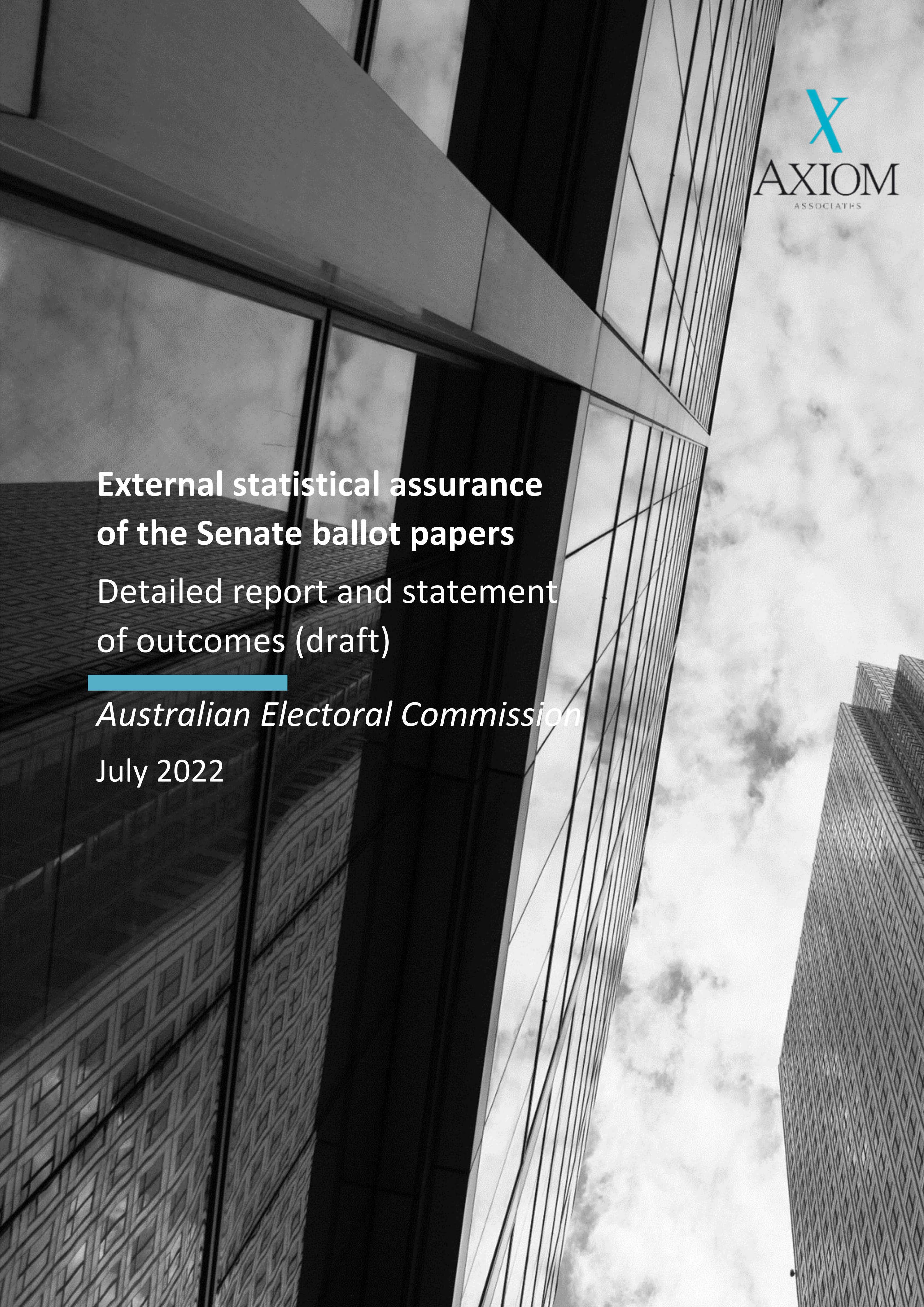


AXIOM
ASSOCIATES

External statistical assurance of the Senate ballot papers

Detailed report and statement of outcomes (draft)

Australian Electoral Commission

July 2022

Contents

Statement of Responsibility

Axiom Associates (Aust) Pty Ltd documents and reports may be relied upon by the Australian Electoral Commission (AEC) for purposes only pursuant to the terms of our engagement outlined in the Services Agreement Contract Number: SRC001297 dated December 2021 and any subsequent amendments. We disclaim all responsibility to any other party for any loss or liability that the other party may suffer or incur arising from or relating to or in any way connected with the contents of our documents or reports, the provision of our documents or reports to the other party or the reliance upon our documents or reports by the other party.

Liability limited by a scheme approved under Professional Standards Legislation.

Document Control

Version Number	Date Provided	Detail
1.0	30 June 2022	Initial draft copy provided to the Australian Electoral Commission for management comments
1.1	1 July 2022	Track changes to the report
1.2	5 July 2022	Report updated to reflect AEC feedback and statistics updated
1.3		

1. STATISTICAL ASSURANCE ACTIVITIES PERFORMED

LEGISLATIVE CONTEXT FOR BALLOT PAPER ASSURANCE

1.1 The *Electoral Legislation Amendment (Assurance of Senate Counting) Act 2021* (the Act) sets out the requirements for ballot paper sampling assurance throughout computerised scrutiny of votes in Senate election.

1.2 Section 273AC requires the Electoral Commissioner to arrange for statistically significant samples of ballot papers to be checked throughout the scrutiny of votes for the election to assure that the electronic data used in counting the votes reflects the data recorded on the ballot papers.

1.3 The ballot paper sampling process is not part of the scrutiny in relation to the election.[1] The ballot paper sampling process may be inspected by the scrutineers at the counting centre where the scrutiny is being conducted.[2]

1.4 Before the polling day for the election, the Electoral Commissioner must publish on the Electoral Commission's website: a methodology to be used for the ballot paper sampling process; and the process to be used for reconciling preferences.[3]

1.5 Within 14 days after the return of the writ for the election, the Electoral Commissioner must publish on the Electoral Commission's website a statement setting out the outcomes of the ballot paper sampling process.[4]

ASSURANCE PROCESSES UNDERTAKEN

1.6 On 16 May 2022, the Australian Electoral Commission (AEC) approved the methodology and operational approach that would be used to deliver the external statistical assurance of the 2022 Senate ballot papers. The assurance process would consist of five components: sample selection, assurance testing, quality assurance, AEC review of potential exceptions and report creation.

1.7 The following sections outline how the approved methodology was implemented for the 2022 Senate ballot paper assurance activities and any deviations from the agreed methodology. It also details any risks of these deviations to the statistical assurance outcomes.

SAMPLE SELECTION

1.8 AEC determined that sample selection must occur at the Central Senate Scrutiny (CSS) centres in each state and territory. A physical ballot paper was at the CSS and this was compared to the electronic data used by the AEC in the counting of the votes (preference data recorded in AEC's Ballot Paper Reconciliation System (BPRS)).

1.9 This approach to sample selection was based on the flow of 'physical ballot paper to electronic data record'. As a result, the assurance was designed to conclude on whether, for each ballot paper tested, there were matching electronic data used in counting the votes.

1.10 The assurance was not designed to conclude on the completeness of the number of physical or electronic ballot papers or the accuracy of the count. The approach to assurance testing was not designed to

[1] Subsection 273AC(4) of the *Electoral Legislation Amendment (Assurance of Senate Counting) Act 2021*.
[2] Subsection 273AC(5) of the *Electoral Legislation Amendment (Assurance of Senate Counting) Act 2021*.
[3] Subsection 273AC(6) of the *Electoral Legislation Amendment (Assurance of Senate Counting) Act 2021*.
[4] Subsection 273AC(7) of the *Electoral Legislation Amendment (Assurance of Senate Counting) Act 2021*.

detect if there were electronic data records held by the AEC that were not associated with a physical ballot paper or whether there were multiple electronic data records associated with a single physical ballot paper.

1.11 The risk of the introduction of additional electronic records and data into AEC systems is the subject of the requirements of section 273AA of the Act related to assurance of security of computer systems for scrutiny of votes in Senate election. This is to be reported separately by the Electoral Commissioner.

1.12 In November 2021, the AEC received advice from the Australian Bureau of Statistics (ABS) to determine the number of ballot papers to be selected for statistical assurance purposes.[5] This advice identified the number of ballot papers nationally to be selected to provide statistical assurance and the breakdown of the number of ballot papers to be selected in each CSS state and territory location. The ballot paper assurance process used the sample sizes identified by the ABS. In all CSS state and territory locations, the assurance process oversampled ballot papers by a small number to allow for any contingencies such as where a batch was rescanned after testing.

1.13 The sample sizes are summarised in **Table 1**.

Table 1: Ballot paper assurance process sample sizes

CSS state and territory location	ABS ballot paper sample size [a]	Actual ballot paper sample size in 2022
NSW	1,733	1,735
VIC	1,377	1,400
QLD	1,272	1,310
WA	1,272	1,285
SA	1,200	1,245
TAS	1,106	1,135
ACT	977	1,028
NT	958	965
Total national sample size	**9,895**	**10,103**

Note a: ABS advice to AEC on sampling methodology, November 2021

1.14 The ABS also identified that cluster sampling of ballot papers would assist with the practical implementation of the assurance process. The approach to cluster sampling would allow the AEC to test five ballot papers from a single batch of 50 ballot papers. This approach was used to reduce the number of batches of ballot papers to be subject to assurance testing.

1.15 In each CSS state and territory location the sample was selected throughout the scanning process. Sample selection was based on schedules of extraction times to select ballot paper transport containers (BPTCs) which had just been scanned. The sampling approach was to use the first batch in each of these selected BPTCs and to test five ballot papers in that batch. BPTCs were not tested until BPRS indicated that the status of the BPTC was ‘confirmed’, indicating that data had been transmitted and reconciled by the AEC.

1.16 The extraction schedules were developed by Axiom to align to the proposed Fujifilm scanning shift schedules provided to the AEC prior to the election. These sampling extraction schedules were updated during the course of scanning to reflect changes in the Fujifilm scanning shift schedules. Due to the late notification of some changes to scanning shifts (not scanning on some planned days, shorter shifts and extended shifts), the sampling extraction schedules were not fully aligned to Fujifilm scanning shifts and extractions could not always

[5] Refer to Appendix A of this document for a full copy of the ABS advice.

be performed as planned. This deviation from the agreed methodology is considered to be low risk as assurance testers independently determined when to select the BPTCs for testing.

1.17 Due to a number of unplanned staff absences in the Northern Territory CSS location, extractions were not performed in line with the extraction schedules and as a result, there was a risk of significant under sampling in the Northern Territory. To overcome this risk and deviation from the agreed methodology, Axiom instructed the Northern Territory assurance staff to extract the sample by selecting two batches from each BPTC selected and testing five ballot papers in each batch. In effect, selecting 10 ballot papers from a single BPTC rather than five as specified in the approved methodology. This allowed for confidence that assurance testing was being performed on BPTCs that had been extracted immediately after scanning. This deviation from the agreed methodology is considered to be low risk because the independence of sample selection was maintained and the principles of the cluster sampling (five in a batch) was preserved.

ASSURANCE TESTING

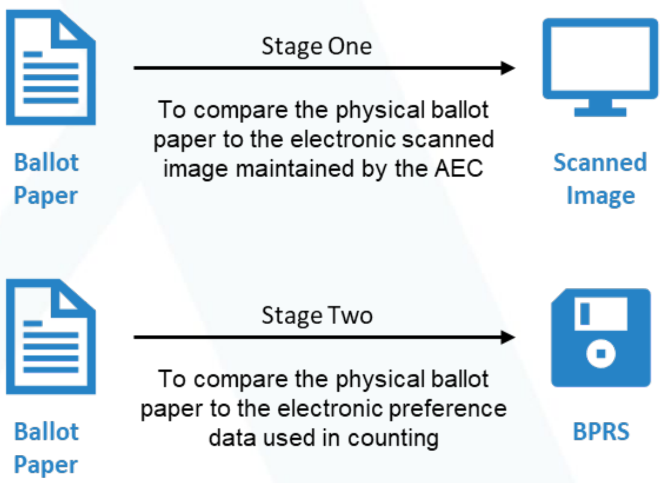
1.18 Assurance testing was undertaken using a two-stage testing approach.

1.19 The first stage tested the physical ballot paper to the scanned image transmitted to the AEC and included in BPRS. The first stage testing assessed the completeness and accuracy of the physical ballot paper to the scanned image. It was used to validate that all marks on the scanned image were an exact replica of the physical ballot paper.

1.20 The second stage tested the physical ballot paper to the electronic data used in counting, contained in BPRS. The second stage testing assessed whether:

- the electronic preference data file was an accurate reflection of the preferences recorded on the physical ballot paper; and
- interpretation of voter intent aligned with AEC business rules as stated in the Ballot Paper Formality Policy and Ballot Paper Formality Guidelines. **Figure 1** outlines the two-stage testing approach.

Figure 1: Two-stage ballot paper assurance testing approach



1.21 The ballot paper assurance testing was designed to detect the following:

- where the scanned image was an incomplete representation of the physical ballot paper (the marks on the scanned image were not an exact replica of the physical ballot paper); and
- where preferences recorded on the physical ballot paper did not match the electronic data in BPRS which was used in counting.

1.22 The assurance testing approach was designed to identify potential exceptions impacting on statistical assurance results. The types of potential exceptions the assurance was designed to identify included[6]:

- where no scanned image that exactly replicates the physical ballot paper can be found;
- voter disenfranchised where a vote was recorded as informal when the voter expressed a valid preference(s);
- above the line voter preference sequence was broken too early within the first six preferences;
- above the line voter preference sequence was broken too early after the first six preferences;
- below the line voter preference sequence was broken too early within the first twelve preferences;
- below the line voter preference sequence was broken too early after the first twelve preferences;
- voter preference sequence was broken too late (additional preferences were recorded); and
- vote recorded as formal when it should have been informal.

1.23 Assurance testing was performed by Temporary Election Workforce (TEW) resources selected and engaged by the AEC to act as assurance testers. To maintain independence of the assurance activities performed, Axiom received representations from the AEC that no assurance testers had performed non-assurance activities in the relevant CSS state and territory location for the 2022 Senate election.

1.24 Assurance testers were trained by Axiom on the assurance approach, detailed methodology, operational approach and expectations prior to the election. For the Northern Territory, the assurance testers were also provided with supplementary training. The AEC recruited additional assurance testers after the election in the Northern Territory and Victoria. These additional assurance staff were also provided with training.

1.25 Any questions or issues related to the application of the assurance methodology and results were required to be referred to, and answered by, Axiom staff.

1.26 With one exception, Axiom staff performed their activities remotely. The assurance testers in the Victorian CSS location identified a number of complex potential exceptions which required an Axiom staff member to visit the CSS location to assist with resolution of the exceptions.

QUALITY ASSURANCE

1.27 Ballot paper assurance testing was subject to quality assurance. Quality assurance activities consisted of re-performance of 30% of stage two testing by different assurance testers.

1.28 Quality assurance activities assessed whether:

- the electronic preference data file was an accurate reflection of the preferences recorded on the electronic scanned image of the ballot paper; and
- interpretation of voter intent aligned with AEC business rules as stated in the Ballot Paper Formality Policy and Ballot Paper Formality Guidelines.

[6] For the 2022 Federal Election, voter instructions on the Senate ballot paper were to number a minimum of six preferences above the line or twelve preferences below the line, across all states and territories. These instructions were used as the cut off for the classification of exceptions.

1.29 In addition, all assurance test sheets were reviewed by Axiom staff and all potential exceptions were confirmed and collated by Axiom. The Axiom review of test sheets and the quality assurance processes did not indicate issues with the quality and consistency of testing performed by any of the assurance testing teams.

AEC REVIEW OF POTENTIAL EXCEPTIONS

1.30 Stage Two potential exceptions frequently arise because of judgements, by Fujifilm and AEC staff, about the interpretation of preferences recorded on ballot papers. These judgements are intended to be determined through the application of the AEC's Ballot Paper Formality Policy and Ballot Paper Formality Guidelines.

1.31 All potential exceptions from Stage Two testing were subject to review by three experienced AEC officers. The purpose of this AEC review was to confirm how AEC experienced officers would apply the AEC business rules for interpretation of voter intent. The AEC chose four experienced officers to perform the AEC review of potential exceptions. The AEC review of potential exceptions was performed by three of the four AEC officers. Three tranches of potential exceptions were sent to and reviewed by the AEC officers as follows:

- tranche one sent on 17 June 2022 to officers 1, 2 and 3;
- tranche two sent on 22 June 2022 to officers 1, 2 and 3; and
- tranche three sent on 28 June 2022 to officers 1, 2 and 4 (to accommodate leave taken by AEC officer 3).

1.32 For each potential exception identified in Stage Two testing by the ballot paper assurance process, the experienced AEC officers were asked to independently 'blind test' the relevant ballot paper and to assess how the voter intent should be interpreted and recorded in BPRS. Where:

- all three AEC experience officers considered that the application of the AEC business rules would result in an interpretation of the ballot paper that differed to that which had been recorded in BPRS, an exception was recorded in the final assurance results; and
- in all other cases, the potential assurance exception was not recorded in the final assurance results.

1.33 Results from the AEC review of potential exceptions were assessed and collated by Axiom staff.

REPORT CREATION – STATISTICAL ANALYSIS

1.34 All exceptions identified from the assurance process (including AEC review of potential exceptions) were collated at the state and territory level and at the national level by Axiom staff. Axiom followed the ABS' guidance for calculating, analysing and reporting the statistical conclusions that may be drawn from the 2022 Senate external statistical assurance activities.[7]

1.35 The assurance results and conclusions have been expressed as state and territory and national exception rates. As the ballot paper assurance approach uses a different sampling rate for each state and territory location, the national exception rate was derived using the weighted number of exceptions in each state and territory location relative to the state and territory proportion of the national vote for 2022.

1.36 Stage One exceptions were not attributed to individual state or territory locations. One element of the ABS methodology was based on the AEC assertion that scanning processes were homogenous across Australia and the methodology provided for a single national sample for the assessment of Stage One exceptions.

[7] Australian Bureau of Statistics, *ABS advice to AEC on sampling methodology*, November 2021, p. 4.

2. FINDINGS, STATISTICAL ANALYSIS AND STATEMENT OF ASSURANCE OUTCOMES

2.1 An analysis of the findings of the Senate ballot paper sampling process has been provided below. Exceptions impacting on the statistical assurance results were considered in each stage of assurance testing.

For Stage One of assurance testing, which compared the physical ballot paper to the scanned image transmitted to the AEC and included in BPRS, a total of seven exceptions were identified. These exceptions were not attributed to individual state or territory locations when calculating the statistical exception rate in order to align the analysis to the agreed methodology based on advice from the Australian Bureau of Statistics. The seven Stage One exceptions are described in **Table 2**.

Table 2: Stage One exceptions identified during assurance testing

CSS state and territory location	Description of the exception
NSW	Two physical ballot papers in a sampled batch could not be matched against scanned images in that batch. The scanned images in the batch included duplicate images for two other ballot papers in the batch.
NT	The scanned image of one physical ballot paper did not reflect the preferences recorded on the physical ballot paper. The physical ballot paper had a light ‘smudge’ over one preference. This ‘smudge’ did not obscure the number entered into the box on the physical ballot paper. The scanned image enhanced the ‘smudge’ to extent that it was not possible to distinguish the preference number on the scanned image of the ballot paper and consequently it was not recorded in BPRS.
VIC	One physical ballot paper in a sampled batch could not be matched against scanned images in that batch. The batch contained 51 physical ballot papers and only 50 scanned images were associated with that batch. Three physical ballot papers in a sampled batch could not be matched against scanned images in that batch. Despite efforts to locate the associated images, these could not be found.

2.2 For Stage Two assurance testing, which compared the physical ballot paper to the electronic data used in counting and contained in BPRS, a total of 33 exceptions were identified. These exceptions related to the accuracy of the electronic preference data file (used in counting) and the interpretation of voter intent. Determination of voter intent should be in line with the AEC business rules as stated in the Ballot Paper Formality Policy and Ballot Paper Formality Guidelines.

2.3 The 33 exceptions have been classified in accordance with the approved assurance methodology. The classification of exceptions has been outlined in **Table 3**. An additional classification of exceptions has been included in the table below to reflect the assurance findings. The categorisation of statistical exceptions now includes where an incorrect preference type (including related preferences) was admitted to the count.

Table 3: Categorisation of exceptions related to the interpretation of voter intent

	CSS state and territory Location									
Categorisation of exceptions used in statistical analysis	**NSW**	**VIC**	**QLD**	**WA**	**SA**	**TAS**	**ACT**	**NT**	**Total**	**% of total**
Voter disenfranchised where a vote is recorded as informal when the voter expressed a valid preference(s)	2	2	1	1		1	2		9	27%
Above the line voter preference sequence was broken too early within the first six preferences	2	2		1	1	2			8	24%
Above the line voter preference sequence was broken too early after the first six preferences		1	1		1				3	9%
Below the line voter preference sequence was broken too early within the first twelve preferences		1		1				1	3	9%
Below the line voter preference sequence was broken too early after the first twelve preferences						1			1	3%
Voter preference sequence was broken too late (additional preferences were recorded)	1				1	1			3	9%
Vote recorded as formal when it should have been informal	1	2			2				5	15%
Incorrect preference type accepted	1								1	3%
Total exceptions related to the interpretation of voter intent	**7**	**8**	**2**	**3**	**5**	**5**	**2**	**1**	**33**	**100%**

2.4 Of the exceptions used in the statistical analysis, 27% related to where a voter was disenfranchised from the election. Another 24% related to where the ATL voter preference sequence was broken too early within the first six preferences. 15% of exceptions related to the counting of an informal vote (i.e., where vote was informal and should have been excluded from the count).

2.5 Stage Two exceptions frequently arose because of the inconsistent application of the AEC's business rules, by either Fujifilm or AEC officers. Each assurance exception reflects an assessment of the marks on a ballot paper that has been identified as being inconsistent with the AEC Ballot Paper Formality Policy and/or the AEC Ballot Paper Formality Guidelines. There was no suggestion of any political or logical bias in the exceptions. The effect of including the exceptions in the count would have resulted in, some increase and /or some reduction to some candidates in the number of preferences counted.

2.6 Similarly, the exceptions identified did not suggest any systematic basis for including or excluding preferences or ballot papers. Many exceptions were the result of the need for interpretation where voters did not follow numbering instructions (including x's and ticks and words rather than numerals) or where numbering sequences for preferences were not clearly written on the ballot paper.

2.7 The statistical analysis of the results of the assurance testing was carried out in line with the ABS methodology. The AEC required the presentation of the statistical results to be in the form of a sample mean with a 95% confidence interval with upper and lower bounds. The proportion of exceptions identified in the sample tested was extrapolated to the population of Senate ballot papers used in the AEC Tally Room. **Table 4** outlines the statistical analysis and results. The results were analysed to provide 95% confidence over the range of possible exceptions in the populations (CSS state and territory location and national).[8] This analysis indicates that there is a 95% likelihood that the true number of exceptions within the populations is within the stated ranges in the table.

[8] The confidence intervals for 95% confidence are intended to provide assurance that, all things being equal, if an additional 100 samples from the population were tested, that for 95 of the 100 samples, the sample the mean would lie in the stated range.

Table 4: Statistical analysis and results

	Statistical sampling characteristics							Statistical sampling outcomes [b]					
CSS state and territory location	Actual number of ballot papers [a]	Location as a % of national total	Assurance sample size	Stage 1 exceptions [c]	Stage 2 exceptions	Exception rate including national Stage 1 exceptions [c]	Extrapolated number of exceptions	AEC confidence level	Upper exception rate as a %	Upper limit: highest number of ballot papers that are 95% likely to have an exception	Lower exception rate as a %	Lower limit: lowest number of ballot papers that are 95% likely to have an exception
NSW	4,996,110	32.1%	1,735		7	0.473%	23,619	95%	0.796%	39,769	0.150%	7494
VIC	3,960,958	25.4%	1,400		8	0.641%	25,378	95%	1.059%	41,947	0.223%	8833
QLD	3,111,034	20.0%	1,310		2	0.222%	6,905	95%	0.477%	14,840	0.000%	0
WA	1,571,899	10.1%	1,285		3	0.303%	4,759	95%	0.604%	9,494	0.002%	31
SA	1,162,472	7.5%	1,245		5	0.471%	5,474	95%	0.851%	9,893	0.091%	1058
TAS	372,973	2.4%	1,135		5	0.510%	1,901	95%	0.924%	3,446	0.096%	358
ACT	290,308	1.9%	1,028		2	0.264%	766	95%	0.577%	1,675	0.000%	0
NT	106,907	0.7%	965		1	0.173%	185	95%	0.434%	464	0.000%	0
National	15,572,661	100.0%	10,103	7								
Totals	**15,572,661**		**10,103**	**7**	**33**	**0.443%**	**68,988**	**95%**	**0.572%**	**89,076**	**0.314%**	**48,898**

Note a - Actual number of ballot papers was sourced from AEC BPRS on 30 June 2022 (Ballot papers on Tally Room)
Note b - Confidence interval calculations from https://www.abs.gov.au/websitedbs/d3310114.nsf/home/sample+size+calculator
Note c - National (Stage 1) exceptions are attributed to each CSS location in line with their proportion of the national vote

2.8 The analysis of the Senate assurance results indicates that there is a 0.45% exception rate in the assurance sample.[9] The AEC can be 95% confident that, nationally, the number of exceptions in 1,000 ballot papers is between 3.2[10] and 5.8. This can be analysed at a CSS state and territory location level. These have been listed below:

- NSW – The AEC can be 95% confident that there are between 1.5 and 8 exceptions in 1,000 ballot papers.
- VIC – The AEC can be 95% confident that there are between 2.3 and 11 exceptions in 1,000 ballot papers.
- QLD – The AEC can be 95% confident that there are between 0 and 4.8 exceptions in 1,000 ballot papers.
- WA – The AEC can be 95% confident that there are between 0 and 6.1 exceptions in 1,000 ballot papers.
- SA – The AEC can be 95% confident that there are between 0.9 and 8.6 exceptions in 1,000 ballot papers.
- TAS – The AEC can be 95% confident that there are between 0.9 and 9.3 exceptions in 1,000 ballot papers.
- ACT – The AEC can be 95% confident that there are between 0 and 5.8 exceptions in 1,000 ballot papers.
- NT – The AEC can be 95% confident that there are between 0 and 4.4 exceptions in 1,000 ballot papers.

[9] This can also be expressed as follows, for 99.55% of ballot papers in the sample there was matching data in the AEC systems for the count, in addition the AEC can be 95% confident that the true population value is between 99.42% (99.55%-0.129%) and 99.68%. (99.55%+0.129%), or 99.55% +/- 0.129%.
[10] All numbers have been rounded to the next whole decimal point.

APPENDICES

APPENDIX A: MARKS ON BALLOT PAPERS NOT RECORDED ACCURATELY IN BPRS WITH NO IMPACT ON THE COUNT

The Stage Two assurance testing compared the physical ballot paper to the electronic data used in counting, contained in BPRS. The second stage testing assessed whether:

- the electronic preference data file was an accurate reflection of the preferences recorded on the physical ballot paper; and
- the interpretation of voter intent aligned with AEC business rules as stated in the Ballot Paper Formality Policy and Ballot Paper Formality Guidelines.

The Stage Two assurance testing identified six instances where BPRS did not accurately record preferences. These findings are described in **Table 1** and **Table 2** below. In five of these cases there was no impact on formality or the recording of valid preferences.

Table 1: Incorrect recording of preferences in BPRS – with impact on the accuracy of the preferences counted

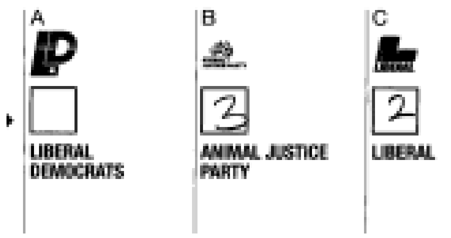
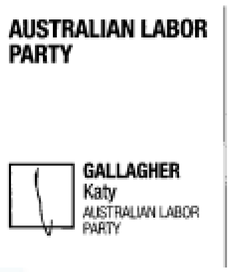
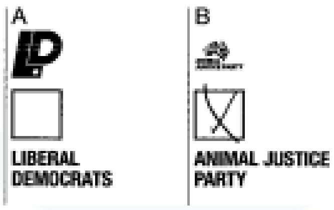
CSS state and territory location	Description of exception	Screen shot of the relevant section of the scanned image of the ballot paper (ballot paper reference number)
SA	A physical ballot paper identified a third preference in column B and a second preference in column C. Data Entry 2 identified the third preference as column C. BPRS did not accurately record these preferences. BPRS recorded a second preference in column B and a third preference in column C. This instance has been included as an exception in the statistical results.	 (27966SA-SPEN032517060160301001)

Table 2: Incorrect recording of preferences in BPRS – with no impact on formality

CSS state and territory location	Description of exception	Screen shot of the relevant section of the scanned image of the ballot paper (ballot paper reference number)
ACT	BPRS did not record a first preference recorded Below the Line (BTL) on the ballot paper. This had no impact on formality as there was not a complete sequence of six preferences recorded BTL. Above the Line (ATL) preferences were accurately recorded in BPRS for this ballot paper.	 (27966ACTCANB010100360046501005)

CSS state and territory location	Description of exception	Screen shot of the relevant section of the scanned image of the ballot paper (ballot paper reference number)
SA	A ballot paper had six crosses recorded above the line. BPRS recorded seven crosses. BPRS recorded a cross in column A where there was no cross recorded on the ballot paper.	 (27966SA-BOOT018203960019201001)
NT	A ballot paper included nine number one preferences. BPRS recorded eight number one preferences and one number seven preference. BPRS inconsistently recorded the numbers. Audit logs in BPRS indicate this ballot paper was reviewed by AEC.	 (27966NT-LING030600310015101001)
NT	A ballot paper included a range of ticks and crosses. No ticks or crosses were recorded in BPRS. Audit logs in BPRS indicate this ballot paper was reviewed by AEC.	 (27966NT-SOLO030701640024301005)
SA	A ballot paper included a range of ticks BTL. Not all of these ticks BTL were recorded in BPRS. The ticks outside preference boxes were not recorded in BPRS.	 (27966SA-STUR019015400078701039)

The AEC may wish to look at the processes and controls over data entry and review by Fujifilm and AEC to ensure more thorough consideration of whether there is a match between the ballot paper and the BPRS data entry.

APPENDIX B: ABS ADVICE TO AEC ON SAMPLING METHODOLOGY

Attached is a copy of the ABS advice to AEC on sampling methodology.

Australian Bureau of Statistics

ABS advice to AEC on sampling methodology

Executive Summary

The Australian Electoral Commission (AEC) has requested advice from the ABS to determine the number of ballots for assurance as part of the elections for the Australian Senate. The number of ballots that are manually checked for errors should be sufficient to demonstrate with a high level of confidence that the possible national error rate is low.

The ABS recommends that Senate ballots should be assured at the following rate:

- 1 in 3,000 ballots in New South Wales and Victoria;
- 1 in 2,500 ballots in Queensland;
- 1 in 1,250 ballots in Western Australia;
- 1 in 1,000 ballots in South Australia;
- 1 in 350 ballots in Tasmania;
- 1 in 300 ballots in Australian Capital Territory;
- 1 in 120 ballots in Northern Territory.

Based on these rates, it is estimated that 9,895 ballots will be assured nationally for the 2021/22 Senate election. A state breakdown is provided in *Table 1*:

This assurance approach will provide a high level of confidence in confirming that the national error rate and error rates in each of the states and territories is low.

In comparison with the internal AEC assurance approach implemented in 2019, the proposed allocation delivers a higher confidence in the national error rate, while requiring fewer ballots to be assured. The proposed approach also allows ballot assurance to be undertaken while processing. This is helpful to speed up the assurance.

Background

The Senate assurance process implements two stages of ballot testing. The first stage of testing checks that the scanned image matches the physical ballot paper. The second stage checks that the scanned image of the ballot paper matches the extracted data file, i.e. that the preferences from the scanned image match the datafile that is used to run the preference allocation process.

An assurance of the 2019 Senate election found **no errors** during the first stage at ballot testing. The national estimate of the proportion of errors during the second stage of ballot testing is **0.45%**. The calculation of the national error rate is discussed here.

The emphasis of this report is to determine an appropriate allocation to assurance for stage 2 errors. Given that no stage 1 errors were detected as part of the 2019 assurance from a sample of 1,368, it is evident that the true stage 1 error rate is very low. For the purposes of stage 1 testing, it should be sufficient to assurance 1 in 10 of the ballots selected for stage 2 testing. The practical implementation is discussed here.

Recommended Allocation

This section details the recommended allocation and diagnostics associated with it Alternate allocations were considered and informed the final recommended allocation. See Appendix.

The allocation utilised the following assumptions.

- While the 2019 assurance indicated that the prevalence of stage 2 errors differed by state, the difference between the state and national proportion of errors was not statistically significant, with the exception of the ACT, which had no errors detected.[1] Therefore, the calculated national stage 2 error rate of **0.45%** was assumed in each state.
- An estimate of 16.095 million Senate forms nationally for the 2021/22 election. The distribution of form by state as provided by the AEC – see Table A1.

The main criterion implemented for designing the target number of ballots to assurance by state was to have 99% confidence that the observed error rate in the sample for each state will be less than 1%, assuming that an error rate of 0.45% (as estimated in 2019) applies for the full population of senate votes.

The minimum sample size to achieve this is to select 828 ballots in each state and territory – see Appendix for details.

The recommended allocation places sample beyond this minimum value into each state. This is a conservative approach to ensure we have enough sample to meet the accuracy targets, and it produces round numbers for the sampling skips to be used, simplifying the implementation of this proposal. It also helps to ensure robustness. The sample allocation will remain statistically valid if the actual number of Senate ballots in a particular state or the error rate differs slightly from what has been assumed.

Table 1: Number of ballots to assure for stage 2 error by state

State	Estimated Forms 2021/22	Estimated Ballots assured (stage 2)	Assurance Rate (1 in X ballots)	95% confidence limit for maximum error rate	99% confidence limit for maximum error rate
NSW	5,200,000	1,733	3,000	0.72%	0.83%
VIC	4,130,000	1,377	3,000	0.75%	0.88%
QLD	3,180,000	1,272	2,500	0.77%	0.89%
SA	1,200,000	1,200	1,000	0.77%	0.91%
WA	1,590,000	1,272	1,250	0.77%	0.89%

2

[1] The 2019 assurance found zero errors in ACT, during stage 2 testing. Consequently, there is over 95% confidence that the true ACT stage 2 error rate is less than the national stage 2 error rate. The national second stage error rate is applied to ACT in the interests of simplicity and to ensure that ACT is not under-allocated.

TAS	387,000	1,106	350	0.79%	0.92%
NT	115,000	958	120	0.81%	0.96%
ACT	293,000	977	300	0.81%	0.95%
AUS	16,095,000	9,895		0.59%	0.65%

Testing conclusions

Based on the observed error rates from the 2019 assurance and the sample sizes in each state the following statistical statements could be made.

- If there is a 0.45% error rate found in the assurance sample, then the AEC can be 95% confident that nationally, there are less than 6 errors per 1,000 ballot papers in the Senate scanning process. It is also true that if the true error rate in the population is 0.45%, then the AEC can be 95% confident that the error rate estimated from the assurance sample will be less than 6 errors per 1,000 ballot papers.
- Similarly, there is 99% confidence that nationally there are less than 6.5 errors per 1,000 ballot papers.
- In any given state, there is 99% confidence that there are less than 10 errors per 1,000 ballot papers.

These statistical statements are illustrative only. They are based on the assumption of a true error rate of 0.45% in the population to give confidence on the size of the estimated error rate from the sample; or similarly on the assumption of an error rate of 0.45% in the assurance sample to give confidence in what the error rate is for the full population. Final confidence intervals will depend on the **actual error rates found** during the 2021/22 assurance.

Comparison with 2019 assurance approach

It is instructive to compare the proposed assurance approach with the assurance approach previously implemented in 2019.

First, it is noted that the total expected number of ballots to assurance (9,895) is slightly lower than in 2019 (10,400).

Secondly, rather than assuring a constant number of ballots in each state, the proposed allocation is assurances of more ballots in the more populous states and less ballots in the less populous states.

Increasing the number of ballots assured in the more populous states allows the proposed allocation to deliver a higher confidence in the national error rate, while assuring a smaller number of ballots.

Third, it is specified to assure at a constant **rate** in each state, rather than a fixed total number of ballots. This is efficient to allow ballots to be assured while processing is ongoing, rather than having to wait for all ballots to be processed before commencing assurance.

Practical implementation of assuring

The AEC arranges senate ballots into bundles of 50. From a logistical perspective, it would be more efficient to first select a number of bundles and then select more than one ballot from each bundle.

Furthermore, selecting bundles at a constant rate allows assurance to be undertaken while processing is ongoing – as it will not be necessary to have every bundle processed for assurance to commence.

This is known as clustered sampling of the ballots. Clustered samples can lead to lower accuracy if errors can also be clustered together, i.e. if errors are not evenly spread across all bundles. We have suggested an approach that we believe balances the risk to accuracy from using a clustered sample with the benefits that it provides, i.e. reducing the number of bundles that need to be selected for the assurance sample. The allocations provided in *Table 1* have already allowed for some 'slack' by selecting more ballots than strictly necessary to obtain a precise national estimate of the stage 2 error.

We propose the assurance selects a certain proportion of 'bundles' (e.g. 1 in every 300 bundles in NSW) and then to select 1/10 of all ballots in the bundle for stage 2 testing (so that overall 1 in every 3,000 ballots is selected in NSW).

Once ballots have been selected for stage 2 testing, select 1 in every 10 of the stage 2 sample for stage 1 testing.

If the sampling rate from *Table 1* is adopted, then the process is described below in *Table 2*.

Table 2: Number of forms to assure by state

State	Estimated Forms 2021/22	Estimated Bundles 2021/22	Assurance Rate (1 in X bundles)	Estimated Bundles selected	Estimated Ballots assured (stage 2)	Assurance Rate (1 in X ballots)	Estimated Ballots assured (stage 1)
NSW	5,200,000	104,000	300	347	1,733	3,000	173
VIC	4,130,000	82,600	300	275	1,377	3,000	138
QLD	3,180,000	63,600	250	254	1,272	2,500	127
SA	1,200,000	24,000	100	240	1,200	1,000	120
WA	1,590,000	31,800	125	254	1,272	1,250	127
TAS	387,000	7,740	35	221	1,106	350	111
NT	115,000	2,300	12	192	958	120	96
ACT	293,000	5,860	30	195	977	300	98
AUS	**16,095,000**	**321,900**		**1,979**	**9,895**		**989**

Calculating the national error rate

If an assurance approach uses a different sampling rate in different states, then in order to calculate the national error rate, it is important to weight the number of errors found in each state by the state's proportion of the national population.

Table 3: 2019 assurance calculation of national error rate

State	Total Senate ballots 2019 (formal + informal)	Proportion of national total	Stage 2 errors 2019	Stage 2 sample 2019	Error rate	Estimated total errors
NSW	4,905,472	32.3%	7	1,300	0.54%	26,414
VIC	3,896,236	25.7%	6	1,300	0.46%	17,983
QLD	2,999,372	19.8%	6	1,300	0.46%	13,843
SA	1,134,556	7.5%	5	1,300	0.38%	4,364
WA	1,497,532	9.9%	4	1,300	0.31%	4,608
TAS	365,272	2.4%	6	1,300	0.46%	1,686
NT	108,994	0.7%	2	1,300	0.15%	168
ACT	276,651	1.8%	0	1,300	0.00%	0
AUS	**15,184,085**				**0.45%**	**69,065**

The error rate in each state is estimated by dividing the number of errors in each state by the assurance sample size. For example, in NSW the assurance for 7 errors from a sample of 1,300, giving an error rate of 0.54%. An error rate of 0.54% would mean that there is a total of 26,414 errors from the full population of 4,905,472 votes in NSW.

After calculating the estimated number of total errors in each state they can be added to produce an estimate of total number of errors in Australia. This total is 69,065 based on the 2019 assurance results.

Dividing the estimate of 69,065 errors by the total national votes of 15,184,085 gives the estimated national error rate of 0.45%.

An alternate approach to calculate this national error rate is to multiply the error rate in each state by the proportion of votes in that state. This gives:
(0.323 x 0.0054) + (0.257 x 0.0046) + (0.198 x 0.0046) + (0.075 x 0.0038) +
(0.099 x 0.0031) + (0.024 x 0.0046) + (0.007 x 0.0015) + (0.018 x 0)
= 0.0045.

Appendix

Table A1: Estimated senate forms by state for 2021/2022 Senate Election – source AEC

State	Estimated Senate Forms
NSW	5,200,000
VIC	4,130,000
QLD	3,180,000
SA	1,200,000
WA	1,590,000
TAS	387,000
NT	115,000
ACT	293,000

Table A2: number of stage 2 errors by state – 2019 Senate assurance – source AEC

State	Stage 2 errors 2019 assurance	2019 Error rate
NSW	7	0.54%
VIC	6	0.46%
QLD	6	0.46%
SA	5	0.38%
WA	4	0.31%
TAS	6	0.46%
NT	2	0.15%
ACT	0	

Alternate allocations

This section outlines various allocation options that were considered, that informed the final recommended approach. These options are presented for technical background and can be skipped.

The allocation described in *Table 1* represents the ABS' main recommendation.

Option A1: Allocation using a constant national sample rate

The first option considered is to apply a constant assurance rate across each state nationally. This would differ from the assurance process from 2019, which assured a constant number of ballots (1,300) in each state as part of stage 2 testing.

The advantages of applying a constant sample rate nationwide, is that it would allow the same assurance procedure to be applied in each state. Furthermore, the estimate of the national error rate would be easier to interpret as no weighting would be required.

The disadvantage of applying a constant sample rate is that the smallest states would have relatively few ballots assured. This would result in a less confidence in the estimate of the state error rate.

Sample allocations

Table A3 shows the national level of accuracy associated with different sample sizes, while applying a constant sample rate nationally.

Table A3: National sample size vs 95% margin of error of estimate

Scenario	National sample size	1 in Rate	One-sided 95% confidence level	One-sided 99% confidence level
A	10,400	1,548	0.56%	0.61%
B	5,810	2,770	0.60%	0.66%
C	6,438	2,500	0.59%	0.65%

Scenario A represents the national sample size that was used for stage 2 testing as part of the 2019 assurance. Scenario B represents the minimum national sample size to be 95% confident that the national error rate is less than 0.6%.

From a practical perspective, it would make sense to use a larger sample size than this. Scenario C represents this, using a 'round' sample rate of 1 in 2,500 dwellings for each state.

Table A4: Number of forms to assurance by state by scenario

State	Estimated Forms 2021/22	Scenario A	Scenario B	Scenario C
NSW	5,200,000	3,360	1,877	2,080
VIC	4,130,000	2,669	1,491	1,652
QLD	3,180,000	2,055	1,148	1,272
SA	1,200,000	775	433	480
WA	1,590,000	1,027	574	636
TAS	387,000	250	140	155
NT	115,000	74	42	46
ACT	293,000	189	106	117
TOTAL	16,095,000	10,400	5,810	6,438

It is evident that if precisely estimating the national error rate is the key objective, than the sample rate required can be significantly lower than what was applied in 2019 (Scenario A).
It is also clear that this approach results in a relatively small number of ballots being sampled in Tasmania, Northern Territory and Australian Capital Territory.

Option A2: Allocation with maximum state margin of error (MOE) constraint

A notable disadvantage of applying a fixed sampling rate across all states is that the number of ballots assured in the smaller states is low. This will result in wide confidence intervals for the state level estimates of proportion of errors in smaller states/territories.

The following two allocations examine the number of ballots required to be assured in each state in order to be 95% or 99% confident that the true state level error rate would be less than 1%

Table A5 : state assurance size required to be 95/99% confident that the true error rate < 1%

State one-sided confidence interval	95%	99%
State sample	413	828
National 95% confidence interval bound	0.71%	0.64%
National 99% confidence interval bound	0.82%	0.71%

Therefore, the state allocation to be 99% confident that the observed error rate is less than 1% in each state (assuming a 0.45% error rate in the population) is as in *Table A6*.

Table A6 : State sample size and rate to be 99% confident that the assurance error rate is less than 1%

State	Estimated Forms 2021/22	State sample	State sample rate (1 in X)
NSW	5,200,000	828	6,280
VIC	4,130,000	828	4,988
QLD	3,180,000	828	3,841
SA	1,200,000	828	1,449
WA	1,590,000	828	1,920
TAS	387,000	828	467
NT	115,000	828	139
ACT	293,000	828	354

Table A6 was used as the basis behind the recommended option in *Table 1*. Additional sample was put into each state, in order to round off the sampling rates, and to allow a small buffer for error (e.g. if total votes in a state is smaller than expected; or if the true population error rate is higher than 0.45%).

Glossary[2]

Confidence Interval

A confidence interval is an interval which has a known and controlled probability (generally 95% or 99%) to contain the true value. In the context of senate assurance, one-sided confidence limits are calculated for the stage 2 error rates, to determine the maximum error rate that could potentially occur, for the given level of confidence.

Margin of Error (MoE)

Margin of Error describes the distance from the population value that the assurance estimate is likely to be within, for a specified given level of confidence. For instance, at the 95% confidence level, the MoE indicates that there are about 19 chances in 20 that the estimate will differ from the population value (the figure obtained if all senate ballots had been assured) by less than the specified MoE. Equivalently it is one chance in 20 that the difference is greater than the specified MoE, i.e. outside the MoE.

Significance testing

To determine whether a difference between two survey estimates is a real difference in the populations to which the estimates relate, or merely the product sampling variability, the statistical significance of the difference can be tested. The test is performed by calculating the standard error of the difference between two estimates and then dividing the actual difference by the standard error of the difference. If the result is greater than 1.96, there are 19 chances in 20 that there is a real difference in the populations to which the estimates relate.

Standard error

The square root of the variance of the sampling distribution of a statistic (square root of variance of state or national error rate in the context of senate assurance)

Variance

The variance is the mean square deviation of the variable around the average value. It reflects the dispersion of the empirical values around its mean.

[2] Glossary definitions have been taken from ABS publications and [The OECD Glossary of Statistical Terms] and modified to fit the context of senate assurance