Pennay, Darren, Sebastian Misson, Dina Neiger, and Paul J Lavrakas. 2023. "How Weighting by Past Vote Can Improve Estimates of Voting Intentions." *Survey Practice* 16 (1). https://doi.org/10.29115/SP-2023-0001.

ARTICLES

How Weighting by Past Vote Can Improve Estimates of Voting Intentions

Darren Pennay[1,2a], Sebastian Misson[2], Dina Neiger[1,2], Paul J Lavrakas[2]

[1] Centre for Social Research and Methods, Australian National University, [2] The Social Research Centre Pty Ltd

Keywords: Online panels, polling, voting intentions, past vote, weighting

https://doi.org/10.29115/SP-2023-0001

Survey Practice

Vol. 16, Issue 1, 2023

Polling error for the 2020 US election was the highest in 40 years and no mode of surveying was unambiguously more accurate. This occurred amid several recent polling failures in other countries. Online panels, as the dominant method now used by pollsters to survey voters, are well-positioned to help reduce the level of bias in pre-election polls. Here, we present a case for those pollsters using online panels for pre-election polling to (re)consider using past vote choice (i.e., whom respondents voted for in the previous election) as a weighting variable capable of reducing bias in their election forecasts under the right circumstances. Our data are from an Australian pre-election poll, conducted on a probability-based online panel one month prior to the 2019 Australian federal election. Three different measures of recalled vote choice for the 2016 election were used in weighting the forecast of the 2019 election outcome. These were (1) a short-term measure of recall for the 2016 vote choice obtained three months after the 2016 election, (2) a long-term measure obtained from the same panelists three years after the 2016 election and (3) a hybrid measure with a random half of panelists allocated their short-term past vote measure for 2016 and the remainder their long-term measure. We then examined the impacts on the bias and variance of the resulting estimates of the 2019 voting intentions. Using the short-term measure of the 2016 recalled vote choice in our weighting significantly reduced the bias of the resulting 2019 voting intentions forecast, with an acceptable impact on variance, and produced less biased estimates than when using either of the other two past vote measures. The short-term recall measure also generally resulted in better estimates than a weighting approach that did not include any past vote adjustment. Implications for panel providers are discussed.

Polling error for the national popular vote for the 2020 US pre-election polls was the highest in 40 years and no mode of interviewing or method of sampling was unambiguously more accurate (AAPOR 2021). This poor result for pollsters at the 2020 US election occurred amid a spate of well documented polling failures in several countries in recent years (Cornesse et al. 2020). Online panels, being the dominant method of surveying voters nowadays, are well-placed to help reduce the level of error in pre-election polls.

Trying to improve survey-based estimates of voting intentions by adjusting (i.e., balancing or weighting) one's sample so that the recalled vote choice of respondents in the previous election is consistent with that of the voting population in the previous election is a commonly used method in pre-election polls in many parts of the world (Cabrera-Álvarez and Escobar 2019; Durand

a Correspondence concerning this article should be addressed to Darren W. Pennay, The Social Research Centre, PO Box 13328 Law Courts VIC 8010 (Email: darren.pennay@anu.edu.au)

and Johnson 2021; Wells 2019) but less common in the US (AAPOR 2021). Despite widespread use of this practice, there has been relatively little recent research into the efficacy of past vote weighting, and the research that has been done tends to show mixed results (Durand, Deslauriers, and Valois 2015).

However attractive balancing or weighting a pre-election poll sample so that it is representative of the voting choices of the population at the last election may appear, such an approach is not without potential respondent-related measurement error. Recall of past vote may be inaccurate due to: (1) memory failure, (2) the tendency of voters to misreport how they previously voted to reconcile it with how they currently intend to vote, and (3) social desirability (Durand, Deslauriers, and Valois 2015). In addition, some respondents will not have voted in the previous election.

Furthermore, the implications of weighting using an unreliable measure of past vote can be substantial. Wells (2019) conducted an experiment in which he reweighted YouGov polling data for the 2017 Brexit election using a reported past vote measure collected immediately after that election (at which time 41% reported voting for Labour) and a reported past vote measure collected from the same respondents two years later in 2019, at which time only 33% reported having voted for Labour in 2017. When used as weighting variables and aligned to past vote benchmarks, the difference that these two measures had on the subsequent estimates of voting intentions was relatively large in the context of pre-election polling. When the 2017 (short-term) measure of recalled past vote was used as the weighting variable, estimated support for Labour was 21%; when the 2019 (long-term) measure of recalled past vote choice was used, estimated support for Labour increased to 24%.

Our research adds to the literature on past vote weighting by discussing how more reliable measures of past vote choice can be gathered and used to improve estimates of voting intentions. The focus of the current research is to compare alternative methods available to online panel providers for measuring respondents past voting behavior with a view to offering some practical guidance as to how panel providers can contribute to efforts to reduce bias in the pre-election polls conducted on their panels.

Data and Methods

The data used for this study are from an Australian pre-election poll undertaken by the Australian National University (ANU) in April 2019, one month prior to the 2019 federal election in Australia. This ANUpoll was conducted on a sample drawn from Australia's only national probability-based online panel, Life in Australia™.[1] A total of 2,686 panel members aged 18 years

1 A description of how Life in Australia™ operates is provided in Appendix 1.

and older were invited to take part in the survey and 2,054 (76.5%) completed the questionnaire with a cumulative response rate (for forming the panel and completing this questionnaire) of 8.6% (Callegaro and DiSogra 2008).

In our view, despite the Australian electoral system using "compulsory" preferential voting,[2] there are sufficient similarities between the Australian electoral system and polling practices and those of many other countries, particularly those countries that have single member multi-party constituencies, to make these results broadly applicable to online pollsters. Despite compulsory voting, the voting age population turnout in Australia's most recent national election, 76%, was not too dissimilar from the 62.6% voting age population turnout in the US and 62.3% in the UK (DeSilver 2022). As such, predicting the proportion of the eligible population that will cast a valid vote is almost as significant an issue for Australian pollsters as it is elsewhere and this suggests a broad applicability of these findings.

Measuring past vote

Three alternative measures of recalled past vote choice were created as weighting variables and added to the survey data: One was based on respondents' short-term recall of their past vote choice (collected three months after the 2016 election), another was based on their long-term recall of their past vote choice (collected 34 months after the 2016 election), and the third was a blended measure whereby a random half of respondents were assigned their short-term measure of their 2016 vote choice and the remainder their long-term measure. These measures of past vote in 2016 were alternatively incorporated into a standard survey weighting solution and aligned to voting benchmarks in the 2016 election. The impact of each weighting solution on the bias and variance on the resultant ANUpoll estimates of voting intentions for the impending 2019 election are then compared to the actual 2019 election outcome.

Short-term recall of past vote: At the initial recruitment stage of the panel in October 2016, panelists were asked about their vote choice in the preceding federal election which was held just a few months earlier on July 2, 2016. The question wording and response options are provided in Appendix 2. The responses to this question were merged with the ANUpoll dataset and formed our short-term recall measure of past vote.

2 Australian elections are conducted under what is known colloquially as "compulsory" voting. What the rules require is that all those aged 18 years and over who are Australian citizens or eligible British subjects (those who arrived in Australia pre-1984 and who are on the electoral roll) and have resided at their current address for at least one month, must be on the electoral rolls and turnout to vote.

Long-term recall of past vote: The long-term recall measure of past vote was collected almost three years after the 2016 election and was available for a large subset of the same panelists. The question asked of these panelists is also provided in Appendix 2. The responses to this question were merged with the ANUpoll data and formed our long-term recall measure of past vote.

Blended recall measure of past vote: This measure was created to simulate a situation very likely to confront online panel providers whereby due to reasons associated with unit and item nonresponse, panel attrition, and panel replenishment, even if a short-term measure of past vote choice is routinely collected from all panelists immediately following each election, it is unlikely that all of these panelists will still be responding to survey requests when the next pre-election polling cycle is underway. This means that the measure of past vote likely to be available for a substantial portion of panelists would be collected at different times during the election-to-election cycle. While we could not simulate this situation exactly, we did approximate it by creating the blended measure of past vote using a combination of the short-term and long-term recall measure and appended this "blended" measure to the data set.

As per Durand, Deslauriers, and Valois (2015) and Dassonneville and Hooghe (2017), we found that for Life in Australia™ panelists, in aggregate, the accuracy of their recalled past vote—measured as the average error between the recalled vote and the actual election result—diminishes over time. The average absolute error of our short-term recall measure relative to the 2016 election outcome was 2.3 pp increasing to 3.0 pp for the long-term measure (data not shown).[3] This level of inconsistency between short-term and long-term past vote recall measures is of a similar magnitude to that reported in the literature cited by Durand, Deslauriers, and Valois (2015).

Approach to weighting

To investigate the relative impact of weighting by these three measures of past vote choice (short-term, long-term, and blended) on estimates of voting intentions, it was necessary to incorporate these measures into an appropriate weighting solution. Seven weighting scenarios were evaluated:

Weight 1: Age by education, sex, and geography (state by capital city/rest of state). This is our baseline weight and main point of comparison for the other six weights. Educational attainment was included in this baseline weighting solution because a failure to do so was seen as one of the contributing factors for the less accurate state-level polls in the 2016 US Presidential Elections (AAPOR 2017).

Weight 2: Age by education, sex, geography, and short-term recall of past vote.

3 The average absolute error is the mean of the absolute errors observed between the recalled measure of past vote and the actual election outcome for the Coalition, Labor, The Greens and Other candidates/parties.

Weight 3: Age by education, sex, geography, and long-term recall of past vote.

Weight 4: Age by education, sex, geography, and blended estimate of past vote.

Weights 5 to 7: Weighting only by short-term recall of past vote (Weight 5), only by the long-term recall measure of past vote (Weight 6), and only by the blended measure of past vote (Weight 7), as single factor weighting solutions.

Each weight was calculated using the rake procedure from the survey package in R (Lumley 2004, 2010, 2020).

The population benchmarks for the demographic variables (age, sex, geography, and educational attainment) were compiled from the Australian Bureau of Statistics 2016 Census counts and the March 2019 Estimated Residential Population figures. The past vote weighting benchmarks were compiled by the Australian Electoral Commission (see Appendix 3, Tables A1 and A2 for details).

Error metrics

Measures of bias. Two measures of bias were used. The first was to calculate the *weighted average absolute error* of the 2019 Australian national election primary vote estimates compared with the election results.[4] The second measure is peculiar to Australian polling and is in response to the preferential voting system used in Australia.[5] This measure captures the bias in the polls-based estimate by measuring the *average absolute error of the two party-preferred vote (2PP).*

Measure of variance. The variance introduced by the weights is measured using the design effect (i.e., *deff*) calculated by Taylor series linearization by the *svymean* procedure in the survey package in R (Lumley 2020).[6]

Overall error measure. Mean square error (MSE) (Korn and Graubard 1999) is a measure which combines bias and variance to assess the impact of weighting on the total survey error as follows:

4 The average absolute error metric gives equal weight to the estimate for each of the parties. However, given preferential voting in Australia, the relative accuracy of the poll estimates for the main parties (Labor and the Coalition) is more consequential in terms of the 2PP vote than the relative accuracy of the estimates for the other parties. As such, a better measure of the overall performance of the polls is a measure of bias that takes into account the average absolute error on the primary vote weighted by vote share. This is calculated by multiplying the primary vote share by the primary vote error for each party.

5 Australians vote for candidates in single-member constituencies. Elections are won by the party or coalition that wins a majority of the 151 seats in the House (as at the 2019 election); failing that, by the party or coalition that can govern with sufficient support from other parties and/or independents. To cast a valid vote, electors number every candidate in order of preference. To win a seat, a candidate must win an absolute majority of the votes cast. Where no candidate achieves this by way of the first preferences, the votes received by the least popular candidate are distributed according to the electors' second preferences, then the votes of the next least popular candidate, and so on, until one candidate secures an absolute majority. Where the final vote comes down to a contest between the Liberal/National Party (the Coalition) and Labor, as it almost always does, the final figure is called "the two-party preferred (2PP)." Where it comes down to a contest between the Liberal/National Party or Labor and some other party or candidate, the final figure is called "the two-candidate preferred (2CP)." The 2PP and 2CP are unique to electoral analysis in Australia.

6 The *deff* reflects the ratio between the standard error calculated using appropriate adjustments for survey design over what the standard error would be if a simple random sample had been used. A higher *deff* indicates more variability in the weights with respect to the statistic being calculated.

$$MSE = B^2 + VMSE = B^2 + V$$

where B is the primary measure of bias (in this case the 2PP bias) and V is a measure of variance estimated from the data set. Korn and Graubard (1999) estimate the *deff* using the variance of the weights. However, given that the assessment of accuracy in this research is focused on a single measure (i.e., 2PP vote), the Taylor series linearized *deff* is used. Root mean squared error (RMSE) is used so that the result is on the original scale of the percentages.

Simulations

To obtain estimates of the degree to which the different voting intentions estimates produced by the different weighting solutions were due to sampling variation, 10,000 samples were obtained by random re-sampling with replacement of the original data to the same sample size. Each weighting scheme was calculated for each re-sample to obtain estimates for all weighting options. The reported standard errors represent the 95% confidence intervals of the re-samples, i.e.:

$$CI = t^* \pm 1.96\ .\ se^* CI = t^* \pm 1.96\ .\ se^*$$

where CI is the confidence interval, t^* is the average estimate, and se^* is the standard deviation of the 10,000 resamples.

Probabilities represent the proportion that one weighting scheme produces superior results to another weighting scheme adjusted to be two-tailed probabilities, i.e.:

$$p = \frac{\begin{cases} p^* \leq\ .5, & p \\ p^* >\ .5, & (1-p) \end{cases}}{2}$$

where p^* is the proportion of re-samples where one weighting scheme is superior to the other. Probabilities were adjusted for multiple comparisons using the technique described by Benjamini and Hochberg (1995) using the p.adjust function in R.

Results

Table 1 shows the results for each error metric based on the original data, as well as the average for each metric from the simulations, alongside their 95% confidence intervals (with the simulated results in brackets). Table A3 (see Appendix 4) shows the probabilities associated with the null hypothesis which in this case tests whether the RMSE of the 2PP vote (i.e., the final column of Table 1) for one weighting solution is equal to the RSME of a comparative weighting solution. Table 1 in conjunction with Table A3 shows whether the survey estimates generated by the various weighting solutions meet the threshold for statistical significance adjusted for two-tailed probabilities.

Table 1. Comparison of weighting solutions by specified error metrics

Primary vote			2PP			
No.	Solutions	Weighted Avge absolute error	Avge Absolute error	Std. error	Design effect	Root Mean square error
1	Age by education, sex, geography	2.58 (2.93, 1.30 – 4.56)	4.08 (4.08, 1.46 – 6.71)	1.34 (1.34, 1.25 – 1.43)	1.41 (1.42, 1.24 – 1.59)	4.29 (4.33, 1.90 – 6.76)
2	Age by education, sex, geography and short-term recall of past vote	1.41 (1.81, 0.47 – 3.15)	2.41 (2.39, 0.12 – 4.67)	1.22 (1.22, 1.07 – 1.36)	1.14 (1.14, 0.88 – 1.41)	2.70 (2.76, 0.84 – 4.68)
3	Age by education, sex, geography and long term	2.95 (3.06, 1.59 – 4.53)	4.45 (4.44, 2.35 – 6.52)	1.07 (1.07, 0.94 – 1.20)	0.90 (0.90, 0.68 – 1.12)	4.57 (4.57, 2.57 – 6.58)
4	Age by education, sex, geography and blended estimate of past vote term	2.54 (2.68, 1.14 – 4.22)	3.92 (3.89, 1.64 – 6.15)	1.16 (1.16, 1.02 – 1.30)	1.04 (1.05, 0.79 – 1.30)	4.09 (4.08, 1.96 – 6.20)
5	Weighting by short-term past vote only	0.20 (0.82, 0.04 – 1.60)	0.64 (0.89, -0.37 – 2.15)	0.90 (0.90, 0.86 – 0.93)	0.61 (0.61, 0.56 – 0.65)	1.10 (1.34, 0.43 – 2.25)
6	Weighting by long-term past vote only	1.76 (1.85, 0.70 – 3.00)	2.80 (2.80, 1.19 – 4.42)	0.82 (0.82, 0.78 – 0.86)	0.52 (0.52, 0.47 – 0.57)	2.92 (2.93, 1.41 – 4.46)
7	Weighted by blended measure of past vote only	1.01 (1.17, 0.09 – 2.25)	1.84 (1.84, 0.19 – 3.49)	0.87 (0.87, 0.83 – 0.91)	0.58 (0.58, 0.53 – 0.62)	2.03 (2.08, 0.66 – 3.51)

Results outside brackets represent the observed estimate based on the original data; results in brackets represent the average estimate from the simulated samples and the upper and lower confidence intervals.
Source: Australian Electoral Commission 2019 and authors' analysis.

Weight 1 (age by education, sex, and geography) produces a primary vote estimate with a weighted average absolute error of 2.58 pp and an average absolute 2PP error of 4.08 pp. This weighting solution increases the *deff* by a factor of 1.41 and has a RMSE (which encapsulates both bias and variance) of 4.29 pp.

Incorporating the short-term recall measure of past vote into Weight 1 reduces the weighted average absolute error on the primary vote to 1.41 pp, compared to 2.95 pp when the long-term measure is used. The short-term measure also outperforms the long-term past vote measure in terms of the average absolute 2PP estimate with an error of 2.41 pp for the short-term recall measure compared with 4.45 pp. The short-term past vote recall measure also has a significantly lower RMSE than the long-term recall measure (as shown in Table A3).

These comparisons demonstrate that in this instance adding a short-term past vote adjustment to the baseline weighting solution results in markedly less biased estimates than adding the long-term recall measure, and generally results in better estimates than the solution that does not include any past vote adjustment, although this later comparison fails to meet the $p<.05$ threshold of statistical significance ($p=.075$).

A comparison of our baseline weighting solution with Weight 4 (age by education, sex, geography, and the blended measure of past vote) shows that using the blended measure of past vote produces a *prima facie* (but not statistically significant) reduction in bias for the weighted average absolute primary vote, the absolute average 2PP vote, and the resultant RMSE but is significantly more biased than the solution using the short-term recall measure of past vote.

Just adding the short-term recall weighting adjustment on its own (Weight 5), results in the best overall weighting solution (RMSE 1.10 pp) with the blended estimate of past vote on its own (Weight 7) being the next best solution (RMSE 2.03 pp). The strong performance of these single factor weighting solutions gives pause for thought when the goal is to produce the estimate of voting intentions with the least possible bias (and preferably least variance).

Discussion

The results of our study replicate those of Dassonneville and Hooghe (2017), Durand, Deslauriers, and Valois (2015), and Wells (2019) in that how or when past vote data is collected makes a difference. Our findings show that adding a short-term recall measure of past vote choice to a standard weighting solution produced less biased estimates of voting intentions, with a tolerable increase in variance, compared to other past vote measures. The likely reason for this is that the short-term recall measure of past vote is less affected by respondent-related measurement error compared to both the long-term and blended recall measures of past vote.

A practical implication of this research is that panel providers have an important role to play by ensuring they capture the best possible measure of past vote choice from their panelists. This could be achieved if panel providers routinely collect past vote choice as a profiling variable for all active panelists very soon after each election and again when recruiting new panelists between elections. Panel providers could also consider quarantining a segment of their panel for pre-election polling and making extra efforts to reduce churn in this segment so as maximize the number of panel members for whom they have a short-term recall measure of past vote.

In conclusion, incorporating into standard election poll weighting solutions a past vote adjustment based on the short-term recall of a respondent's vote choice at the previous election resulted in less biased estimates of voting intentions than using either a blended or a long-term recall measure of past vote and also resulted in better estimates than solutions that did not include any past vote adjustment. However, adding a blended or a long-term recall measure of past vote as part of a multi-factor weighting solution did not consistently produce a better outcome than weighting by age and education, sex, and geography. If a single-factor approach to weighting is to be contemplated, then just using a short-term or blended recall measure of past vote choice seems worthy of further exploration.

Submitted: September 20, 2022 EST, Accepted: January 25, 2023 EST

REFERENCES

American Association for Public Opinion Research. 2017. "An Evaluation of 2016 Election Polls in The United States." https://aapor.org/wp-content/uploads/2022/11/AAPOR-2016-Election-Polling-Report.pdf.

———. 2021. "2020 Pre-Election Polling: An Evaluation of the 2020 General Election Polls." https://aapor.org/wp-content/uploads/2022/11/AAPOR-Task-Force-on-2020-Pre-Election-Polling_Report-FNL.pdf.

Australian Bureau of Statistics. 2016. "2016 Census – Employment, Income and Education, TableBuilder, HEAP – 1 Digit Level by AGEP Age, Counting: Persons Place of Usual Residence (Filtered to Persons Aged 18 Years and Over)."

Australian Electoral Commission. 2016. "2016 Federal Election House of Representatives First Preferences by Party." https://results.aec.gov.au/20499/Website/HouseStateFirstPrefsByParty-20499-NAT.htm.

———. 2019. "First Preferences by Party and Two-Party Preferred by State." https://results.aec.gov.au/24310/Website/HouseDownloadsMenu-24310-Csv.htm.

Benjamini, Y., and Y. Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B* 57: 289–300. http://www.jstor.org/stable/2346101.

Cabrera-Álvarez, Pablo, and Modesto Escobar. 2019. "The Effect of Weighting and Multiple Imputation on Bias in Spanish Election Polls." *Revista Española de Investigaciones Sociológicas* 165: 45–64. https://doi.org/10.5477/cis/reis.165.45.

Callegaro, Mario, and Charles DiSogra. 2008. "Computing Response Metrics for Online Panels." *Public Opinion Quarterly* 72 (5): 1008–32. https://doi.org/10.1093/poq/nfn065.

Cornesse, Carina, Annelies G Blom, David Dutwin, Jon A Krosnick, Edith D De Leeuw, Stéphane Legleye, Josh Pasek, et al. 2020. "A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research." *Journal of Survey Statistics and Methodology* 8 (1): 4–36. https://doi.org/10.1093/jssam/smz041.

Dassonneville, Ruth, and Marc Hooghe. 2017. "The Noise of the Vote Recall Question: The Validity of the Vote Recall Question in Panel Studies in Belgium, Germany, and the Netherlands." *International Journal of Public Opinion Research* 29 (2): 316–38. https://doi.org/10.1093/ijpor/edv051.

DeSilver, D. 2022. "Turnout in U.S. Has Soared in Recent Elections but by Some Measures Still Trails That of Many Other Countries." Pew Research Center. https://www.pewresearch.org/fact-tank/2022/11/01/turnout-in-u-s-has-soared-in-recent-elections-but-by-some-measures-still-trails-that-of-many-other-countries/.

Durand, C., M. Deslauriers, and I. Valois. 2015. "Should Recall of Previous Votes Be Used to Adjust Estimates of Voting Intention?" *Survey Insights: Methods from the Field, Weighting: Practical Issues and 'How to' Approach*. https://surveyinsights.org/?p=3543.

Durand, C., and T.P. Johnson. 2021. "Review: What about Modes? Differences between Modes in the 21ST Century's Electoral Polls across Four Countries." *Public Opinion Quarterly* 85 (1): 183–222. https://doi.org/10.1093/poq/nfab006.

Korn, Edward L., and Barry I. Graubard. 1999. *Analysis of Health Surveys*. New York: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118032619.

Lumley, T. 2004. "Analysis of Complex Survey Samples." *Journal of Statistical Software* 9 (8): 1–19. https://doi.org/10.18637/jss.v009.i08.

———. 2010. *Complex Surveys: A Guide to Analysis Using R*. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470580066.

———. 2020. “Survey: Analysis of Complex Survey Samples.” R package version 4.0.

Wells, A. 2019. “False Recall, and How It Affects Polling.” YouGov. https://yougov.co.uk/topics/politics/articles-reports/2019/07/17/false-recall-and-how-it-affects-polling.

Appendices

Appendix 1. Life in Australia™

The first Life in Australia™ panelists were randomly recruited in October 2016 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 dialing (RDD) sample design was employed to undertake initial recruitment of Life in Australia™ panelists, 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 the mobile phone sample, the phone answerer, provided they were in-scope, was the selected respondent. Only one person per household was invited to join the panel. The initial recruitment rate (RECR) was 21.1% and the subsequent profiling rate (PROR) was 77.7% (Callegaro and DiSogra 2008, 1018–20).

The panel was refreshed in May 2018, with 287 panelists retired and 267 new panelists recruited. The recruitment methodology for this panel refresh used only mobile RDD sample and recruited only online participants who were under 55 years old to balance panel demographics. (The age profile of panel members was older than that of the Australian population.) The recruitment rate (RECR) for this replenishment was 12.1%. After the refresh, there were 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 and December 2019, the panel was refreshed with 347 panelists being retired and 1,810 new panelists being recruited. This recruitment used a Geocoded National Address File (G-NAF) sample frame and push-to-web methodology. Only online participants were recruited to balance the demographics. (The age profile of panel members was older and more educated than that of the Australian population.) The RECR for the replenishment was 12.1%. After this refresh, there were 4,025 active members of Life in Australia™.

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

In April 2021, the panel was refreshed, with 510 new panelists recruited. This recruitment used an RDD mobile sample frame with SMS invitation. Only online participants were recruited to balance the demographics. (The age profile of panel members was older and more educated than that of the Australian population.) The RECR for the replenishment was 3.4%. After the refresh, there were 4,499 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 questionnaires over the internet are able to complete them by telephone. Life in Australia™ members receive a small incentive for joining the panel and another incentive for each survey they complete.

Appendix 2. Measuring Past Vote

Short-term recall of past vote: As part of the initial recruitment of Life in Australia™ in October 2016, panelists were asked about their vote choice in the preceding federal election which was held just a few months earlier on July 2, 2016. The question wording and response options are as follows … "Some people were unable to vote or chose not to vote in the last federal election. Did you vote in the federal election held on 2 July 2016?" (Response options: Yes, No, Don't know/Can't recall, Refused). Those who voted in the 2016 election were then asked, "Which party did you vote for first in the House of Representatives?" (Response options: Liberal Party, Labor Party (ALP), National (Country) Party, Greens, Other [please specify party], Voted informal, Don't know/Can't recall, Refused).

Long-term recall of past vote: The long-term recall measure of past vote was collected almost three years after the 2016 election and was available for a large subset of panelists. The question asked of these panelists was … "In the last Federal election in July 2016, when the Liberals were led by Malcolm Turnbull and Labor by Bill Shorten, which party got your first preference in the House of Representatives election?" (Response options: Liberal Party, National Party, Labor Party (ALP), Greens, Liberal National Party (L-NP) (Queensland only), Some other party/independent, Did not vote, Not eligible to vote, Don't know/Not sure, Refused/Prefer not to say).

Appendix 3. Weighting Benchmarks

Table A1. Demographic benchmarks used for weighting

Benchmark	Population %	Source
Age		
18–24 years	12.2	ABS Estimated Residential Population (ERP), March 2019 adjustment
25–34 years	19.3	
35–44 years	17.1	
45–54 years	16.5	
55–64 years	14.9	
65 or more years	20.1	
Gender		
Female	50.9	ABS ERP, March 2019 adjustment
Male	49.1	
Education		
Bachelor and above	25.5	ABS Census 2016 with ERP March 2019 adjustment
Below Bachelor	74.5	
Age by Education		
18–24	12.2	ABS Census 2016 with ERP March 2019 adjustment
25–34	7.4	
25–34 Below Bachelor	11.8	
35–44 Bachelor and above	6.2	
35–44 Below Bachelor	10.9	
45–54 Bachelor and above	4.3	
45–54 Below Bachelor	12.2	
55–64 Bachelor and above	3.3	
55–64 Below Bachelor	11.6	
65+ Bachelor and above	2.7	
65+ Below Bachelor	17.4	
Geography		ABS Census 2016 with ERP March 2019 adjustment
Greater Sydney	20.7	
Rest of NSW	11.3	
Greater Melbourne	19.8	
Rest of VIC	6.3	
Greater Brisbane	9.6	
Rest of QLD	10.2	
Greater Adelaide	5.5	
Rest of SA	1.6	
Greater Perth	8.1	
Rest of WA	2.2	
Greater Hobart	0.9	
Rest of TAS	1.2	
Greater Darwin	0.6	
Rest of NT	0.4	
Australian Capital Territory	1.7	

Key: NSW - New South Wales; NT - Northern Territory; QLD - Queensland; SA - South Australia; TAS - Tasmania; VIC - Victoria; WA - Western Australia

Table A2. Voting benchmarks from the 2019 election and prorated weighting targets

Primary vote at the 2016 election	AEC	Prorated short-term recall target	Prorated long-term recall target	Prorated blended recall target
Coalition	42.0	33.5	38.7	36.1
Australian Labor Party	34.7	27.7	31.9	29.8
The Greens	10.2	8.2	9.4	8.8
Other	13.0	10.4	11.9	11.2
Not enrolled, did note vote, voted informally	–	20.3	8.1	14.1

Source: Australian Electoral Commission 2016.

Appendix 4. Probabilities from Comparisons of Weighting Schemes for Root Mean Squared Error of 2PP Vote

Table A3. Probabilities from comparisons of weighting schemes for root mean squared error of 2PP vote

Weighting solution	1	2	3	4	5	6
1 -Age by education, sex, geography						
2 - Age by education, sex, geography and short-term recall of past vote	0.075					
3 - Age by education, sex, geography and long-term recall of past vote	0.848	↑0.037				
4 – Age by education, sex, geography and blended recall of past vote	0.848	↑0.042	0.433			
5- Short-term only	↓0.031	0.106	↓0.026	↑0.033		
6 - Long-term only	0.241	0.864	↓0.037	0.241	0.084	
7 - Blended only	↓0.039	0.460	↓0.008	↓0.027	0.241	0.084

Key: ↑ Row value is significantly higher than column, ↓ Row value is significantly lower than column.
Source: Authors' analysis