The Polls in the 2002 French Presidential Election : An Autopsy.

By Claire Durand, dept. sociology, Université de Montréal
André Blais, dept. political science, Université de Montréal
Mylène Larochelle, dept. sociology, Université de Montréal

It is not exaggerated to say that the results of the first round of the 2002 French presidential election were a shock to French voters and people in other countries. In the months preceding the election, each and every poll had asked respondents whether they would vote for Chirac or for Jospin, the presumed candidates of the second round. What happened on election day had not been forecast by the polls : Contrary to predictions, the extreme-right leader Le Pen finished second with 16.9% of the vote and moved to the second round. In the famous newspaper Le Monde, one could read "France is hurt. And for many French people, humiliated". (Le Monde, April 23, 2003).

No poll had even hinted that such a result was a possibility. The twelve polls published during the week preceding election day gave an average of 18.0% to Jospin and 12.7% to Le Pen, an average difference of 5.3 in favor of Jospin; each and every poll put Jospin ahead of Le Pen by at least 4 points..

This paper first reviews the various explanations put forward to account for "errors" in the polls.. The paper then discusses the various criteria that have been used to assess the quality of poll estimates and present those that are utilized in this study. The third section examines the French results in order to assess to what point and how polls, collectively or individually, erred in their estimations. We finally compare the French results with those of the US 2000 election in order to conclude on the quality of the criteria used to evaluate the adequacy of polls and we investigate whether the methodology used may explain at least part of the errors in estimation.

Why do polls go wrong?

When polls do not correctly estimate the vote, a number of explanations are proposed. Durand, Blais and Vachon (2001, 2002) synthesized the various explanations proposed. In short, when polls go wrong, the fault is attributed either to respondents or to the methodology used.

Respondents may lie to pollsters; they may also change their mind and decide to vote for a different candidate than they report supporting or decide not to participate in the vote. Jowell et coll., 1993 attribute part of the discrepancy between the polls and the vote in 1992 in England to this situation while Durand et al., 2002 conclude that there was no effect in the Quebec 1998 election.

Respondents who refuse to reveal their vote intention or who say that they do not know for whom they intend to vote may vote differently than those who reveal their intention. There is a

general agreement that such respondents are more generally more likely to be conservative (Curtice, 1998; Jowell, et al, 1993, Durand et al., 2002).

Respondents who refuse to answer polls or are more difficult to reach may also vote differently than poll respondents, inducing bias in the polls. Consequently, polls conducted on a shorter period, where less effort is put into reaching the whole sample and convincing respondents to cooperate may be more likely to have a conservative bias. According to most authors (Lau, 1994; Curtice, 1998; Vachon et al. 1999; Durand et al., 2002), this is a likely possibility.

Finally, sampling itself may induce bias. People who ask that their phone number remains unpublished may have specific characteristics linked to electoral preferences. Similarly, people who live in institutions, mostly old-age pensioners, do not usually have a personal listed phone number so that they cannot be reached by pollsters. Again, most authors (Durand et al., 2002; Curtice, 1998, Jowell et al., 1993) attribute part of the underestimation of the conservative vote in various elections to these types of sampling-related problems.

These explanations are not mutually exclusive and may in fact all be true at the same time. In general, pollsters tend to prefer respondent-based explanations while academics often find that sampling frames and sample management may explain a substantial part of the discrepancy between the polls and the vote.

Though various catastrophes in the history of polls in different countries have brought important self-examination within the polling industry and improvements in the methodology, it seems that some methodologies work fine in most situations but that, in some circumstances, they fail. In short, when polls fail, they all fail in the same direction. We thus have to ask why they sometimes fail and whether it is possible to improve the methodology so that the polls almost never fail.

How do you know that polls went wrong? Assessing the accuracy of the polls

How to ascertain the accuracy of the polls? A related question in the present case is the possibility to have comparable measures of accuracy that would allow to compare the French election with other elections. How is it possible to compare the results of polls in one-round presidential elections (US), in legislative elections where the chief of government is the leader of the party who gets the majority of seats (Canada/Quebec, Great Britain) and in a two-round presidential election. Does the varying number of major candidates hinder comparability?

Jowell et al. (1993) and Curtice et al. (1997) measure the accuracy of the British polls by comparing the estimates of the final polls of the campaign with the vote for each party. Durand et al. (2001, 2002), in the Quebec situation, focussed on the collective performance of pollsters and thus measured accuracy by a) comparing the average estimates of the polls published during the last week and the actual vote and b) comparing the forecast time-series analysis of the campaign polls with the vote. Lau (1994), Vachon et al. (1999) and Durand (2002) measured individual polls' accuracy using the difference between the estimates of each poll and estimates from time-series analysis.

For the U.S. presidential elections, Mitofsky (1998) summarized the methods used by the Social Science Research Council (SSRC) study of the 1948 pre-election polls, the goal being to devise a single measure of accuracy for individual polls published as close to the end of the campaign as possible. Eight different methods had been devised at that time, all with virtues and drawbacks. He chose to retain four methods. Two of them, methods 1 and 2, assess the accuracy of the estimate of the leading party's share of either the total vote or of the two leading parties' vote, while method 3 is based on the sum of the absolute difference between the estimates of voter intent for each candidate and the final vote and method 5 evaluates this same difference but only for the two leading candidates. Mitofsky concludes that methods 3 and 5 give the most unbiased estimates of the accuracy of the polls. A debate on the issue of inclusion of third party estimates and on allocation of undecided followed, in which Panagakis (1999) argued that a modified method 3, including third party candidates (except the other category) and after allocation of undecided, was more appropriate. Traugott (2001) used these same two methods (3 and 5) with and without the third-party candidate Nader, in order to evaluate the polls' accuracy during the November 2000 US presidential election.

In the French situation, a number of problems arise. First, the election is neither a legislative nor a one-round presidential election. Second, the notion of leading candidates and third party candidates does not apply. In the 2002 election, 16 candidates ran in the first round, half of them considered to be on the left and half on the right. Among these, three candidates ended up with between 15% and 20% of the vote (Chirac, right, Jospin, left and Le Pen, extreme-right), four got between 5 and 10% (three from the left, one from the right) and the other nine candidates, four from the left and five from the right got between 0.5% and 4.2%. In such a situation, using method 3, as suggested by Mitofsky (1998) and Traugott (2001), would certainly inflate error. Furthermore, in order to use method 5, one has to decide who are the two leading candidates : those leading in the polls or those who finally lead in the vote? There are in fact three leading candidates.

The second problem lies in using the final vote as the only criteria. The final vote can be used only to assess the quality of the surveys published in the last days of the campaign, and this, only if one can be confident that no substantial movement occurred in voter intent during those last days. In the French situation, there is a general agreement that movement occurred during the campaign, including during the last week. It is then imperative to assess the evolution of vote intentions during the campaign.

In order to assess the evolution of vote intention, Vachon, Durand and Blais (2000) devised a method using time-series analysis inspired by Lau (1994). This same method was used in Durand et al. (2001, 2002) and Durand (2002). For each party, the estimate of voter intent produced by each poll is distributed on the days each poll was conducted. A weighted mean[1] is then computed for each day using all the estimates produced on a given day. Time-series analysis is then used to estimate the evolution of voter intent, taking into account the time-dependency of the measures. These analyses allow to forecast the results on election day and compare it to the

[1] In the present case, an unweighted mean is used since the samples size were all similar (N total = 1000)

real vote.

This paper looks at the polls' performance globally. In order to do this, two methods are used. First, the mean estimates from the polls published during the last week of the campaign for each candidate and for each side (left vs right) are compared to the election results. Second, the evolution of the campaign for the three leading contenders is examined using time-series analysis.

The paper then examines the performance of individual pollsters in order to see if differences in the methodology used may explain part of the discrepancies between the polls and the vote. Again two methods are used. First, the estimates from the last poll of each pollster are compared to the vote for the three leading candidates and second, the difference between the pollsters' estimates throughout the campaign and the evolution as measured by the time-series will be examined.

In order to compare the French situation with the US situation, a comparison of the estimates for the total left-wing and right-wing vote intention is performed; it is considered akin to comparing two leading candidates.

How did the polls fare globally?

Table 1. presents the mean estimation of the surveys published during the last week of the electoral campaign – from April 14 to April 19 – with the results of the vote held on April 21, 2002, this for the 8 candidates considered to be on the left-wing side and for the 8 candidates on the right wing side. If the results for the candidates on each side are added, the comparison shows that overall, the left was overestimated by 5.6 percentage points (48.5% estimated compared to an actual vote of 42.9%). The table shows that, on the left, overestimation is not concentrated, 5 out of 8 candidates being slightly overestimated. The highest overestimation is for Hue, the Communist party candidate (2 points on a vote of 3.4%). On the other hand, on the right, only one candidate, Le Pen, appears grossly underestimated by 4.2 percentage point for a vote of 16.9%. Le Pen is responsible for most of the underestimation of the right.

Table 2 focuses on the three leading candidates. It compares the difference between the estimation of vote intention during the last week of the campaign for Jospin, Chirac and Le Pen and the vote. It clearly shows that, while the difference between Jospin and Chirac is only slightly underestimated, the difference between Chirac and LePen is overestimated by 4 points and the difference between Jospin and LePen by 6 points.

Graph 1 presents the estimation based on time-series analysis of the compared evolution of voter intent for Jospin and Le Pen while Graph 2 presents comparative figures for Chirac and LePen. The time-series analysis allows to forecast the vote by projecting the evolution through election day, April 21. Graph 1 shows that the best estimate of the evolution of vote intention for Jospin is a linear downward trend from 21.2% at the beginning of the official campaign to 17.1 on election day, a total drop of 4.1 percentage points. A concurrent increase in vote intention for Le Pen of 3.4 points, from 10.2% to 13.6% was estimated by the polls. Graph 2 shows a decrease of

3.2 points for Chirac from 22.6% to 19.4%. The time-series correctly forecasts the vote for Jospin (16.2%), within a classical margin of error[2] (represented on the graph by the upper and lower dotted lines) as well as for Chirac, (19.9% of the vote). However the prediction for Le Pen is 3.3 points below his actual score of 16.9, outside the margin of error.

It can thus be concluded that the polls conducted during the presidential campaign erred in their overestimation of the left-wing vote. As for individual candidates, they erred mostly in their prediction of the vote for Le Pen.

How did the pollsters fare individually?

One way to look at the pollsters' performance is to compare their last published poll with the vote, focusing on the three leading candidates. Table 3. shows each pollster's last estimate of the campaign. Two observations can be made. First, as with the average estimates calculated for the last week and the forecast from the time-series, the estimates for Jospin and for Chirac are quite accurate. All the pollsters' estimates for Chirac and Jospin are within a classical margin of error for the sample size and all the estimates for Le Pen are inaccurate, the differences between the estimates and the vote varying from 2.9 to 6.4 points. However, another observation has to be made : five out of six estimates for Jospin are equal at 18%; three estimates for Chirac are at 20% and two at 19.5%.

If we examine the difference between the estimates for the three leading candidates, it remains quite obvious that the problem lies with Le Pen's estimates and more so with the estimates of the difference between Jospin and Le Pen. The difference between Le Pen and Chirac is overestimated by 2 to 6.5 points while the difference between Le Pen and Jospin is overestimated by 4.7 to 6.7 points.

Another way to examine the accuracy of the various pollsters is to compare their estimates during the last month with the estimated evolution from the time-series analysis (see Graph 1). Two estimates are well below the confidence interval of the series for Le Pen. Both these estimates belong to IFOP. Since the series underestimates the vote for Le Pen, these are well beyond the margin of error. On the opposite, three estimates are over the confidence interval which means that they are better than the others. The first of these belongs to IFOP while the two others belong to BVA.

Synthetically then, an examination of the performance of various pollsters shows that they all erred in their estimation of Le Pen's vote but that some were worse than others. It also shows that the estimates are much more similar than what we expect from sampling theory.

Comparing of the 2002 French presidential election and the 2000 US presidential election

---

[2] Though the first step in the sampling procedure used by French pollsters is a random sample, the data collection is quota-based. Therefore, the concept of a margin of error does not strictly apply. It is used here only for comparative purposes.

It is relevant to compare the polls of the two elections. Not only does it give a "bareme" but also, it allows to see if the measures of accuracy used for the US presidential elections are also adequate for other types of presidential elections. A number of differences in the two elections may impact on the measurement (two-round versus one-round elections, number of candidates, proportion of the vote held by leading candidates, etc.).

The last column of table 3 presents the average absolute candidate deviation for all the candidates who get more than 15% of the vote, i.e. the three leading candidates. The problem is that the sum of the vote for the three candidates equals 53% and not 100%. If the proportions would be recalculated so that the total equals 100%, it would artificially inflate the measure of error. Since the increase in the margin of error is not linear, a 2% error on an estimate of 20% cannot be doubled to a 4% error on an estimate of 40%.

A way to get round the problem of comparability is to evaluate how well both sides (left vs right) were evaluated. We end up with the equivalent of two candidates. In such a situation Method 3 equals half of method 5. Therefore, Table 4 only presents the results for the absolute difference in the differences (akin to method 5). In order to have a comparable figure, we use the deviation in the assessment of the candidates grouped according to their ideological family (right or left). The mean absolute difference is 10.9, compared to 3.5 for the US 2000 election. Table 4 also allows to see that if ideological families are divided into extreme and traditional (which is usual in the presentation of results in the French media), the misestimation is concentrated in two groups : the extreme-right (minus 4%) and the traditional left (4.8%).

We conclude that using method 3 a) is possibly not appropriate when the situation is not one of two leading candidates and b) tends to minimize bias. Averaging deviations has allowed to make the problem of the misestimation of Le Pen disappear while, in this instance, it was the main problem. In order to assess the global quality of the estimates, it appears relevant to examine also the proportion of polls who deviate in the same direction for a given candidate. In the US 2000 election, 16 out of 19 polls underestimated Gore. In the French election, all the 6 polls underestimated Le Pen.

When the estimates grouped according to ideological family are used, the measure of poll adequacy, whether using Method 5 or 3, portrays the French polls of the 2002 election as much worse than those of the US election.

Why did the polls go wrong in their estimation of Le Pen?

Two groups of reasons are given for erroneous poll estimates, one related to respondents, another related to sampling and coverage. In order to assess the plausibility of the various explanations, we rely on the methodological information that is provided upon publication either in the media or to the *Commission des sondages*, an instance where pollsters are requested by law to file all the relevant methodological information. Table 5 presents the relevant information regarding the methodology used by French pollsters.

This information may be examined in order to assess whether all the appropriate and reasonable measures have been taken in order to avoid bias as much as possible.

Table 1

Comparison between the polls published during the last week of the campaign and election results

Left-wing candidates									
	Jospin	**Laguiller**	**Chevène-ment**	**Mamère**	**Besan-cenot**	**Hue**	**Taubira**	**Gluck-stein**	**Total left**
Election	16,2	5,7	5,3	5,2	4,3	3,4	2,3	0,5	**42. 9**
Polls (n=12, N=8134[3])	18	7,6	6,6	6	2,9	5,4	1,5	0,5	**48,5**
Difference	**1,8**	**1,9**	**1,3**	**0,8**	**-1,4**	**2**	**-0,8**	**0**	**5,6**
Right-wing candidates									
	Chirac	**Le Pen**	**Bayrou**	**Saint-Josse**	**Madelin**	**Mégret**	**Lepage**	**Boutin**	**Total right**
Election	19,9	16,9	6,8	4,2	3,9	2,3	1,9	1,2	**57,1**
Polls (n=12; N=8134)	20	12,7	5,9	3,7	4,3	2,4	1,4	1,1	**51,5**
Difference	**0,1**	**-4,2**	**-0,8**	**-0,5**	**0,4**	**0,1**	**-0,5**	**-0,1**	**-5,6**

Table 2

Average difference between the three leading candidates

	Chirac vs Jospin	**Chirac vs Le Pen**	**Jospin vs Le Pen**
Election	3.7	3.0	-0.7
Average estimates - polls published during the last week	2.0	7.3	5.3
Difference polls- vote	**-1.7**	**4.3**	**6.0**

[3] The total sample is calculated as....

Table 3
Estimation of individual pollsters- last poll - three leading candidates

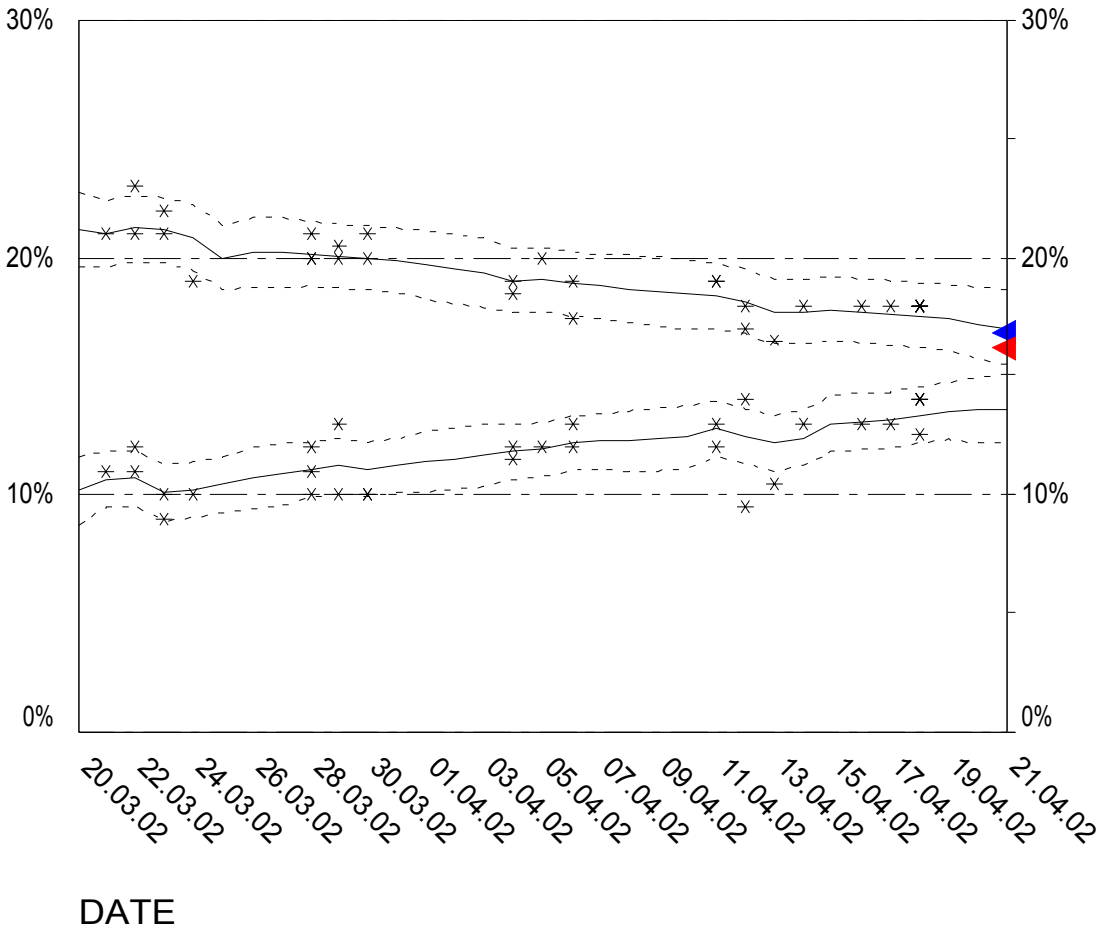
	Chirac		***Jospin***		***Le Pen***		***Average absolute difference (Method 3)***
		Difference		Difference		Difference	
Election	**19,9**		**16,2**		**16,7**		
BVA	19	-0,9	18	1.8	14.0	-2.9	1,9
CSA	19,5	-0,4	18.0	1.8	14.0	-2.9	1.7
IFOP	20	0,1	16.5	0.3	10.5	-6.4	2.3
IPSOS	20	0,1	18.0	1.8	14.0	-2.9	1.6
LOUIS-HARRIS SOFRES	20 19,5	0,1 -0,4	18.0 18.0	1.8 1.8	13.0 12.5	-3.9 -4.4	1.9 2.2
Mean	19,7	-0.2	17.8	1.6	13	-3.9	1.9

Table 4
Estimation of individual pollsters - Last Poll - Total Left versus Right

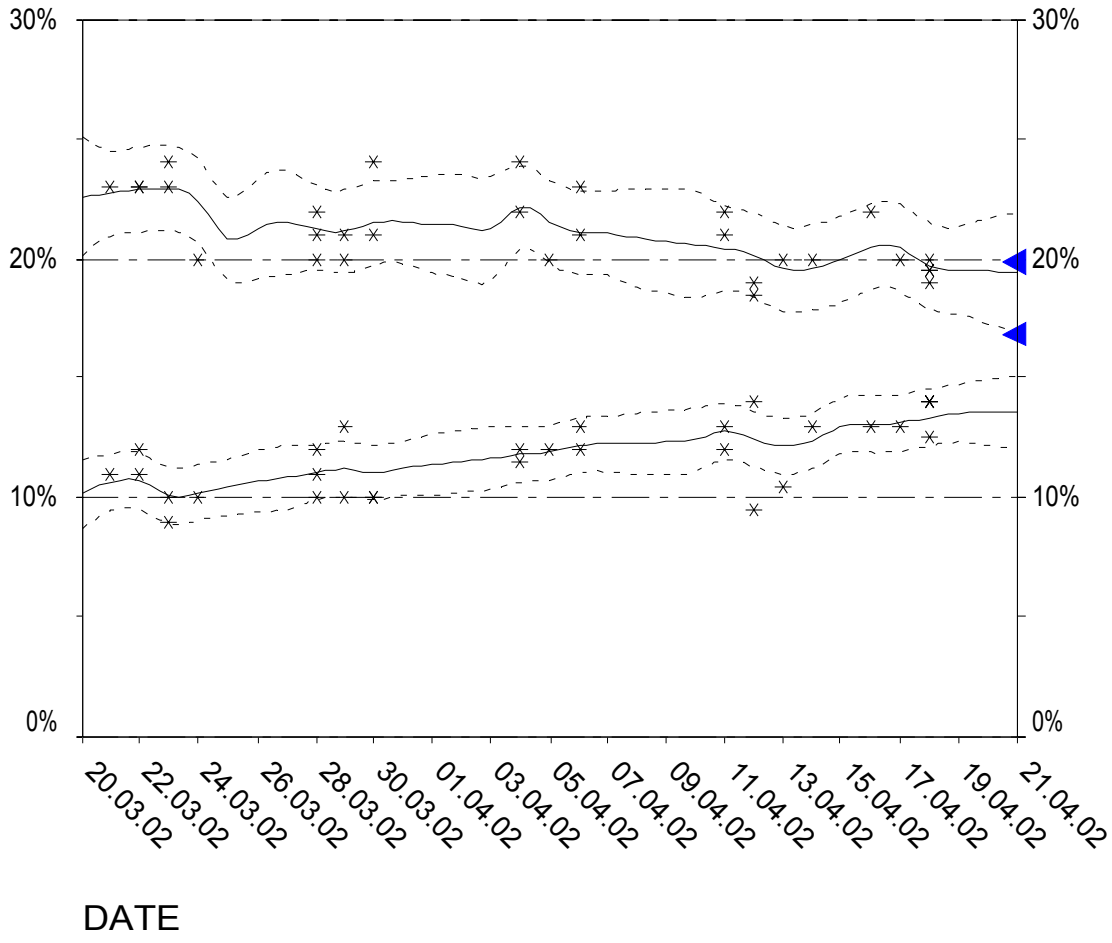
Vote	***Total Right***	***Total Left***	***Absolute difference in the differences (M5)***	***Trad. Right***	***Extreme Right***	***Trad. Left***	***Extreme Left***
Election	**57,1**	**42.9**		**37.9**	**19.2**	**32.4**	**10.4**
BVA	52.0	48.0	10.2	36.0	16.0	36.0	12.0
CSA	52,5	47.5	9.2	36.0	16.5	37.0	10.5
IFOP	50	50.0	14.2	36.5	13,5	38.5	11.5
IPSOS	51,5	48.5	11.2	35.0	16.5	37.0	11.5
LOUIS-HARRIS SOFRES	53 51	47.0 49.0	8.2 12.2	38.0 37.0	15.0 14.0	37.0 38.0	10.0 11.0
Mean	**51.7**	**48.3**	**10.9**	**36.4**	**15,2**	**37.2**	**11.1**

Table 5 Methodological information

	France - April	Other
Average length of surveys	2.2 days	
Sampling and	Quota-based : sex age socio-professional status of the household chief, region, category	
Weighting		
Referent population		
Sample Size		
Question on vote intention		
Non disclosers - proportion		
- attribution of vote		
Error of estimates		



Graph 1 : Evolution of vote intention for Jospin and Le Pen



Graph 2 : Evolution of vote intention for Chirac and Le Pen