (eg combination of field and desktop) • Uses existing tool functionality • No source cost	• Design work required to determine exact methods for selecting a list of addresses • Some risk of confirmation bias impacting on quality

Appendix 3: Difficulty of finding dwellings

Work completed as part of 2018 PES investigated the idea that dwellings exist on a difficulty spectrum – with some dwellings much easier to find than others. This work explored potential predictors of where the 2018 PES dwelling enumeration had the highest undercount of dwellings compared with the 2018 Census dwelling list.

Several variables of interest were identified. Those associated with an increased likelihood of being missed by the 2018 PES (hard to find) were:

- proportion of NPDs in the area
- proportion of private dwellings associated with an NPD in the area
- proportion of many-to-one situations (multiple dwellings at one address) in the area.

Those associated with a decreased likelihood of being missed by the 2018 PES (easy-to-find) were:

- urban areas
- areas with a high proportion of addresses associated with census dwellings
- areas with a high proportion of addresses classified as residential in NPAD (New Zealand Post Address Database).

Appendix 4: Methods evaluation for adjusting for differences in reporting

Three international approaches to correcting for differences in reporting of population characteristics during population estimation were investigated and evaluated for suitability for the New Zealand PES situation. The approaches investigated were those used or under development by:

- Central Statistics Office (CSO) of Ireland (pending print)
- Office of National Statistics (ONS) of England and Wales (Račinskij & Hammond, 2019)
- Australian Bureau of Statistics (ABS) of Australia (Chipperfield, Brown, & Bell, 2017)

Additionally, an experimental approach being developed by the Stats NZ Census Transformation programme was also assessed. In all cases, the studied countries carry out a full-field enumeration census. However, they are attempting to move towards either a combined-census model or a register-based system.

Each of these methods were assessed on their viability to work with either ethnicity or location misclassification and their suitability to the NZ context. Suitability was assessed as ability to:

- work in a top-down census coverage model (holding national coverage stable while allowing sub-groups to vary)
- easily incorporated into a Bayesian framework, and
- provide information on both gross overcount and gross undercount.

Central Statistics Office of Ireland

The method developed by the CSO is fundamentally a calibrated DSE. While the uncalibrated DSE could work when there is no difference in reporting or on a national level, coverage bias is introduced when records are incorrectly reported in the wrong location. The model adjusts for this bias by calibrating the DSE.

The model assumes that one of the two lists has the correct location for all records in the list, as the model only accounts for errors in reporting in one list. The assumption is made that reporting differences between all domains do not have an impact on the total population of the country. Therefore, the number of incorrect exclusions in all domains must be equal to the number of incorrect inclusions. The calibrated DSE includes an adjustment term to remove the incorrect inclusions in a domain while the traditional DSE calculation estimates the under-coverage and the incorrect exclusions from the domain.

Office of National Statistics of England and Wales

The ONS approach treats differences in reporting of location as a form of local level over-coverage. Račinskij & Hammond (2019) describe “overcoverage within the census occurs when an element of interest (i.e. individual, household etc.) is either enumerated more than once, or enumerated in the wrong location, or does not exist, or does not belong to the target population” (p1). The ONS approach to over-coverage estimation for the 2021 Census is to estimate the probability of correct census enumeration using the following record scenarios in the estimator:

- counted once in the correct location
- duplicate record – the one in the correct location
- duplicate record – the one in the wrong location
- counted once in the wrong location (this can be disaggregated to multiple levels).

Several assumptions underpin this model:

- location reported in the Census Coverage Survey (CCS) is the correct location
- the CCS does not contain overcount records
- the census does not contain any erroneous inclusions (records outside the target population)
- the census does not contain any triplicate records or duplicate records that are both in the incorrect location.

Australian Bureau of Statistics

The ABS uses an adapted DSE method for estimating coverage and the population. Their approach goes beyond the typical DSE and has been adapted to handle differences in reporting – specifically reporting of Aboriginal and Torres Strait Islanders status (Chipperfield, Brown, & Bell, 2017). Like the ONS approach – the response reported to the coverage survey (the ABS operate a PES) is assumed correct over the census response. Accounting for difference in reporting is embedded in the over-all approach to coverage estimation. After linking to the census, each PES record is grouped into one of the following coverage outcomes:

- valid coverage: the PES and census counts a person once in the same subgroup
- duplicate: the PES and census counts a person once and twice in the same subgroup
- missed: the PES counts a person by the census does not
- misclassified: the PES and census counts a person once but in different subgroups
- duplicate + misclassified: the PES and census counts a person once in the same subgroup, and the census also counts them again in a different subgroup.

The ABS uses a generalised regression estimator with a prediction model (PReg). The inclusion of the prediction model helps account for non-ignorable non-response in the PES and reduces the bias that would result from using a standard generalised regression estimator.

Stats NZ

Stats NZ has been exploring alternative population estimation methods under the long running Census Transformation programme. This research has focused on methods for producing population estimates in the absence of traditional census data. Following the 2018 Census and the lower-than-expected response, this research was adapted to consider a DSE with 2018 Census data and a large admin population forming the two lists. There were several challenges to such a large scale DSE – the solution was a multiple component Gibbs Sampler framework incorporating Bayesian models to estimate coverage error, linkage error, and misclassification error (differences in reporting – in this application the census reported value was accepted as the true value over that obtained from admin data). The approach to misclassification was to treat the problem as a missing data problem – impute the ‘true’ value of the unobserved census variable.

This research scenario is slightly different to the PES as it is a full DSE with two full lists, rather than the rearranged ‘one-way’ DSE. However, it is possible to adapt the same approach – carrying the underlying principle to treat the differences in reporting as a missing data problem.

Table 30: Evaluation of methods approaches and suitability to the NZ context

	CSO	ONS	ABS	NZ
Top-down coverage model	No - assumes no impact at the national level from sub-group differences	Maybe – simple form operates as sub-group level and collapses up, but complex version allows for multiple levels	Uncertain – documentation is unclear.	Yes – if a reporting difference model can be incorporated into the existing PES Bayesian framework
Bayesian framework	Yes – can be adapted	Yes	No	Yes
Gross measures of reporting differences	Partly – available for sub-groups	No – reporting differences is just one contributor to a wider over-coverage model	No – reporting differences is just one contributor to a wider over-coverage model	Yes – if a separate model is developed in addition to the existing under and over-coverage models

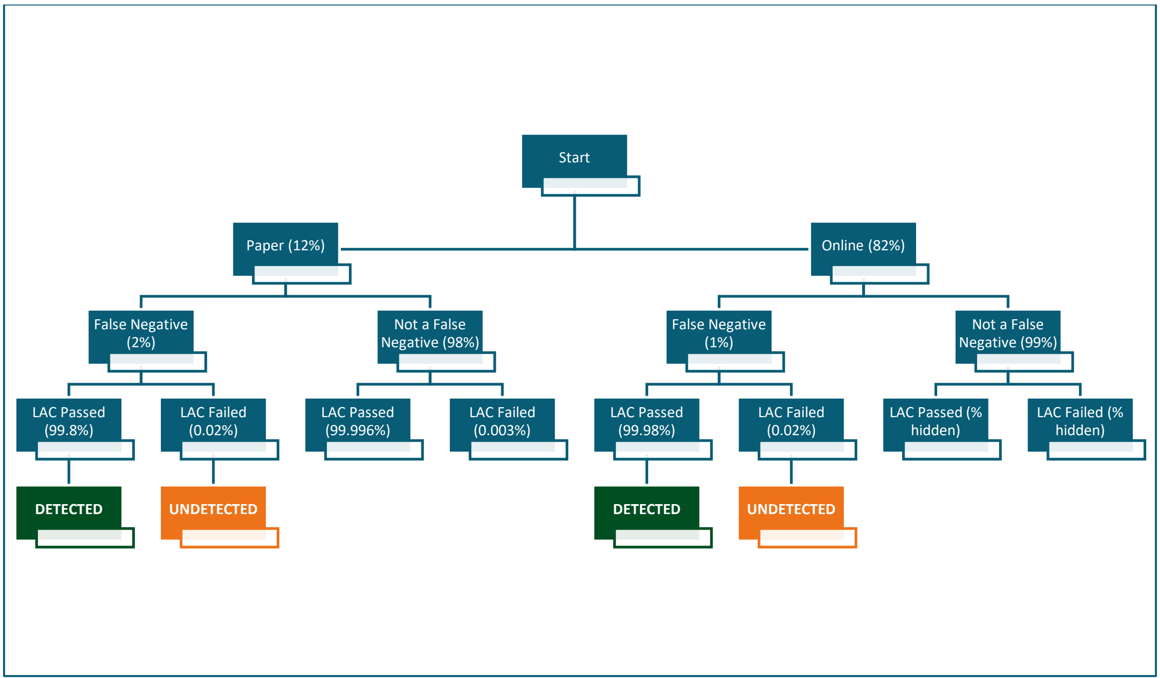
Appendix 5: Sensitivity testing bias resulting from undetected false negative duplicates

While the interaction between false negatives and the linkage acceptance criteria was negligible in 2018, we were unsure if this conclusion could be extrapolated to 2023. In particular, the use of paper forms in 2018 was much lower than what is planned in 2023. As paper forms are over-represented in the false negatives cohort, an increase in the proportion of paper census forms may result in a non-negligible number of undetectable false negatives.

The decision tree shown in was used to evaluate the impact that an increase in the proportion of paper responses would have on the rate of undetected false negatives caused by use of the linkage acceptance criteria. The probabilities at each step were taken from the 2018 data. A binomial distribution was used at each step of the decision tree:

- probability that the record is a paper form (a multinomial distribution was used for this when household summary/dwelling form records were included)
- probability that the record is a false negative
- probability that the record passes the linkage acceptance criteria.

Figure 88: Decision tree to evaluate impact of increased proportion paper forms on rate of undetected false negatives



A 1,000 iteration Monte-Carlo simulation was run for each scenario of 12 percent, 22 percent, 32 percent, 42 percent and 100 percent of all census responses being paper forms. From this it was determined what proportion of all false negatives were estimated to be undetectable at a given level of paper responses, and what number of all census responses (estimated as 5,000,000 records) were undetectable false negatives.

The decision tree in Figure 8 and results shown in Table 31 assume that only records that appear in M* and did not link to the spine are a source of false negatives. We were concerned that this process disregarded the higher rates of linkage error in household summary/dwelling forms. The results in include records that appeared on a household summary/dwelling form but did not link to the spine (set at a constant 6 percent of all responses). As these records are ineligible for inclusion in M*, we could not easily determine whether they were a false negative. However, we used information about the rates of unlinked records in this group to estimate that they were twice as likely as paper forms to be a false negative and to fail linkage acceptance criteria.

Table 31: Number of undetectable false negatives at increasing proportions of paper forms (individual forms only)

Percentage of paper forms	Rate of undetected false negatives (compared with all false negatives) (median [5%, 95%])	Number of undetected false negatives in the combined census file (median [5%, 95%])
12%	0.06 [0.04, 0.07] %	32 [23, 42]
22%	0.09 [0.07, 0.10] %	52 [40, 65]
32%	0.11 [0.09, 0.13] %	71 [57, 86]

42%	0.13 [0.11, 0.15] %	89 [75, 106]
100%	0.20 [0.18, 0.22] %	200 [179, 224]

As can be seen in , even with 100 percent of all census responses being from paper forms, the simulations estimated that there would still be a negligible amount of undetected false negatives in the census file (200 [179, 224] (*median [5 percent, 95 percent]*)). Including household summary/dwelling forms in this simulation resulted in higher error rates () with up to 0.24 percent of all false negatives being undetectable due to use of linkage acceptance criteria when 94 percent of all census responses are paper-based. However, this only translated to 235 [212, 261] (*median [5 percent, 95 percent]*) undetectable duplicates. It is more realistic that 30–40 percent of census responses will be from paper, so the overall rate is likely to be much lower than these worst-case scenarios.

Consequently, we can conclude that the interaction between false negatives within the census file and the PES use of linkage acceptance criteria will remain the source of very little coverage bias even with a significant increase in the proportion of paper forms.

Table 32: Number of undetectable false negatives at increasing proportions of paper forms (including household and dwelling forms)

Percentage of paper forms	Rate of undetected false negatives (compared with all false negatives) (median [5%, 95%])	Number of undetected false negatives in the combined census file (median [5%, 95%])
12%	0.12 [0.10, 0.15] %	80 [67, 42]
22%	0.14 [0.12, 0.16] %	99 [83, 116]
32%	0.16 [0.14, 0.18] %	118 [101, 136]
42%	0.17 [0.15, 0.20] %	136 [117, 157]
94%	0.22 [0.29, 0.24] %	235 [212, 261]

Appendix 6: Response rate calculation

The survey response rate indicates what percentage of eligible households responded to the survey. Response rates measure one component of non-response bias, which is a contributory factor in non-sampling error.

Survey response rate is calculated as:

$$Response\ Rate\ = \frac{Eligible\ responding\ households}{Eligible\ households} = \frac{D}{C\ +\ D\ +\ E\ (\frac{C\ +D}{B\ +\ C\ +\ D})}$$

Where:

- *A* = sum of selection weights of all **ineligible pre-contact** households
- *B* = sum of the selection weights of all **ineligible post-contact** households
- *C* = sum of selection weights of all **eligible non-responding** households
- *D* = sum of the selection weights of all **eligible responding** households

- E = sum of the selection weights of **all unknown eligible** households.

Appendix 7: Sample design evaluation results

Table 33. Detailed results summary of design options against desired design outcomes

	Design details				Design outcome					
					100-120	>85	>5%	>5%	>5%	>5%
Design	Base	Region group	Mode groups	Sample size	# Strata	# TALB, average of 88	Iwi-led P1 % of 171 PSU	Iwi-led P2 % of 166 PSU	Iwi-led P3 % of 194 PSU	Emergency housing % of 48 PSU
1a	Urban / Rural NZDep (HLFS groups)	Broad (12)	Mail-out Deliver + Assist	15000	112	84	5.6	4.8	6.8	5.7
1b				16500		84	6.0	5.3	7.6	6.2
1c				18000		84	6.5	5.7	8.1	6.9
2a			Mail-out Deliver Assist	15000	118	84	5.2	3.7	6.4	5.6
2b				16500		84	5.7	4.3	7.1	6.3
2c				18000		84	6.1	4.8	7.5	6.9
3a		Regional council (18)	Mail-out Deliver + Assist	15000	115	84	4.7	5.0	8.1	6.5
3b				16500		84	5.3	5.2	8.7	7.2
3c				18000		84	5.9	5.8	9.6	7.6
4a			Mail-out Deliver Assist	15000	120	84	4.7	4.7	8.2	6.4
4b				16500		84	5.2	5.3	8.5	7.3
4c				18000		84	5.9	5.9	9.5	7.9

Appendix 8: Administrative sources used in PES imputation via the IDI spine

Table 34. Administrative data sources by variable

Estimation variable	Admin sources
Age	DIA births, MBIE Visas, IR Tax
Ethnicity	DIA births, MOH NHI, MOE, MSD benefits, ACC
Māori Descent (Census Concept)	DIA births since 1998; Electoral enrolments
Country of Birth	DIA births, MBIE Border Movement
Territorial authority and local board (TALB) - Usual Residence	Addresses from MOH (NHI, PHO), IR tax, MOE schools, MSD benefits, ACC (1), NZTA

Data providers:

- Department of Internal Affairs (DIA)
- Ministry of Business, Innovation and Employment (MBIE)
- Inland Revenue (IR)
- Ministry of Health (MOH)
  - National Health Index Number (NHI)
  - Primary Health Organisation (PHO)
- Ministry of Education (MOE)
- Ministry of Social Development (MSD)
- Accident Compensation Commission (ACC)
- Waka Kotahi (NZTA)
- Electoral Commission.

Appendix 9: Estimation derivations

Scope and coverage

Scope refers to whether a person is a member of the target population. Coverage rules are there to ensure that a person has only one chance of being selected into the PES.

Someone can be selected into the PES if they are usually resident or visiting (staying in) another private residence at the time of the PES. The coverage rules ensure that if they are visiting, they cannot be also selected at their usual residence. Both scope and coverage information feed into the person scope code.

Person scope code is derived from information collected through the questionnaire. People in scope 1 are those who should have been counted in the census and meet the coverage rules in the PES. People in scopes 2, 3, and 5 are those who may have been counted (as they are New Zealand residents) but should not have been (because they were overseas on census night). Scopes 1, 2, and 5 are included in PES coverage estimation. Scope 3 is used to identify possible RTOs and excluded from processing along with scopes 4 and 6.

Table 35: PES person scope code descriptions

Person scope code	Description
1	In scope: usual resident in NZ on census night
2	Babies born after census night
3	Usual resident, overseas on census night and returned during PES period
4	Out of scope due to coverage rules
5	Visitor usually resident overseas, in NZ on census night
6	Usual resident, overseas on census night and not returned during PES period

These are the coverage rules.

1. Someone overseas on census night who usually lives in a New Zealand private dwelling will be included at their usual residence if they have returned to New Zealand by the time the PES fieldwork starts.
2. If a person is away from their usual New Zealand residence, (that is they normally live in a New Zealand private dwelling), for all the weeks covering the census period and the PES survey period then they will be included at the address where they are a visitor, if it is the first private dwelling address they have stayed at during the PES survey period.
3. If a person is not away from their usual residence for this entire period, then they will be included at their usual residence if there is someone at their usual residence during the PES survey period who will be able to proxy for this person. Otherwise, they are included at the address where they are contacted for the PES.
4. If a person is a New Zealand resident but does not normally live in a private dwelling, then they are included only at the first private dwelling they are staying at during the PES survey period.

The scope and coverage rules are intended to ensure that each person in the survey population has one chance and only one chance of being included in the PES. Rule 1 is intended to include those people who have been mistakenly counted in the census, and this will contribute to the level of overcount. Covered by rule 4 are people who usually live in non-private dwellings and people who have no fixed abode. It is expected that there will not be many people included in the PES under rule 4.

Person eligibility

Each person from a sampled dwelling where full- or part-contact has been made is assigned one of the following person eligibility codes:

2 = Ineligible post-contact (i.e. out of scope or coverage - scopes 4 and 6)

3 = Eligible non-responding (scopes 1, 3 and insufficient quality of linking information provided)

4 = Eligible responding (scopes 1, 2, 3, 5 and high-quality linking information provided)

5 = Unknown Eligibility (if the person could not be contacted, refused to take part in the survey, or had insufficient information to determine a scope code).

Person inclusion in the estimation is based on whether the person is in scope, their eligibility, and their response to the PES. Partial responding households are included, that is, if only some of the people in a household are 'eligible responding' they will be included in the estimation. We do also assign a household eligibility code based on the person eligibility codes within the household, but this is only used for calculating household response rates.

The person eligibility is derived from the scope code and person and household participation codes. These codes are completed by the interviewer and reflect the contact with the household.

Table 36: Person participation codes

1	Complete response
2	Complete refusal
3	Complete part refusal
4	Complete unable to contact
5	Complete out of scope

Table 37: Household participation codes

1	Complete response
2	Complete full refusal
3	Complete part refusal
4	Complete confirmed occupants away for survey period
5	Complete unable to contact household
6	Complete unable to re-contact household
7	Complete no interview due to illness/death/language
8	Complete all persons out on scope
9	Complete vacant dwelling
10	Complete dwelling under construction
11	Complete non-dwelling / non-private
12	Complete dwelling derelict / demolished
13	Complete non-contact due to natural disaster or health and safety

Appendix 10: 2018 PES model specifications

Under-coverage model

Each record in the PES dataset has a binary under-coverage indicator Y_{10} which states whether the record is present in the census file or absent from it. Each record is also characterised by a set of demographic characteristics $\boldsymbol{X}$ and location information. The under-coverage indicator Y_{10} is regressed against these demographic covariates in a multilevel Bernouilli-logistic model. A detailed description of demographic covariates is given in Table 38. A multilevel intercept reflects the hierarchical structure of the sampling design, where 11 whole households are sampled at random within PSUs. PSUs were selected within pre-defined socio-economic strata created based on New Zealand Deprivation indexes. Each sampled PSU is contained within a single stratum. We needed to add an additional geographical level, territorial authorities and Auckland Local Boards (TALBs), as this is the geographic resolution needed for the ERP published output. In our case the variable T refers to 88 TALBs. As each stratum can span several TALBs, the hierarchy of strata within TALBs is controlled by an occurrence matrix $\boldsymbol{W}$ where each row corresponds to a stratum and each column corresponds to a TALB. Each matrix cell $\boldsymbol{W}(s,t)$ therefore represents $Pr(T = t|strat = s)$, interpreted as the proportion of the stratum **s** population that is in TALB *t.* The idea is that TALB effects are modelled as impacting on stratum effects in a manner proportional to the share of the stratum population in each TALB. Thus if 99 percent of stratum **s** population falls into TALB **t** and 1 percent falls into TALB **u**, the stratum **s** effect will be most influenced by the effects of TALB **t** and only slightly influenced by the effect of TALB *u*. The proportions $\boldsymbol{W}(s,t)$ are estimated based on individual counts in the full census file. Therefore, what they really estimate for a given cell is $Pr(T = t \mid strat = s, census = 1)$, where "*census=1*" correspond to individuals that are included in the census file. We assume census record proportions are an acceptable approximation of total population proportions.

For a given household *h*, the nested hierarchical structure is represented by N_h^{resp}, the number of respondents within household *h*, N_p^{hhld}, the number of responding households in PSU *p*, N_s^{psu}, the number of PSUs in stratum *s*, N_t^{strat}, the number of strata in TALB *t*, and N^{ta}, the total number of TALBs. The effect of different TALBs is modelled through covariates $\boldsymbol{X_{ta}}$, which correspond to four socio-economic predictors of TALB effects that are calculated at the TALB level from New Zealand Deprivation indexes (see table 35).

For a given PES record *j* within household *h*, the model is specified as follows:

$$[Y_{10_{hj}} | \mathrm{p}_{ucov_{hj}}] \overset{\text{indep}}{\sim} Bernoulli(\mathrm{p}_{ucov_{hj}}); j = 1, \dots, N_h^{resp}; h = 1, \dots, N_p^{hhld}$$

$$logit(\mathrm{p}_{ucov_{hj}}) = \alpha_{h[i]}^{hhld} + \beta \boldsymbol{X}_{hj}$$

$$\left[\alpha_h^{hhld} \middle| \alpha_{psu[h]}^{psu}, \sigma_{hhld}^2\right] \sim Normal\left(\mu + \alpha_{psu[h]}^{psu}, \sigma_{hhld}^2\right); h{:}\, psu[h] = p; p = 1, \dots, N_s^{psu}$$

$$[\alpha_p^{psu} | \alpha_{strat[p]}^{strat}, \sigma_{psu}^2] \sim Normal(\alpha_{strat[p]}^{strat}, \sigma_{psu}^2); p{:}\, strat[p] = s; s = 1, \dots, N_t^{strat}$$

$$[\alpha_s^{strat} | \alpha^{ta}, \sigma_{strat}^2] \sim Normal(W \quad \alpha_{ta}, \sigma_{strat}^2), s{:}\, ta[s] = t; t = 1, \dots, N^{ta}$$

$$[\alpha_t^{ta} | \beta_{ta}, \sigma_{ta}^2] \sim t_3(\boldsymbol{X_{ta}} \beta_{ta}, \sigma_{ta}^2), t = 1, \dots, N^{ta}$$

where notation such as $psu[h]$ refers to the psu to which household h belongs with similar conventions holding for $strat[p]$ and $ta[s]$.

Parameter priors are selected to be weakly informative to allow for both regularisation and occasional high covariate effects and variances. We choose $Cauchy(0,2.5)$ as a prior for μ. Covariate effects β and β_{ta} are drawn from a $Normal(0,1)$. Multilevel variances

$\sigma^2_{ta}, \sigma^2_{strat}, \sigma^2_{psu}, and\ \sigma^2_{hhld}$ are drawn from $Cauchy^+(0{,}2.5)$. The notation t_3 corresponds to a Student's t-distribution with 3 degrees of freedom.

Table 38: Covariates used in the under-coverage and over-coverage models
Only a subset of covariates are used in the over-coverage model (see 'Used in ocov.' column). The column 'n. of param' shows the number of parameters estimated for each of the covariates and covariate interactions.

Variable	Coding	Description	n. of param	Used in ocov.
individual covariates				
sex	binary	0=male, 1=female	1	N
Age	10 age splines	quadratic age splines with 8 knots at ages 10,20,30,40,51,61,71,81	10	Y
Ethnicity		Independent variables with same limit on number of ethnicities as census form		
Māori	binary		1	Y
Pacific	binary		1	Y
Asian	binary		1	Y
Other Ethnicity	binary		1	Y
NZ born	binary	0=born abroad, 1=born in New Zealand	1	Y
Māori descent	binary	0=non-Māori descent, 1=Māori descent	1	N
Covariate interactions				
Māori * Other Ethnicity	binary		1	N
sex* age	binary	sex and all 10 age splines	10	N
Ethnicity * age	binary	5 first age splines with each ethnicity, as well as with Māori * Other Ethnicity (3-way interaction)	25	N
TALB covariates (X_{ta})				
communication	continuous	Access to internet communication	1	N
income	continuous	Average income	1	N
qualification	continuous	Average level of professional qualification	1	N
internet response	continuous	proportion of online responses among all census responses	1	N

Once the logistic model is fitted and the joint parameter posterior sampled, each of the 1,000 draws from the posterior can be used to predict the under-coverage probability associated with each combination of covariate values that exist in the census. However, we need to integrate over parameters related to the sampling design (household, PSU, and stratum effects) to obtain a prediction at the TALB level, by age sex and ethnic group. For each combination of covariate values $X = x$ within a TALB t, and each sample of the parameter vector θ we have

$$
\begin{aligned}
&Pr(Y_{10} = 1|ta = t, \boldsymbol{X} = \boldsymbol{x}, \theta) \\
&\quad = \sum_{s=1}^{N_t^{strat}} Pr(Y_{10} = 1|ta = t, \boldsymbol{X} = \boldsymbol{x}, strat = s, \theta) Pr(strat = s|ta = t, \boldsymbol{X} = \boldsymbol{x}) \\
&\quad = \sum_{s=1}^{N_t^{strat}} Pr(Y_{10} = 1|, \boldsymbol{X} = \boldsymbol{x}, strat = s, \theta) Pr(strat = s|ta = t, \boldsymbol{X} = \boldsymbol{x}) \\
&\quad = \sum_{s=1}^{N_t^{strat}} \int Pr(Y_{10} = 1|\alpha_{hhld}, \boldsymbol{X} = \boldsymbol{x}, strat = s, \theta) p(\alpha_{hhld}|\boldsymbol{X} = \boldsymbol{x}, strat = s, \theta) d\alpha_{hhld} \\
&\qquad \times Pr(strat = s|ta = t, \boldsymbol{X} = \boldsymbol{x}) \\
&\quad = \sum_{s=1}^{N_t^{strat}} \int expit(\mu + \alpha_{hhld} + \boldsymbol{x}\beta) N\left(\alpha_{hhld}|\alpha_s^{strat}, \sigma_{hhld}^2 + \sigma_{psu}^2\right) d\alpha_{hhld} \\
&\qquad \times Pr(strat = s|ta = t, \boldsymbol{X} = \boldsymbol{x}),
\end{aligned}
$$

where $expit(.)$ is the inverse logit function and $N(.|\mu, \sigma^2)$ is the normal density function with mean μ and variance σ^2. $Pr(strat = s|ta = t, X = x)$ is estimated using an occurrence matrix constructed the same way as $\boldsymbol{W}$, using the census file to approximate the population proportion of strata in each TALB. Again here, $Pr(strat = s|ta = t, census = 1)$ is used to approximate $Pr(strat = s|ta = t, \boldsymbol{X} = \boldsymbol{x})$. After integration, we obtain $Pr(Y_{10} = 1|ta = t, \boldsymbol{X} = \boldsymbol{x}, \theta)$ which is the marginal posterior of $\widehat{p_{ucov}}(\boldsymbol{x}, t)$ as needed for the calculation of the coverage adjustment ratio $R(t, \boldsymbol{x})$. Two TALBs were not included in the PES estimation: Chatham Islands Territory and Great Barrier Local Board Area. For records within these TAs as well as records from areas outside of TALBs, α_{hhld} was drawn from $N(0, \sigma_{hhld}^2 + \sigma_{psu}^2)$, the rationale being that as strata do not cover these TALBs, no sample was taken from them for the PES model estimation, a non-informative mean of 0 is appropriate. For Chatham Islands Territory and Great Barrier Local Board Area, we could have drawn α_{hhld} from $N(\alpha_t^{ta}, \sigma_{hhld}^2 + \sigma_{psu}^2)$ but this would give a disproportionate importance to their values of covariates $\boldsymbol{X_{ta}}$, which are extreme for these TALBs and were not used in the model fitting. By using 0 as a mean for the normal distribution, we assign them a coverage adjustment appropriate for their demographic structure in a 'typical' TALB.

Over-coverage model

The over-coverage regression was performed for the PES dataset without under-coverage records using the same modelling strategy as used for under-coverage. First, we created the over-coverage indicator Y_{01}, which was assigned to each PES record: 1, if the person was counted more than once or included in the census by mistake, otherwise 0. The small number of over-coverage records in the PES dataset required a reduction of the number of variables used in the regression model. The initial model run was performed with the same geographical hierarchy as the under-coverage (TA, strata, PSUs, households) and all individual covariates described in the under-coverage section. Nevertheless, calculated Pareto *k* estimates suggested that the resulting model may have been overfitting the dataset. Therefore, further exploration of model variables was performed to choose the ones with significant effect on over-coverage.

The geographic component of the model was reduced to contain only TALB-level, where all individuals from the same TALB were assigned the same TALB-specific intercept. Included individual covariates were 4 ethnicity variables, 10 age-splines and a binary NZ-born indicator – all as per under-coverage model (see Table 38). No interactions between covariates were specified.

The model is specified as follows:

$$\left[Y_{01_{tj}} \middle| p_{ocov_{tj}}\right] \overset{indep}{\sim} Bernoulli\left(p_{ocov_{tj}}\right), j = 1, \dots, N_t^{resp}; t = 1, \dots, N^{ta}$$

$$logit\left(p_{ocov_{tj}}\right) = \alpha_t + \boldsymbol{X}_{tj}\beta$$

$$[\alpha_t | \sigma_t^2] \overset{indep}{\sim} N\,(\mu, \sigma_t^2)$$

, where N_t^{resp} is the number of PES respondents in the t^{th} TALB, $Y_{01_{tj}}$ a response for j^{th} individual in t^{th} TALB, $\boldsymbol{X}_{tj}$ – corresponding vector of individual covariate values, β – vector of coefficients corresponding to individual parameters, σ_t – between TALB standard deviation, α_t – TALB-specific intercept, μ – general intercept.

The following priors were specified for the model:

$$\mu \sim Cauchy(0, 2.5)$$

$$\beta \sim N(0, 1)$$

$$\sigma_t \sim Cauchy^{+}(0, 2.5)$$

Calculations of over-coverage probabilities were performed, according to the following equation for a person with individual covariate values x in TALB t and sampled model parameters θ:

$$\Pr(Y_{01} = 1 |\, \boldsymbol{X} = \boldsymbol{x}, ta = t, \theta) = expit(\alpha_t + \boldsymbol{x}\beta)$$

For the people outside TALB, probabilities were predicted as below:

$$\Pr(Y_{01} |\, \boldsymbol{X} = \boldsymbol{x}, \theta) = \int expit(\mu + \alpha_t + \boldsymbol{x}\beta) N(\alpha_t | 0, \sigma_t^2) d\alpha_t$$

, where $N(.\,|0, \sigma_t^2)$ – normal distribution density function with mean 0 and variance σ_t^2.

As for under-coverage, 1000 over-coverage probabilities were predicted for each of individual variable and TALB value combination found in the census, to be used for the adjustment-ratio estimation.

Appendix 11: Incorporating the sample design into the estimation models

The PES has a complex survey design, and the final estimation models take into account the sample stratification and over-sampling of some groups. There are three components in the modelling process which allow us to account for the sample design.

1. Weights are applied to the census records
2. Household and PSU-level covariates
3. Strata – TALB integration.

In the classic design-based estimation, weighting, and calibration is done for the counts observed in the sample. In our case, we apply a correct coverage adjustment ratio (weight) to every census record, and then the aggregation over census records adds the values up in the right proportion.

PSU-level covariates help us to account for the sampling mechanisms like proportional to size sampling of some target groups. We also may introduce household level covariates to account for some heterogeneity of the coverage patterns (this was not done in the 2018 PES model but was investigated as a possible tool). The household and PSU level covariate values are included in the prediction at the individual level, and the correct weighting of these effects is achieved in the same way to the individual effects – at the step of the census adjustment. The proof of it is detailed in Appendix B of (Ellequet, Graham, Kondratev, Morgan, & Green, 2022).

In the model, household-level and PSU-level random effects are collapsed to the strata effects. Then, the strata are weighted by the proportion of their occurrence in TALBs ($\Pr(strat = s|TA = t)$) as shown in the Equation 10 in (Elleouet, Graham, Kondratev, Morgan, & Green, 2022). Therefore, the TALB layer of the model includes post-stratification to account for the sample design.

Appendix 12: *Near-exact* link indicators

Near exact links

For a first name to agree, we allow:

- One of the following:
    - insertion
    - deletion
    - replacement
    - double
    - single
    - swap.
- Two of the following:
    - truncations
    - appendments.
- **That the name is three or more characters in length.**

For a last name to agree, we allow one of the following:

- insertion
- deletion
- replacement
- double
- single
- swap
- truncation
- appendment.
- **That the name is three or more characters in length**

For date of birth to agree, we allow one of the following:

- replacement
- swap
- transposition of the day and month (if the year is the same).

A *near-exact* link satisfies the conditions for first name, last name, and date of birth. A non-exact link does not satisfy the *near-exact* conditions.