UNIVERSITY OF CALGARY
Press

ORANGE CHINOOK: Politics in the New Alberta

Edited by Duane Bratt, Keith Brownsey, Richard Sutherland, and David Taras

ISBN 978-1-77385-026-9

UNIVERSITY OF CALGARY
Press

press.ucalgary.ca

Marginally Better: Polling in the 2015 Alberta Election

Janet Brown and John B. Santos[1]

Public opinion polling is a fixture in the politics of Western democracies, particularly during the course of an election campaign. Since Gallup predicted Franklin Roosevelt would be re-elected in the 1936 American presidential election, polling has grown to become its own industry that, in addition to pollsters, now also includes polling aggregators and election forecasters. Canada is no exception to this trend, and the number of polls conducted during Canadian elections has steadily increased since the 1988 federal election.[2] This trend has since trickled down to the provincial level— in Alberta, 4 polls were published during the 2004 election campaign, 8 in 2008, 23 in 2012, and 17 in 2015.

Polls are important in that they inform the actions of parties, campaign, interest groups, the media, and voters. Moreover, polling itself is increasingly becoming the subject of media coverage over and above substantive election issues, leading to the rise of what some have called "horserace journalism."[3] Despite the importance and proliferation of polling, the polling industry in Alberta faced a credibility problem going into the 2015 Alberta election campaign. The polls were widely off the mark in the province's 2012 election, leading to such post-election headlines as " 'We were wrong': Alberta Election pollsters red faced as Tories crush Wildrose."[4] Alberta is not alone in this respect, and other notable examples of polling failures include the 2013 British Columbia, the 2014 Quebec, and the 2014 Ontario provincial elections. The 2015 Alberta election was a chance for pollsters to redeem themselves, and, at least at first blush, they did. The tone of the

headlines was different this time. "Pollsters relieved at getting it right in Alberta's unlikely swing to the left," read one such headline in *Maclean's*.[5] But is that a correct assessment?

To answer that, we must first ask a different question: What are the criteria for "getting it right?" The easy answer is accuracy, but that then raises the question of what constitutes accuracy. In the 2015 Alberta election, all but one poll published after the 23 April leaders' debate was "accurate" in the sense that they showed the New Democrats ahead of all other parties, and the New Democrats eventually won the election. However, polling is about more than just predicting who will cross the finish line in first place. Polls make claims about the support of all major political parties in the race. They also include a "margin of error," which provides an upper and lower range within which actual public opinion should be—nineteen times out of twenty, of course. As such, accuracy entails more than just identifying the winner correctly. An accurate poll should also identify the correct ordering of the parties in terms of their proportion of the popular vote. As well, the difference between each party's measured level of support and their actual level of support should not exceed the size of the poll's stated margin of error.[6]

However, accuracy is difficult to assess, given that the only time we can actually verify how the public intends to vote is when they vote on election day. A pre-election poll may be different from the actual election result because it is a poorly executed poll, or because it was accurate at the time but last-minute events caused shifts in public opinion. With this in mind, pollsters, politicians, and pundits alike use qualifiers when commenting on polls, saying they are only "snapshots in time" or that "the only poll that matters is election day."[7] Yet pollsters eagerly take credit when their polls are in line with the actual election results, and—as evidenced by the previously mentioned headlines—the news media can be eager to accept pollsters' claims. In fact, at least for some polling firms, election polling is a service done free of charge as a demonstration of their capabilities and accuracy to prospective clients. As such, there is an implicit predictive value in polls.

Johnston and Pickup describe polls as "trial heats," or preliminary tests between the parties contesting elections that anticipate the eventual result.[8] While there is evidence that polls conducted closer to election day tend to more closely mirror the actual election result,[9] the pattern exhibited by the polling in the 2015 Alberta election suggests most shifts in public opinion

occurred after the leaders' debate. This means that even though polls become more accurate the closer they are to election day, all of the polls conducted in Alberta after the leaders' debate should have been reasonably accurate. As this chapter will show, this was not necessarily the case. Polls did perform better in the 2015 election campaign than they did in the 2012 campaign, but they were only marginally better than other recent Canadian provincial elections that are widely regarded as polling failures. This is because most polls did not predict the correct order of the parties in 2015, and because there were systematic errors (i.e., bias) in that the polls overestimated support for political change.

Data and Methods

To facilitate this analysis, we compiled a list of all publicly available polling released during the campaign period.[10] This excludes any proprietary polling conducted for political parties, candidates, or third-party groups, the results of which would not be made available to the news media or general public. This dataset contains seventeen polls in total conducted by ten companies, using a variety of sampling sizes, sampling methods, and interview modes. These are summarized in Table 4.1. The most prolific polling firm was Mainstreet Technologies, which released five polls. Forum Research was similarly prolific, releasing four polls. The only other firm to release more than one poll was EKOS, which released two. Pantheon Research, Leger Marketing, ThinkHQ Public Affairs, Return on Insight, Ipsos-Reid, and Insights West all released one poll apiece.

The most prevalent interview mode was interactive voice response (also known as IVR, or robo-polling), whereby telephone numbers are called at random and those answering are invited by a pre-recorded voice to answer questions by pressing numbers on their phone keypad or saying their answers aloud. Mainstreet, Forum, Pantheon, and EKOS used IVR. Leger and Insights West fielded their polls through online panels, which involve sending surveys via the Internet to people who have agreed to become a member their survey panel. Only Return on Insight used the traditional method of live telephone interviews with a random sample of the population. ThinkHQ and Ipsos-Reid used a mix of live telephone interviews and interviews conducted through their online panels.

Table 4.1. Polling Summary by Firm

FIRM	# OF POLLS	SAMPLE SIZES (N)	MOE (±PP)	TYPE OF MOE	RANDOM SAMPLE?	INTERVIEW MODE
Mainstreet Tech.	5	2,013–4,295	1.5–1.9	Claimed	Yes	IVR
Forum Research	4	801–1,661	2.0–3.0	Claimed	Yes	IVR
Pantheon Research	1	4,131	1.5	Claimed	Yes	IVR
Leger Marketing	1	1,180	2.8	Equiv.	No	Online
ThinkHQ Public Affairs	1	2,114	2.1	Equiv.	No	Online/ phone
Return on Insight	1	750	3.6	Claimed	Yes	Phone
EKOS	2	721–823	3.4–3.7	Claimed	Yes	IVR
Ipsos-Reid	1	761	4.1	Equiv.	No	Online/ phone
Insights West	1	1,003	3.1	Equiv.	No	Online

Sources: Data from polling firm news releases, www.threehundredeight.com, and Election Almanac (www.electionalmanac.com).

Sample sizes varied from around or just under 1,000 respondents for most firms to 2,000 or greater in the case of Mainstreet, Pantheon, and ThinkHQ. Correspondingly, claimed margins of error ranged from plus or minus 4.1 percentage points, 19 times out of 20 for Ipsos-Reid's sample of 761 respondents, to plus or minus 1.5 points for Mainstreet and Pantheon's polls, with samples of more than 4,000 respondents. According to the guidelines of the Marketing Research and Intelligence Agency, the industry association of market research professionals, it is only appropriate to calculate

a margin of error for random probability samples such as telephone surveys.[11] Online panel surveys are considered to be convenience rather than random samples, and as such, it is not appropriate to report a margin of error. That said, polling firms that use online panels do strive to ensure their panel sample is demographically representative of the general population, so an "equivalent margin of error" usually accompanies the results of an online panel survey; this indicates what the margin of error *would* be for a true random probability sample of the same size. Surveys that use a hybrid method involving an online panel sample and live telephone interviewing face the same limitation.

Using the dataset described above, this chapter will evaluate the accuracy of each poll based on the following criteria:[12]

1. The poll correctly anticipates the winner of the election.

2. The poll correctly anticipates the order of the parties in terms of the proportion of the popular vote won by each party.

3. The predicted vote for each party falls within the poll's stated margin of error.

4. The poll's *total absolute polling error*[13] is comparable to accurate polls in other elections.

The first three criteria compare a poll to the final election result. The fourth criterion relies on comparisons with polls conducted during other elections in Canada.

How the Horse Race Unfolded

Before analysing each poll, a simple visual examination helps set the stage for the analysis and provides some preliminary confirmation for the argument. Figure 4.1 shows all seventeen polls that comprise the dataset, plotted by the last date in field. The large symbols on 5 May indicate the actual election result. The trendlines are fitted using the LOWESS smoothing procedure[14] and illustrate the trajectory of each party's support over the course

Figure 4.1. 2015 Alberta Election Polls
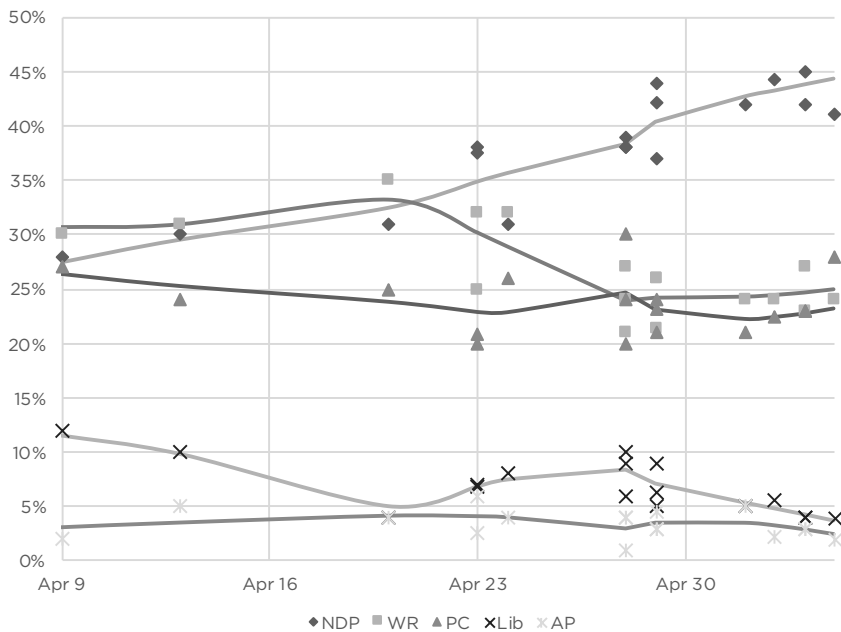
of the campaign. Two things are readily apparent in Figure 4.1. First, most polls were done in the final week of the campaign. Second, the debate serves as a turning point in the campaign as the fluctuations of party support within the pre- or post-debate periods are less than the shift in support patterns from one period to the other.

PC support does not change very much over the course of the campaign—the party went into the campaign period in an unprecedented and severely weakened state (see Bratt, this volume). The most interesting aspect of the race was the surge in support for the NDP and a decrease in support for the Wildrose, and to a lesser extent, the Liberals. The PCs did not lose

the election to the NDP over the course of the campaign (see Thomas, this volume); if anything, they lost it even before the campaign began. The NDP surge was a function of anti-PC voters consolidating around Rachel Notley and the NDP after the leaders' debate.

The final thing to note in Figure 4.1 is the vertical distance between each poll's measurement of a party's level of support and the actual level of support that party receives. While the polls closer to election day are closer to the final result, PC support is consistently underestimated by all but one poll—the Leger poll that finished on 28 April. The numbers for the NDP and the Wildrose tend to be higher than the actual level of support they received, though not as marked as the PCs. The polls were accurate for the two minor parties, the Liberals and the Alberta Party.

How Accurate Were the Polls?

Table 4.2 summarizes, for each poll, the error between the poll's measured level of support for a party and the actual proportion of the vote received by that party in percentage points, the total absolute error, and which of the first three criteria the poll meets. Polls marked with an asterisk (*) denote a firm's final (or only) poll. For the column, "correct order," rows marked as "close" mean the poll incorrectly anticipated the Wildrose to be ahead of the PCs in terms of the popular vote, but that the difference between the two parties is within the poll's stated margin of error. Table 4.2 confirms the conventional wisdom that polls closer to election day tend to be more accurate.[15] However, there is still variation in the total absolute error of polls within periods that must be accounted for, especially since most of the movement in party support was between the pre- and post-debate periods, not within periods.

Almost all of the eleven polls conducted exclusively within the post-debate period correctly anticipated that the NDP would win the popular vote. The only poll that did not was the Mainstreet poll ending on 24 April, which had the Wildrose at 32 per cent and the NDP at 31 per cent—a difference within their stated margin of error. By the first criteria, the polls in 2015 were accurate.

Meeting the second criterion is more difficult. Of the same eleven post-debate polls, only three correctly anticipated that the PCs would

Table 4.2. Polling Error in the 2015 Alberta Election

Final field date	Poll	Moe (±pp)	Party support errors (±pp)					Criteria 1: correct winner	Criteria 2: correct order	Criteria 3: # of parties within moe	Criteria 4: total abs. error
			PC	Wr	NDP	Lib	AP				
7/04/	Mainstreet	1.8	-1.0	7.0	-15.0	8.0	1.0	No	No	2	32.0
9/04	Forum	2.0	-1.0	6.0	-13.0	8.0	0.0	No	No	2	28.0
13/04	Mainstreet	1.8	-4.0	7.0	-11.0	6.0	3.0	No	No	-	31.0
20/04	Mainstreet	1.8	-3.0	11.0	-10.0	0.0	2.0	No	No	1	26.0
23/04	Forum	3.0	-8.0	1.0	-3.0	3.0	4.0	Yes	No	1	19.0
23/04	Pantheon *	1.5	-7.1	8.0	-3.5	2.9	0.6	Yes	No	1	22.1
24/04	Mainstreet	1.5	-2.0	8.0	-10.0	4.0	2.0	No	No	1	26.0
28/04	Leger *	2.8	2.0	0.0	-3.0	2.0	-1.0	Yes	Yes	4	8.0
28/04	Thinkhq *	2.1	-8.0	3.0	-2.0	5.0	2.0	Yes	No	2	20.0
30/04	Roi *	3.6	-4.0	-3.0	-3.0	6.0	2.0	Yes	Yes	3	18.0
29/04	Ekos	3.7	-4.9	-2.7	1.2	2.3	2.6	Yes	Yes	4	13.7
29/04	Mainstreet *	1.9	-7.0	2.0	3.0	1.0	1.0	Yes	No	2	14.0
30/04	Ipsos *	4.1	-4.0	2.0	-4.0	5.0	1.0	Yes	Close	4	16.0
2/05	Forum	3.0	-7.0	0.0	1.0	1.0	3.0	Yes	Close	4	12.0
3/05	Ekos *	3.4	-5.5	0.0	3.3	1.6	0.2	Yes	Close	4	10.6
4/05	Forum *	3.0	-5.0	-1.0	4.0	0.0	1.0	Yes	Close	3	11.0
4/05	Insights west *	3.1	-5.0	3.0	1.0	0.0	1.0	Yes	No	4	10.0
Election result (for reference)			28%	24%	41%	4%	2%				

Sources: Data from Elections Alberta, "Provincial Results—Provincial General Election May 5, 2015," polling firm news releases, threehundredeight.com, and Election Almanac (www.electionalmanac.com).

receive a greater proportion of the popular vote than the Wildrose (Leger, Return on Insight, and the first EKOS poll). Four polls were close, or had the gap between the PCs and Wildrose within their claimed or equivalent margin of error (Ipsos-Reid, the last two Forum polls, and the second EKOS poll). Four polls (both post-debate Mainstreet polls, ThinkHQ, and Insights West) showed the Wildrose ahead of the PCs with a gap greater than their stated margin of error. By the stricter standards set by the second criterion, the polls are less consistent in their accuracy. Note that the second criterion is only concerned with order, and not the size of the gaps, and yet the polls are already coming up short. Interestingly, the first EKOS poll actually outperforms the second EKOS poll in terms of this criterion.

When it comes to correctly anticipating the level of support for each party within the poll's stated margin of error, no poll gets it right for all five parties that won seats. Six polls had four out of five parties within their stated margin of error (Leger, both EKOS polls, Ipsos, Insights West, and the penultimate Forum poll), whereas Return on Insight and the final Forum poll had three out of five parties within their margin of error. ThinkHQ and the final Mainstreet poll got two out of five correct, and the Mainstreet poll immediately following the debate only got one party within the stated margin of error. Perhaps more concerning is that the errors have a consistent direction—namely, PC support is consistently underestimated. All but one post-debate poll (ten in total) showed the PCs at a level of support lower than what they actually received on election day, and of these ten, only one (Ipsos) was within its stated margin of error. Leger was the only firm that showed the PCs at a level of support higher than what they actually received, and Leger's poll was within the stated margin of error.

On the basis of total absolute error, the polls exhibit a wide range of total absolute errors within each time period. Among the post-debate polls, the total absolute error ranges from eight points to twenty-six points. Only when the time horizon is narrowed to polls conducted exclusively within the first four days of May does the total absolute error decrease to the low double digits. Yet, even those final four polls have total absolute errors greater than the lone Leger poll, which was finished fielding almost a week before the election and was the only poll to have a total absolute error in the single digits.

Thus, when the polls conducted during the 2015 Alberta election campaign are compared against one another on the basis of the first three criteria—correctly anticipating the winner, correctly anticipating the order, and correctly measuring each party's support within their stated margin of error—the polls become less consistent in fulfilling the criteria as the criteria become more stringent.

Not only are there clear issues with these polls when comparing them to one another, but these issues become even more clear when they are compared to polls in other elections. Using the metric of average total absolute error for the final batch of polls conducted and released in a given election, Coletto found the final polls in the 2015 Canadian federal election were very accurate and had an average total absolute error of 6.7 points, which is 10.3 points lower than the error in the 2013 British Columbia provincial election (17 points) and 16.3 points lower than the error in the 2012 Alberta provincial election (23 points).[16] Table 4.3 presents average total absolute errors for various time periods during the 2015 Alberta provincial election campaign alongside Coletto's data for comparison. The rows are the average total absolute polling errors for the respective period. The row labelled "final polls average" calculates the average based on the final—or only—poll released by each firm, which makes it an effective subset of the post-debate polls.

The 2012 Alberta election and the 2013 British Columbia election represent well-known poll failures,[17] and the total absolute error, averaged for the final election polls in those elections, was 23 points and 17 points, respectively.[18] In the case of the 2015 Alberta election, while there is a difference between the pre- and post-debate polls, there is no substantial difference between the post-debate polls and the final polls conducted by each firm; both measures have average total absolute errors of about 14.4 points. While that is an improvement over the average total absolute error of the final polls in the 2012 Alberta election, it is an improvement of less than 3 points over the average error of the polls in the 2013 British Columbia election. How can it be that the 2015 Alberta election was a vindication of the beleaguered polling industry when the polls this time around were only marginally better (2.6 points) than the "polling failure" that was the 2013 British Columbia election?

The shortcomings are even more apparent when the 2015 Alberta election polls are compared to an election in which polling was quite accurate—in this case, the 2015 Canadian federal election, in which the average total

Table 4.3. Average Error in the 2015 Alberta Election (By Time Period)

	Time Period	Avg. Total Error
Alberta Election	All polls	18.7
	Pre-debate	26.4
	Post-debate	14.5
	Final polls average	14.4
Comparators, calculated by Coletto and Breguet, 2015	**Election**	**Avg. Total Error**
	Alberta 2012	23.0
	British Columbia 2013	17.0
	Canada 2015	6.7

Sources: Data from Elections Alberta, "Provincial Results—Provincial General Election May 5, 2015," polling firm news releases, www.threehundredeight.com, and Election Almanac (www.electionalmanac.com). Comparators: Data from David Coletto and Bryan Breguet, "The Accuracy of Public Polls in Provincial Elections," *Canadian Political Science Review* 9 (2015): 41–54.

absolute error was only 6.7 points across the final polls released, or less than half that of the 2015 Alberta election. Even if the sample of polls in the 2015 Alberta election were reduced to the final four polls, the average total absolute error would still be 10.9 points, which would only close half the distance (3.5 points) between the average error of the final polls in the 2015 Alberta election and the federal election of the same year. Moreover, the best-performing poll in terms of total absolute error, the lone Leger poll, had a total error of 8.0 points, which beats the average total error of the final four polls, and is much closer to the average total error from the 2015 federal election.

Discussion

These findings should give pause to the conventional wisdom that the polling companies "got it right" in 2015. While the polls were closer in 2015 than in 2012 in Alberta, they were only marginally better than the polling failure that was the 2013 British Columbia election. Moreover, the polling errors in 2015 were in a consistent direction (i.e., they were biased in a way that underestimated PC support). With the dominant narrative of the

election being the David-versus-Goliath story of the NDP taking down the PCs, perhaps it was simply convenient for the commentariat to ignore the reality that, in terms of the popular vote, the PCs actually came in second. Thus, the polling companies got a pass for underestimating PC support because to look too closely at the discrepancies between polls and the actual vote would undermine the prevailing narrative. But, as has been shown with this analysis, the post-debate polls met the four criteria for accuracy either inconsistently, incompletely, or not at all.

Unlike other analyses that have used similar criteria[19] this chapter does not make a judgement about which criteria are more important in evaluating the accuracy of a poll, other than to point out that predicting the winner is too low of a bar to set for accuracy. This is especially true, given the multi-party systems that exist in Canada at both the federal and the provincial levels, and the frequency with which close electoral contests occur. Being off by five points when the claimed margin of error is two points is easier to wave away when the gap between the first- and second-place parties is over twelve points, as it was in this election. If Alberta has transitioned away from a one-party dominant system (see Sayers and Stewart, this volume), and competitive elections will become the norm, polls will have to live up to the margins of error that they claim. The uncertain prospects for the merging of the PCs and the Wildrose mean that, at least for the foreseeable future, polls will also need to worry about correctly ordering multiple parties, rather than just predicting a winner and a loser.

Finally, the bias, or systematic error, exhibited by polling in Alberta calls into question the validity of aggregating multiple polls, as several analysts and organizations do, such as ThreeHundredEight and VoxPopLabs in Canada and FiveThirtyEight in the United States. Trusting that the aggregation of multiple data points converges on the truth rests on the assumption that polling errors are normally distributed.[20] In the figures presented in this chapter, that would mean that there are as many dots above the actual result as there are below the actual result. As has been demonstrated in the 2015, anticipated levels of public support for the PCs were consistently below the proportion of the popular vote the PCs actually garnered, so the necessary conditions for effective aggregation are not met in Alberta. Therefore, aggregating polls when they are biased would just give a false sense of the actual accuracy of the data. Before we can further explore why this issue with

Figure 4.2. 2012 Alberta Election Polls
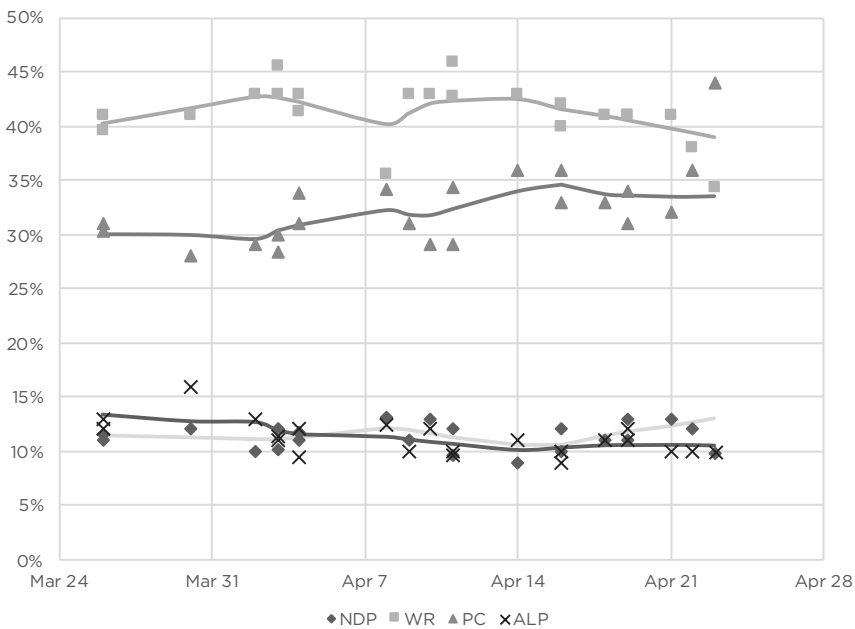
Plotted by last day in field
Final data points indicate actual election result
Lines indicate LOWESS curve; a=0.5

Sources: Data from Elections Alberta, "Provincial Results—Provincial General Election April 23, 2012," polling firm news releases, www.threehundredeight.com, and Election Almanac (www.electionalmanac.com).

accuracy occurs, it is important to note that the 2015 provincial election is just one particular instance of a larger trend of polling issues in Alberta. [21]

As alluded to earlier, the 2012 Alberta provincial election is one of the most well-known cases of polling failure in Canada. Danielle Smith's Wildrose Party was widely expected to defeat Alison Redford's PCs, and with good reason—all the polls released during the campaign said so, as seen in Figure 4.2. Throughout the course of the 2012 election campaign, not a single poll showed the PCs ahead of the Wildrose, despite the PCs eventually winning the election by a margin of 9.7 points. The PCs performed better than all but two polls anticipated, and the Wildrose performed worse than all polls anticipated. Further, the difference between each poll's estimated

versus actual PC or Wildrose support was consistently above its margin of error. As stated in the previous section, average total absolute polling error was also very high. Polling in the 2012 Alberta provincial election fails to meet any of the four criteria outlined at the beginning of this chapter.

What differs between 2012 and 2015 is that the party subject to over-estimation of support changes from the Wildrose to the NDP. These two parties sit at opposite ends of the political spectrum, but both were the party around which opposition to the PCs coalesced. This suggests that polling bias in Alberta has less to do with ideology and more to do with opposing the status quo. Alberta is not alone in this phenomenon, as the governing parties during the 2013 British Columbia and 2012 Quebec provincial elections also defied campaign-period polls, which tended to say that they would be defeated.

This pattern is not just limited to provincial politics in Alberta, but federal politics in Alberta as well. Figure 4.3 shows the Alberta subsamples from polls conducted during the federal election campaign. While the errors are not as stark as in provincial election data, the same pattern can be seen where the federal Conservative Party of Canada outperforms the polls. The Liberals performed at around the middle of the range anticipated by the polls, and the NDP performed at the lower end of what the polls anticipated. Using the LOWESS curve to analyse this data is particularly helpful, since it both averages and calculates trends in the data. The final data point in the LOWESS curve provide further proof of the systematic underestimation of CPC support (by 5.2 points) and overestimation of NDP support (by 7.0 points).

In 2012, the dominant narrative was was of a last-minute shift in vote intentions away from the Wildrose and towards the PCs,[22] and this is one of the shortcomings of any pre-election poll, regardless of its accuracy. The assumption in 2012 was that several pollsters reaching the same conclusion using different methodologies could not all be wrong. The final poll of that campaign, conducted by Forum Research, gives some support to this argument—it showed the closest race out of all the polls, with the Wildrose at 38 per cent and the PCs at 36 per cent. However, PC strategists maintained that their internal polling consistently showed them ahead of the Wildrose, which suggests the possibility that the Wildrose were never really as far ahead as all of the other polls suggested.[23] In the 2015 federal election campaign in Alberta, most of the movement occurred among progressive

Figure 4.3. 2015 Canadian Federal Election Polls in Alberta
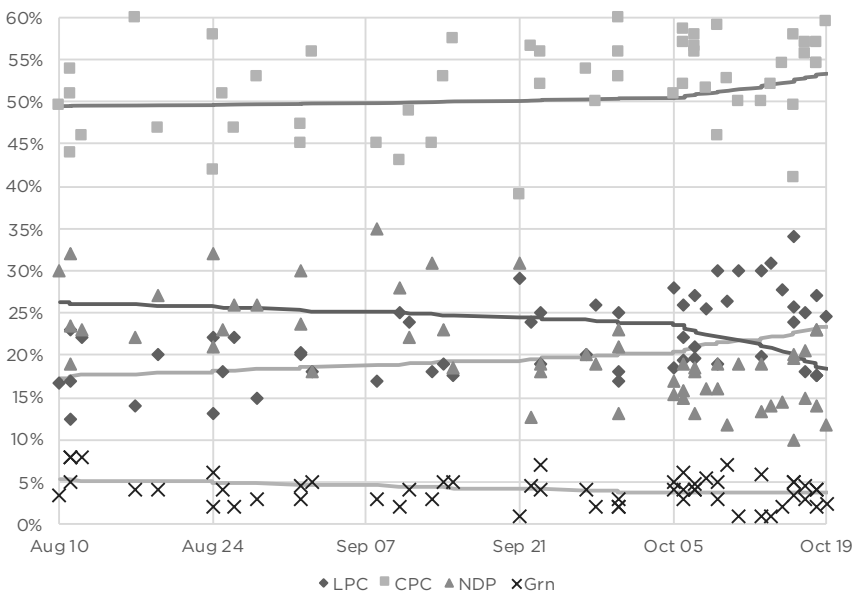
Plotted by last day in field
Final data points indicate actual election result
Lines indicate LOWESS curve; a=0.5

Sources: Data from Wikipedia, "Results of the Canadian federal election, 2015—Results by Province," polling firm news releases, www.threehundredeight.com, and Election Almanac (www.electionalmanac.com).

voters, who moved away from the NDP and to the Liberals. *On the whole*, progressive vote intentions were overestimated and conservative vote intentions were underestimated.

Despite the numerous examples of poll failures, it is important to note those elections—aside from the 2015 federal election—in which polling has been very accurate. More recently, the polls performed very well in the 2017 British Columbia provincial election, with all four polls released in the final week having total absolute errors of less than five points.[24] That the previous British Columbia election is one of the examples of poll failure demonstrates that just because a jurisdiction has a history of inaccurate polling does not mean that all future polls in that jurisdiction are condemned to the same

fate. If polling methods in British Columbia can be improved between elections, there is no reason to think that the same could not happen in Alberta. However, in order for improvement to occur, pollsters will need to continue to refine their methods, and consumers of research need to demand more transparency and accountability from pollsters.

Having established that there is a trend in Alberta whereby support for political change is overestimated and support for the status quo is underestimated, the next question is why. As said previously, timing is a factor. The final Forum and EKOS polls in 2015 were among the last polls conducted during the campaign and were also among the most accurate. However, there is still variability between polls conducted around the same time, and the lone Leger poll out-performed all other polls despite being conducted almost a week prior to the final EKOS, Forum, and Insights West polls. It is possible that the Leger poll was the outlier and that all the other polls around the same time were correct, and that support patterns merely shifted in such a way that made them seem more accurate after the fact. However, most of the movement in intention occurred after the leaders' debate, and the 2015 campaign period lacked any events that could have precipitated a last-minute shift in vote intentions, which suggests that vote intentions had more or less coalesced in the final week.

Another possible explanation is methodology. When polling methodology is discussed in the media, the focus tends to be on interview mode (i.e., live-telephone, interactive voice response, or online) and sample size, to the exclusion of other aspects of methodology. In terms of interview mode, the trends are difficult to identify. Leger's was fielded through an online panel. However, other surveys that used online panels (including those that used online panels in conjunction with live telephone interviews) did not fare as well, with ThinkHQ having a total absolute error of 20.0 points and Ipsos having a total absolute error of 16.0. The most common survey mode was IVR, and those polls had a range of total absolute errors. On the higher end, Pantheon and Mainstreet's final polls had total polling errors of 22.1 and 14.0 points, respectively. On the lower end of the range, Forum's and EKOS's final polls had total absolute errors of 11.0 and 10.6 points, respectively. Live telephone interviews, considered the gold standard in polling, were only used in one poll (conducted by Return on Insight), and that poll had a total

absolute error of 18.0. Thus, interview mode is not a consistent predictor of a poll's final accuracy.

One methodological aspect that does not seem to have had a bearing on accuracy is sample size. While it is true that margin of error decreases as sample size increases, this is only true *if* the sample is truly representative of the population. If the sample is biased, a larger sample size will only give the illusion of increased accuracy. Just as driving faster when one is lost will only make someone even more lost, increasing sample size when there are flaws in either the construction of the sample or in the execution of contacting that sample will only further contribute to error. In 2015, the polls with the largest sample sizes had some of the highest total polling errors. In the post-debate period, the average total absolute error for all polls with samples greater than 2,000 was 20.0 points, whereas the average total absolute error of polls with samples less than 2,000 was 12.4 points.[25] The four most accurate polls, in terms of total absolute error, all had samples that used less than 1,200 respondents (Leger, EKOS, Forum, and Insights West). Thus, the accuracy of a poll has less to do with its size and more to do with the quality of its sample. If a sample is representative, and if a polling firm takes the necessary steps to contact as many people in that sample without being too ready to replace hard-to-reach individuals, then increasing the sample beyond a certain number does not substantially decrease the margin of error, but it does substantially increase the cost of conducting that poll. This is why most public opinion polls have a sample of around 1,000 respondents—that is the "sweet spot" in terms of balancing accuracy and cost, and it is better to make sure that that sample of 1,000 is representative of the population than it would be to increase its size. The lack of a consistent effect on accuracy of either sample size or methodology in the 2015 Alberta provincial election mirrors previous findings by Coletto and Breguet.[26]

Another methodological aspect worth considering is the length of time a poll is in the field. Field length must balance the competing priorities of allowing adequate time to fully reach the targeted population sample while not taking so long that the poll is no longer a snapshot of a given moment in time. Well-executed polling, regardless of the interview mode, should make multiple attempts to contact a sampled respondent before "dropping" that respondent and re-sampling another respondent. This is because not all segments of the population are as easy to get a hold of as others. Thus, if a

polling firm did not make a concerted effort to contact hard-to-reach indi-
viduals, there is a danger of introducing selection bias and only speaking to
those who want to answer polls—which could be individuals who have an
axe to grind against the government. Looking at the top-performing polls
in terms of total absolute error, three out of four of them (Leger, EKOS, and
Insights West) were fielded over periods of three to five days. The exception
is Forum's poll conducted and released on 2 May. Thus, while it does not
give a perfect explanation, length of time in field gives a more consistent
explanation than either interview mode or sample size.

Lessons for the Future

Clearly, polling in Alberta has room to improve, and the analysis in this
chapter shows that polling errors often exceed polls' stated margins of error,
and are biased in a way that underestimates support for the status quo and
overestimates support for change. And, while polling was better in 2015
than in 2012, it still does not come close to the accuracy of national polling
in the 2015 federal election. Based on this analysis, we offer three lessons
that can be learned from 2015.

The first lesson is that, while methodology is important, we must move
beyond simply discussing interview mode and sample size. How well a
poll's sample is constructed and the effort a firm makes to reach a wide
cross-section of survey participants may be more important than how a
firm interviews those respondents. Sample "stratification" and the use of
sample quotas are aspects of methodology that are not often discussed.
To ensure representativeness, key demographic subgroups are identified
within the population. In order to create a well-constructed sample, efforts
must be made to ensure that the demographic composition of the survey
sample matches the actual population. This means creating a sample that,
at minimum, matches the actual population in terms of age, gender, and
region. Efforts should also be made to include hard-to-reach respondents.
For a telephone survey, this means making multiple calls to a telephone
number chosen at random, before classifying it as "unreachable." In the age
of online surveys, this means sending multiple email reminders. If firms are
too eager to drop a hard-to-reach respondent and simply sample another,
easier-to-reach respondent, then the sample may be biased, and this bias

could manifest itself in an under- or overestimation of certain opinions. Dialing 10,000 numbers to complete 1,000 interviews is different from dialing 50,000 numbers to complete 1,000 interviews. The data cannot prove or disprove that selection bias is the reason that PC support is consistently underestimated in Alberta, but the possibility exists that people who want political change are more motivated to share their political opinions, and make themselves more readily available to pollsters by joining online panels, or picking up the telephone when a polling firm calls. All that said, firms are loathe to reveal the details of their sampling and fielding methods, and these other aspects of polling are more difficult to discuss and critique in the media than the more readily understood concepts of interview mode and sample size. However, an honest discussion about which polls are methodologically rigorous cannot occur without this information.[27]

The second lesson is the problematic nature of polling aggregation in Alberta. As popularized by sites such as FiveThirtyEight in the United States and ThreeHundredEight in Canada, some analysts and commentators have taken to aggregating polls in the hopes that more information leads to more accuracy. Statistically speaking, aggregating polling data only works if the estimates of party support provided by polling data are normally distributed around actual public opinion, which, as this analysis has shown, is not the case in Alberta. In fact, in the 2015 Alberta election, a single poll out-performed the aggregation of all polls! Until the accuracy issues of polling in Alberta are resolved across the industry, it would be better to trust selected, well-executed polls than the "collective wisdom" of all polls. The 2016 US presidential election provides further evidence of this. On average, the polls were only off by a couple of points, but they systematically underestimated Donald Trump's support and overestimated Hillary Clinton's support in key battleground states with close races where the election was ultimately decided.[28]

The third point is a warning for the future. The overriding debate in Alberta in 2012 and 2015 was whether or not the PCs should be deposed. With that having happened, it is difficult to say if there is still a systematic bias in polling in Alberta, and if there is, in what way it will manifest itself. Have the NDP become the "new status quo" and will polls underestimate support for them? Or, is the NDP victory an aberration in a streak of small-*c* conservative governments, and will polls continue to underestimate support

for one or the other or both conservative parties in Alberta? Further complicating things is the discussion of a merger between the PC and Wildrose Parties in Alberta, the outcome of which will affect the Alberta party system and the electoral dynamics in subsequent elections.

As this book goes to press, the polling industry has less than a year to resolve the general issue of accuracy and the specific issue of overestimating the desire for change. To the industry's credit, some pollsters readily acknowledge this. Frank Graves, for example, CEO of EKOS, noted the overestimation of NDP support and underestimation of PC support and the need for "better yardsticks" to gauge the effectiveness of polling.[29] It bears reiterating that, in spite of the shortcomings identified in the polling during the 2015 election campaign, there have been improvements since 2012. But there is still much room for improvement, and given the increasingly important role that polling plays in political discourse, it is vitally important that improvement continues.

Notes

1   Janet Brown operates Janet Brown Opinion Research, a public opinion polling firm based in Calgary. John B. Santos is a project manager at Janet Brown Opinion Research, and an MA student in political science at the University of Calgary. Janet Brown Opinion Research conducts polling in Alberta, but did not release any polls during the 2015 Alberta provincial election campaign.

2   Mark Pickup and Richard Johnston, "Campaign Trail Heats as Election Forecasts: Evidence from the 2004 and 2006 Canadian Elections," *Electoral Studies* 26 (2007): 460–76. Pickup and Johnston surveyed the literature on polling in Canadian federal elections and found that 22 polls were published during the 1988 election campaign, 14 in 1993, 14 in 1997, and 23 in 2000. Their analysis focused on the 2004 and 2006 elections, in which there were, respectively, 26 and 66 polls published.

3   Elizabeth Goodyear-Grant, Antonia Maioni, and Stuart Soroka, "The Role of the Media: A Campaign Saved by a Horserace." *Policy Options* 25 (2004): 86–91; J. Scott Matthews, Mark Pickup, and Fred Cutler, "The Mediated Horserace: Campaign Polls and Poll Reporting," *Canadian Journal of Political Science* 45 (2012): 261–87.

4   National Post Wire Services, " 'We were wrong': Alberta Election pollsters red-faced as Tories crush Wildrose," *National Post* (Toronto), 24 April 2012, Accessed October 31, 2016: http://news.nationalpost.com/news/canada/we-were-wrong-alberta-election-pollsters-red-faced-as-tories-crush-wildrose (accessed 31 October 2016).

5   Bruce Cheadle, "Pollsters relieved at getting it right in Alberta's unlikely swing to the left," *Macleans*, 6 May 2015, http://www.macleans.ca/politics/pollsters-relieved-at-getting-it-right-in-albertas-unlikely-swing-to-the-left/ (accessed 31 October 2016).

6   While almost all polls report a margin of error (or an equivalent margin of error, in the case of online panels, which use non-random, or convenience, samples), polls rarely state that their reported margin of error is actually the "maximum margin of error," which applies to proportions of 50 per cent. Actual margins of error for smaller proportions are less than the maximum margin of error (see Francois Petry and Frederick Bastien, "Following the Pollsters: Inaccuracies in Media Coverage of the Horse-race during the 2008 Canadian Election," *Canadian Journal of Political Science* 46, no. 1 (2013): 1–26 for a full discussion of the reporting and misunderstanding of margins of error). To illustrate this, the maximum margin of error on a typical Alberta poll with a sample of 900 would be ±3.27 percentage points, 19 times out of 20. For a proportion of 25 per cent, the actual margin of error decreases to ±2.83 percentage points. Another difficulty for assessment arises due to polls typically excluding unlikely or undecided voters, which would increase the margin of error (due to lowering the sample size being analyzed). For ease of interpretation, this chapter uses the margin of error reported on the standard methodology "boilerplate" included with most press releases, reports, and news stories.

      Pickup and Johnston, in an analysis of the 2004 and 2006 Canadian federal elections, found evidence of bias, or systematic error, in polling at the federal level (see Pickup and Johnson, "Campaign Trail Heats as Election Forecasts"). The most recent American presidential election, the most recent United Kingdom election, and the Brexit Referendum demonstrate that bias in polling may not be confined to Alberta, or even Canada in particular.

7   For example, Forum Research includes a disclaimer with their poll reports that reads, "This research is not necessarily predictive of future outcomes, but rather, captures opinion at one point in time. Election outcomes will depend on the success of the parties in getting out their vote."

8   Pickup and Johnston, "Campaign Trail Heats as Election Forecasts."

9   Elias Walsh, Sarah Dolfin, and John DiNardo, "Lies, Damn Lies, and Pre-Election Polling." *The American Economic Review* 99, no. 2 (2009): 316–22.

10   The dataset was assembled over the course of the election campaign from news releases from polling firms, stories posted on mainstream news media websites, and "poll aggregators" such as ThreeHundredEight (www.threehundredeight.com) and Election Almanac (www.electionalmanac.com). No polling firms, media outlets, or bloggers are responsible for the analysis in this chapter.

11   See MRIA's *Code of Conduct for Members*, http://mria-arim.ca/about-mria/standards/code-of-conduct-for-members (accessed 31 October 2016).

12   While these criteria make use of quantitative data (i.e., numbers), they are essentially qualitative in nature, and this chapter does not construct an "index of accuracy" using the criteria. For such an attempt, see Elizabeth A. Martin, Michael W. Traugott, and Courtney Kennedy, "A Review and Proposal For a New Measure of Poll Accuracy,"

Public Opinion Quarterly 69 (2005): 342–69, which uses similar criteria in quantitative evaluations of polling accuracy.

13 Total absolute polling error is the sum of the absolute differences between each party's level of support as measured by a poll and the actual level of support that each respective party receives at election time. See David Coletto, "Polling and the 2015 Federal Election," in *The Canadian Federal Election of 2015*, ed. Jon H. Pammett and Christopher Dornan, 305–26 (Toronto: Dundurn, 2016).

14 LOWESS stands for "locally weighted scatterplot smoothing," which is a smoothing algorithm that fits a curve through a set of data points by weighting data points closer to a point in time greater than data points further away from that point in time. This is as opposed to trendlines that use functions, or moving averages that use a moving set of data points surrounding a given point in time that are equally weighted (see Pickup and Johnston, "Campaign Trail Heats as Election Forecasts"). Pickup and Johnston use a more complex procedure that calculates the LOWESS curve based on a "poll of polls" that distributes polling data orthogonally along each date a poll was conducted. That method did not produce substantively different results than simply using the last date in field, so the latter was used to facilitate ease of interpretation.

15 Walsh, Dolfin, and DiNardo, "Lies, Damn Lies, and Pre-Election Polling."

16 Coletto, "Polling and the 2015 Election."

17 J. Scott Matthews, "Horserace Journalism under Stress?" *Canadian Election Analysis 2015*: *Communication, Strategy, and Democracy*, 2015, http://www.ubcpress.ca/ CanadianElectionAnalysis2015 (accessed 31 October 2016).

18 Coletto, "Polling and the 2015 Election."

19 Martin, Traugott, and Kennedy, "A Review and Proposal For a New Measure of Poll Accuracy."

20 Pickup and Johnston, "Campaign Trail Heats as Election Forecasts."

21 Ibid.

22 Tu Thanh Ha, " 'Entire environment shifted': Pollsters seek answers following Alberta Election," *Globe and Mail* (Toronto), 24 April 2012, http://www.theglobeandmail.com/ news/politics/entire-environment-shifted-pollsters-seek-answers-following-alberta-election/article1390916/ (accessed 31 October 2016).

23 Karen Kleiss, "Alberta Election 2012: Smith takes the reins as front-runner, poll reveals," *Edmonton Journal*, 18 April 2012, http://www.edmontonjournal.com/news/ Alberta+Election+2012+Smith+takes+reins+front+runner+poll+reveals/6477884/story. html (accessed 31 October 2016). PC strategist Stephen Carter was on record saying that he was not worried about either losing the election or only winning minority government in 2012.

24 The four polls were Ipsos (conducted 4–6 May), Mainstreet Research (conducted 5–6 May), Insights West (conducted 5–8 May), and Forum Research (conducted 8 May). The total absolute errors for each poll, calculated on the basis of parties that won seats, were

1.8 (Ipsos), 4.8 (Mainstreet), 1.5 (Insights West), and 2.2 (Forum) points. See Elections British Columbia, https://catalogue.data.gov.bc.ca/dataset/44914a35-de9a-4830-ac48-870001ef8935 (accessed 8 August 2018).

25 The post-debate polls with sample sizes greater than 2,000 were the three Mainstreet polls. See Table 4.2 for the list of the relevant polls.

26 David Coletto and Bryan Breguet, "The Accuracy of Public Polls in Provincial Elections," *Canadian Political Science Review* 9 (2015): 41–54.

27 André Turcotte calls for the reporting of similar criteria. See his post-mortem on polling in the 2011 Canadian federal election, "Polls: Seeing Through the Glass Darkly," in *The Canadian Federal Election of 2011*, ed. Jon H. Pammett and Christopher Dornan, 195–218 (Toronto: Dundurn, 2012).

28 Carl Bialik and Harry Enten, "The Polls Missed Trump. We Asked Pollsters Why," *FiveThirtyEight*, 9 November 2016, https://fivethirtyeight.com/features/the-polls-missed-trump-we-asked-pollsters-why/ (accessed 31 October 2016).

29 Frank Graves, "EKOS Accurately Predicts NDP Majority Victory in Alberta . . . but should we have better polling yardsticks?" 7 May 2015, http://www.ekospolitics.com/index.php/2015/05/ekos-accurately-predicts-ndp-majority-victory-in-alberta/ (accessed 31 October 2016).