DRAFT – DO NOT CITE OR DISTRIBUTE WITHOUT PERMISSION

Political polls in New Zealand

Assessing and understanding one poll's bias

Rob Salmond*

This paper examines political polls released since the 2002 election by New Zealand's main polling organisations. Colmar Brunton's poll is found to report significantly higher support for the right-leaning National Party than both the UMR Insight poll and the TNS Global poll. Examination of differences in the survey designs reveals that a deficiency in the way Colmar Brunton draws its sample may be an important reason for this difference in reported support levels.

Media coverage of politics in New Zealand is heavily influenced by political opinion polls. Whose political fortunes are rising? Which party is on the decline? If an election were to be held tomorrow, who would win? By how much? Television and newspapers alike put huge emphasis on this 'horse race' aspect of inter-election manoeuvring by New Zealand's political parties. Despite protestations to the contrary, politicians undoubtedly pay close attention to polls and may even alter their policy proposals or partisan strategy after seeing the results of one or more polls. Anecdotal evidence suggests that politicians privately create 'pecking orders' for polls – some Labour party activists, for example, believe that the polls broadcast on One News are always the worst for them, while some minor party politicians believe that they consistently get the best results from the National Business Review (NBR) poll.

In order to make accurate inferences about the comparative fortunes of the parties, media companies need to be very confident in the reliability of their opinion polls. Three companies which provide regular polls to New Zealand news organisations are Colmar Brunton, which conducts surveys for the One News poll; the New Zealand branch of the TNS Global corporation, which supplies TV3 with its political poll data; and UMR Insight, which does the research for NBR's opinion poll.[1]

Do these three polls tend to provide us with the same picture about politics in New Zealand? If not, how reliable is each of the polls? The analysis in this paper will show that the Colmar Brunton poll provides consistently and significantly higher estimates of the right-leaning National party's support than either of the other two polls, and it will suggest that this difference is due to deficiencies in Colmar Brunton's survey methods. This analysis will lead to the conclusion that both the TNS Global and the UMR Insight poll provide a more accurate and reliable picture of New Zealand's political landscape than the Colmar Brunton poll.

[1] There are other market research companies and other news organizations that conduct surveys of New Zealanders' political opinions, but these three are the most consistently taken and widely reported polls.

Do the polls paint different pictures?

Since the 2002 election Colmar Brunton, TNS, and UMR have collectively run 68 polls on New Zealand's political opinions, reporting the opinions of 61,000 New Zealanders between them. Colmar Brunton has run 26 polls, each of which surveys 1,000 people; TNS has run 14 1,000-person polls; and UMR has run 28 polls of 750 people each.

[TABLE 1 ABOUT HERE]

Table 1 summarizes the support in each of these polls for the left-leaning Labour and Green parties, the right-leaning National and ACT parties, the centrist / populist New Zealand First party, and the economically centrist but morally conservative United Future party.[2] The clearest difference between the polls is that Colmar Brunton consistently estimates higher levels of National party support than both the other polls – almost five percentage points higher, in fact. Also, Colmar Brunton appears to consistently estimate lower levels of support for New Zealand First than UMR. These are the only three cases where all reported differences share the same sign – that is, one company A has ranked a party higher than company B in *all* polls since the 2002 election. While there are differences between the polls in other areas (some of which are in line with informal parliamentary rumours about each poll's tendencies), none of them appear at first glance consistent or large enough to be concerned about.

[FIGURE 1 ABOUT HERE]

Figure 1 charts the level of support for the National party as estimated by the three polling organizations. It shows that Colmar Brunton has estimated the highest or equal highest level of National support for each of the 26 polling months since the election, with the differences reaching as high as 9.5%.[3] It also shows no indication that the UMR and TNS polls consistently differ from each other in terms of National's reported support.[4] Another way to consider the observed differences in estimated National party support levels is to first *suppose* that the three polls really do all draw randomly from the same population, and then assess the probability of one poll coming out ahead of the other two on at least 25 of 26 observed occasions (for the purposes of this exercise I will count the solitary 'tie' between Colmar Brunton and TNS as an occurrence of Colmar Brunton *not* coming first). The probability that a given poll will estimate the highest level of National party support in any one month (given the assumption above) is one third. Therefore the probability of coming first on at least 25 of 26 occasions is given by the formula:

$$\Pr = \left(\frac{1}{3}\right)^{26} + 26 \times \left(\frac{1}{3}\right)^{25} \times \left(\frac{2}{3}\right)^{1} \qquad \text{(Eq 1)}$$

---

[2] The polls generally are not taken during the same week, and therefore direct comparison of one month's polls from multiple polling companies could be misleading. To alleviate this concern, I interpolate poll results for all companies, and then compare one poll's actual result with the interpolated results of other polls. For example, if a poll indicated that New Zealand First enjoyed 10% one month, and then four weeks later is estimated to have 8% support, then the support is interpolated to be 9.5% in the first intervening week, 9% in the second week, and 8.5% in the third week. The results are not dependent on this technique, however, and comparing all actual polls taken in the same month (regardless of the specific week when the poll was taken) provides results broadly the same as those reported in Table 1.

[3] Colmar Brunton takes eleven polls each year. It does not take a poll in January, the height of the New Zealand summer.

[4] Those unfamiliar with New Zealand politics will be astonished by the rapidity of National's rise in the polls at the start of 2004. So were New Zealanders. The rise followed a major speech – now known as the 'Orewa speech' – by the party's new leader, Don Brash, which highlighted the party's stand on racial issues. This proved to be a 'hot button' issue for many voters, who shifted their allegiance to National immediately on hearing about the speech.

The probability from equation 1 works out to be approximately 21 chances per trillion, or 0.000000000021. Put either way this is a prohibitively small number, leading us to reject the initial assumption that the three polls actually draw randomly from the same population. These results mirror those of an internal UMR study conducted in 1999, which noted that: 'Colmar Brunton's National vote is 5% higher on average than UMR's. The range is between 9% higher and no difference. None of the 24 Colmar Brunton polls showed a lower National vote than the comparable UMR poll.' (UMR 1999, p 5)

A more robust way to estimate the differences in vote shares between the polls is with a regression model. The model estimates poll ratings as a function of the party being rated, the company doing the rating, the time that the poll was taken, and two interactions – the first between party and time, and the second between party and company. All of the independent variables are operationalized using a series of dummy variables in order to avoid making any assumptions about the functional form of relationships between independent and dependent variables. The independent variables of interest are the interactions between the polling company and the party being rated – if one company rates a party significantly higher than another then this pattern will emerge in these interaction terms. There are too many variables in this regression to report in a table – Table 2 reports estimates only the appropriate linear combinations of the independent variables of interest.[5]

[TABLE 2 ABOUT HERE]

The regression results in Table 2 provide strong support for the proposition that the Colmar Brunton poll is more favourable to right-leaning parties and less favourable to left-leaning parties than either of the other two polls. The top half of columns (1) and (2), comparing Colmar Brunton with UMR, shows that UMR is *less* favourable than Colmar Brunton to the right-leaning National and *more* favourable to populist New Zealand First and the left-leaning Greens. The bottom half of the table, comparing Colmar Brunton and TNS, shows almost exactly the same pattern of results – with the additional finding that TNS is consistently *more* favourable to left-leaning Labour than Colmar Brunton.

The combined evidence from Figure 1and the two tables provides a strong indication that the Colmar Brunton poll is significantly more favourable to right-leaning parties than either the UMR poll or the TNS poll.[6]

Why are Colmar Brunton's results different?

Most aspects of the polling methods used by the three large market research firms are the same. All operate their surveys by telephone and use some form of random digit dialling to select phone numbers. All weight their sample by age, gender and geography so as to get a representative sample of the population – Colmar Brunton also weights the results for household size, while UMR also weights by vote choice in the previous election. And all

---

[5] In order to avoid over-specification of the model – since the sum of all the party shares in this data set were extremely close to 100% data for the United Future party (the smallest share in general) were not used. Type III Wald tests on this model indicate that the four party dummy variables jointly had the vast majority of the power in the regression, and that the combined impact of the company-by-party interactions and the party by time interactions were also significant. The main effects of the polling company and time variables were not jointly significant. This is the expected pattern and provides confidence in the results.

[6] While results for National have been strong, we might also expect (using the same argument) Colmar Brunton to estimate higher levels of support for the ACT party than the other polls. Results have not revealed this pattern. Due to ACT's generally low levels of support (the lowest average support of any party included in the Table 2 analysis), however, it is much harder to find inter-poll patterns in support for ACT than it is for National. This is because only a very few interviews are responsible for ACT's monthly poll rating, meaning that the central limit theorem is not able to take hold and therefore a completely random poll is less likely to estimate the actual underlying level of support for ACT than for a large party like National. This increased likelihood of random 'bouncing around' in the ACT rating makes finding systematic inter-poll differences very difficult.

ask almost the same question in their polls, normally a variant the question 'if an election were to be held tomorrow, to which political party would you give your party vote?'[7] There are two important differences between the polls; one that matters for the question at hand and one that does not.

One difference is that the UMR poll surveys only 750 people at a time while both the other polls survey 1,000. This difference in sample size is important for the margin of error on the UMR poll, which is bigger than the margins of error for the other polls, but it does not introduce any systematic bias into the poll and is therefore unimportant to the question 'why does the National party do better in the Colmar Brunton poll?'

The other difference between the polls is that Colmar Brunton conduct their surveys in the evenings, from Monday to Thursday, while both TNS and UMR start their polling late one week and survey people all through the weekend and into the start of the following week.[8] This is important because there is a bias in the set of people who are most likely to be polled on weekday evenings (see Traugott 1987; Lau 1994). Weekday polling biases the sample in favour of wealthy citizens because those people whose jobs require them to work, without access to a telephone, outside of 'normal' weekday working hours tend to have medium or low incomes. Examples of people who exhibit this kind of work pattern are restaurant and bar workers, factory workers doing shift work, and commercial cleaners.

Colmar Brunton does try to alleviate the problem of people not being home with an aggressive callback policy. If a randomly generated phone number produces no answer on the first attempt, Colmar Brunton call the same number back up to six times on the same and subsequent evenings in order to get a response. While this is much better than simply picking another randomly drawn phone number, it still does not alleviate the bias mentioned above. Indeed, this call back policy is likely to result in the company successfully contacting the wealthy businesswoman who was out at a corporate dinner on Tuesday night, but not the waiter who served her or the line cook who prepared her meal – they work every night. And while the policy will contact the office worker who worked late one night, it won't catch up with the person who cleans the offices every night. The bias remains.

By polling over the weekend as well as during the week, other companies are able to alleviate this potential much more successfully. This is because they are able to call back the numbers which had no answer during he week at times when restaurant workers, bar workers, factory shift workers and commercial cleaners are less likely to be at work – during the daylight hours of weekend days. To the (almost certainly incomplete) extent that those people are at home during the weekend, the initial bias from the weekday evening polling is undone.[9]

Political science established long ago that a person's income is a good predictor of their vote choice – poorer people tend to vote left and richer people tend to vote right (the foundational reference is Campbell et al 1964). Therefore if a survey has the pro-wealthy bias of weekday polling without also incorporating the countervailing effects of weekend polling, it is likely to have a sample of the electorate skewed to the ideological right. This is the

---

[7] The 'party vote' is one of the two votes that New Zealand voters have cast in each election following the 1996 introduction of a mixed member electoral system. The other vote is the local 'electorate vote'.

[8] Digipoll, which runs the quarterly New Zealand Herald poll, and BRC, the Sunday Star Times' pollsters since early 2004, also poll over the weekend and typically also report lower levels of National party support than Colmar Brunton.

[9] There may also be a small bias at the weekend away from polling wealthy citizens, because people with greater financial means are more likely than their poorer compatriots to be out of town over a normal weekend. Wealthy people can afford short-break airfares, holiday homes or resort hotel accommodation much more frequently than poor people. Such a bias is likely to have limited impact on the overall sample, however, given that the overall contact rate at weekends is considerably lower than on weekdays. The weekend 'fix' for the weekday bias is also likely to work imperfectly, and therefore the practice of polling at weekends is most unlikely to lead to an overall bias in the sample towards left-leaning voters.

situation that appears to have beset Colmar Brunton, whose weekday-only polling method results in their overstating National's level of support.

This analysis opens the question as to why a company would poll only on weekday evenings, given that it can produce these biased results. The answer may be that weekday evenings are the times when the most people in a large community are near a phone and have the time to complete a survey (Traugott 1987, p 53; Statistical Assessment Service 2004). Therefore surveys conducted during this period have the smaller number of unanswered calls and the highest overall contact rate, meaning that polling companies get to their contracted sample size quicker and (presumably) at a lower monetary cost to the client. What the client saves in money, however, they may lose in reliability. There is, of course, no publicly available information on the contract between Television New Zealand and Colmar Brunton, so it is not possible to know whether there is a financial incentive towards weekday polling in this particular case.

It is possible that there are other differences between the polls that might lead to significantly different reported support levels – such as question wording effects or question order effects. These details of the polling company's questionnaires are commercially sensitive, remain unpublished, and therefore those possible effects cannot be examined in this paper. Anecdotal evidence suggests, however, that the question wording has become fairly standard in New Zealand, as noted earlier, and also that a norm has developed towards asking the vote choice question near the start of the survey.

Discussion.

The analysis is this short paper has shown that the results of Colmar Brunton's political polling differ significantly from that of its two main competitors, with consistently higher levels of support reported for the National party. With only one exception, Colmar Brunton has reported higher levels of National support than comparable polls in each of its 26 post election polls. In addition, Colmar Brunton's polling estimated a higher and more sustained 'bounce' for National following Don Brash's nationhood speech at Orewa than either of the other polling companies, and estimated that National lead Labour for six continuous months following the Orewa speech. UMR and TNS both estimated that this lead lasted one month only.[10]

On their website, Colmar Brunton claim that their results are highly accurate, and point to their polls' excellent performance in predicting the results of both the 2002 and 1999 elections (Colmar Brunton 2004). It is certainly true that their poll performed very well, predicting the vote shares of many of the major parties exactly right and forecasting the vote shares of most parties to within the poll's margin of error. The final poll released by TNS before the 2002 election, however, also predicted the vote shares of all parties to within the margin of error, and since the election it has differed markedly from Colmar Brunton.[11] And UMR's final poll before the 1999 election also performed every bit as well as Colmar Brunton's comparable poll (although in fairness it should be noted that TNS's previous incarnation – CM Research – ran a poll that differed significantly from the final 1999 election results). So if the polls are to be judged by the question 'who is right on the night?' all three organisations have some credible claims to reliability. The question that immediately arises from that observation, however, is: 'if all the companies were right on the night, what can explain the consistent differences in their polling results *since* the election? Which of the companies are *more reliably* accurate than the others?'

---

[10] It should be noted, however, that UMR's July 2004 poll also showed a small lead for National.

[11] Unlike Colmar Brunton and TNS, UMR Insight's last pre-election poll was taken before the 'corngate' scandal - the event which some argue defined the 2002 campaign (see Williams 2003) - made the headlines and therefore it isn't surprising that its poll is a poor predictor of the final outcome.

This paper has provided an argument that supports the intuitive conclusion that the two polls saying A are more likely to be right than the one poll saying B. The dates on which the various samples are drawn provide the key to understanding the difference – Colmar Brunton draws a consistently biased weekday sample, while the others draw a more balanced sample by mixing weekday and weekend polling. Thus the Colmar Brunton results are, I argue, not as accurate or reliable as the other results on offer. The results are consistent with findings in the American context that 'trial heat' polls leading up to the 1992 presidential election were consistently more favourable to the Republican candidate (George H W Bush) if the poll was conducted on weekdays only (Lau 1994, p 17).

There are two ways that Colmar Brunton and One News can fix this problem. The first is to poll over the weekend like the other polls. Even if this is more expensive for Television New Zealand, it will increase the reliability of their political polling data. The second method is for Colmar Brunton to weight their sample by income or previous vote choice in addition to the other statistical correction used currently. The first method is preferable to the second because it improves the quality of raw data going in to the poll – which is always better than post-hoc statistical corrections, but either method will result in improved estimates of the relative standings of political parties in New Zealand. The public, and their agents the news media, should demand nothing less.

DRAFT

References.

Campbell, Angus, Philip Converse, Warren Miller and Donald Stokes (1964) _The American Voter: An Abridgement_ (New York: Wiley)

Colmar Brunton (2004) 'How did we get it so right?' available online at: www.colmarbrunton.com

Lau, Richard (1994) 'An analysis of the accuracy of "trial heat" polls during the 1992 presidential election' in _Public Opinion Quarterly_ (58:1), pp 2-20.

Statistical Assessment Service (2000) 'Understanding the mechanics of polling' available online through: http://www.stats.org/record.jsp?type=news&ID=378

TNS Global (2004) 'TV3 / TNS political poll' available online at: www.tns-global.co.nz

Traugott, Michael (1985) 'The importance of persistence in respondent selection for preelection surveys' in _Public Opinion Quarterly_ (51:1), pp 48-57.

UMR Insight (1999) 'New Zealand political polls' (Internal company report)

Williams, Mike (2003) 'Oddidy or new paradigm? A Labour view of the 2002 election' in Boston, Jonathan, Stephen Church, Stephen Levine, Elizabeth McLeay and Nigel Roberts (eds) _New Zealand Votes_ (Wellington: Victoria University Press)

Table 1: Inter-poll differences in partisan support levels since July 2002

Party	(1) TNS v Colmar Brunton			(2) UMR v Colmar Brunton			(3) TNS v UMR		
	Max.	Mean	Min.	Max.	Mean	Min.	Max.	Mean	Min.
Labour	6.75	2.38	-2.00	4.50	-0.44	-6.00	6.75	2.82	-1.00
National	0.00	-4.92	-10.50	-1.38	-4.83	-9.50	4.38	-0.10	-5.75
NZ First	3.00	0.85	-2.20	5.00	2.27	0.00	0.33	-1.48	-4.10
Green	3.30	1.25	-1.20	4.00	1.85	-0.25	2.00	-0.50	-4.63
ACT	2.75	-0.26	-1.80	2.00	0.15	-1.35	1.43	-0.37	-2.35
United Future	2.00	0.25	-3.60	3.90	1.10	-2.03	0.3	-0.91	-2.13

Notes:

Positive numbers indicate that the first named poll has a higher average estimate of a party's support than the second named poll. Negative numbers indicate the opposite. Thus the observation of 2.38 near the top left indicates that TNS polls indicate, on average, 2.38% *higher* support for Labour than Colmar Brunton.

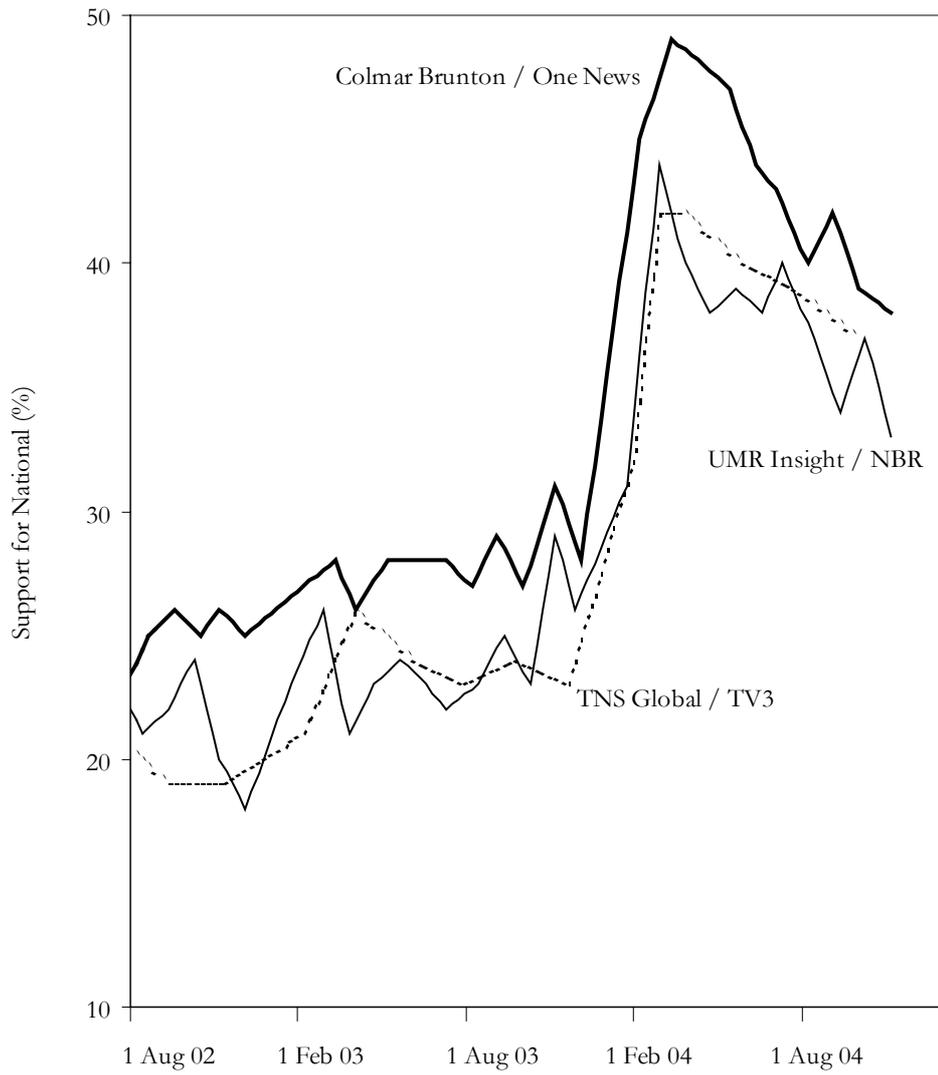Figure 1: Support for National since the 2002 election

Table 2: Testing for differences in poll ratings since August 2002.

Party	Company *(comparison with Colmar Brunton)*	(Model 1) Dependent variable = poll ratings Reference party = Labour	(Model 2) Reference party = ACT
Labour	UMR	-0.58† (0.260)††	-
National	UMR	-4.85*** (0.000)	-4.85*** (0.000)
NZ First	UMR	2.42*** (0.000)	2.42*** (0.000)
Green	UMR	1.92** (0.000)	1.92*** (0.000)
ACT	UMR	-	0.10 (0.627)
Labour	TNS	2.49*** (0.000)	-
National	TNS	-5.50*** (0.000)	-5.50*** (0.000)
NZ First	TNS	1.02** (0.003)	1.02*** (0.003)
Green	TNS	1.29** (0.001)	1.29*** (0.001)
ACT	TNS	-	-0.16 (0.634)
N		420	

Notes:

† Point estimates are the linear combination of the main effect of the polling company in the regression and the interaction term between the named company and the named party. The interpretation of these estimates is: "On average, Company X estimates the support level for party Y to be [point estimate] higher (if point estimate is positive) than Colmar Brunton. This difference is significant at [combined p-value -see below]."

†† Combined p-values (obtained using a Wald test with the null hypothesis 'that the linear combination of the two variables that combine for the point estimate equals zero') appear in parentheses.

* indicates p<0.05; ** indicates p<0.01; *** indicates p<0.001 (all tests two tailed)

Colmar Brunton is the reference category for company, which is why it does not appear in the main part of the Table. United Future is excluded from the analysis to remove the potential problem of linear constraints on the dependent variable.