



Estimating Marine Le Pen's vote in the 2012 presidentials: an experiment

Who knows better than a political expert what the outcome of an election will be? Perhaps a roomful of political experts. Not individually, of course – their views are no more or less likely to be accurate than the single expert. But, take the average position of the roomful of experts, and many of their inaccuracies should cancel out. Those dreaming of a landslide victory will be counterbalanced by others thinking mournfully of unfulfilled promise. Those of a more objective bent will still fall short in roughly similar numbers to those who overshoot by a margin. Thus is the statistical rule of so-called 'normal distributions', if our room is big enough, and more colloquially what we think of as a law of averages.

This method has a wide range of potential applications, of course. Check out the old US artillery guide to range-finding, for example. Nothing to do with use of range-finding equipment, and everything to do with a battalion commander treating his men like a roomful of experts, allowing for their foibles and inaccuracies, and averaging their guessed distances to target with good effect.¹

We decided to test this theory on a 'hard case', and so via e-mail asked around 100 experts on French elections and Extreme Right parties to give, anonymously, their own prediction of perhaps one of the most speculated-upon results in French electoral history: the first-round presidential vote for the *Front National* candidate, Marine Le Pen. We asked them to answer the following question:

“What is your own estimate, today, of Marine Le Pen's score in the first round of the Presidential election in April 2012?”

57 did so. Individual commentators in the past have proved remarkably incapable of forecasting the vote for her father, Jean Marie Le Pen, in previous presidential races, and actually the law of averages wouldn't have been much help either. In 2002, most people went low. In 2007, the majority went high. Popularity polls have always overestimated his electoral appeal; voting intentions have revealed – well, we don't really know what voting intentions have revealed, as some polling institutes, stung by previous failures in getting even ball-park close to what Jean-Marie Le Pen would poll, took to 'adjusting' their polls to what 'should' be the outcome.²

¹ Hofstee, W. and H. Schaapman (1990) “Bets beat polls: averaged predictions of election outcomes”, *Acta Politica* 25(3): 194-207.

² One polling expert recently noted that the corrections applied to Marine Le Pen's estimates are now similar to those of other candidates, given people's willingness to state they would vote for her. A very different scenario to her father's. < http://www.lemonde.fr/election-presidentielle-2012/article/2011/11/19/les-electeurs-fn-sont-plus-nombreux-a-assumer-leur-vote_1606334_1471069.html >

So, what was their average prediction? As near as makes no difference, 17%.

In one sense, the full results (reported below) surprised us. Any expectation that the experts would agree on a likely score, or even on a small spread of scores, were clearly wrong: the range of estimates covered 20%, from 10% at the minimum to 30% at the maximum. Even at five months' distance from the election, there is a remarkable lack of agreement. Still, 17% is not revelatory – it isn't as high as many estimates currently in circulation, but equally it is not the lowest. It is – about average.

We promised our experts anonymity, and only asked them to answer the one question (experts get asked to complete lengthy surveys far too often). The only other data we collected was the time they submitted their answer. That allowed us to run one more test – was there any difference in their estimates according to when they replied? There are many reasons why some might have replied later than others, most of them uninteresting (hadn't checked e-mail, forgot to reply, etc.). But those who responded more quickly, other things being equal, had less time to think about their response, or to look for external pointers and validation before replying. We deliberately asked them to give their estimate 'today' (9 November) not only because we wanted a relatively rapid response to make the data timely in the pre-campaign period, but also precisely because we didn't want our experts to think too hard about their answer, or spend time researching it.

28 of the valid responses came within 'today', i.e. 24 hours, and 28 after. In terms of their estimate, the difference was negligible – a measly 0.14% higher for the early respondents. Moreover, those that took their time to respond – for whatever reason – had a much larger standard deviation. If they were mulling over their response, or searching for information to help, it was not bringing them any closer together in their prediction. Even when we broke the respondents down into three groups – those that replied within the first hour, those between 1 and 24 hours, and those who replied later, there were no statistically significant differences³, although the very early respondents estimated a higher score of 18.14%. Overall, then, we can't say that our 'slower' experts were thinking it over. But those who didn't think it over were no more inconsistent – quite the reverse.

There is another interpretation which, as election forecasters, we find somewhat discomfiting, namely that there is a whole range of scores which we cannot yet identify as 'wrong' or 'right', covering a wide array of eventualities. Whilst we would assume our experts know, what we assume they know is context, issues, trends – but they don't know the future. As the histogram in the appendix shows, within the bounds of 10%-30% the distribution of the scores cluster towards the centre, diminish towards the tails. If we permit ourselves one data adjustment, namely exclusion of the 30% prediction – a score almost double that of Jean-Marie Le Pen in

³ Given the relatively small number of respondents, this is not that surprising.

2002, and triple that of 2007 – the data show no significant skewness.⁴ Run random number generators using a normal distribution with the same mean and standard deviation, and for the same number of scores, and the results are noticeably similar. So, overall our experts may well have no better chance of getting the right answer than a set of random picks, beyond knowing roughly the range of scores the more successful presidential candidates normally get in France. There are too many possible outcomes.

On election day, one (or a few) of our experts are likely to be right.⁵ Whether this will be through good luck or good judgement is largely in the eye of the beholder. But the majority will be wrong, a number by 1 or 2%, and many more by much larger margins. In fairness to our experts, whilst the divergence in forecast may be unusually high, there are good reasons for thinking that many of these predictions could bear out. Campaigns change elections. Many of our experts may be wrong for the right reasons. Marine Le Pen could score a range, if not all, of these values. The inability to forecast precisely every time is one of the reasons we still hold elections.

17% would constitute a good performance by Marine Le Pen, but no better. In 2002, it would have been sufficient to ensure her progression to the second round of the elections. But, in 2002, the political context was different. The Left of the spectrum in particular was highly fragmented, each additional candidate eating in to the Socialist candidate, Lionel Jospin's, vote. Abstention by Left-wing supporters was also high, inflating the weight of each valid vote. Finally, so soon after the split with Bruno Mégret and his *Mouvement National Républicain*, almost no-one gave Jean-Marie Le Pen a chance of even a good performance. Five years later, in 2007, 17% would have placed the FN candidate fourth, some nine points off second place.

A very good performance for Marine Le Pen would be the 19.2% which Le Pen père and Mégret garnered between them in 2002. By that standard, she could claim to have revived and reunited a divided Far Right. An outstanding performance, and one she would likely need to guarantee the progression into the second round many were considering likely not six months ago, would be over 20%. Ten of our experts believe a score of this magnitude to be likely (though we cannot tell whether they think she will make the second round). But another ten don't think she will make the 15% mark, let alone the run-off.

⁴ If we exclude the two highest and two lowest scores, the distribution of scores shows almost no skew whatsoever, and the mean remains relatively stable, at 16.9%.

⁵ If we use the modal response as our predictor, Marine Le Pen will score 18%.

17%, then, is the average expert estimate. In many ways, this is a middling score – middle of the estimates, middling performance from, presumably, a relatively good campaign. Average, as in “nothing special”. We will find out soon enough how well the theory of averages works for hard cases in practice. Just finally, though, let us give ourselves a salutary reminder of a similar experiment in 2000 which asked 50 US political scientists who would win the US presidential elections.⁶

49 said – “Al Gore”.

Jocelyn Evans
(University of Salford)
J.A.Evans@salford.ac.uk

Gilles Ivaldi
(University of Nice Sophia Antipolis)
gilles.ivaldi@unice.fr

⁶ Michael Lewis-Beck and Jean-Dominique Lafay "Qui sera le nouveau Président des Etats-Unis?", *Le Figaro*, 20 October 2000, p. 13.

Appendix

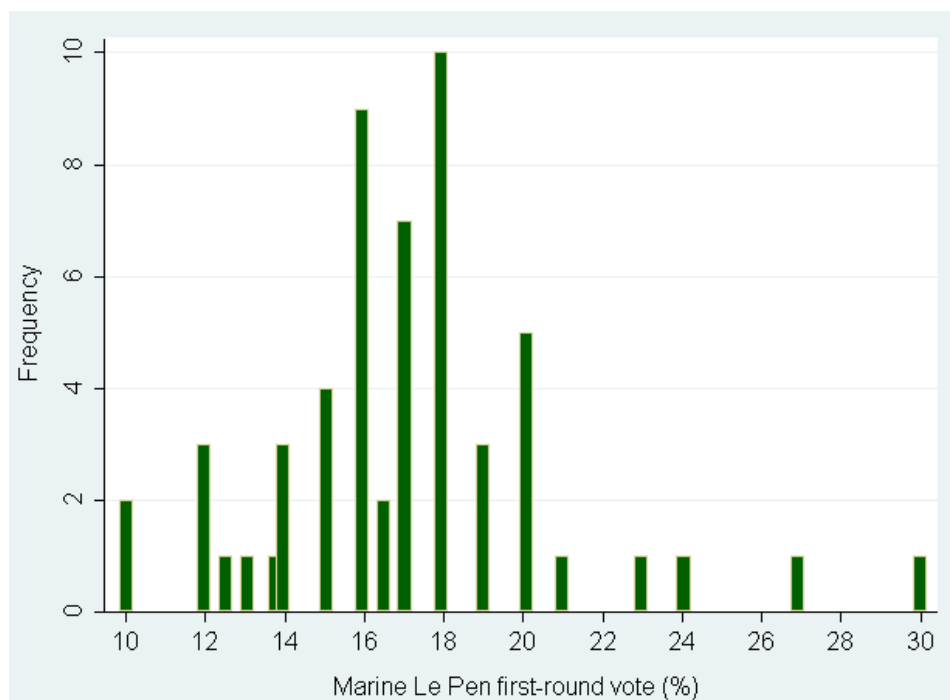
Note: full dataset is available at <http://www.pratiquesciencesociales.net/forecast/>

Question

“What is your own estimate, today, of Marine Le Pen's score in the first round of the Presidential election in April 2012?”

Number of responses: 57

Number of valid responses: 56 [one respondent gave vote range of 15-25%]



Mean	17.06	Standard deviation	3.59
Median	17.00	Percentiles	
Mode	18.00	25th	15.00
		50th	17.00
		75th	18.00

Skewness (s.e.) 1.02 (0.32) *Excluding maximum* 0.43 (0.32)

Response-time subset

	Mean	Standard deviation
'Today'	17.13	2.90
More than 24 hours	16.99	4.22
Within an hour	18.14	2.78
Over an hour, under a day	16.47	2.87