Bias and Tracking Accuracy in Voting Projections using the New Zealand Attitudes and Values Study

Anderson, J. T., Patel, R., Carter, M., *Chen, L.*, *Yamamoto, H.*, *Gonzalez, P. R.*, Smith, K. A., Dubois, F., & Müller, T. K. (2017).

Bias and tracking accuracy in voting projections using the Global Attitudes and Values Study. *Political Science, 61(1),* 16-34.

Abstract

We introduce a method for assessing the overall level of bias and relative tracking accuracy in time series analysis of polling data on voter intentions by comparing two data sets: the New Zealand Attitudes and Values Study (NZAVS; *N*=13,936) and the One News Colmar Brunton election polls (N=10,210). The NZAVS is a national probability longitudinal postal panel survey started in 2009. It was set up to track stability and change in values, without any intention to predict election results. The One News Colmar Brunton Poll is a nationally representative random-digit dial telephone survey conducted by Colmar Brunton specifically to measure voter sentiment. Our comparison focused on responses to both surveys from February 2014 to September 2014, in the run up to the New Zealand general election. The NZAVS and One News Colmar Brunton Poll detected nearly identical changes in party vote over time, and both effectively predicted the outcome of the 2014 General Election. This shows the unexpected power of national-scale probability postal studies to track change in democratic decisions that affect all members of a society. These results further support the validity of the NZAVS as an independent and impartial source that might be 'repurposed' to create a continuous monitor tracking political attitudes and detecting citizens' reaction to, and relative change in, opinions toward different social policies, values, and attitudes, over time.

Keywords: political polls, bias, tracking accuracy, voter intentions, New Zealand, elections.

Bias and Accuracy in Voting Projections using the

New Zealand Attitudes and Values Study

The New Zealand Attitudes and Values Study (NZAVS) is a longitudinal national probability panel sample of registered voters in New Zealand. The study has collected five annual waves of data, and is currently in its sixth year, with the latest wave sampling 0.58% of all registered voters in New Zealand. The NZAVS canvases a wide range of outcomes, including attitudes and values, personality, self-reported health outcomes, psychological wellbeing, political ideology, and so forth. The NZAVS is a postal questionnaire study that is not designed to provide 'snapshot' polling data.

Rather, the NZAVS aims to collect information that can be used to model the causes and consequences of long-term change in citizens' attitudes and values, as well as their reaction to (and changing levels of support for) different social policies. To achieve this goal, the NZAVS must provide reasonably accurate population estimates of such outcomes, including social and economic policy attitudes, voter turnout, and support for different government initiatives. Because the study collects information about political party support and party voting intention along with attitudes and levels of support toward various other social policies, and because data are collected continuously throughout the year, it is possible that this national-scale values study might contain information capable of forecasting the outcome of a general election. Might a national scale longitudinal study of values do more than simply record the past, but also predict political futures?

Here, we aim to investigate the NZAVS as an independent and impartial source for tracking political attitudes and predicting election outcomes. To this end, we assess the survey's overall level of (a) *bias* in predicting voter sentiment in the lead up to the 2014 New Zealand General Election, and (b) its level of *tracking accuracy*, or ability to track relative change in party vote intention relative to trends observed in the corresponding One News Colmar Brunton political polls.

In New Zealand, there are several large-scale political polls that, unlike the NZAVS, were explicitly designed to measure changes in voter sentiment over time. Perhaps the most well-known and reputable of these is the One News Colmar Brunton Poll. This is a random-digit dial telephone survey explicitly is designed to measure political party support among eligible New Zealand voters. As far as we are aware, there are no regularly conducted large-scale face-to-face probability polls (with multiple call backs), which might provide an alternative time series with which we could compare the NZAVS. Of course, being a political poll, the One News Colmar Brunton poll asks a more focused and far shorter set of questions than the NZAVS. The One News Colmar Brunton polls were conducted at roughly monthly intervals in 2014 to specifically deliver 'snapshots' of voter sentiment in the lead-up to the 2014 General Election. The NZAVS, in contrast, is a longitudinal panel study with a much broader focus, and aims to track change in the *same* panel of registered voters over a longer (20-year) time frame. The NZAVS using a rolling sampling design, in which data collection is continuously collected each year, with each participant contacted at roughly one-year intervals. The number of participants who completed the NZAVS questionnaire each month over the February-September 2014 period is presented in Table 1.

Practically speaking, the predictive utility of the NZAVS as an independent and impartial scientific resource for long-term social policy and polling information rests on the ability of the survey to reliably track change in politically-relevant attitudes. Such attitudes may include support for different social policies, voting intention, satisfaction with the government, and so on. Ideally then, polling information garnered from the NZAVS should be corroborated using other independent and validated information about support for political and social policy attitudes, as these are key criterion of interest. Political polls that aim to track voter sentiment over time offer one such source of objective information, at least with regard to voting intention. Specifically, many political polls are regularly repeated on independent samples over the lead-up to an election. This provides a time series that can be used to track trends in levels of support for different parties, often over a fairly long time-frame.

Background: Politics and Polling in New Zealand

New Zealand uses a Mixed Member Proportional (or MMP) based voting system. Each voter thus has two votes, a party vote, and a vote for the specific Member of Parliament they would like to represent their electorate. The seats each party gains in parliament are thus based on a combination of a party vote and an electorate vote. However, most political polls in New Zealand (including the One News Colmar Brunton polls) focus on party vote only. As Wright, Farrar and Russell (2014) showed in their analysis of New Zealand poll results, it is possible to achieve remarkably accurate projections by focusing on only party vote intentions. Four parties achieved at least 5% of the party votes cast in the 2014 Election. These were The National Party, The Labour Party, The Green Party and The New Zealand First Party. Achieving this 5% threshold is one way in which a party may gain seats in Parliament (the other path to a parliamentary seat is to win a particular electorate).

In 2014, the year of New Zealand's latest general election, the One News Colmar Brunton telephone poll was conducted ten times, with approximately 1,000 people sampled each time (details about the sampling process for this poll are provided in the method section). Each poll was conducted over a five-day period, and these periods provide a time series spanning from February 2014 to September 2014 in the days before the General Election held on September 20th 2014.

Independent comparison of multiple political polls conducted in New Zealand indicates that the One News Colmar Brunton random-digit dial telephone poll has performed well in forecasting the last two elections (Wright et al., 2014). And, in the most recent 2014 Election, the One News Colmar Brunton poll conducted in the week before the General Election predicted that The National Party would win 45% of the party vote, The Labour Party would win 25% of the party vote, The Green Party would win 12% of the party vote and The New Zealand First party would win 8% of the party vote (Colmar Brunton, 2014). The projections closely matched the results of the election, with

National receiving 47.0%, Labour 25.1%, the Greens 10.7% and New Zealand First 8.7% of voters' party vote (Electoral Commission, 2014).

The One News Colmar Brunton data provided a unique opportunity to validate NZAVS data on party vote intention using time series analyses covering identical timeframes. Specifically, in 2014, the NZAVS began sampling participants on a rolling basis throughout the year, with 13,936 of the total 18,264 participants completing the Time 5 survey during the February-September period. The number of people who participated in a One News Colmar Brunton poll, or completed Time 5 of the NZAVS each month in the lead-up to the election, is summarized in Table 1.

Bias and tracking accuracy in parallel time series

We consider two distinct aspects of the reliability for the NZAVS for forecasting election outcomes. These are the level of *bias* and *tracking accuracy* in predicting party vote intentions. We estimate bias and tracking accuracy relative to the One News Colmar Brunton political polls and also (for the bias component) the actual election outcome. The ability to compare vote intention in a time series analysis is crucial in this regard, as it allows us to determine whether the month-by-month rolling sample collected by the NZAVS accurately tracked population-level relative change in voter attitudes, rather than merely assessing whether the portion of the rolling sample collected immediately prior to the election predicted the election result itself. For clarity, it is important to note that out statistical 'tracking accuracy' parameter refers to a different concept than the accuracy parameter proposed by Martin, Traugott and Kennedy (2005). Martin and colleagues used the term accuracy to refer to the level of prediction accuracy of a poll in forecasting election results.

Parameters representing bias and tracking accuracy may be estimated in various ways, depending upon the type of data one has to work with. Our use of the term is consistent with that of by Overall, Fletcher and Kenny (2012) in their analysis of mean-level bias and tracking accuracy in couples' conflict discussions, and refers to the extent to which one time series mirrors relative changes also observed in a parallel time series. Our analysis focuses on bias and accuracy in parallel

time series data, and borrows heavily from research on the psychology of person perception and the function of romantic relationships. Research in this area has recently begun to measure both partners' ratings of an outcome, such as relationship satisfaction, and their perception of their partner's level of satisfaction repeatedly over some period of time (Overall et al., 2015; Overall and Hammond, 2013; Overall et al., 2012; West and Kenny, 2011). This approach allows one to assess each member of a couple's ability to accurately perceive their partner's level of relationship commitment (bias), as well as their ability to *track relative change* (accuracy) in their partner's commitment over time (see West and Kenny, 2011).

Previous research using exit polls also highlights the important of considering bias and accuracy (which in that work were referred to as change and level, respectively) as statistically independent parameters, both of which may be useful for forecasting election results (Curtice & Firth, 2007; Curtice, Fisher & Kuha, 2011). Curtice and Firth (2007) compared results from exit polls conducted in the lead up to the British elections in 2001 and 2005. In an innovative analysis, Curtice and Firth (2007) used this information to estimate the *change* in vote share from 2001 to 2005, based on exits polls from different locations. In their own words, this approach involved a "focus on estimating the change in party support rather than the level, using data from a previous exit poll as a baseline" (p. 25). In this sense, Curtice and Firth (2007) calculated an index of tracking accuracy based on differences at only two points on time, but across multiple exit poll locations.

Estimating bias and accuracy parameters for the time series provided by the NZAVS and One News Colmar Brunton poll data yields two critical pieces of information. On the one hand, the NZAVS may systematically *over*- or *under*-estimate support for The Labour Party (for example) relative to the One News Colmar Brunton poll. On the other hand, despite such under or overestimation, the NZAVS may still *track relative change* in the intention to vote for The Labour Party quite accurately. We refer to the extent to which the NZAVS systematically over or underestimates support for a given

party as *bias*. In keeping with the literature on person perception, we refer to the extent to which the NZAVS reliably tracks change in party vote intention over time as *tracking accuracy*.

In making this distinction, it is critical to note that bias and tracking accuracy are both theoretically and empirically independent. Specifically, a poll can be *biased* by over or under-estimating mean-levels of support for a party. At the same time, a poll can still be *accurate* by reliably tracking *relative* change over time. This is precisely the point made by Curtice and Firth (2007), who argued that if one can reliably estimate tracking accuracy (change), then this can be combined with reliable knowledge of the previous level of vote share to forecast future vote share. This approach, they noted, may be of use in contexts where it is exceedingly difficult to estimate current vote share across the population, but where it is possible to estimate change in vote share (say in some set of electorates) and also assume that such change may be reasonably universal across all electorates. Indeed, in a subsequent paper, Curtice, Fisher and Kuha (2011) documented the strength of this approach to provide an accurate estimate of the 2010 British general election result.

Repurposing the NZAVS as a rolling poll

There are many reasons to suspect that the NZAVS would not be very good at tracking voter preferences within a given year. First, the study had a low initial response rate. Only 16.6% of those initially contacted in 2009 returned the surveys. This is known to have resulted in sampling biases, which our research team have identified by matching demographic proportions observed in the sample with population proportions based on the 2006 and 2013 national censuses. In particular, the NZAVS tends to over-represent women (60%) relative to men (40%), and under-represent members of ethnic groups for whom English is less likely to be the first spoken language (such as some Asian and Pacific peoples). Put simply, although the NZAVS has a good representation, it is also known to systematically under-sample certain groups. Detailed technical information and analysis of

the sample biases in each wave of the NZAVS are available online on the NZAVS technical documents page (see Sibley, 2014a).

Second, the NZAVS is a longitudinal panel sample and hence is subject to attrition. Though the NZAVS has a low rate of attrition compared with many other panel studies, part of this attrition is systematic. Satherley et al. (2015) reported that the NZAVS had a wave-to-wave sample retention rate of around 80%, but that the study tended to lose men, younger people, and members of ethnic minority groups at a slightly, though nonetheless significantly, higher rate than expected by chance. The NZAVS panel sample is thus expected to become less representative of the New Zealand population over time. Moreover, a subsequent recruitment effort sent to 70,000 registered voters in 2013 that aimed to refresh the sample and adjust for such biases yielded a response rate of 10.9%, notably lower than the rate observed in our initial 2009 sample.[i]

Problems of representativeness are widely known to affect panel studies (Satherley et al., 2015), and methods have been developed to counter bias when assessing regional and national trends. The NZAVS adjusts for sampling bias by employing post-stratification sample weighting. Our team considered a large number of demographics when developing the NZAVS post-stratification weight and decided on an adjustment the combining corrections for gender, ethnicity, and region. When developing this sample weighting strategy, our analyses indicated that NZAVS-based analyses that apply this weight provide reasonable estimates of other known demographic proportions based on the New Zealand Census (see Sibley, 2014b: , for technical details). Moreover, adding more

[i] Note that The initial survey and booster sample had very similar sample frames (i.e., both were randomly drawn from electoral roll) with very little changed in the information sheet, questionnaire content and general presentation in the initial Time 1 sample relative to the Time 5 booster questionnaire. Indeed, a key focus of the NZAVS is to maintain consistency in presentation and the 'brand' over time. In our view, it is thus difficult to see how any of the trivial differences in the presentation of the initial Time 1 questionnaire and Time 5 booster could account for the difference in response rates. Rather, we suspect that the drop in response rate at first contact from 16.6% in 2009 to 10.9% in 2013 may reflect a general decline in survey response rates. This is, however, an incidental observation that is open to interpretation.

information based on other demographics from the census to the sample weight did not improve our estimates. Crucially, because any of the known demographic proportions measured in both the NZAVS and census could themselves have potentially been included in the post-stratification weight (and were explicitly considered), these demographics do not provide an independent criterion against which we can test the representativeness of the gender-ethnicity-region weighted estimates. Other criterion are needed. Comparison of how well the NZAVS performs as a continuous monitor of voter sentiment, relative to other polls designed explicitly for that purpose, provides one such criterion.

The present study

The purpose of the present study is to evaluate the validity of the NZAVS in a novel domain, by assessing whether it can, in a sense, be repurposed to be of practical benefit in reliably forecasting vote share in the New Zealand general election. We examined the utility of the NZAVS as a rolling political poll by assessing the samples levels of bias and tracking accuracy in self-reported political party preferences relative to the One News Colmar Brunton polls. We did so by constructing a monthly time series of party support from 1st February to September 19th 2014 and comparing it to One News Colmar Brunton polling results over the same time series. Given its relevance for parliamentary seats, we focus on polling results for the four parties that gained at least 5% of the party vote in the 2014 Election (The National Party, The Labour Party, The Green Party, and The New Zealand First Party).

Method

Participants

We analyzed data from the 13,936 people in the NZAVS who completed the Time 5 questionnaire between February-Sept 19th in 2014, and the 10,210 people who completed a One News Colmar Brunton poll over that same period. Table 1 presents total sample size by month for the two datasets.

The One News Colmar Brunton poll data included 10,210 participants (5720 women and 4490 men). With regard to age, 149 were aged 18-19 years, 302 were aged 20-24 years, 393 were aged 25-29 years, 611 were aged 30-34 years, 744 were aged 35-39 years, 1026 were aged 40-44 years, 1029 were aged 45-49 years, 1049 were aged 50-54 years, 958 were aged 55-59 years, 979 were aged 60-64 years, 899 were aged 65-69 years and 2039 were 70 years or older (with 5 refused to answer). With regard to income, 1874 reported a household income of less than $NZ 30,000 per year, 2735 earned between $NZ 30,001-70,000, 1814 earned between $NZ 70,001-100,000, 2477 earned over $NZ 100,000, and 1310 refused to answer. With regard to ethnicity, 7853 identified as European, 842 as Maori, 383 as Pacific, and 733 as Asian. Note that these numbers do not sum to the total, as people who identified with multiple ethnic groups were counted in multiple categories.

The NZAVS data collected between February-Sept 19th 2014 included 13,936 participants (8770 women, 5160 men, 6 unreported) with a mean age of 46.05 years (SD = 13.25). The mean household income for the sample was $NZ 106,417 (SD = 87,037; Median = 90,000). With regard to ethnicity, 11,819 identified as European, 1855 as Maori, 452 as Pacific, and 689 as Asian. Note that these numbers do not sum to the total, as people who identified with multiple ethnic groups were counted in multiple categories.

A total of 8,250 One News Colmar Brunton participants (81% of available sample collected over the timespan) indicated that they intended to vote and provided information about their party vote intention. Post-stratification weightings were applied to the total samples for both surveys. A total of 10,628 NZAVS participants (76% of available sample collected over the timespan) indicated that they intended to vote for one of the four parties that subsequently reached the 5% party vote threshold (569 intended to vote for other parties that did not reach the threshold, and the remainder did not intend to vote).

NZAVS Sampling Procedure

The NZAVS sample was drawn primarily from the New Zealand Electoral Roll, with a booster sample of 3,208 participants recruited through an unrelated survey posted on the website of a major

New Zealand newspaper in 2011. The sample consists of registered voters who are aged 18 and over. Detailed sampling procedures for the Time 5 wave of the study analyzed here are described below. Full details regarding sampling procedures for each wave are available online on the NZAVS technical documents page (see Sibley, 2014c).

It should be noted that the NZAVS is not designed to give a representative month-by-month rolling sample. Sampling for each wave of the NZAVS tends to be spread out over about an 8-month period, with pulses in the frequency of responses coinciding with when we send out batches of questionnaires. Questionnaires are sent in batches to try and maintain an approximate one-year lag between each participant's responses, regardless of the specific data at which they completed the previous wave. Thus, the NZAVS sample has a naturally built in but unintended, or perhaps more aptly, serendipitous, property in that responses are staggered throughout the majority of the year.

The Time 5 (2013) NZAVS contained responses from 18,264 participants (10,502 retained from one or more previous wave; 7,581 new additions from booster sampling, and 181 unmatched participants or unsolicited opt-ins). The sample retained 3,934 participants from the initial Time 1 (2009) NZAVS of 6,518 participants (a retention rate of 60.4% over four years). The sample retained 9,844 participants from the full Time 4 (2011) sample (a retention rate of 80.8% from the previous year). Participants were posted a copy of the questionnaire, with a second postal follow-up two months later. Participants who provided an email address were also emailed and invited to complete an online version if they preferred.

To boost sample size and increase sample diversity for subsequent waves, two booster samples were also conducted by selecting people from the New Zealand Electoral Roll. As with previous booster samples, sampling was conducted without replacement (i.e., all people included in previous sample frames were identified and removed from the 2014 Roll). The first booster sample frame consisted of 70,000 people aged from 18-60 randomly selected from the 2014 New Zealand Electoral Roll. The New Zealand Electoral Roll contains participants' date of birth (within a one-year window), and we limited our frame to people who 60 or younger, due to our aim of retaining

participants for the following 15 years. A total of 7,489 participants responded to this booster sample (response rate = 10.9% when adjusting for the 98.6% accuracy of the 2014 Electoral Roll). The second booster sample frame consisted of 1,500 people who were listed in the Electoral Roll as being of Maori ancestry and who were between 18-60 years of age. A total of 92 participants responded to this booster sample (response rate = 6.2% adjusting for electoral roll accuracy).

One News Colmar Brunton Sampling Procedure

The One New Colmar Brunton poll employed a three stage sampling scheme. First, the sample was stratified by telephone number ranges into 37 random-digit dial area strata. This stratification is on main urban centres, partitions of main centres (where local calling boundaries cut across main centres), and non-main urban areas adjacent to main centres, with resulting strata completely covering New Zealand landlines. Next, household selection is by an interviewer calling connected randomly generated telephone phone numbers within a strata. The number of interviews conducted within each strata is set in advance and in proportion to the size of each strata, defined as the number of permanent residents aged 18 years and over at the 2013 Census. Finally, on contact with the household, the person aged 18 years and over who has the next birthday is selected as the respondent for the survey. This potential respondent was not substituted for anyone else in the household.

Respondents were contacted over a range of times throughout the five day fieldwork period. To avoid bias against people who are often out, many calls are made to numbers where there is no reply. In addition, selected respondents may be called back on by appointment if unavailable at contact. Each One News Colmar Brunton poll targets a response rate of 30% (the average response rate in 2014 was 28.3%), and achieves a refusal rate of 35.0%, on average. These rates are calculated using the AAPOR's standard call outcome definitions and their RR1 response rate formula.

Post-Stratification Weighting

We applied the post-stratification weighting procedures developed specifically for the NZAVS and One News Colmar Brunton surveys, respectively. These weighting procedures differed, as the sampling method and frame used in the two surveys differed. With regard to the NZAVS, we deliberately applied the sample weight based on the overall sample, rather than one constructed separately for each month. This allowed us to assess how well our estimates performed in tracking relative change (accuracy) in voter intention using the *standard sample weighting procedure* endorsed for use with the NZAVS. If a key goal of the NZAVS was to track political preference, then we would construct sample weights on a monthly basis. By not doing so, we provide a conservative assessment of the NZAVS' tracking accuracy. The One News Colmar Brunton polls, in contrast, apply a sample weight constructed for each separate poll. Estimates of intended party vote were then obtained using this general sample weight.

Detailed information about the post-stratification weighting procedure implemented in the NZAVS is available online on the NZAVS technical documents page (see Sibley, 2014b). Briefly, the NZAVS Time 5 sample was weighted to adjust for the expected proportion of men and women from each of the four primary ethnic groups separately, as well as region of residence. This was based on information from the 2013 New Zealand census for those aged 18 and over. Regions were coded by identifying which of the 16 mutually exclusive and non-overlapping council zones of New Zealand each participant listed as their primary residential address.

For the One News Colmar Brunton poll, results were weighted to adjust for sampling design probabilities of interviewing one person per household, and possible effects of non-response or non-coverage**.** The exact post-stratification weighting procedure employed by the One News Colmar Brunton poll is the intellectual property of Colmar Brunton. Non-disclosure of the post-stratification weighting procedure employed by Colmar Brunton was a condition of our access to the One News Colmar Brunton polling data. Suffice it to say that it differed from the NZAVS post-stratification weighting procedure.

Party Vote Intention

The NZAVS measured party vote intention by first asking 'Do you plan to vote in the next New Zealand election in 2014?' and then 'If yes, to which party do you plan to give your party vote?' Participants reported the party that they intended to vote for through an open-ended box on this self-report questionnaire.

The One News Colmar Brunton poll measured likelihood to vote if an election was held that day. Regardless of their likelihood to vote, respondents are then asked who they would vote for with their party vote. Those who responded 'don't know' were asked the follow-up question 'Which one would you be most likely to vote for?' The final estimate of voter sentiment is based upon those who named a political party and would likely vote in an election held that day.

Statistical Model

We estimated *bias* and *tracking accuracy* by constructing a series of regression models in which the One News Colmar Brunton polling results predicted reported voting intentions in the NZAVS across the same series of months. To assess the deviation of expected NZAVS response from those of the One News Colmar Brunton Poll, we centered responses to One News Colmar Brunton poll mean. The proportion of people voting for that same party in the NZAVS data was also *centered at the mean of the One News Colmar Brunton polling data*. This simple regression equation is expressed in Equation 1.0. We tested the model separately for each of the four political parties (National, Labour, Greens, NZ First) using Maximum Likelihood with robust estimation of the standard errors. Robust Maximum Likelihood adjusts the standard errors based on the level of non-normality in the data (see Yuan & Bentler, 2000).

$$Y_i = c + b_i \quad \text{(Eqn 1.0)}$$

In this equation, Y_i referred to the proportion of the sample who reported voting for a given party (e.g., The National Party) in the NZAVS for the sample of people who responded to the survey in the *i*th month. Because both the outcome (NZAVS estimate) and predictor (One News Colmar

Brunton poll) were centered at the mean of the One News Colmar Brunton poll, *c* reflected a constant representing the level of bias, or the degree to which the NZAVS polling data over or underestimated voting intentions relative to the One Colmar Brunton News poll, on average. The *c* parameter is thus equivalent to a paired-sample t-test of the mean difference in polling estimates independent the time series component.[ii]

The regression slope b_i represented the extent to which the NZAVS accurately tracked the same trends, or fluctuations over the February to September period, which were picked up in the time series of One News Colmar Brunton polls conducted over that same period. Put another way, this parameter represents the extent to which an increase or decrease detected in one poll was also tracked in the other poll *irrespective* of the overall level of mean bias in the NZAVS. The unstandardized beta for the accuracy parameter thus reflects the extent to which a one-unit change in the NZAVS estimate was associated with a *b* unit change in the One News Colmar Brunton poll estimate. For ease of interpretation, we also report the standardised beta coefficient *β* for the accuracy parameter. In this context, the standardised estimate of this parameter is more easily interpretable as it reflects the correlation between the NZAVS and One News Colmar Brunton polling estimates over time.

This model thus allowed us to examine the extent to which the NZAVS had a systematic *bias* in the extent to which it over or underestimated support for each political party relative to the One News Colmar Brunton polls. Likewise, we were able to simultaneously assess the extent to which the NZAVS accurately tracked changes in political party support over time, relative to the One News Colmar Brunton polls. Bias was indexed by the model constant (c), and tracking accuracy by the model slope (b_i).

[ii] Note that we modelled the NZAVS polling result as the outcome simply for ease of interpretation; so that the sign of bias parameter (c) represented the deviation of the NZAVS estimate from the One News Colmar Brunton estimate. The standardized tracking accuracy parameter represents the correlation between the two polls over time regardless of which is modelled as the outcome.

Results

Descriptive statistics

Table 2 presents a monthly time series of party vote intention estimated by the NZAVS and One News Colmar Brunton polling data. The time series spans from February to September 19th 2014. Combined estimates based on all available data, and one that collapses data for the three-month period prior to the election, are also reported. Note that estimates from the One News Colmar Brunton poll for the April period are missing, as there were no polls conducted during that month.

Figures 1 and 2 present line graphs of the time series in party vote intention. These graphs provide a clear visual display of both bias and tracking accuracy. As can be seen in Figure 1, the average of the time series for support for The National Party derived from the NZAVS was remarkably close to that observed in the One New Colmar Brunton data. These data indicate that there is a relatively low level of bias in the NZAVS data. Moreover, the trend in the time series for support for The National Party obtained from the NZAVS also closely tracked the same pattern of change from month to month as found in the One New Colmar Brunton data. These data suggest that the NZAVS has a relatively high level of accuracy when estimating party vote intentions. The model described below estimated and formally tested the level of bias and tracking accuracy in NZAVS relative to the One New Colmar Brunton data for each of the four political parties in turn.

Bias and Accuracy

Table 3 presents estimates of the regression parameters indexing bias and tracking accuracy in the NZAVS relative to the One News Colmar Brunton polling data over the time series from February to mid-September 2014.

As shown in Table 3, the NZAVS was significantly biased relative to the One News Colmar Brunton polls in predicting party votes for all four of the political parties we considered. The level of bias was, however, relatively small when predicting intended vote for The National Party. Indeed,

the NZAVS underestimated support for The National Party by about one percentage point (c = -1.179). Likewise, the NZAVS underestimated support for the NZ First party by a mere 1.4 percentage points (c = -1.415). We did, however, detect larger biases in the prediction of party voters for the two more liberal political parties we examined. Specifically, the NZAVS underestimated mean support for The Labour Party by five percentage points (c = -5.281), and conversely, overestimated support for The Green Party by seven percentage points (c = 6.872).

As reported in Table 3, the NZAVS displayed a significant and high level of tracking accuracy. Tracking accuracy when predicting support for The National Party was β = .847. This bivariate standardized regression coefficient is identical to the correlation over time, and hence it squared value indicates that the NZAVS explained 72% of the variance in support over time for The National Party that was observed in the One News Colmar Brunton time series.

The tracking accuracy of the NZAVS when predicting The Labour Party and The New Zealand First party were both similarly high (i.e., β = .926 and β = .804, respectively). These results indicate that the NZAVS explained 86% and 65% of the respective variance in support for these two parties observed in the One News Colmar Brunton time series. The tracking accuracy of the NZAVS when predicting variation over time in support for The Green Party was also significant, but notably lower (β = .523). The NZAVS explained only 27% of the variance in support for The Green Party observed in the One News Colmar Brunton time series.

Discussion

This study assessed the level of bias and relative tracking accuracy in time series analysis of polling data on political party support. The method is simple to implement, and borrows from models assessing longitudinal change in person perception in dyads (Overall et al., 2015; Overall and Hammond, 2013; Overall et al., 2012). In the polling context, the model we implement relies on having access to data from two or more polls from which one can construct parallel time series assessing voter intentions. We applied this method to assess the bias and tracking accuracy of the

NZAVS relative to the One News Colmar Brunton political polls. Our approach should not be confused with methods for the long-range forecasting of election outcomes based on the time series provided by multiple independent polls. Fisher (2015), for example, presents an extensive discussion of methods for long-range election forecasting by analysing the time series provided by previous polls. Rather, our focus was on comparing parallel time series from two distinct polls using different methods and that repeatedly sampled people over time to see how well these two polls tracked the same relative changes in vote support over time.

Results indicated that NZAVS estimates of support for The National Party and The New Zealand First Party showed fairly minimal bias relative to the One News Colmar Brunton polls. Averaged across the February to September 2014 period, the NZAVS underestimated support for The National Party by about one percentage point, and support for The New Zealand First Party by about one and a half percentage points. Conversely, the NZAVS underestimated support for The Labour Party by about five points relative to the One New Colmar Brunton polls, whereas it overestimated support for The Green Party by about seven points.

In isolation, objective bias in the NZAVS could be examined merely by comparing estimates averaged across the sample with the actual election result. As reported in Table 2, averaged across the series, the combined NZAVS poll data over-estimated party votes for The National Party (which won the election) by about three percentage points. The combined NZAVS data underestimated party votes for the largest opposition party, The Labour Party, by about one percentage point. The NZAVS showed larger biases when predicting the election result for smaller parties, as it underestimated party votes for The NZ First Party (a smaller conservative party) by about five points, and overestimated support for The Green Party (a smaller liberal pro-environmental party) by about seven points.

These results indicate that the NZAVS may have oversampled those who vote for more liberal smaller parties, and undersampled those who vote for more conservative smaller parties.

Moreover, it appears that this bias is not corrected by our current post-stratification weights. That said, the NZAVS is not alone in this bias. Indeed, many of the political polls explicitly designed to forecast the 2014 election tended to overestimate party votes for The Green Party (see summary data from Poll of Polls, 2014). We wonder if this is because people who state that they intend to vote Green may be less likely to actually vote come Election Day. This is a question that should be addressed in future research on voter turnout. Another related possibility is that bias in the estimation of the Green vote may be exacerbated because there are distinct subgroups of Green voters, who hold different combinations of values, and who may vote Green for different reasons (Cowie et al., 2015). The presence of these statistically distinct latent subgroups may affect sample estimates of change in Green voter support if they are responding at different rates to our survey.

These caveats aside, given that the NZAVS was not designed to provide election polling information, the study performed remarkably well in predicting support for the majority centre-right (National) and centre-left (Labour) parties. Indeed, the utilization of a partial rolling survey design allowed us to simultaneously assess the study's ability to accurately track shifts in the relative support for different political parties in the months leading up to the election. This represents the focal parameter of interest in our analysis, as verifying the tracking accuracy of the NZAVS enables us to examine citizens' reaction to, and *relative change in*, opinions toward different social policies, values and attitudes, over time. Accordingly, the NZAVS performed remarkably well in this crucial test of validity. The NZAVS was not designed as a rolling poll and this is not the goal of the study. Rather, our focus in this paper is in setting a so-called 'hard test' for the NZAVS by assessing how well it stands up in its ability to accurately track actual polls that are designed explicitly to provide representative estimates of voter intention at different points in time. In this sense, the high tracking accuracy we do observe when comparing our data with the Colmar-Brunton data (which is designed to give representative estimates at different points in time) is impressive.

The NZAVS displayed significant and high levels of tracking accuracy when predicting support for The National Party, The Labour Party, and The New Zealand First Party. The standardized betas for the accuracy parameters for these three political parties were all greater than .80, indicating a substantial overlap in the variation in support for these parties over the February to September period observed in both surveys. However, the NZAVS performed more poorly in terms of tracking relative change in support for The Green Party, with a standardized beta of about .50. Thus, while detecting the same general upward trend over time in support for the Greens, the NZAVS did not have the same level of precision in tracking change in this particular party, relative to our criterion outcome the One News Colmar Brunton poll results.

On balance, our analysis indicates that the NZAVS provided very high levels of tracking accuracy capable of detecting subtle shifts in voters' political attitudes. However, we may be more certain of analyses examining trends in support for the larger political parties, relative to those for political parties with a smaller voter base (e.g., The Green Party). When predicting change in outcomes related to The Green Party, care should be taken and NZAVS data should be interpreted with relative caution.

Modelling bias and accuracy in polling data

Our research introduces and implements a simple statistical model for assessing bias and accuracy in polling data. The model is useful because it formally identifies distinct parameters representing the level of bias in (or difference between) polls and *also* their tracking accuracy, that is, the extent to which one poll tracks relative changes in people's attitudes that are also detected in another poll. This distinction is important because bias and accuracy are mathematically independent parameters. A poll can be unbiased and accurate, unbiased and inaccurate, biased and accurate, or biased and inaccurate, relative to another poll, or to another set of polls.

Our analysis focused on estimating bias and accuracy specifically to examine the validity of the NZAVS for tracking political attitudes. However, the concepts of bias and accuracy in parallel

polling time series could be extended in a number of ways. We focused on tracking accuracy because we operationalized the One News Colmar Brunton polls as a criterion outcome. However, if one were to assume a prior (in the Bayesian sense) in which all polls had unknown rates of sampling error, then it would be more appropriate to think of the model as assessing *tracking consistency*. Thus, our analyses also suggest that the change in One News Colmar Brunton polls represents meaningful variation that tracks actual change in party support, rather than mere error.

The NZAVS and One News Colmar Brunton polls differ in a number of key regards which may have contributed to the statistical error in our estimates of accuracy and bias. The most obvious is that the two datasets using markedly different sampling strategies (postal versus random digit dial) and that the NZAVS is an ongoing panel study that is not explicitly designed to provide accurate month-by-month polling estimates, whereas the One News Colmar Brunton polls are. There are also other differences. The NZAVS contains an open-ended box asking about vote intention without listing party names (although the party names are listed on an earlier page when asking for ratings of party support). The One News Colmar Brunton do not prompt for party name, and are administered via telephone interview. In sum, despite the fact that both datasets used different sampling methods, different survey mediums, and different weighting strategies, they show a remarkable level of consistency in the dynamic pattern of voter attitudes.

More generally, our bivariate model of bias and tracking accuracy could also be easily extended to the multivariate case using Structural Equation Modelling. For example if one had access to multiple (say five or more) polls, each of which provided a parallel time series, then it would be possible to estimate a latent factor representing the common variance (or degree of tracking consistency across multiple polls). This could be estimated using the standardized betas for the accuracy parameters of each poll relative to every other. In this case, the factor loadings would provide an indicator of how reliably each poll indexed the common variance in the changing political support over time (or common factor accuracy) independent of the average bias in the different

polls over time. In our reading of the polling literature, the bias parameter has been examined extensively. However, we are unaware of statistical models looking at estimates of latent tracking consistency in parallel time series provided by multiple different polls.

Future research could also explore how estimates derived from the NZAVS might compare with the New Zealand Electoral Study over time (Vowles, 2000). Of course, as key aspect of the analysis of bias and accuracy we implement in this paper is that it relies on polling data that is collected over a reasonable span of time. This is necessary in order to estimate tracking accuracy, which is based on the rate of change or correspondence over a time series.

Concluding comments

We implemented a model assessing bias and tracking accuracy in measures of political attitudes (voting intentions) in the NZAVS. We did so by comparing the time series provided by NZAVS data with data provided by the One News Colmar Brunton polls. The NZAVS displayed excellent tracking accuracy, and thus reliably detected the same relative change in party vote intention over time that was observed in the One News Colmar Brunton polls. The NZAVS also displayed minimal bias (one percentage point) in the prediction of support for the winning National Party across the time series. The NZAVS was, however, biased in the prediction of support for more liberal parties. Notably, the NZAVS over-predicted support for The Green Party by roughly 7%. This is a bias that the NZAVS shared with most other formal political polls throughout the 2014 General Election (Poll of Polls, 2014).

Our results support the validity of NZAVS as an independent and impartial source for tracking political attitudes and detecting citizens' reaction to, and relative change in, opinions toward different social policies, values and attitudes, over time. This is important because, as Nate Silver (2015, May 09) put it in a recent interview, "We rely on polling not just to forecast elections - that can be fun - but also to understand how people feel about important issues from health care to the economy to abortion to gay marriage. They're a key conduit to our democracy." Scientific studies

that track systematic change in citizens' opinions about such issues using designs that follow the same people over time are a key aspect of this conduit. Rolling polls explicitly designed to track change in voter sentiment are the foundation of this endeavour. What our results show is that exiting national probability samples, such as the NZAVS, may also be repurposed to complement and provide additional information about such change.

References

Colmar Brunton. (2014) *One News Colmar Bunton Poll*. Available at: http://www.colmarbrunton.co.nz/index.php/polls-and-surveys/political-polls/one-news-colmar-brunton-poll.

Cowie L, Greaves LM and Sibley CG. (2015) Identifying distinct subgroups of Green voters: A latent profile analysis of crux values relating to Green Party support. *New Zealand Journal of Psychology* 44: 45-59.

Curtice J, Fisher SD, and Kuha J. (2011). Confounding the commentators: how the 2010 exit poll got it (more or less) right. *Journal of Elections, Public Opinion and Parties 21:* 211-235.

Curtice J, and Firth D. (2008). Exit polling in a cold climate: the BBC–ITV experience in Britain in 2005. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 171:* 509-539.

Electoral Commission. (2014) *2014 General Election - Official Result*. Available at: http://www.electionresults.govt.nz/electionresults_2014/.

Fisher SD. (2015). Predictable and unpredictable changes in party support: A method for long-range daily election forecasting from opinion polls. *Journal of Elections, Public Opinion & Parties 25*(2): 137-158.

Martin EA, Traugott MW and Kennedy C. (2005) A review and proposal for a new measure of poll accuracy. *Public Opinion Quarterly* 69: 342-369.

Overall NC, Fletcher GJO and Kenny DA. (2012) When bias and insecurity promote accuracy: Mean-level bias and tracking accuracy in couples' conflict discussions. *Personality and Social Psychology Bulletin* 38: 642-655.

Overall NC, Fletcher GJO, Simpson JA, et al. (2015) Attachment insecurity, biased perceptions of romantic partners' negative emotions, and hostile relationship behavior. *Journal of Personality and Social Psychology* 108: 730-749.

Overall NC and Hammond MD. (2013) Biased and accurate: Depressive symptoms and daily perceptions within intimate relationships. *Personality and Social Psychology Bulletin* 39: 636-650.

Poll of Polls. (2014) *Poll of polls: The large parties*. Available at: http://pundit.co.nz/content/poll-of-polls.

Satherley N, Milojev P, Greaves LM, et al. (2015) Demographic and psychological predictors of panel attrition: Evidence from the New Zealand Attitudes and Values Study. *PLoS One* 10: e0121950.

Sibley CG. (2014a) Comparison of demographics in the NZAVS and New Zealand Census. *NZAVS Technical Documents*: e22.

Sibley CG. (2014b) Procedures for estimating post-stratification NZAVS sample weights. *NZAVS Technical Documents*: e08.

Sibley CG. (2014c) Sampling procedure and sample details for the New Zealand Attitudes and Values Study. *NZAVS Technical Documents*: e01.

Silver N. (2015, May 09) *Nate Silver on missed UK forecasts: We flubbed the margin of error*. Available at: http://www.npr.org/2015/05/09/405415749/nate-silver-on-missed-uk-forecasts-we-flubbed-the-margin-of-error.

Vowles J. (2000). The New Zealand Election Study. Political Science 52: 150-157.

Yuan K-H, and Bentler PM (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. Sociological Methodology 30: 165–200. DOI: 10.1111/0081-1750.00078/abstract

West TV and Kenny DA. (2011) The truth and bias model of judgment. *Psychological Review* 118: 357-378.

Wright MJ, Farrar DP and Russell DF. (2014) Polling accuracy in a multiparty election. *International Journal of Public Opinion Research* 26: 113-124.

Table 1. Sample size by month for the NZAVS and One News Colmar Brunton polling data.

Month	NZAVS (n)	ONCB (n)	Dates of One New Polls
Feb-14	1,213	1,024	15-19 Feb
Mar-14	2,352	1,029	22-26 March
Apr-14	1,192	-	
May-14	343	1,016	17-21 May
June-14	437	1,022	21-25 June
July-14	6,427	1,024	19-23 July
Aug-14	1,533	2,555	9-13 Aug; 23-27 Aug; 30 Aug-3 Sept
Sept-14	439	2,540	30 Aug-3 Sept; 6-10 Sept; 13-17 Sept
Total	13,936	10,210	

Note. The 30-Aug to 3 Sept poll was split across August and September to construct a monthly time-series.

Table 2.

Month-based time series of the proportion of party vote for the New Zealand Attitudes and Values Study (NZAVS) and One News Colmar Brunton (ONCB) polling data.

	The National Party		The Labour Party		The Green Party		The New Zealand First Party	
Month	NZAVS	ONCB	NZAVS	ONCB	NZAVS	ONCB	NZAVS	ONCB
Feb-14	47.28	50.80	29.25	34.44	15.31	8.36	2.49	3.20
Mar-14	45.09	46.73	27.87	31.38	16.39	11.48	4.62	6.54
Apr-14	45.61	-	25.77	-	17.78	-	5.02	-
May-14	51.03	50.67	22.63	29.73	18.93	10.80	4.12	4.80
June-14	50.00	50.38	24.83	28.85	15.56	11.79	2.98	3.72
July-14	51.40	51.58	22.28	28.12	17.06	10.34	3.07	4.41
Aug-14	49.68	48.97	20.27	27.03	21.18	11.36	2.90	6.12
Sept-14	43.35	46.95	20.23	24.79	20.52	12.70	6.07	7.36
July-Sept Polls	50.62	48.53	21.80	26.24	18.01	11.77	3.21	6.37
Combined Polls	48.95	49.00	23.93	28.12	17.43	11.31	3.55	5.66
Election Result	47.0		25.1		10.7		8.7	

Note. Reported election results represent the percentage of party votes received by each party in the 2014 New Zealand general election. New Zealand Attitudes and Values Study n = 13,936. One News Colmar Brunton n = 10,210.

Table 3.

Regression parameters indexing bias and tracking accuracy in the New Zealand Attitudes and Values Study (NZAVS) relative to the One News Colmar Brunton (ONCB) polling data over the time series from February to September 2014.

	Bias			Accuracy			
	c	se_c	t_c	b	se_b	β	t_b
The National Party	-1.179	.577	-2.04*	1.354	.260	.847	5.20**
The Labour Party	-5.281	.471	-11.20**	1.061	.114	.926	9.33**
The Green Party	6.872	.708	9.71**	.905	.311	.523	2.91**
The NZ First Party	-1.415	.264	-5.37**	.660	.169	.804	3.90**

Note. Bias (c) reflects the extent to which the NZAVS systematically over or underestimates support for a given party relative to the One News Colmar Brunton polls. Accuracy (*b*) reflects the extent to which the NZAVS reliably tracks change in political party support over time relative to the One News Colmar Brunton polls. * $p < .05$. ** $p < .01$.

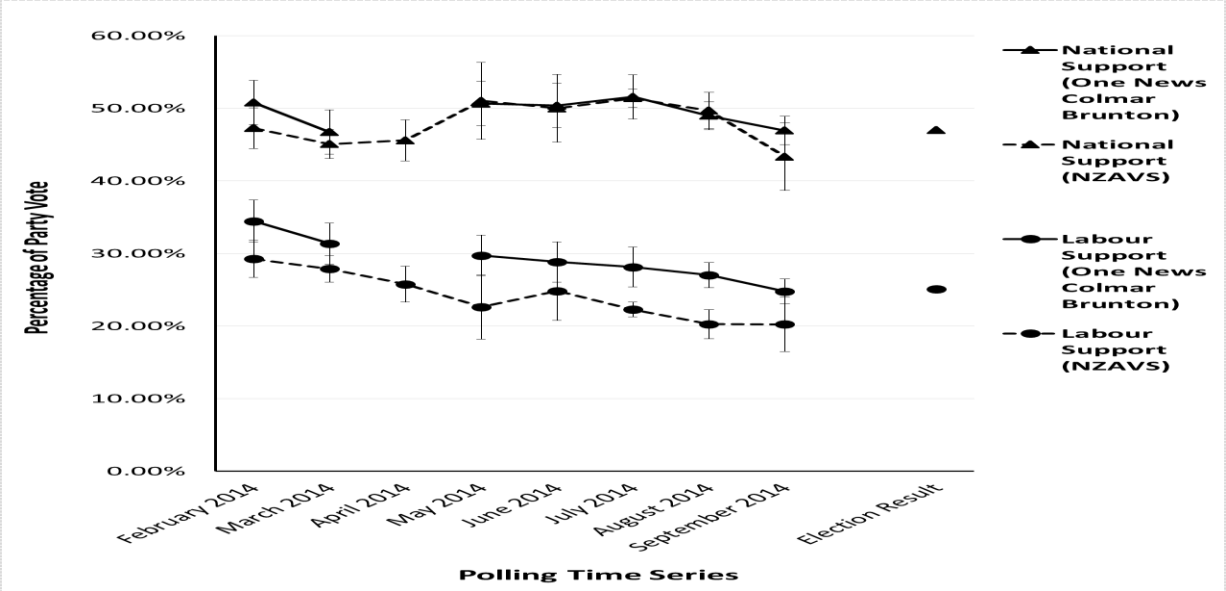

Figure 1.

Time series of NZAVS and One News Colmar Brunton polling data for The National Party and The Labour Party for the February-September 2014 period leading up to the national election held on September 20, 2014 (error bars represent the 95% margin of error for each monthly estimate).

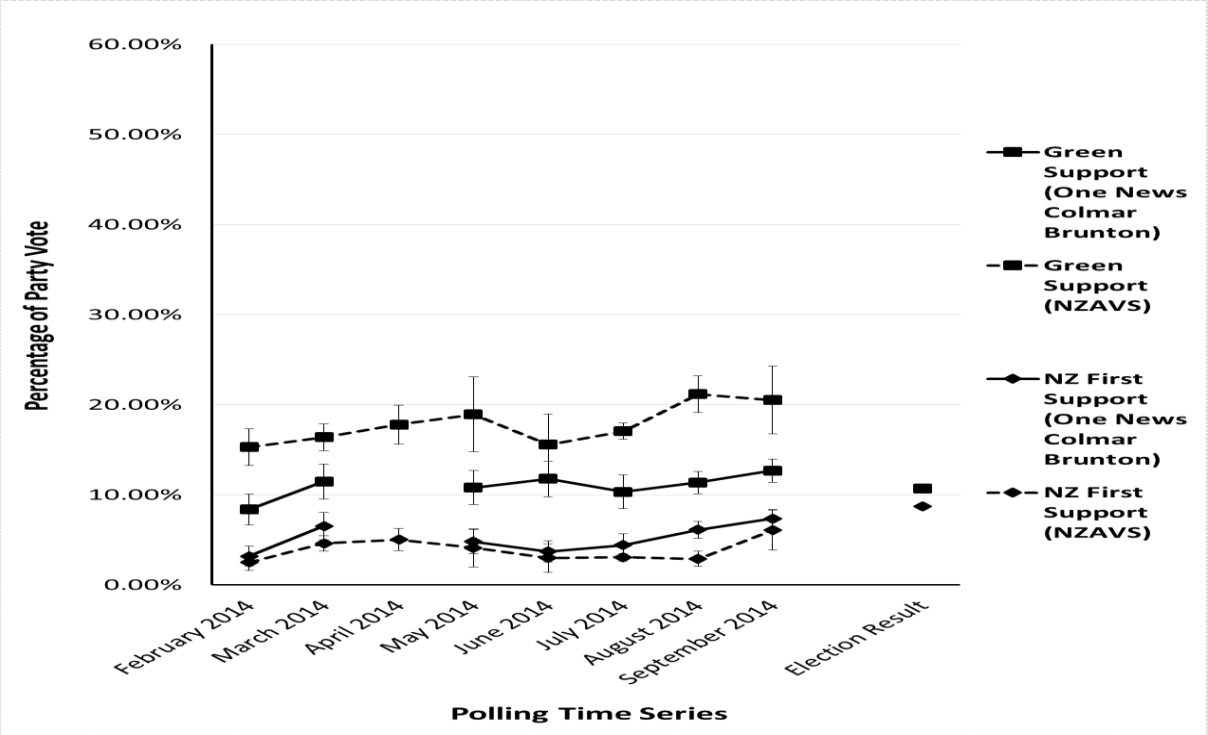

Figure 2.

Time series of NZAVS and One News Colmar Brunton polling data for The Green Party and The New Zealand First Party for the February-September 2014 period leading up to the national election held on September 20, 2014 (error bars represent the 95% margin of error for each monthly estimate).