

The Polls and the 1995 Quebec Referendum

John Fox, Robert Andersen, and Joseph DuBonnet*

May 1998

Abstract

We examine 23 voter polls reported in the news media during the two-month period preceding the 1995 Quebec sovereignty referendum. Contrary to common interpretations of the course of the referendum campaign, we find that there was a smooth and general increase in support for the sovereignty option during this period, and that Lucien Bouchard's dramatic intervention on behalf of the 'yes' campaign on October 9 apparently had little effect on voters' stated intentions. The same was true of other putatively consequential events that transpired during the referendum campaign. We conclude that the pre-referendum polls slightly, but systematically, exaggerated the level of support for separation. An analysis of undecided voters suggests that there was a general decline in indecision during the course of the campaign. There was also substantial variation in percentages of undecided voters among polls, and among polling firms, only partly accounted for by differences that we were able to identify in the firms' polling practices.

*Department of Sociology, McMaster University, Hamilton, Ontario, Canada L8S 4M4. The research reported here was supported by a grant from the McMaster University Arts Research Board. This paper is a combined and revised version of papers read at the 1997 meetings of the Canadian Sociology and Anthropology Association and the American Sociological Association. Please address correspondence to the first author.

The October 30, 1995 Quebec sovereignty referendum is surely the most dramatic electoral event in recent Canadian political history. Substantial attention has been paid to the referendum in the popular press, both during and after the event. The scholarly literature has, however, been relatively (if not entirely) silent on this issue. In particular, while there is substantial commentary on the polls, we are aware of no careful statistical analysis of the dynamics of the campaign as revealed in pre-referendum polls.

The period leading up to the 1995 Quebec referendum was notable for the intensity of political polling — 23 general polls of voters were reported in the news media in the two-month period preceding the referendum, and by all accounts many more private polls were commissioned by interested parties.¹ In this paper, we analyze the published polling data for what they reveal about the course of the referendum campaign, testing interpretations of the campaign that have appeared in the popular media and in the academic literature. We also examine the relationship between polling practices and the undecided vote. The latter issue is important because of the closeness of the outcome (in the event, of course, the ‘no’ side won the referendum, by a margin of about one percent.²), and because of wide-spread speculation about the meaning of the undecided vote in pre-referendum polls. Finally, we briefly assess the accuracy of the polls in predicting the outcome of the referendum.

1 Characterizations of the Referendum Campaign

In an analysis of the 1992 Canadian constitutional referendum, LeDuc and Pammett (1995: 5) argue that while referenda in parliamentary democracies can be understood in much the same terms as elections, “their outcome is even more dependent on the short-term elements of the campaign.” At the start of the 1995 Quebec referendum campaign, the ‘yes’ forces were under the formal leadership of the then-premier of Quebec, Jacques Parizeau. The ‘no’ forces were captained by the leader of the Quebec Liberal Party, Daniel Johnson. Parizeau’s leadership was widely regarded as ineffective. On October 9, in a dramatic gesture, Parizeau ceded direction of the ‘yes’ campaign to Lucien Bouchard, then the leader of the Bloc Québécois, and, of course, the current premier of the Province of Quebec.

Just as Parizeau’s leadership was considered ineffective, Lucien Bouchard’s takeover of the campaign was widely interpreted as giving the ‘yes’ forces an important boost — an interpretation that persists to this day. For example, on the day after the referendum, the Toronto *Globe and Mail* ran an article with the headline, “Bloc Quebecois Leader Lucien Bouchard brought big gains when he took over the reigns of the Yes committee” (Picard, 1995).³ Nearly a year after the referendum, the *Globe* reported, “The federal government and the No forces were slow to respond to the emergence of Lucien Bouchard in last year’s Quebec referendum campaign, and this tardiness was largely responsible for the narrowness of the No victory, former prime minister Pierre Trudeau said in an interview broadcast last night” (Canadian Press, 1996). It would not be difficult to produce literally dozens of similar statements.

While dominant, the efficacy of Bouchard’s intervention on the referendum campaign is not universally accepted. Writing in the Montreal newspaper *La Presse*, Quebec sociologist Pierre Drouilly (1995b) pointed to a generally rising trend of support for the ‘yes’ option in polls conducted during the pre-referendum period. Likewise, in a paper published shortly after the referendum, Université de Montréal political scientist Edouard Cloutier identified four potential “turning points” in the referendum campaign (Cloutier, 1995: 37):

On September 24, Claude Garcia, Chief Executive Officer of the Standard Life Insurance Co., and a key official of the NO Committee, declared in a well-publicized

speech that the NO must not only win the vote, but crush the YES side. A very loud uproar followed this pronouncement.

On October 3, Laurent Beaudoin, Chief Executive Officer of Bombardier, and a key spokesman for the NO Committee, announced that following a majority YES vote, he would have to consider moving some of his business outside of Quebec, a statement which also produced a strong reaction on the YES side.

On October 7, Lucien Bouchard was officially named by Jacques Parizeau as the Chief negotiator for Quebec in the talks to be held with Canada after a YES vote.⁴

On October 17, Paul Martin, the Federal Minister of Finance, said that one million jobs could be jeopardized if the YES side won the referendum.

Cloutier went on to argue that the second and third events were too close in time for their effects to be separated. He then examined essentially the same polling data that we employ in this paper, arguing (as we shall) that Bouchard's intervention had little effect, but claiming that Garcia's and Martin's pronouncements served to increase support for the 'yes' option. It is unclear whether this interpretation was formed before or after examining the polling data, and it was not, in any event, put to a test by Cloutier.

Similarly, making reference to commentary by political scientist Pierre Serre in *Le Devoir*, Bernard (1996) argues that support for the sovereignty option in Quebec increased smoothly during the pre-referendum period (although he concedes a short-term impact of Bouchard's intervention). Bernard also argues that the polls accurately reflected voter sentiment. We shall see that the first claim is essentially correct (though there is no evidence for a short-term impact), but the second is not. In contrast, Feavor (1996: 56) maintains that support for sovereignty in the polls "remained remarkably steady throughout the entirety of the campaign," a claim that is not supported by a careful analysis of the polling data.

There was, finally, a great deal of speculation in the news media about the significance of the substantial undecided vote for the accuracy of predictions derived from pre-referendum polls. The conventional wisdom was that, in a close contest, most undecided voters would vote against sovereignty. In a more-or-less typical statement, for example, Pierre Drouilly (1995a) predicted that 75 percent of undecided individuals would vote 'no.' In addition to exploring variations in indecision among the polls, we therefore examine how the undecided vote is related to the 'yes' vote, reasoning that if undecided voters are more likely than not to vote 'no,' then — other factors held constant — polls with relatively large percentages undecided should also show relatively large percentages of voters intending to vote 'yes.'

2 Data and Methods

The principal sources of data for this paper are published pre-referendum polls. Table 1 displays the results of the 23 polls published during the two months preceding the referendum, a summary of which appeared in the *Toronto Globe and Mail* (Winsor, 1995).⁵ Along with the completion date of each poll, we show the percentage intending to vote 'yes' among decided voters and the percentage undecided.

These polls were conducted by nine firms (counting the SOM/Enviro-nics poll as one of four SOM polls⁶), each including between about 500 and 2,000 respondents. Léger and Léger, a Quebec polling firm generally identified with the sovereigntist cause (see, e.g., Cobb 1995; Drouilly, 1995a)⁷, was the most active, conducting six polls. Treating the polls as independent random samples, standard errors for individual percentages range from about 1.1% to about

2.2%. Conventional margins of error (i.e., the half-width of 95% confidence intervals) are roughly twice these figures. Margins of error reported in the media are consistent with these calculations.

In addition to examining the published polling data, we conducted face-to-face or telephone interviews with knowledgeable individuals in each of the firms conducting published polls, collecting information about the firms' polling practices. Finally, each of the firms, with the exception of Decima (which conducted only one poll), returned a questionnaire (based on the one employed by Crespi, 1988) detailing various aspects of their practices. Follow-up telephone calls were employed to resolve ambiguities. Because some of the firms requested that their responses remain anonymous, we report information on the impact of polling practices in such a manner that individual firms cannot be identified.

In reviewing our interview and questionnaire data, we looked for factors that differentiate the firms and that might reasonably be expected to have an impact either on the 'yes' vote or on the percentage undecided. There are more similarities than differences among the firms: All, for example, employed telephone surveys and used probability-sampling methods. In many other instances, there is little variation among the firms: For example, all but one firm employed a 'leaning' question as a follow-up to an initial question about anticipated vote. The polling results reported in the media took account of this follow-up question.⁸ We do not include in our statistical analysis factors that are unique to individual firms, primarily because these factors are completely confounded with the identities of the firms.

With these considerations in mind, we identified the following six factors as worth pursuing in our analysis of the polling data. The first three factors are available for all nine firms; the last three are derived from our questionnaire data, and therefore are unavailable for Decima:

- Some of the polls focused specifically on the upcoming referendum, while others were part of more general surveys. In the latter event, political content either appeared first or after other material; in the former event, political content necessarily appeared first. In all, six of the nine firms placed the political content of their questionnaires first.
- Two of the firms employed the official referendum question in their polling; the remaining firms asked, with small variations, "How do you plan to vote in the upcoming referendum?"⁹
- Seven of the firms positioned the question on anticipated vote near the beginning of the political questions, one firm in the middle of these questions, and one at the end.¹⁰
- Four of eight firms employed demographic quotas for age, gender, or region in sample selection.
- The average length of the interviews for these eight firms varied from about two minutes to about 15 minutes.
- Two of eight firms employed random-digit dialing for sample selection; these are, it turns out, the same firms that used the official referendum question, so the effects of these two factors cannot be separated.

Our statistical data analysis constitutes a simple meta-analysis, employing binomial logit models (e.g., Collett, 1991: Ch.3; Fox, 1997: Ch.15), of the data from the 23 polls. The response variables in these models are the number of voters intending to vote 'yes' among decided voters and the number undecided among all respondents; explanatory variables include characteristics of the polls, such as the date that the poll was conducted, the firm that conducted the poll, and various practices employed by this firm. Although there are only 23 binomial observations, these

represent more than 20,000 respondents, and the binomial logit models fit to the data are formally equivalent to binary logit models based on individuals’ survey responses.¹¹ We employ both linear logit models, and nonparametric (and semiparametric) logit models in which the functional form of the relationship between anticipated vote and (some of) the explanatory variables is not specified in advance (Hastie and Tibshirani, 1990).

3 Decided Voters

Figure 1 plots the percentage of decided voters intending to vote ‘yes’ against the completion date of the poll, expressed in days after August 31 (i.e., September 1 = 1). The line on the plot results from fitting a logit model to the polling data, as described below. With the exception of the first Léger and Léger poll, conducted on September 9, there is an apparent general tendency for the percentage of decided voters planning to vote ‘yes’ to increase with time.

Several binomial logit models fit to the 23 polls are summarized in Table 2. The table shows the predictors included in each model, the deviance under the model (that is, $-2 \times$ the maximized log-likelihood for the model), and the degrees of freedom associated with the deviance. The dependent variable is voting intention, among decided voters. Model 0 is a null model, fitting a constant only, and therefore incorporating the hypothesis of no differences among the polls. The deviance for each model is distributed as chi-square under the hypothesis that the model is correct (and implicitly contrasts the model with a ‘saturated’ model, which fits a separate parameter to each poll). The null model, therefore, fits the data poorly. The remaining models in the table are classified by topic: Models 1 and 2 will be used to test for trends; models 3 through 7 will be used to test for inter-firm differences; and so on.

All of the six Léger and Léger polls produced relatively large percentages of ‘yes’ voters, but the September 9 poll stands out particularly from other polls conducted during the early days of the referendum campaign. A likelihood-ratio test of the hypothesis that this poll is not an outlier produces the chi-square test statistic $G^2 = 11.201$, with one degree of freedom, yielding the p -value .000725 (which, upon Bonferroni adjustment, becomes $23 \times .000725 = .0167$). On the basis of this statistically significant outlier test, we set the first poll aside in most of our reported analysis of decided voters, although we show results both including and excluding the suspect poll.¹²

The analysis of deviance in Table 3 shows likelihood-ratio tests for a variety of hypotheses concerning the course of the referendum campaign as reflected in the polls. This table is analogous to an analysis of variance table for a linear model: The likelihood-ratio test statistics are calculated by contrasting pairs of models from Table 2. As in Table 2, the tests in the analysis of deviance table are classified by topic.

3.1 The Trend in the ‘Yes’ Vote

It is apparent from the first panel of Table 3 that there was a strong linear trend in the percentage of decided voters intending to vote ‘yes’ during the two-month referendum campaign.¹³ Moreover, this model fits the data well: The departure from linearity is small and non-significant (particularly when the first Léger and Léger poll is excluded); and the deviance for model 1 is smaller than its degrees of freedom. Because of the highly significant linear trend in support for the ‘yes’ option, date is controlled in all of the tests reported below.

3.2 Effects of Firms and Polling Practices on the ‘Yes’ Vote

The tests in the second panel of Table 3 examine differences among polling firms (i.e., “house effects”). Recall that the firm of Léger and Léger was associated (perhaps inaccurately) with the sovereigntist cause. Lepage and Créatec were, respectively, the official pollsters of the ‘yes’ and ‘no’ campaigns.

A test of general differences among the nine polling firms produces a non-significant result ($p = .15$) when the first Léger poll is excluded; but a test that focuses specifically on the three firms of Léger and Léger, Lepage, and Créatec turns up a nearly significant result ($p = .061$). Follow-up tests for each of the three firms show just-significant differences from the other firms for Léger and Lepage ($p = .046$ and $p = .045$, respectively), but not for Créatec ($p = .72$).

Table 4 reports estimates for a model of ‘yes’ support that includes a linear trend in time along with dummy regressors for Léger and Lepage (corresponding to model 7 in Table 2). Coefficient estimates are given in the column labelled B and estimated asymptotic standard errors in the column labelled $SE(B)$; coefficients that are more than twice their standard errors are statistically significant at the five-percent level, two-tailed, by a Wald test. Because the ‘yes’ vote stayed relatively close to 50 percent, dividing each coefficient by four and multiplying by 100 gives approximate effects of each regressor on the percentage scale. Thus, on average, the decided ‘yes’ vote increased by 0.17 percent per day, and — holding date constant — Léger and Léger polls tended to produce results about 1.8 percent higher, and Lepage polls results about 2.3 percent higher, than those conducted by other firms.

The third and fourth panels of Table 3 test for effects on the anticipated ‘yes’ vote of interfirm differences in specific polling practices. Political, Question, and Position are 0/1 dummy regressors coded one, consecutively, if the political content of the questionnaire came first, if the official referendum question was employed, and if the anticipated-vote question appeared near the beginning of the political questions; Quotas is a dummy regressor coded one if demographic quotas were employed in sample selection; and Length is the average length of the interview, in minutes. Recall that the last two variables are unavailable for the September 25 Decima poll. When the discrepant September 9 Léger and Léger poll is omitted, there is little evidence that any of these polling practices had an impact on support for the ‘yes’ option.

3.3 The Dynamics of the Campaign

We turn now to characterizations of the referendum campaign. A test for a change in level or trend of support for the ‘yes’ option following Bouchard’s October 9 intervention fails to find support for these effects; indeed, omitting the first Léger poll, there is virtually no difference in the level or trend prior to and subsequent to this date ($p = .99$). We model possible changes in level and slope by including a dummy regressor coded one after October 9 and zero before this date, along with the product of the dummy regressor and date.

There is also little evidence for Cloutier’s characterization of the referendum campaign. We test his characterization by constructing dummy regressors and interactions with date for each of the three interventions identified as potentially significant. Figure 2 shows the polling data divided into Cloutier’s four periods: The fit of a logit model permitting different levels and slopes (model 12 in Table 2, omitting the first Léger poll) is graphed along with the simple average of percentages ‘yes’ in each of the four periods. The averages are included because the periods are narrow, and consequently the individual slope estimates are relatively unstable.

Table 3 reports likelihood-ratio tests for each intervention and for all three interventions simultaneously. When the discrepant first poll is included, there is apparent evidence for the

impact of Garcia’s intervention: A decline in ‘yes’ support in the first period reverses itself in the second period. This effect largely disappears, however, when the September 9 Léger poll is omitted.

The last panel of Table 3 reports a test of the relationship between the percentage intending to vote ‘yes’ and the percentage ‘undecided,’ controlling for the other factors that appear to have affected the ‘yes’ vote — in particular, a linear trend in date and dummy regressors for the firms of Léger and Lepage. Although there is, as predicted, a positive relationship between the percentages ‘yes’ and undecided, this relationship is not statistically significant ($p = .21$).

On the eve of the referendum, the outcome of the vote was widely regarded as too close to call. Polls conducted close to the date of the referendum indicated that about 52 percent to 53 percent of decided voters intended to vote ‘yes,’ but these same polls indicated an undecided vote in excess of 10 percent (to be discussed presently), and the individual polls had margins of error of about ± 3 percent. Leaving aside the question of undecided voters, however, combining data from the nearly 24,000 respondents to the 22 polls permits a much sharper prediction: Using model 1 (which — recall — fits a linear trend on the logit scale) to project the percentage ‘yes’ vote to the October 30 date of the referendum, produces the prediction of a 53.1 percent ‘yes’ vote, with a 95-percent confidence interval running from 51.7 percent to 54.5 percent. This prediction is, of course, mistaken.¹⁴

4 The Undecided

A striking characteristic of most of the pre-referendum polls — even those conducted close to October 30 — is the large proportion of voters who gave noncommittal responses, ranging from eight percent, in a Créatec poll conducted on October 11 and an Angus Reid poll conducted on October 25, to 24 percent in a COMPAS poll conducted on September 14 (see Table 1 and Figure 3). Because the reports of many of the polls failed to distinguish ‘don’t know’ responses from refusals to answer, we combine these two categories of response, terming them “undecided.” When the two categories of noncommittal response were distinguished, most represent ‘don’t know’ responses. Of course, neither category is unambiguous, since ‘don’t know’ can mask a refusal to respond, and a refusal to respond can be motivated by the failure to make up one’s mind. Ironically, despite the substantial percentages of undecided voters in polls conducted close to October 30, the turnout in the referendum was in excess of 90 percent.

Our analysis of decided voters, described above, produced an essentially simple result: Although there was some indication of differences among polling firms, a linear trend in the logit of ‘yes’ voters fits the data very well. The situation with respect to undecided voters is more complex. Figure 3 suggests that there may be a nonlinear relationship between the undecided vote and time, with an initial decline in the percentage undecided, followed by a period of relative stability, followed by another decline towards the end of the referendum campaign. Several individual polls, however, are relatively remote from this general trend: This is particularly clear for the September 24 COMPAS poll, but several other polls also depart significantly from the general trend.

Table 5 summarizes the fit of several logit models to the referendum data on indecision. Table 6, derived from Table 5, presents an analysis of deviance for factors that may affect indecision.

4.1 The Trend in the Undecided Vote

The first panel in Table 5 includes models representing general trends: Model 1 fits a linear trend on the logit scale; model 2 fits date categorically, therefore capturing any trend; and model 3 is a nonparametric logit model fitting a smooth trend in date.¹⁵ The first panel of Table 6 shows a test for the linear trend in date, along with two tests for nonlinearity — contrasting model 1 respectively with models 2 and 3. All of these effects are highly statistically significant, but none of the models fits the data well. Even model 3, which fits a unique effect for each date, and from which the lack of fit therefore derives only from dates with more than one poll, has a statistically significant lack of fit ($p \ll .0001$). There is, therefore, substantial room for finding other factors — such as inter-firm differences and the polling practices upon which such differences potentially depend — that influence indecision.

4.2 Effects of Firms and Polling Practices on the Undecided

The second panels of Tables 5 and 6 pertain to interfirm differences in the rate of indecision. These differences turn out to be highly statistically significant, even when we only focus on polls conducted by different firms on the same date (the test that contrasts models 2 and 5). Model 5, which treats date categorically and includes dummy regressors for firm effects, fits the data well, but this model uses 22 independent parameters for the 23 polls. It would, therefore, be desirable to account for interfirm differences more parsimoniously on the basis of differences in their polling practices.

The final two panels of Table 5 examine the effects of various polling practices on indecision, along with a nonparametric trend in the date of the polls. The effects of the position of the political content in the questionnaire, of the question employed, and of the position of the anticipated-vote question are all highly statistically significant. There is some suggestion that either the use of quotas or the length of the questionnaire has a significant effect on indecision, but neither factor tested individually is statistically significant. None of these models achieves a satisfactory fit to the data: All have deviances far in excess of their degrees of freedom.

Table 7 displays the coefficients for model 6. This is a semiparametric model, with dummy regressors for position of political content (coded one if political content appears first), question (one if the official referendum question is employed), and position of voting question (one if the voting question is near the beginning of the political questions), along with a nonparametric trend in date.¹⁶ The coefficients for the parametric part of the model are given in the table. Because indecision was typically near about 15 percent, the logit fit is substantially nonlinear; we therefore opt to show $\exp(B) = e^B$ for each coefficient, which gives multiplicative effects on the *odds* of indecision.

Putting the political content first decreases the odds of indecision by about one-third; using the official question instead of simply asking for anticipated vote increases the odds of indecision by about 20 percent; and putting the voting question before other political content decreases the odds of indecision by nearly 20 percent. The first two coefficients are in the direction that we anticipated, but the third coefficient is not. In addition, model 6 does not account completely for interfirm differences, as can be seen by contrasting models 6 and 4 ($G^2 = 18.627, df = 5, p = .002$).

5 Discussion

Our analysis of the pre-referendum polling data for decided voters suggests that common interpretations of the course of the campaign are likely in error. Lucien Bouchard is widely credited with giving a substantial boost to the ‘yes’ forces, and Jacques Parizeau’s leadership is generally regarded as having been ineffective, yet there is hardly a hint of the transition in leadership of the ‘yes’ campaign in the polling data. We cannot know, of course, what would have happened if Bouchard had not intervened: Perhaps, in Bouchard’s absence, ‘yes’ support in the polls would have gone into a tailspin, rather than continuing to increase gradually. Although we cannot rule out this possibility, we see no reason to suppose that it would have materialized. Cloutier’s more detailed analysis of ‘turning points’ in the referendum campaign is likewise not supported by the data.

The polls point towards a clear, linear increase in ‘yes’ support during the two-month campaign period preceding the referendum, with some suggestion of differences among polling firms — but no indication that these differences are due to the specific polling practices that we examined. What accounts for the trend in the pre-referendum polls? Here we can only speculate: The pre-referendum period witnessed an apparent increase in the polarization of Quebec society along linguistic and ethnic lines, and between Montreal and the rest of the province. A majority of French-speakers, and a majority of the Quebec population outside of Montreal, voted for independence. Almost all English-speakers and “allophones” (individuals whose native language is neither French nor English), and a majority of residents of Montreal, voted “no” (see Drouilly, 1995c; and Lachapelle, 1995).

A prediction of the referendum outcome based on the trend among decided voters suggests a clear — if narrow — ‘yes’ victory. The discrepancy between the polls and the referendum result has been attributed variously to undecided voters opting to vote ‘no’ and (implicitly) to sampling error in the polls. But our analysis indicates little positive impact of the percent undecided on the ‘yes’ vote in the polls (as seems to be implied by the first argument), and our meta-analysis of the polling data yields a sharp prediction (contradicting the second argument).¹⁷

We can also only speculate about the sources of nonsampling error in the poll results. It is important to recall, first, that the error is not large. The ‘no’ side, after all, won only the narrowest of victories. Second, it is conceivable that voters who remained undecided in the polls conducted near the end of the campaign period did in fact opt disproportionately to vote ‘no,’ even though the level of support for the ‘yes’ side during the campaign was not strongly related to the rate of indecision. Third, although in only one instance did a poll report a *global* rate of nonresponse (39 percent, for the October 12 poll conducted by the Gallup organization), it is safe to assume that this rate was high in all of the polls, producing potentially substantial biases in the results. Given the heated atmosphere of the referendum campaign, voters who opposed independence may have been differentially disposed not to respond to polls, and, when they did respond, may have misreported their voting intentions. Some of these speculations could have been addressed by conducting re-interviews with respondents after the referendum. To our knowledge, however, no data of this kind have been collected.

It is, finally, conceivable — if unlikely — that the discrepancy between the polls and the election result represents neither sampling nor nonsampling error: It is possible that some event immediately before the referendum, and after the final polls, substantially influenced the outcome. For example, a very large federalist rally, attended by many Canadians from outside of Quebec, took place in Montreal just two days prior to the vote.

Our analysis of sources of indecision in the pre-referendum polling data produced less generally satisfactory results. There was an apparent tendency for indecision to decrease during

the course of the referendum campaign. There were also substantial variations in rates of indecision among polling firms. To a degree, inter-firm differences could be attributed to some of the practices employed by these firms, but the relatively small number of firms, and the limited variation in the practices that they employed, make interpretation of these differences difficult. Moreover, no parsimonious model for indecision fits the data well.

Beyond the particularities of the 1995 Quebec referendum campaign, this paper demonstrates how a statistical meta-analysis of polling data can address questions about the trajectory of public opinion. Our analysis transcends the typical informal, descriptive examination of polling results, to test specific hypotheses about the dynamics of the campaign, such as Bouchard's role and the impact of nonresponse. Moreover, by combining data from several sources, we show how it is possible to control for — and to determine the effects of — variations in survey practices, producing more definitive conclusions than would otherwise be possible.

Notes

¹The prevalence of private polls was confirmed in our interviews with the polling firms.

²Some claim that the margin would have been perhaps two percent in the absence of voter fraud, but the nature and extent of fraud is the subject of dispute. Bernard (1996: 245), for example, states that "[f]raud was almost negligible."

³Curiously, on the same page, another article ran with the headline, "Analysis of Quebec-based surveys shows rise and fall of Yes and No support rarely varied more than plus or minus 3 per cent" (Winsor, 1995). We shall see that neither of these contradictory characterizations is accurate.

⁴The October 7 date differs slightly from the October 9 date of Bouchard's assumption of control of the 'yes' campaign. Because no polls were completed between October 7 and October 9, however, the difference is moot.

⁵Media reports of the October 25 Angus Reid poll indicated that 16 percent of respondents were undecided, and that 52 percent of decided voters intended to vote 'yes.' These reports, however, did not distribute initially undecided voters based on a followup 'leaning' question that was also asked in this poll, as was standard practice in other Reid polls. We have opted to make the data for the October 25 poll consistent with the other Reid polls, using information supplied to us by that firm. Although the percentage undecided declines substantially when the followup question is employed, the percentage of decided voters intending to vote 'yes' does not change.

⁶An interview with Environics revealed that this poll was conducted by SOM and employed that firm's standard methods.

⁷In an interview, a representative of the firm disputed this characterization, and the official pollster for the 'yes' side in the referendum campaign was Lepage, not Léger.

⁸As noted, reported results for the October 21 Angus Reid poll did not employ the leaning question, even though such a question was asked; this appears to have been an oversight.

⁹Not all of the firms were willing to reveal to us the precise wording of the question that they employed.

¹⁰One of firms asked for anticipated vote both near the beginning and at the end of the political questions. Because the former question was used in published results, we classified this firm with those asking about voting intentions near the beginning.

¹¹Access to the individual data would, of course, permit us to address how individual characteristics — such as language, ethnicity, and place of residence — affected anticipated vote in the referendum, something that it is not possible from the aggregate poll results.

¹²This likelihood-ratio test compares the deviance for model 1 in Table 2, excluding the first observation, with the deviance for the same model fit to the data including the first observation. We do not, by the way, know why the first Léger and Léger poll produced apparently discrepant results: The firm reports using the same methods in all of their pre-referendum polls.

¹³Because the model is a logit model, the trend is, strictly speaking, linear on the logit scale. 'Yes' support never strays far from 50 percent, however, and so the trend is also very nearly linear on the percentage scale.

¹⁴An alternative here is to base a 'prediction' on model 7 (the coefficients for which are given in Table 4), discounting the possibly positively biased Lepage and Léger polls. This approach produces a predicted 'yes' vote of 52.3 percent, with a 95-percent confidence interval running from 50.7 percent to 53.9 percent — still a clear prediction of a 'yes' victory. To discount the Lepage and Léger polls in this manner, moreover, requires hindsight knowledge of the referendum result.

¹⁵We used a locally linear (loess) smoother, with a span of 60 percent of the observations.

¹⁶It is not possible to add polling-practices effects to model 2, which treats date categorically. Since model 2 has almost as many parameters as polls, adding polling effects produces many aliased coefficients, making it impossible to separate these effects from the effect of date.

¹⁷It is possible that the sampling errors of the polls is understated by treating them as independent random samples, but we do not expect substantial inflation of sampling errors in telephone surveys, which do not employ clustering, and we have insufficient information to try to correct for design effects.

References

- Bernard, P.
1996 Canada as a social experiment. *The Canadian Journal of Sociology*, 21:245–258.
- Canadian Press
1996 Trudeau faults tardiness. *The Globe and Mail (Toronto)*, Oct. 24:A2.
- Cloutier, E.
1995 The Quebec referendum: From polls to ballots. *Canada Watch*, 4(1):37,39.
- Cobb, C.
1995 Oui the people. *The Sun (Vancouver)*, Sept. 9:B1.
- Collett, D.
1991 *Modelling Binary Data*. Chapman and Hall, London.
- Crespi, I.
1988 *Pre-election Polling: Sources of Accuracy and Error*. Russell Sage Foundation, New York.
- Drouilly, P.
1995a Référendum: un parfum de défaite. *La Presse (Montreal)*, Sept. 2:B3.

1995b La progression du OUI dans les sondages. *La Presse*, Oct. 21:B10.

1995c An exemplary referendum. *Canada Watch*, 4(1):25–27.
- Feavor, G.
1996 Canadian political miscalculation? Quebec's referendum 95. *Government and Opposition*, 31:45–61.
- Fenwick, I., Wiseman, F., Becker, J. F., and Heiman, J. R.
1982 Classifying undecided voters in pre-election polls. *Public Opinion Quarterly*, 46:383–391.
- Fox, J.
1997 *Applied Regression, Linear Models, and Related Methods*. Sage, Thousand Oaks, CA.
- Hastie, T. and Tibshirani, R.
1990 *Generalized Additive Models*. Chapman and Hall, London.
- Jowell, R., Hedges, B., Lynn, P., Farrant, G., and Heath, A.
1993 The 1992 British election: The failure of the polls. *Public Opinion Quarterly*, 59:238–263.
- Lachapelle, G.
1995 The support for sovereignty among decided voters: The picture. *Canada Watch*, 4(1):3–4.
- LeDuc, L. and Pammett, J.
1995 Referendum voting: Attitudes and behaviour in the 1992 constitutional referendum. *Canadian Journal of Political Science*, 28:3–33.

Mackie, R.

1995a Yes 50.2, No 49.8, poll suggests. *The Globe and Mail (Toronto)*, Oct. 21:A1, A4.

1995b Dead heat: Quebec vote is a toss-up. *The Globe and Mail (Toronto)*, Oct. 28:A1, A7.

Picard, A.

1995 A chronology of events leading to yesterday's vote. *The Globe and Mail (Toronto)*, Oct. 31:A4.

Winsor, H.

1995a Campaign hoopla belies calm pattern of opinion polls. *The Globe and Mail (Toronto)*, Oct. 31:A4.

1995b Sovereignists' lead jolted money markets. *The Globe and Mail (Toronto)*, Oct. 24:A4.

<i>Date</i>	<i>Firm</i>	<i>Percent</i>		<i>Sample Size</i>
		<i>Yes (Among Decided)</i>	<i>Undecided</i>	
Sep. 9	Léger and Léger	51	13	959
Sep. 12	SOM	45	18	1003
Sep. 14	COMPAS	47	24	500
Sep. 19	Créatec	46	15	1004
Sep. 25	Decima	49	18	750
Sep. 25	CROP	45	14	2020
Sep. 25	SOM/Environics	45	13	1820
Sep. 27	Angus Reid	48	14	1000
Sep. 28	Léger and Léger	49	11	1006
Sep. 29	Lepage	49	10	1369
Oct. 4	Léger and Léger	49	13	1015
Oct. 9	Lepage	52	13	1285
Oct. 11	Créatec	47	8	470
Oct. 12	Gallup	48	18	1013
Oct. 12	Léger and Léger	52	13	1002
Oct. 16	SOM	50	14	981
Oct. 16	CROP	49	14	1151
Oct. 18	Angus Reid	51	11	1012
Oct. 20	Léger and Léger	52	12	1005
Oct. 23	CROP	51	13	1072
Oct. 25	Angus Reid	52	8	1029
Oct. 25	SOM	53	14	1115
Oct. 27	Léger and Léger	53	12	1003

Table 1: Results of 23 pre-referendum polls, showing the completion date of each poll, the percentage of decided voters intending to vote ‘yes,’ and the percentage undecided. The percentage undecided includes “don’t know” and no-answer responses. *Source of Data:* Winsor (1995); the results from the Oct. 25 Angus Reid poll are corrected on the basis of information provided by Reid.

<i>Model</i>		<i>Including First Poll</i>		<i>Excluding First Poll</i>	
		<i>Deviance</i>	<i>d.f.</i>	<i>Deviance</i>	<i>d.f.</i>
0	Constant Only	61.133	22	60.381	21
<i>Trend</i>					
1	Date	28.756	21	17.555	20
2	Date (categorical)	6.608	5	6.608	5
<i>Firm Effects</i>					
3	Date, Firms	12.111	13	5.556	12
4	Date, Léger, Lepage, Créatec	15.830	18	10.195	17
5	Date, Lepage, Créatec	25.604	19	14.162	18
6	Date, Léger, Créatec	19.570	19	14.222	18
7	Date, Léger, Lepage	16.150	19	10.328	18
<i>Survey Practices I (All Firms)</i>					
8	Date, Political, Question, Position	20.471	18	12.975	17
<i>Survey Practices II (Omitting Decima)</i>					
9	Date	28.408	20	16.721	19
10	Date, Quotas, Length	22.900	18	14.837	17
<i>“Bouchard Effect”</i>					
11	Date (D), Bouchard (B), $D \times B$	25.921	19	17.530	18
<i>Cloutier’s Characterization of the Campaign</i>					
12	Date, Garcia (G), Bouchard, Martin (M), $D \times G, D \times B, D \times M$	15.642	15	12.658	14
13	Date, Bouchard, Martin, $D \times B, D \times M$	23.970	17	15.580	16
14	Date, Garcia, Martin, $D \times G, D \times M$	19.926	17	16.942	16
15	Date, Garcia, Bouchard, $D \times G, D \times B$	17.592	17	14.608	16
<i>Effect of Nonresponse on the Reported ‘Yes’ Vote</i>					
16	Date, Leger, Lepage, Undecided	15.405	18	8.728	17

Table 2: Deviances and degrees of freedom for binomial logit models for decided voters intending to vote ‘yes’ in 23 pre-referendum polls. Models including and excluding the first Léger and Léger poll are shown. The terms in the models are as follows: Date is a linear term in the dates of the polls; Date (categorical) is a set of dummy regressors for poll dates; Firms is a set of dummy regressors for polling firms; Léger, Lepage, and Créatec are dummy regressors for specific firms; Political, Question, Position, and Quotas are dummy regressors coding various polling practices (see the text for details); Length is the average length of the interview in minutes; Bouchard, Garcia, and Martin are dummy regressors coded 1 after potentially significant events (discussed in the text); and Undecided is the percentage undecided. Terms with a times sign (\times) represent interactions.

<i>Models</i>	<i>Source</i>	<i>Including First Poll</i>			<i>Excluding First Poll</i>		
		G^2	<i>d.f.</i>	<i>p</i>	G^2	<i>d.f.</i>	<i>p</i>
<i>Trend</i>							
0 – 1	Date (linear)	32.377	1	≪ .0001	42.826	1	≪ .0001
1 – 2	Date (nonlinear)	22.148	16	.14	10.947	15	.76
<i>Firm Effects</i>							
1 – 3	Firms	16.645	8	.034	11.999	8	.15
1 – 4	Léger, Lepage, Créatec	12.926	3	.0048	7.360	3	.061
5 – 4	Léger	9.774	1	.0018	3.967	1	.046
6 – 4	Lepage	3.740	1	.053	4.027	1	.045
7 – 4	Créatec	0.320	1	.57	0.133	1	.72
4 – 3	Other Firms	3.719	5	.59	4.639	5	.46
<i>Survey Practices I (All Firms)</i>							
1 – 8	Political, Question, Position	8.285	3	.040	4.580	3	.21
<i>Survey Practices II (Omitting Decima)</i>							
9 – 10	Quotas, Length	5.508	2	.064	1.884	2	.39
<i>“Bouchard Effect”</i>							
1 – 11	Bouchard, Date×Bouchard	2.835	2	.24	0.025	2	.99
<i>Cloutier’s Characterization of the Campaign</i>							
1 – 12	Garcia, Bouchard, Martin, $D \times G, D \times B, D \times M$	13.114	6	.041	4.897	6	.56
13 – 12	Garcia, Date×Garcia	8.328	2	.016	2.922	2	.23
14 – 12	Bouchard, Date×Bouchard	4.284	2	.12	4.284	2	.12
15 – 12	Martin, Date×Martin	1.950	2	.38	1.950	2	.38
<i>Effect of Nonresponse on the Reported ‘Yes’ Vote</i>							
7 – 16	Undecided	0.745	1	.39	1.600	1	.21

Table 3: Analysis of deviance table for effects on the ‘yes’ vote in binomial logit models fit to the pre-referendum polls (see Table 2). Two sets of tests are shown, including and excluding the first Léger and Léger poll. G^2 is the likelihood-ratio chi-square test statistic.

<i>Coefficient</i>	B	$SE(B)$	$100 \times B/4$
Constant	-0.3116	0.0429	
Date	0.006735	0.001103	0.1684
Léger	0.07249	0.03538	1.812
Lepage	0.09275	0.04472	2.319
Deviance	10.328		
df	18		

Table 4: Coefficients for a logit model for the ‘yes’ vote. The model includes a linear trend in time and dummy regressors for polls conducted by Léger and Léger and Lepage. Data from the first Léger and Léger poll are omitted. The column labelled $100 \times B/4$ gives approximate effects on the percentage scale near 50 percent.

<i>Model</i>	<i>Deviance</i>	<i>d.f.</i>
0 Constant Only	164.134	22
<i>Trend</i>		
1 Date	135.162	21
2 Date (categorical)	40.135	5
3 Date (nonparametric)	116.721	18.8
<i>Firm Effects</i>		
4 Date (n), Firms	40.526	10.8
5 Date (categorical), Firms*	0.382	1
<i>Survey Practices I (All Firms)</i>		
6 Date (n), Political, Question, Position	59.153	15.8
7 Date (n), Question, Position	108.350	16.8
8 Date (n), Political, Position	68.621	16.8
9 Date (n), Political, Question	71.289	16.8
<i>Survey Practices II (Omitting Decima)</i>		
10 Date (n), Political, Question, Position, Quotas, Length	45.653	12.8
11 Date (n), Political, Question, Position	52.168	14.8
12 Date (n), Political, Question, Position, Length	47.846	13.8
13 Date (n), Political, Question, Position, Quotas	45.872	13.8

Table 5: Deviances and degrees of freedom for binomial logit models for the undecided in 23 pre-referendum polls. The terms in the models are explained in Table 1 and in the text. Model 4, marked with an asterisk, includes aliased effects. Date (n) is an abbreviation for a nonparametric trend in the dates of the polls.

Models	Source	G^2	<i>d.f.</i>	<i>p</i>
<i>Trend</i>				
0 – 1	Date (linear)	28.972	1	\ll .0001
1 – 2	Date (nonlinear) ^a	95.027	16	\ll .0001
1 – 3	Date (nonlinear) ^b	18.441	2.2	.0001
<i>Firm Effects</i>				
3 – 4	Firms (after Date, nonparametric)	76.195	8	\ll .0001
2 – 5	Firms (after Date, categorical)	39.753	4	\ll .0001
<i>Survey Practices I (All Firms)</i>				
3 – 6	Political, Question, Position	57.568	3	\ll .0001
7 – 6	Political	49.197	1	\ll .0001
8 – 6	Question	9.468	1	.002
9 – 6	Position	12.136	1	.0005
<i>Survey Practices II (Omitting Decima)</i>				
11 – 10	Quotas, Length	6.515	2	.038
12 – 10	Quotas	2.193	1	.14
13 – 10	Length	0.219	1	.64

Table 6: Analysis of deviance table for effects on the undecided vote, based on logit models fit to the pre-referendum polls (see Table 5).. Two tests of nonlinearity of the date effect are shown: Test (a) contrasts the linear logit model with a model in which date is treated as categorical; test (b) contrasts the linear logit model with a nonparametric logit model. G^2 is the likelihood-ratio chi-square test statistic.

<i>Coefficient</i>	B	$\exp(B)$
Political	-0.3906	0.6767
Question	0.1807	1.1981
Position	-0.1973	0.8209

Table 7: Coefficients for a semi-parametric model (model 6 in Table 5) for the ‘undecided’ vote. The model includes a nonparametric trend in time and dummy regressors for the following survey practices: the position of political content in the questionnaire (1 = last); whether the official referendum question was used (1 = yes); and the position of the anticipated vote question relative to other political questions (1 = first). The column labelled $\exp(B)$ gives multiplicative effects on the odds of indecision.

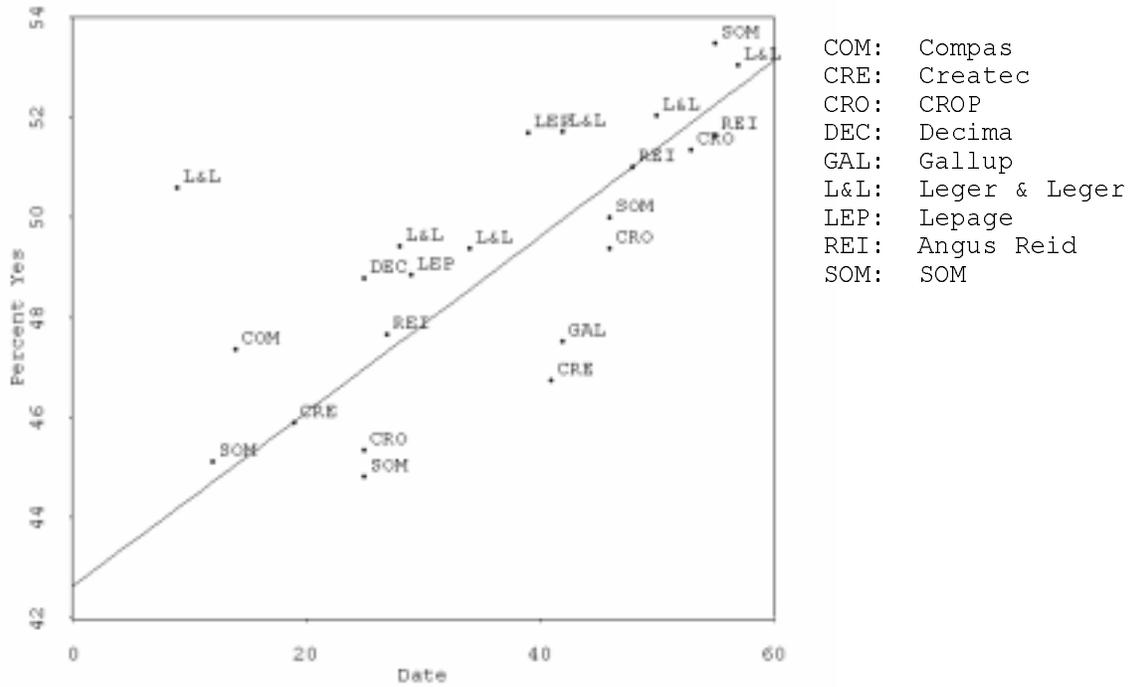


Figure 1: Percentage intending to vote 'yes' among decided voters in 23 pre-referendum polls. Date is number of days after Aug. 31. The three-character codes representing polling firms are given to the right of the graph. The line on the plot is from a logit model (model 1 in Table 2) fit to the data omitting the first Léger and Léger poll.

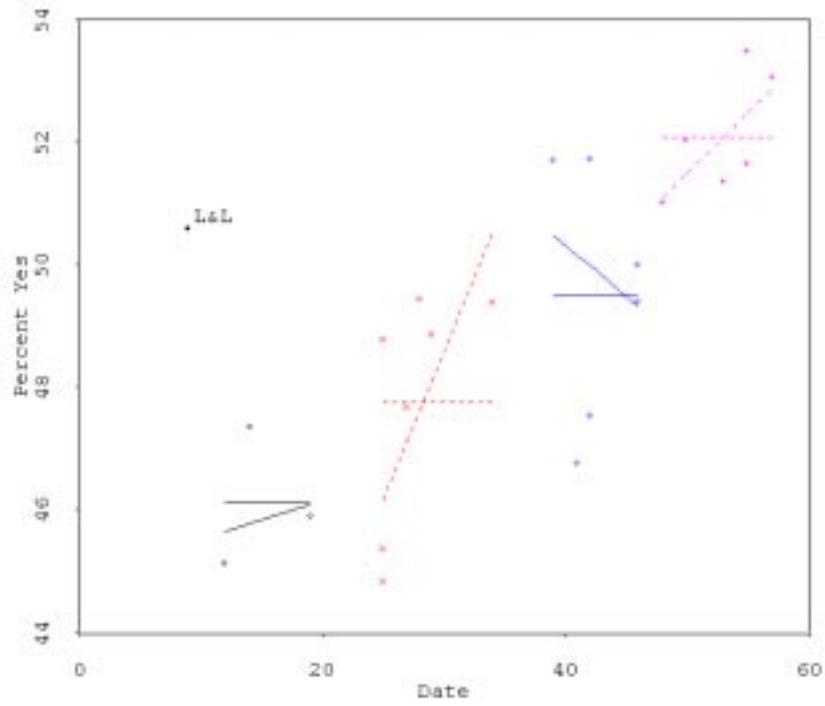


Figure 2: Edouard Cloutier's 'turning points' in the referendum campaign: The horizontal lines represent the mean percentage 'yes' vote for the polls in each of four periods; the other lines show the linear logit-model fit to the polls in each period. Both lines in the first period exclude the September 9 Léger and Léger poll.

