Comparative Study of Electoral Systems (CSES) Module 6 Design Report

Version: December 2022

Country/territory: Australia
Date of election: 21 May 2022
Prepared by: Ian McAllister
Date of preparation: 8 September 2024

The answers provided in this Design Report are used to create variables in the CSES dataset, enable the CSES Secretariat to evaluate the eligibility of election studies for inclusion in the cross-national dataset, and provide users with important information for their analyses.

Notes to collaborators…

- All sections of the Design Report must be filled in for an election study deposit to be considered complete. The CSES Secretariat is available to provide clarification and support.
- Where brackets [ ] appear, answer by placing an "X" within the appropriate bracket(s).
- If more space or iterations are required to answer any question, please lengthen the document and duplicate items as necessary.
- In your deposited dataset, please be sure to include any weight variables, as well a variable that indicates the mode of interview for each observation (respondent) that is included.
- If you have an existing methodology report for your study, we would appreciate receiving an electronic copy of it, to supplement the information in your completed Design Report. We will make both your Design Report and any supplemental methodological report available for download from the CSES website. We encourage you to cut-and-paste information from your existing methodology report into your answers below, as appropriate.

Section A1. Collaborator(s)

Collaborators are the contact persons for election studies that appear in CSES. Collaborators are not necessarily who collected the data. This collaborators list will appear on the CSES website.

Collaborator name: Ian McAllister
Affiliation: ANU
Email: ian.mcallister@anu.edu.au

Collaborator name: Sarah Cameron
Affiliation Griffith University:
Email: s.cameron@griffithedu.au

Collaborator name: Simon Jackman
Affiliation: Sydney University
Email: simonjackman@icloud.com

Collaborator name: Jill Sheppard
Affiliation: ANU
Email: jill.sheppard@anu.edu.au

Section A2. Data Collection Organization

Organization that conducted the survey field work/data collection…

Organization name: Survey Research Centre
Website: www.srcentre.com.au

Section A3. Funding Organization(s)

Organization(s) that funded the data collection…

Organization name: ANU
Website: www.anu.edu.au

Organization name:
Website:

Section A4. Archiving Organization

If appropriate, please indicate the primary location where the full, original election study dataset (not just the CSES portion) will archived.

Organization name: Australian Data Archive
Website: https://ada.edu.au/

Section B. Study Design

1. Timing of the study that the CSES Module was included in:
    [x] Post-Election Study
    [ ] Pre-Election/Post-Election Panel Study
    [ ] Between Rounds
    [ ] Other, please specify: _____

2. For the post-election survey in which the CSES Module appeared, please provide the following three pieces of information…

The date that interviewing began: 23 May 22

The date that interviewing ended: 4 Jun 22

The number of days between the date of the election and when interviewing began: 10

3. Was the survey that included the CSES Module part of a panel study?
    [] Yes

[x] No

If yes, please describe the design of the panel study, including the date at which interviewing for each prior wave began and ended:

4. How many respondents answered the post-election survey in which the CSES Module appeared? That is, how many records/observations are there in the CSES portion of your study?

3,556

5. Did respondents give their consent to share their responses as part of the CSES dataset, in accordance with local human rights regulations and data protection laws?

[x] Yes
[ ] No

Section C. Sample frame, selection, and eligibility

6. Please describe the population that your study is meant to be representative of.

National population aged 18 and over. Note that only citizens can vote, and this sample includes a small number of non-citizens:

p_citizen Are you an Australian citizen?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-99 Refused	1	.0	.0	.0
	-98 Don't know	1	.0	.0	.1
	1 Yes	3269	91.9	91.9	92.0
	2 No	285	8.0	8.0	100.0
	Total	3556	100.0	100.0	

For analyses involving vote, they should be excluded as they are ineligible to vote.

Please note:
For CSES Module 6, all non-citizens and those for whom we had no information (287 respondents from overall 3,556 respondents) were dropped from the dataset.

7. What steps were taken as part of the sampling and/or data collection process to ensure that the sample is representative of the target population? In what ways were those steps successful, and in what ways were they not (please provide evidence wherever possible)?

The majority of the respondents were online panel members (96.8%), with a smaller portion of offline members (3.2%). The full sample profile is shown below in Table 1.

Table 1 Sample profile (unweighted)

Subgroup	Completed survey	Online members (completed)	Offline members (completed)	Benchmark[1]
	%	%	%	%
Base (n)	**3,556**	**3,441**	**115**	
Gender				
Male	43.3	43.7	31.3	49.0
Female	56.2	55.8	68.7	51.0
Age				
18-24 years	4.1	4.2	0.0	11.0
25-34 years	11.6	12.0	0.0	18.8
35-44 years	16.4	16.8	2.6	17.5
45-54 years	16.3	16.7	6.1	16.2
55-64 years	19.7	19.8	15.7	14.9
65-74 years	21.2	20.6	40.0	12.0
75 years or more	10.7	9.8	35.7	9.5
Location				
Sydney	20.1	20.3	13.0	20.6
Rest of NSW	11.8	11.7	15.7	11.3
Melbourne	19.4	19.5	16.5	19.7
Rest of VIC	6.5	6.4	9.6	6.3
Brisbane	10.0	10.0	8.7	9.7
Rest of QLD	9.2	9.1	11.3	10.4
Adelaide	6.9	6.8	9.6	5.4
Rest of SA	1.5	1.4	3.5	1.6
Perth	7.4	7.4	5.2	8.1
Rest of WA	1.8	1.8	1.7	2.2
Hobart	1.4	1.4	0.9	0.9
Rest of TAS	0.8	0.8	0.9	1.2
Darwin	0.6	0.6	0.9	0.6
Rest of NT	0.2	0.2	0.9	0.3
ACT	2.6	2.6	1.7	1.7

8. Please list the criteria for an individual to be eligible to be interviewed in your study. For example minimum age, citizenship, voter registration, etc.

National adult population aged 18 and over

9. Please describe the sample frame, its type and source, and its coverage of the population.

[1] Australian Bureau of Statistics (September 2021 ERPS).

The in-scope population for the survey was all active Life in Australia™ members included in the ANU longitudinal sample. A total of 4,338 active panel members were invited to take part in the survey and 3,556 (82.0%) completed the survey.

Table 2 Summary of key statistics

Field	Total	Online panel members	Offline panel members
Invited to complete survey	4,338	4,183	155
Total Interviews achieved	3,556	3,441	115
Average interview duration (mins)	18.7	18.4	26.8
Completion rate (%)	82.0	82.3	74.2
Main fieldwork start date	23-May-22	24-May-22	23-May-22
Main fieldwork finish date	5-Jun-22	5-Jun-22	4-Jun-22

Life in Australia™ members were randomly recruited via their landline or mobile phone and provided their contact details so that they could take part in surveys on a regular basis. This means that the population covered by the panel is all Australian adults contactable via either a landline or mobile phone.

10. Were any of the following excluded from the sample frame? Mark all that apply.

[ ] Specific regions of the country
[ ] Institutionalized persons
[ ] Military personnel
[ ] Other, please specify: _____

None excluded

For each group that was excluded, please provide additional details about the exclusion (for instance, the regions excluded), why the exclusion, and what percent of the total eligible population was excluded from the sample frame in each instance. Please also indicate the total percentage of the eligible population excluded from the sample frame.

11. Please describe in detail how the sample for the study was selected from the sampling frame. For multi-stage samples, please sure to list all sampling stages, and describe the sampling units and selection methods at each stage. Furthermore, please describe how individual respondents were identified and selected in the final stage. If the survey is part of a panel study, please describe not just the current wave but also the original sample. For telephone samples, please indicate whether the sample was a random digit dial (RDD) sample, listed sample, or dual frame sample. For surveys conducted by mail, indicate whether a listed sample.

The contact methodology adopted for online Life in Australia™ members is an initial survey invitation via email and SMS (where available), followed by multiple email reminders and a reminder SMS. Up to five reminders in different modes (including email, SMS, and telephone) were administered within the fieldwork period. Telephone non-response of online panel members who have not yet competed the survey commences in the second week of fieldwork and consists of reminder calls encouraging

completion of the online survey. Offline members with a valid mobile telephone number were also sent a short SMS invitation that contained a link to the survey as well as the reminder SMS halfway through fieldwork. The exact contact dates are shown below in Table 3.

Table 1 Summary of contact schedule

Contact type	Date	Population
Phone interviews	23-May-22 - 5-Jun-22	Offline only
Email	24-May-22	Online only
SMS	24-May-22	Both
Email	26-May-22	Online only
Email	28-May-22	Online only
SMS	30-May-22	Both
Reminder calls	31-May-22 - 4-Jun-22	Online only
Email	3-Jun-22	Online only

Telephone

The following call procedures were implemented:

- A four-call regime for mobile sample with an upper limit of six calls and a six-call regime for landline sample, with an upper limit of eight call attempts
- For mobile phones, capping the maximum number of unanswered call attempts to no more than four so as to avoid appearing overzealous in our attempts to achieve interviews
- Contact attempts were spread over weekday evenings (6:30 pm to 8:30 pm), weekday late afternoon/early evening (4:30 pm to 6:30 pm), Saturdays (11 am to 5 pm) and Sundays (11 am to 5 pm) (weekdays between 9 am to 4:30 pm are typically reserved for appointment management)
- Appointments were set for any time that the call centre is operational (weekdays between 9 am to 8:30 pm; weekends 11 am to 5 pm)
- 1800 number operation to address sample member queries and support the response maximisation effort and the establishment of a respondent page on our website (with responses to frequently asked questions).

Life in Australia™ members were able to request an email to complete the survey online.

There was no interviewing in languages other than English. Messages were left on answering machines and voicemails.

12. Were probability-based methods used for all stages of the selection process? That is, units were randomly selected throughout the process, including at initial recruitment (if a panel).

[x] Yes

[ ] No

If no, please describe all instances where non-probability methods were used and provide a justification for each as to why random selection was not used.

13. Did your sampling design make use of one or more of the following techniques? Please mark all that apply.

[ ] Stratification
[ ] Cluster sampling
[ ] Quota sampling
[ ] Interviewing more than one respondent from a single household
[ ] Respondent or household substitution
[ ] Non-sample replacement methods

None

Definitions:

- Stratification involves the division of the population into groups according to certain characteristics (for instance by demographic characteristics). Random selection then occurs within each of the groups that result to ensure their adequate inclusion.
- Cluster sampling divides the population into groups and then uses random selection to include some of the groups but not others. For example, a cluster sample might divide a large country into geographic areas and then select only some of the areas while excluding others, to avoid the need (and cost) of interviewer travel to all areas.
- In quota sampling, a respondent is sometimes selected based on demographic characteristics, rather than randomly, to ensure certain target distributions are met.
- Substitution is replacing one sampling unit with another when the first sampled unit is difficult to reach and/or interview.

Section D1. Interviewing Mode(s)

14. Please indicate the mode(s) of interviewing for the post-election survey in which the CSES Module appeared. Mark all that apply.

[ ] Interviewer-administered: In person (sometimes called face-to-face)
[ ] Interviewer-administered: Video
[x] Interviewer-administered: Telephone
[ ] Self-completion: Paper (by mail, supplement, etc.)
[x] Self-completion: Internet
[ ] Other, please specify: _____

Section D2. Interviewing Mode(s): Mixed-Mode

15. If the survey in which the CSES Module appeared was mixed-mode (that is, made use of more than one mode type)…

Was the mixed-mode design intended to address a particular problem or problems?

[x] Yes
[ ] No

If yes, please elaborate:

Some respondents did not have internet access

Were all modes available to respondents from the beginning, or was there a mode progression (that is, if the respondent did not respond by one mode, another mode was subsequently offered to them)? If the latter, please describe the mode progression.

Were individuals assigned a mode after recruitment or was this part of the recruitment?
[x] Assigned a mode after recruitment
[ ] Assigned a mode as part of the recruitment

Did respondents self-select the mode by which they would respond?
[ ] Yes
[x] No

Did the design include a mode change *within* interviews (e.g., selected self-completion elements within the same questionnaire)?
[ ] Yes
[x] No

If yes, please provide details:

If there were substantial differences in the distribution of key demographics across modes, please note them here.

Section D3. Interviewing Mode(s): Telephone

16. If the post-election survey in which the CSES Module appeared was entirely or partly conducted by telephone…

What is the estimated percentage of households without a phone? ____3___ %

Were unlisted telephone numbers included in the population sampled?
[ ] Yes
[x] No

If no, what percent of the total eligible population did this exclude from the sample frame? ___3____ %

Section D4. Interviewing Mode(s): Internet

17. If the post-election survey in which the CSES Module appeared was entirely or partly conducted via the Internet…

What is the estimated percentage of households without access to the Internet (that is, the percent of the eligible population excluded due to lack of Internet access)? __10__ %

Were provisions taken to include members of the population without access to the Internet?

[x] Yes
[ ] No

If "Yes", please explain:

Telephone as already explained

If "No", what percent of the total eligible population did this exclude from the sample frame? _____ %

Did the survey make use of an access panel (i.e. respondents were selected from a group of pre-screened panelists)?

[x] Yes
[] No

If yes, please describe the access panel (company, population [does it include persons without initial access to the Internet and how are they interviewed], method of recruiting members, total size of access panel, method of selecting survey respondents from the panel):

Life in Australia™ members were randomly recruited via their landline or mobile phone and provided their contact details so that they could take part in surveys on a regular basis. This means that the population covered by the panel is all Australian adults contactable via either a landline or mobile phone.

A dual-frame random digit dialling (RDD) sample design was employed to undertake recruitment of Life in Australia™ in 2016, with a 30:70 split between the landline RDD sample frame and mobile phone RDD sample frame. For the landline sample, an alternating next / last birthday method was used to randomly select respondents from households where two or more in-scope persons were present. For mobile sample, the phone answerer was the selected respondent. Only one member per household was invited to join the panel.

In May 2018, the panel was refreshed with *n* = 287 panellists being retired and *n* = 267 new panellists being recruited. The recruitment methodology used only mobile RDD sample and recruited only online participants that were under 55 years old in order to balance the demographics (the age profile of panel members was older than that of the Australian population). The recruitment rate (RECR) for the replenishment was 12.1%. After the refresh, there were *n* = 2,839 active members of Life in Australia™. For both the recruitment in 2016 and panel refreshment in 2018, the RDD sample was provided by SamplePages.

Between October-December 2019, the panel was refreshed with *n* = 347 panellists being retired and *n* = 1,810 new panellists being recruited. This recruitment used a G-NAF (Geocoded National Address File) sample frame and push-to-web methodology. Only online participants were recruited in order to balance the demographics (the age profile of panel members was older and more educated than that of the Australian population). The recruitment rate (RECR) for the replenishment was 12.1%. After the refresh, there were *n* = 4,025 active members of Life in Australia™.

Between November 2020 and January 2021, the panel was refreshed with *n* = 385 panellists being retired and *n* = 612 new panellists being recruited. This recruitment used a combination of recruitment methodologies: G-NAF (Geocoded National Address File) sample frame and push-to-web, mobile sample frame IVR (interactive voice response) push-to-web, and mobile sample frame SMS invitation. Only online participants were recruited in order to balance the demographics (the age profile of panel members was older and more educated than that of the Australian population). The recruitment rate (RECR) for the replenishment was 3.1%. After the refresh, there were *n* = 4,060 active members of Life in Australia™.

In April 2021, the panel was refreshed with *n* = 510 new panellists being recruited. This recruitment used an RDD mobile sample frame with SMS invitation. Only online participants were recruited in order to balance the demographics (the age profile of panel members was older and more educated than that of the Australian population). The recruitment rate (RECR) for the replenishment was 3.4%. After the refresh, there were *n* = 4,499 active members of Life in Australia™.

In August and September 2021, the panel was expanded with *n* = 3,715 new panellists being recruited. This recruitment used the G-NAF (Geocoded National Address File) sample frame sample frame and push-to-web methodology. The recruitment rate (RECR) for the replenishment was 7.7%. After the refresh, there were *n* =7,645 active members of Life in Australia™.

Unlike other research panels, Life in Australia™ includes people both with and without internet access. Those without internet access or those who are not comfortable completing surveys over the internet are able to complete surveys by telephone. Life in Australia™ members receive a small incentive for joining the panel and another incentive for each survey they complete.

Did respondents self-select into the survey, at any stage?

[ ] Yes
[x] No

If yes, please explain:

Section E. Respondent Persuasion

18. What was the average number of contact attempts made per household, for the entire sample (not just for survey completions)?

2

19. What was the maximum number of contact attempts made per household, for the entire sample (not just for survey completions)?

4

20. Did respondents receive incentive payments? Please mark all that apply.
[ ] Yes, during the screening process
[ ] Yes, in advance of their interview
[x] Yes, after their interview
[ ] No

21. Were special efforts made to persuade respondents who were reluctant to be interviewed?
[ ] Yes
[x] No

Section F: Interview/Survey Verification

22. Was interview/survey verification used?
[ ] Yes
[x] No

If yes, please indicate the percent of completed surveys that were verified: _____ %

If yes, please describe the method(s) used for verification:

Definition: Interview/survey verification is the process of verifying that an interview was conducted and that the survey was administered to the correct respondent, for quality control purposes.

23. What steps, if any, were taken to ensure that respondents were providing truthful answers to the questions? Were any respondents remove from the final dataset (e.g. due to speeding, satisficing, completing multiple surveys, etc.)? Please provide details.

All interviewers and supervisors selected to work on the survey attended a two-hour briefing session, which focused on all aspects of survey administration, including:

- Survey context and background, including a detailed explanation of Life in Australia™
- Survey procedures and sample management protocols
- The importance of respondent liaison procedures
- Strategies to maintain co-operation
- Detailed examination of the survey questionnaire, with a focus on the use of pre-coded response lists and item-specific data quality issues.

After the initial briefing session, interviewers engaged in comprehensive practice interviewing. A total of 24 interviewers were briefed on the survey.

The in-field quality monitoring techniques applied to this project included:

- Validation of 18.8% of the telephone surveys conducted via remote monitoring (covering the interviewers' approach and commitment-gaining skills, as well as the conduct of the interviews)
- Field team de-briefing after the first shift and, thereafter, whenever there was important information to impart to the field team in relation to data quality, consistency of interview administration, techniques to avoid refusals, appointment-making conventions, or project performance
- Examination of 'Other (specify)' responses
- Monitoring of timestamps for segments of the survey and overall time taken to complete the survey

Section G1. Response Rate

Notes to collaborators:

- If multiple modes of interviewing were used for the post-election survey in which the CSES Module appeared, please repeat the following questions as appropriate for each mode used.
- Standard definitions for the below items, as well as a helpful response rate calculator, can be accessed on the website of the American Association for Public Opinion Research at: https://www.aapor.org/Communications/AAPOR-Journals/Standard-Definitions.aspx

24. For each of the following categories, please indicate the number of cases/records from the sample that fall into each. (If the CSES Module appeared in a panel study, please report the numbers for the first wave of the study, even if the CSES Module did not appear in that wave.)

Interviews	
I. Completed interviews	I=__3556_____
P. Partial interviews	P =_____
Non-interviews	
R. Refusals and break-offs	R = _57_____
NC. Non-Contact	NC=_574_____
O. Other	O = _151_____
Unknown Eligibility	
UH. Unknown if household/occupied household unit	UH = _0_____
UO. Unknown, other	UO = _0_____

25. Please calculate the response rate (RR2) by inserting the answers from Question 24 into the following formula.

$$RR2 = \frac{(I + P)}{(I + P) + (R + NC + O) + (UH + UO)}$$

82%

Table 2 Summary of survey completion rate

AAPOR		Total		Online members		Offline members	
code	Outcome categories	n	%	n	%	n	%
	Total invited	**4,338**	**100.0**	**4,183**	**100.0**	**155**	**100.0**
1.1	Completed interview	3,556	82.0	3,441	82.3	115	74.2
2.10	Refusals and mid-survey terminations	57	1.3	55	1.3	2	1.3
2.20	Non-contacts	574	13.2	543	13.0	31	20.0
2.30	Other	151	3.5	144	3.4	7	4.5
COMR	**Completion Rate (%)**	**-**	**82.0**	**-**	**82.3**	**-**	**74.2**

Please note:
The updates response rate without non-citizens and those for whom we have no information (new I = 3269) is 81%.

26. If the number of cases with unknown eligibility (UH and/or UO in Question 24) is greater than zero, please estimate what proportion of cases of unknown eligibility is actually eligible. Please indicate the basis, evidence, and any scientific justification for this estimate.

Section G2. Response Rate: Panel Studies

27. If the CSES Module appeared in a panel study…

How many waves were conducted prior to the wave that included the CSES Module?

What was the response rate at the initial recruitment stage? Please show your calculations.

What percent of recruited respondents participated in the first wave of the study? Please show your calculations.

What was the total panel attrition between the first wave of the study and the wave that included the CSES Module? Please show your calculations.

If there were substantial differences in the distribution of key demographics (for example, age, gender, education, etc.) between the first wave of the study and the wave that included CSES, please note them here.

Please provide the number of completed interviews for the wave that included the CSES Module.

Section H. Post-Survey Adjustment Weights

28. Are weights included in the data file?

[x] Yes
[ ] No

If No, please explain why no weights are provided:

If Yes, please describe in detail: their construction, what they correct for, and how their use makes the survey dataset more representative of the population.

Approach to weighting

Overview

A subset of Life in Australia™ cases has been created, comprising the 4,375 panellists who were selected through the process described in "*CSRM 2022 Longitudinal Sample Establishment Technical Report*". This subset will serve as a longitudinal sample of Life in Australia™ for the purposes of tracking changes in attitudes and behaviours across ANU Poll waves in the succeeding year. This document outlines the approach used to derive establishment weights for this group and how weights will be calculated for subsequent waves.

Establishment weights

Members of Life in Australia™ were originally recruited through a national dual-frame telephone survey in 2015 (refer to Kaczmirek *et al.*, 2019 for details on the establishment of Life in Australia™). Subsequent replenishment rounds were conducted through mobile telephone in 2017 and then through a combination of SMS and address-based sampling in 2019, 2020 and 2021. Design weights for original recruits were derived as the inverse of their probability of selection, based on the approach of Best (2010), and then adjusted through generalised regression (Deville *et al.,* 1993) to reflect the population distributions for key demographic characteristics. This is a standard approach for design-based surveys in which records are selected at random from an explicit sampling frame, inclusion probabilities are calculated and then inferences are made on the basis of repeated sampling from a population (see Valliant *et al.*, 2013, especially Ch 13 and 14).

As more panellists were recruited through multiple rounds and mechanisms and from multiple sampling frames, the calculation of selection probabilities grew increasingly cumbersome. Differential response and attrition rates compounded the complexity of weighting, leading to weights with large variations and to survey estimates with consequently large design effects. The combination of these factors produced a panel for which the assumptions and methods of traditional design-based approaches were no longer appropriate. To reduce the complexity of the weighting process and to arrest the ever-degrading quality of the weights, the calculation of panel weights was simplified to use a model-based approach (Valliant *et al.*, 2000; Elliott and Valliant, 2017).

Such methods avoid the assumptions of a random selection mechanism and the calculation of inclusion probabilities, instead using a model to "project" the responding sample to the population. When using a model-based approach to estimate a population total, the number of sample cases is added to the predicted total for non-sample cases (Valliant, 2020). Covariates used in the predictive model need to be known for sample cases but only totals need to be available for non-sample cases. In the situation where the number of sample cases is negligible compared to the number of non-sample cases, which is true for Life in Australia™, this approach is equivalent to generalised regression if the inverse selection probabilities are all set to 1. Typically, the same set of model covariates is used to make predictions for all survey variables.

Life in Australia™ collects more than 20 characteristics about respondents for which population totals can be obtained from the Australian Bureau of Statistics or from other official sources. Having such a wealth of available data meant that we could try a range of model covariates with a view to aligning the weighted sample as closely as possible with the population totals for available characteristics. Refer to Kreuter *et al.* (2009) and Peytchev *et al.* (2018) for more details about the choice and use of characteristics for non-response and weighting adjustments.

For the CSRM 2022 longitudinal sample, many different sets of covariates were used in the generalised regression model in an effort to align the responding sample as closely as possible with the population of Australian adults. The final choice of covariates was determined from two considerations:

1. Minimising the bias, defined here as the absolute percentage point difference between the model predictions and the population totals, averaged across all available characteristics (primary consideration); and

2. Maximising the effective sample size, measured here by the weighting efficiency[2] (Kish, 1965, 1992) (secondary consideration).

Ideally, the responding sample would be perfectly aligned with the population on all characteristics (average absolute bias = 0%) and there would be no variability introduced through weighting (efficiency = 100%). In reality, though, these two measures work against each other – bias is reduced through the inclusion of more model covariates, but the variability in the weights then increases. A compromise between the two is required to choose the optimal set of covariates.

For each set of covariates, the bias and efficiency were calculated for the model solution. An excerpt from the results is shown in Table 3. Included are an unweighted solution (option A), for reference purposes, along with a range of other solutions at different points in the bias-efficiency space (options B–E). It is evident that including more covariates, especially those where the cases differ most from the population (age and education), reduces the bias but also reduces the weighting efficiency.

The final set of covariates selected for the model (option E) was as follows:

- Adults in the household

[2] Defined as $n^{-1}\sum_s {w_i}^2/(\sum_s w_i)^2 \times 100\%$, where n is the size of sample s and w_i is the vector of weights. Ranges from 0% to 100%, representing maximally different and all identical weights, respectively.

- Age group by highest education
- Gender
- Language spoken at home
- Remoteness
- State

Table 3. Bias and efficiency for selected model covariates

Option	Description	Model covariates	Bias (pp)	Efficiency (%)
A	Unweighted	-	3.19	100.00
B	Lowest bias overall	Age group Adults in the household Country of birth Dwelling tenure Gender Highest education Time at current address	1.18	61.14
C	Improved bias, high efficiency	Adults in the household Dwelling tenure Part of State State Telephone status	2.43	91.00
D	Improved bias, contains education	Highest education Marital status Part of state State Telephone status	2.89	81.36
E	Final solution	Adults in the household Age group by highest education Gender Language spoken at home Remoteness State	2.26	69.49

This set was chosen since it yielded a satisfactory balance between bias and efficiency and contained a wide range of model covariates, including the two that were most different between the sample and the population (age and education, as noted above). The categories and population totals corresponding to these covariates are shown in Table 4.

Wave weights

As is typical for a panel survey, we expect that not all members of the CSRM 2022 longitudinal sample will respond to subsequent waves and that response will not be uniform across member characteristics. To limit the impact of these events on the representativeness of estimates made from a wave, we will adjust the establishment weights through the use of propensity scores (Rosenbaum and Rubin, 1983). These will be calculated by means of a logistic regression model predicting the likelihood of a longitudinal sample member participating in the given wave, conditional on characteristics available for both

respondents and non-respondents (Valliant *et al.,* 2013). To reduce the impact of very low or very high values, the predicted probabilities will be collapsed into classes (after Cochran, 1968), with propensity scores assigned as the mean probability within each class. The propensity-adjusted establishment weights will then be re-aligned with the population totals from Table 4. These steps will ensure that estimates from each wave represent longitudinal sample members and the target population as closely as possible.

Table 4. Covariates used in model for establishment weights, with population distributions and data sources

Characteristic	Benchmark target (#)	Benchmark target (%)	Source
Age group by Highest education			(A)
18-24 years	2211718	11.02	
25-34 years x Bachelor or higher	1451360	7.23	
25-34 years x Below Bachelor	2313290	11.53	
35-44 years x Bachelor or higher	1270203	6.33	
35-44 years x Below Bachelor	2243060	11.18	
45-54 years x Bachelor or higher	842814	4.20	
45-54 years x Below Bachelor	2405971	11.99	
55-64 years x Bachelor or higher	665397	3.32	
55-64 years x Below Bachelor	2332532	11.63	
65+ years x Bachelor or higher	587158	2.93	
65+ years x Below Bachelor	3740734	18.64	
Gender			(A)
Male	9838431	49.03	
Female	10225806	50.97	
Number of adults in the household			(B)
One	3422389	17.06	
Two	11076916	55.21	
Three or more	5564933	27.74	
Language spoken at home			(A)
Speaks a language other than English at home	4652110	23.19	
Speaks only English at home	15412127	76.81	
Remoteness			(A)
Major Cities of Australia	14325247	71.40	
Inner Regional Australia	3643015	18.16	
Rest of Australia	2095975	10.45	
State or Territory of residence			(A)
New South Wales	6394859	31.87	
Victoria	5225575	26.04	
Queensland	4025464	20.06	
South Australia	1403746	7.00	
Western Australia	2066413	10.30	

Tasmania	429081	2.14
Northern Territory	184344	0.92
Australian Capital Territory	334755	1.67

(A) Census 2016 (ABS, 2016) with population updates (ABS, 2021).

(B) National Health Survey 2017-18 (ABS, 2017-18) with population updates (ABS, 2021).

29. Were any other steps taken to mitigate the impact of non-response in the dataset? If so, how do the adjustments affect the survey results?

No

Please note:
For CSES Module 6, all non-citizens and those for whom we had no information (287 respondents from overall 3,556 respondents) were dropped from the dataset.
The explanation from the collaborators is that when they started their work on the file, they asked their statistician who devised the weights about this issue. His view was that they should leave the already calculated weight and a new weight was not necessary. The already supplied weight is adequate.

Section I. Translation

Notes to collaborators…
- As part of your deposit, please provide the original questionnaire(s) in each language used.
- For questionnaires in a language other than English, if a back-translation to English happens to be available, we would appreciate receiving the back-translation, also.
- The below questions on translation are adapted from those developed for the ISSP.

30. Was the questionnaire translated?
[ ] Yes, translated by member(s) of research team
[ ] Yes, by translation bureau
[ ] Yes, by specially trained translator(s)
[x] No, not translated

31. Please list all languages used for the fielded module.

32. If the questionnaire was translated, was the translated questionnaire assessed/checked or evaluated? Please mark all that apply.
[ ] Yes, a group worked together on it and reconciled their differences through discussion
[ ] Yes, an expert checked it
[ ] Yes, by back translation
[ ] Other; please specify: _____
[ ] No
[ ] Not applicable

33. If the questionnaire was translated, was the questionnaire pre-tested?

[ ] Yes
[ ] No
[ ] Not applicable

34. If the questionnaire was translated, were there any questions which caused problems when translating?

[ ] Yes
[ ] No
[ ] Not applicable

35. If the questionnaire was translated, please provide a list of all questions which caused problems when translating. For each question listed, describe what problems were encountered and how they were solved:

Section J. Other

36. Please list any additional advice you have for analysts about how the survey should be used, and any remaining special considerations in that regard.