

FORECAST ERROR: 2022 FRENCH PRESIDENTIAL ELECTION PREDICTORS

By Timothy Martyn Hill

PART 0: TAGLINE

In 2022 Emmanuel Macron, the President of the French Republic, sought re-election against opponents of various political opinions. The challenge for Macron was to win another five-year term against a background of post-Brexit readjustment and the Ukraine crisis. The challenge for the author was to see whether techniques adapted to Anglo-American partisan elections would cope with a Francophone election with different traditions. Here Timothy Martyn Hill reviews the predictions - and the errors - that were made.

PART 1: THE ELECTION

All politics is local and temporal. You can discern basic themes – here is an election, there are candidates, there are votes which are counted and there is a winner – but in every case there are quirks and features – what are the rules, where is *here*, how are things *organised* exactly? For this election the post is President, the system of government is the Fifth French Republic and the sovereign state is France, which includes the French mainland and its overseas states and territories.

In order to become a candidate for the election, an individual must get signed nominations known as *parrainages* or sponsorship from more than 500 elected local officials (known as *parrains* or sponsors), mostly mayors. The candidates are usually members of political parties, although those parties may have been formed around that individual for election purposes. A popular vote is held across France and the candidates with the top two popular votes go thru to a runoff election some weeks later. The winner of that runoff election is the one with the most popular votes. The two major differences to US Presidential elections are the shorter life-spans of the parties and the two-stage nature of the election.

During this process there are many predictors: modellers had predicted, bookies had taken bets, pollsters had polled. Which of them had predicted the outcome and how far out had they done so? This article sets out to answer that question, by analysing the performance of pollsters, seat and vote modellers, and betting firms all the way up to round two election day 2022.

PART 2: WHICH DATES SHALL WE USE?

To examine predictors we must know what they were at the time. But this begs the question: at what time exactly? At the time of writing it is March 6th 2022, several weeks before round one of 2022, so we don't know the predictors at the time of round one. But we do know the predictors on/immediately before March 1st – ie the first day of the month before round one. Even better, in many cases we know that datum for past elections as well. So we will nominate the first day of the month, the day before round one, and the day before round two as our initial datum dates, with others introduced on an ad-hoc basis

The datum dates we will select are given below. We will capture our predictors at these dates.

- First datum date: the first day of the month before round one: 01 March 2022
- Second datum date: the day on which the official round one candidates were selected: 07 March 2022
- Third datum date: ad-hoc: 28 March 2022
- Fourth datum date: ad-hoc: 05 April 2022
- Fifth datum date: the day before round one election: 09 April 2022
- Sixth datum date: ad-hoc: 17 April 2022
- Seventh datum date: the day before round two election: 23 April 2022

PART 3: ASSESSMENT

In line with our previous articles we will use two metrics: Mean Absolute Error (MAE) and WIN (a metric that scores a prediction on whether or not it predicted the eventual winner). But the shorter life span of French elections poses a problem: how can we use an X-party-forced prediction metric if parties are not stable from election to election? To see the problem, consider this list.

Parties scoring 5% or over

Below we see the parties scoring 5% or over from each presidential election in the Fifth Republic since 1958.

Table: parties scoring 5% or over in round one in presidential elections in the Fifth Republic[0502a]

| Pos'n | 1958[2] | 1965 | 1969 | 1974 | 1981 | 1988 | 1995 | 2002 | 2009 | 2012 | 2017 |
|--------|---------|------|------|------|------|------|------|------|------|-------|------|
| 1st[1] | UNR, | UNR, | UDR, | PS, | UDF, | PS, | PS, | RPR, | UMP, | PS, | EM, |
| 2nd | PCF, | CIR, | CD, | RI, | PS, | RPR, | RPR, | FN, | PS, | UMP, | FN, |
| 3rd | DVG | MRP, | PCF, | UDR | RPR, | UDF, | RPR, | PS, | UDF, | FN, | LR |
| 4th | | CTV | SFIO | | PCF | FN, | FN, | UDF, | FN, | FG, | LFI |
| 5th | | | | | | PCF | PCF, | LO, | | MoDem | PS |
| 6th | | | | | | | LO | MDC, | | | |
| 7th | | | | | | | | LV | | | |

- [1] Parties are in order of popular vote in round one. In 1981 the party that won round one lost in round two
- [2] The 1958 election was the first of the Fifth Republic and used an electoral college and one round

The colours are chosen by the author in an attempt to categorise the parties into political families (PCF, FG are Communists, VO, LO are Trotskyists, grey is unsure, and so on). But we are falling into a trap here: we are attempting to place an overarching long-term structure on a country that just hasn't got one, because France has parties with shorter lifespans than we are used to. Even worse, we are doing it from the author's best guesses which while defensible are still a subjective interpretation and might be wrong - where would one put UDF? En Marche? Are you sure? So, we need a better solution. But what would that solution be?

Chosen method of assessment

If we can't reliably do it by political *party* or by political *strand*, then the only thing remaining is by political *candidate*. So instead of "X-party-forced" we will use "X-person-forced", and so on. Specifically, we will select the top five candidates on datum date one and use five-person-unforced (Individual A/Individual B/Individual C/Individual C/Individual D/Individual E) to assess accuracy. For a fuller discussion of X-party forced format, our metrics and our choice of predictors, see our article on the 2016 election here[0313f].

Following our previous practice we will also use the WIN indicator: 1 if predicted the winner, 0 if not.

Threshold

The question now arises as to the threshold: at what point do we deem a prediction to be a failure? After consideration (see Appendix 3) we settled on a MAE threshold of 2% for five-person-unforced. For the WIN indicator it was simple: if a majority of predictors accurately produced the winner then it was a success.

PART 4: WHICH CANDIDATES WILL WE FOLLOW?

The five people we will select will be the top five in the polls on the first datum date. They are:

- Emmanuel Macron (En Marche!)
- Valerie Pécresse (The Republicans)
- Marine LePen (National Rally)
- Jean-Luc Mélenchon (La France insoumise)
- Eric Zemmour (Reconquete)

They went on to exceed the 500 parrainages threshold by the deadline of March 4th 2022 and so were on the list of twelve official candidates published on March 7th 2022.[0313b]

PART 5: WHICH RESULTS SHALL WE MEASURE PREDICTIONS AGAINST?

As regular readers of the “Forecast Error” series will remember, an oft-overlooked problem is working out what the final result actually *was*. Initial public commentary is based on ad-hoc data collected by newspapers or television channels. But the actual final results are usually published later by an official organisation, and these final results may differ slightly to the contemporary media reports. The differences are usually slight but it’s sometimes enough to make the error rate change a decimal point.

For the 2022 French Presidential election, we will use the following sources

- Initial media reports: France24 (France), the Guardian (UK). For 2017 examples, see[0313d]
- Final results: Constitutional Council of France. For 2017 examples, see[0313e]

PART 6: WHICH PREDICTORS SHALL WE USE?

The predictors we will use will be nationwide opinion polls, betting odds and spreads, and models. We discuss each type below

Nationwide Opinion Polls

Unlike British elections where poll data can be obtained relatively easily, it may not be so easy for elections in other countries and France is no exception. So we are reduced to getting polling data from Wikipedia, which does not cover opinion polling for elections before 1995. The opinion polls we can easily access are given below.

Table: opinion pollsters polling on presidential elections in the Fifth Republic at datum date 1[1]

| Category | 1995 | 2002 | 2007 | 2012 | 2017 | 2022 | Source | See also | |
|-------------------|------|--------------|---------------|--------------------|--------------------|--------------------|---------|----------|---------|
| Opinion pollsters | | BVA | BVA | BVA | BVA | BVA | [0502b] | [0601a] | |
| | | CSA | CSA | CSA | | | [0502b] | [0601b] | |
| | | | | | | Cluster17 | [0502b] | [0601c] | |
| | | | | | Dedicated Research | | [0502b] | [0601d] | |
| | | | | | Elabe | Elabe | [0502b] | [0601e] | |
| | | | | Harris-Interactive | Harris-Interactive | Harris-Interactive | [0502b] | [0601f] | |
| | | Ifop | Ifop | Ifop | | | [0502b] | [0601g] | |
| | | | Ifop-Fiducial | Ifop-Fiducial | Ifop-Fiducial | Ifop-Fiducial | [0502b] | [0601h] | |
| | | Ipsos | Ipsos | Ipsos | Ipsos | Ipsos | [0502b] | [0601i] | |
| | | Louis Harris | LH2 | LH2 | | | [0502b] | [0601j] | |
| | | Sofres | Sofres | TNS Sofres | TNS Sofres | Kantar Sofres | Kantar | [0502b] | [0601k] |
| | | | | | | Odoxa | Odoxa | [0502b] | [0601l] |
| | | | | OpinionWay | OpinionWay | OpinionWay | [0502b] | [0601m] | |

[1] Datum date 1 is the first day of month before round one (eg if round one is in April, it’s March 1)

We selected the following pollsters to retain continuity and give us the greatest choice (six in case one dropped out during the campaign). Note that all are online polls, although that is not why they were selected.

- BVA (an Online poll)
- Elabe (an Online poll)
- Harris-Interactive (an Online poll)
- Ifop-Fiducial (an Online poll)
- Ipsos (an Online poll)
- Opinion Way, now OpinionWay-Kéa (an Online poll)

Betting Odds And Spreads

We do not have historical data for French election odds and spreads, but we do have access to 2022 odds via Oddschecker (<https://www.oddschecker.com/>) Since Ladbrokes and Coral have rejoined Oddschecker after a hiatus we will not need to augment Oddschecker with Ladbrokes and Coral data.

Table: list of bookmakers for this presidential election at datum date 1[1]

| Bookmaker | Predictions | Link |
|--------------------|---|---------|
| 10Bet | Probability of winning round two | [0326b] |
| 888sport | Probability of winning round two | [0326b] |
| Bet Victor | Probability of winning round two | [0326b] |
| Betfair Exchange | Probability of winning round two, Probability of being the winning party | [0326b] |
| Betfair Sportsbook | Probability of winning round two | [0326b] |
| Betfred | Probability of winning round two, Probability of top two finish | [0326b] |
| Betway | Probability of winning round two, Probability of being the winning party | [0326b] |
| Boyle Sports | Probability of winning round two | [0326b] |
| Coral | Probability of winning round two, Probability of top two finish | [0326b] |
| Ladbrokes | Probability of winning round two, Probability of top two finish | [0326b] |
| Mansion Bet | Probability of winning round two | [0326b] |
| Matchbook | Probability of winning round two | [0326b] |
| Paddy Power | Probability of winning round two | [0326b] |
| Parimatch | Probability of winning round two | [0326b] |
| QuinnBet | Probability of winning round two | [0326b] |
| SBK | Probability of winning round two, Probability of top two finish, Probability of winning round one | [0326b] |
| Skybet | Probability of winning round two | [0326b] |
| Smarmets | Probability of winning round two, Probability of top two finish, Probability of winning round one | [0326b] |
| Sport Nation | Probability of winning round two | [0326b] |
| Sporting Index | Probability of winning round two | [0326b] |
| Spreadex | Probability of winning round two | [0326b] |
| Unibet | Probability of winning round two | [0326b] |
| VBet | Probability of winning round two | [0326b] |
| William Hill | Probability of winning round two | [0326b] |

[1] Datum date 1 is the first day of month before round one (eg if round one is in April, it's March 1)

We will select the five bookmakers that will enable us to track the odds of being the next president and the odds of top two finish

- Coral: high-street fixed-odds bookmaker. Merged with Ladbrokes
- Betfred: high-street fixed-odds bookmaker
- SBK. mobile only sportsbook established by Smarmets in 2019[0313a].
- Ladbrokes: high-street fixed-odds bookmaker.
- Smarmets: online betting exchange.

Modellers And Other Predictors: academic predictions

The political science journal "PS: Political Science & Politics" from the American Political Science Association invited papers as part of its symposium entitled "Symposium: Forecasting the 2022 French Presidential Election", similar to the one in 2021 for the German federal election[0313h]. The papers[0515a] are as follows:

Table: list of academic predictive papers for this presidential election at datum dates

| Paper | Authors | Predictions | Link |
|--|--|--|---------|
| Citizen Forecasting: The 2022 French Presidential Elections | Yannick Dufresne; Bruno Jérôme; Michael S. Lewis-Beck; Andreas Murr; Justin Savoie, 2022 | %age of round one vote | [0326d] |
| Combining forecasts for the 2022 French presidential election: The PollyVote | Graefe, Andreas, 2022 | %age of round one vote %age of round two vote | [0326e] |
| Forecasting the 2022 French Presidential Election with a SUR Regionalized Model | Jerome-Speziari, Veronique, 2022 | %age of round one vote %age of round two vote | [0326f] |
| Designing Prediction Markets to Forecast Multi-Stage Elections: The 2022 French Presidential Election | Berg, Joyce E.; Gruca, Thomas S.; Rietz, Thomas A., 2022 | None: insufficient data | [0326g] |
| Forecasting the 2022 French Presidential Election: From a Left-Right Logic to the Quadripolarization of Politics | Jerome, Bruno, 2022 | %age of round one vote %age of round two vote | [0326h] |
| Which Historical Forecast Model Performs Best? An Analysis of 1965-2017 French Presidential Elections | Bélanger, Éric; Feitosa, Fernando; Turgeon, Mathieu, 2022 | %age of round one vote %age of round two vote | [0326i] |
| Party's rating and electoral forecasting: the case of French Presidential in 2022 | Facchini, Francois, 2022 | %age of round one vote %age of round two vote | [0326j] |

By their nature the academic models are more static than dynamic: requiring consensus, in-person feedback and write-ups they tend to make single “one-and-done” or monthly predictions rather than weekly, daily or moment-to-moment ones like other predictors. This is both an advantage – if you only make one prediction it’s easy to tell if you were right – and a disadvantage – things may change between your prediction and election. We selected the five models that we could most easily extract the predictions for, with a bias towards those that updated most often.

That gave us the following five selections:

- Dufresne et al: Citizen forecasts
- Dufresne et al: Vote Intentions
- PollyVote/Graefe: expert panel
- PollyVote/Graefe: SUR regional or Jérôme-Speziari & Bélanger: SUR regional
- PollyVote/Graefe via Graef via Bélanger: Iowa model

Modellers And Other Predictors: non-academic predictions

Other predictors offering on the election include media predictions (usually polls-of-polls or poll-based models), prediction markets, or the recent introduction of social media interrogation. We found the following:

Table: list of other predictors for this presidential election at the datum dates

| Predictor | Type | Predictions | Link |
|------------------|----------------------------|--|---------|
| Economist: | Poll-based model | Probability of top two finish, Probability of winning round two | [0327a] |
| IEM | Prediction market | Probability of getting over 50% in round one, Probability of top two finish Probability of winning round two*1 | [0327b] |
| Politico | Poll of polls | %age of round one vote %age of round two vote | [0327c] |
| PredictIt | Prediction market | Probability of top two finish, Probability of winning round two | [0327d] |
| Hypermind | Prediction market | Probability of top two finish, Probability of winning round two %age of round one vote %age of round two vote | [0327e] |
| crowdwisdom.live | Google Search Trends | Google Search Share | [0327f] |
| crowdwisdom.live | Social Media Net Sentiment | Social Media Net Sentiment | [0327f] |

*1 IEM could not provide sensible data for probability of winning round two: too few options offered and too few bets laid. It was therefore omitted from our numbers for probability of winning round two

We will select the five non-academic predictors that we can most easily extract data from as follows:

- Economist, a poll-based model
- IEM, a prediction market
- Politico, a poll of polls
- PredictIt, a prediction market
- Hypermind, a prediction market

PART 7: WHAT DID OUR PREDICTORS PREDICT BEFORE ROUND ONE?

Having decided on our predictors, we now need to record how well they did prior to round one

Vote shares in round one per candidate

We see below the values over time for vote shares in round one.

Table: Average pollster vote shares in round one per candidate at datum dates 1-5 [0502d]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 |
|-----------|---------|--------|--------|--------|--------|--------|
| LePen | Polling | 17.17% | 17.17% | 19.58% | 21.92% | 23.42% |
| Macron | Polling | 25.75% | 28.50% | 28.00% | 27.33% | 26.25% |
| Mélenchon | Polling | 11.42% | 11.83% | 14.08% | 15.50% | 17.42% |
| Pécresse | Polling | 12.83% | 12.50% | 10.67% | 9.17% | 8.50% |
| Zemmour | Polling | 14.25% | 12.83% | 10.75% | 9.58% | 8.92% |

Table: Average non-academic vote shares in round one per candidate at datum dates 1-5 [0508a]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 |
|-----------|--------------|--------|--------|--------|--------|--------|
| LePen | Non-academic | 16.55% | 16.60% | 17.00% | 20.00% | 21.15% |
| Macron | Non-academic | 25.70% | 27.45% | 27.30% | 26.65% | 26.15% |
| Mélenchon | Non-academic | 11.95% | 12.25% | 13.70% | 14.90% | 16.05% |
| Pécresse | Non-academic | 13.20% | 12.65% | 11.85% | 10.60% | 9.60% |
| Zemmour | Non-academic | 14.40% | 12.85% | 11.70% | 10.90% | 10.00% |

Table: Average academic vote shares in round one per candidate at datum dates 1-5 [0508d]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 |
|-----------|----------|--------|-------|--------|-------|-------|
| LePen | Academic | 17.96% | ditto | 18.93% | ditto | ditto |
| Macron | Academic | 23.88% | ditto | 25.91% | ditto | ditto |
| Mélenchon | Academic | 11.06% | ditto | 11.22% | ditto | ditto |
| Pécresse | Academic | 14.96% | ditto | 18.62% | ditto | ditto |
| Zemmour | Academic | 10.22% | ditto | 7.98% | ditto | ditto |

Academic predictions were initially made and uploaded in January or February and then preprints were updated or March 23rd 2022. Most were the same but some were changed slightly, hence the change between datum date 1 and 3.

Probability of a top two finish

We see below the values over time for the probability of a top two finish.

Table: Average bookmakers probability of a top two finish per candidate at datum dates 1-5 [0508e]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 |
|-----------|-----------|--------|--------|--------|--------|--------|
| LePen | Bookmaker | 52.22% | 70.77% | 87.81% | 93.08% | 97.70% |
| Macron | Bookmaker | 95.99% | 97.87% | 99.01% | 99.01% | 99.01% |
| Mélenchon | Bookmaker | 8.40% | 9.66% | 12.21% | 8.56% | 5.99% |
| Pécresse | Bookmaker | 31.57% | 14.85% | 3.45% | 1.71% | 1.17% |
| Zemmour | Bookmaker | 29.28% | 23.88% | 9.38% | 6.16% | 3.92% |

Table: Average non-academic probability of a top two finish per candidate at datum dates 1-5 [0508f]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 |
|-----------|--------------|--------|--------|--------|--------|--------|
| LePen | Non-academic | 46.33% | 62.33% | 80.67% | 86.67% | 90.03% |
| Macron | Non-academic | 97.25% | 98.50% | 98.25% | 98.75% | 98.78% |
| Mélenchon | Non-academic | 3.33% | 5.00% | 7.33% | 5.00% | 3.67% |
| Pécresse | Non-academic | 12.33% | 9.00% | 2.33% | 1.33% | 1.00% |
| Zemmour | Non-academic | 23.67% | 11.33% | 4.00% | 2.00% | 1.00% |

Observation

On these numbers Macron and LePen would easily go thru to the second round, Mélenchon would again have a respectable third place, and Zemmour and Pécresse would battle for last place.

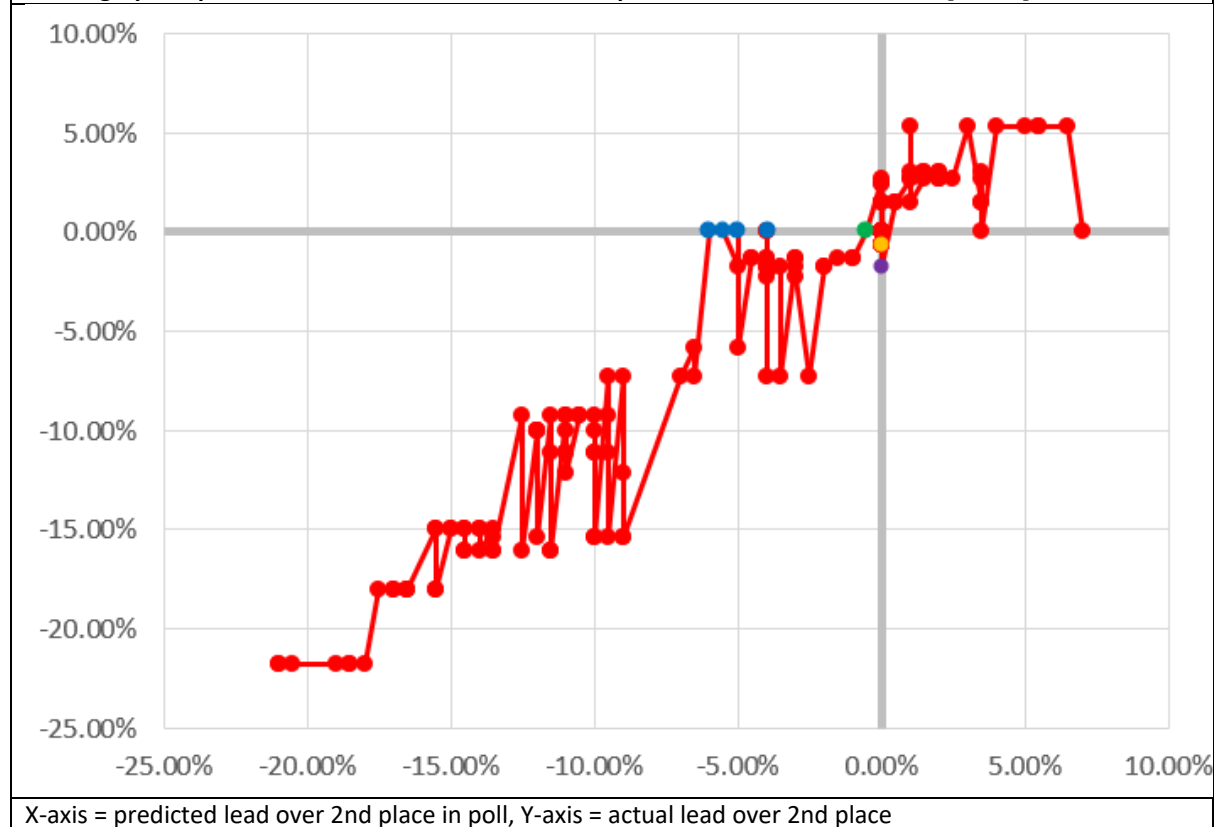
PART 8: INVESTIGATION OF ROUND ONE

Above we looked at the predictions that were made by others. But are there any predictive tools we can use? Here we present a brief investigation using simple tools

Prediction scatter graph for round one

Let's take all the polls within a week of the first-round election since 1995 and build a prediction scatter graph. On the X axis, the lead for each candidate over the 2nd-placed person in that opinion poll (so >0 is 1st place, 0 is 2nd place, and <0 is 3rd place or less). On the Y axis, the achieved lead over the second place in the first-round election. If it's a line that goes diagonally thru the axis bottom-left to top-right, then all the predictions for a place in the top two is correct. But if any point touches the axes outside the origin or is in the top left or bottom right quadrant, then something has gone. Here's the graph

Table: graph of predicted vs actual leads over second place in round one 1995-2017[0522a]



The blue dots are from Jean-Marie LePen in 2002, where all the polls on the last week had LePen in third place behind Jospin by 4-6 points, but he came in second in the election and went thru to the runoff. The green dot is a single poll in 2017 where Marine LePen was narrowly third. The yellow dots (there are several on top of each other) are the Jospin losses to LePen, and the single purple dot is Mélenchon in 2017 where he was briefly second in one poll.

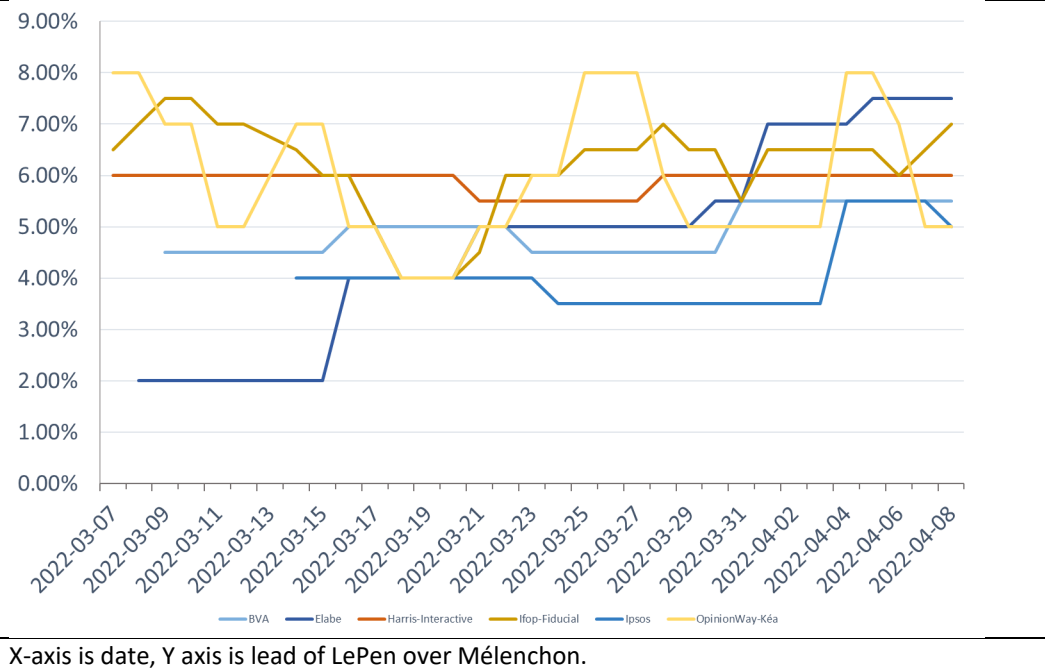
So what does it tell us? There are thirty points between $x=-0.05$ to $x=-6\%$, which means thirty polls with a gap between third and second place of 6 points or less. Of those 30 points, seven were for people in third place who actually came second in the election, which is nearly a quarter of them. So the probability is very approx. 25% that a person polling third in the last week will actually come second. Now since all of them have been surnamed "LePen", that might not have meant much, but it could mean that Mélenchon, trailing LePen by about 6-7%, was not entirely without hope...

Within-polls polling graph for round one

...but not a lot.

It's a commonplace when depicting polls to place each poll value on a graph and use the software's ability to draw a curve between them, and online discussion frequently revolves around which curve to draw. But this method mistakes changes *between-polls* for the pattern, where actually we need to track changes *within-polls*. This can be depicted and we do so in the table below for all polls since the start of the campaign on March 7th

Table: the lead of LePen over Mélenchon for each poll between March 7th and April 8th 2022

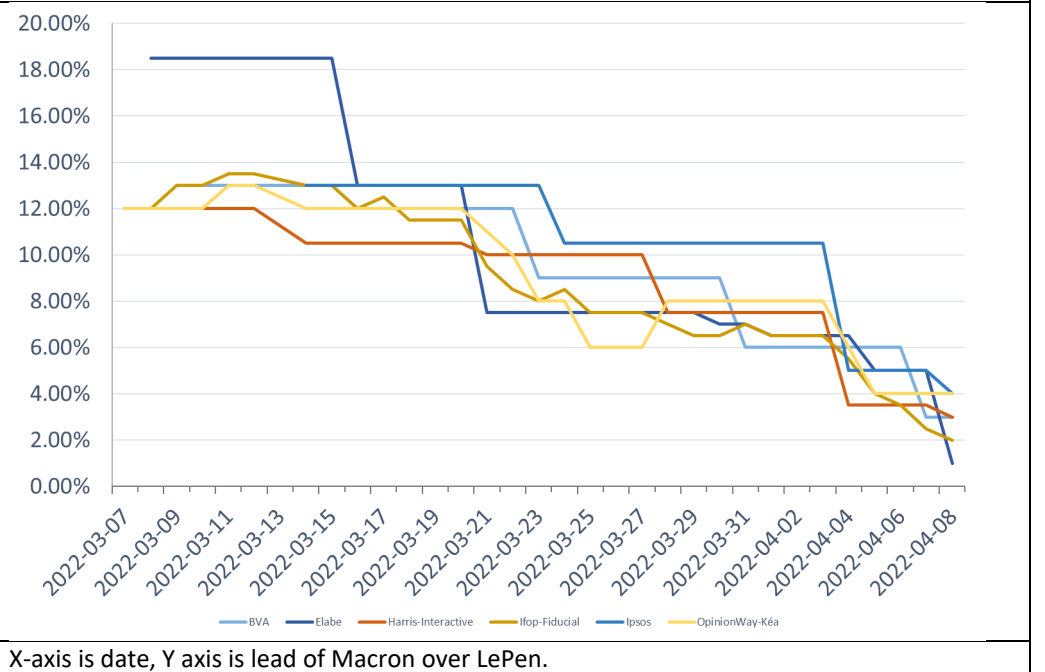


X-axis is date, Y axis is lead of LePen over Mélenchon.

This graph is not helpful. There is no obvious mode effect here, with our five polls executing a drunkard’s walk between 2% and 8% behind LePen. We know from the scatter graph that it was actually possible in theory to make up that deficit (because LePen père did it in 2002), but on this evidence it was not likely.

As for LePen... well that’s where it got interesting. The graph looks like this:

Table: the lead of Macron over LePen for each poll between March 7th and April 8th 2022



X-axis is date, Y axis is lead of Macron over LePen.

There was a definite trend. By the last week of round one LePen was pulling the deficit back by about 4% per week(!), and by April 8th her deficit over Macron was between 1% and 4%, with two days to go. On these figures it was still likely Macron would win round one, but it would be close and - if factors such as social

satisficing masking her support come into play - LePen may have come first on the round one election on April 10th.

INTERLUDE: SUMMARY AT 2022-04-04

This paper is intended for publication by Significance after the election, but early versions have to be sent to the editor for inspections. Periodically therefore we tidied the working version of the paper and summarised the state of play at a given date. This is the verbatim summary for April 4th 2022.

The pollsters are unanimous that Macro will win, although LePen is experiencing a late surge. The bookmakers reflect this (broadly speaking bookmaker odds track opinion polls in direction). The academic analysts are finding French political modelling difficult despite many being French. The instability in French political parties means that they have to aggregate their data into political families, make predictions based on those families, then assign their familial predictions to specific candidates, which is a subjective process. This, combined with the two-stage process and lack of interest in French politics makes academic analysis more difficult and less frequent: we could find few academic analyses updated to pass February. The non-academic analysts are similarly handicapped by the lack of interest, with IEM non-useful – for wisdom of crowds to work there must be a crowd, and there just isn't one. Interestingly Hypermind is French and predicts a better outcome for Pécresse vs_Macron than LePen_vs_Macron in the second round, albeit with less chance of getting there.

At the moment the only prediction we can confidently make for round one is that Macron will face-off against LePen in round two. We are not confident of the outcome in round two other than it will be a Macron win or a LePen win. We present two narratives for the final result as follows

Narrative for a Macron win: “...In retrospect it was obvious. All the indicators pointed to Macron and despite a late surge LePen lost. As the war in Ukraine progressed, the events meant that Macron was seen as the better candidate and French national security voters plumped for his approach to the war. He was seen as having the best answers to the issues of the day (Covid-19, energy bills etc) and secured victory on April 24th. The modellers were chastened and tried to reconcile their findings to their results, eventually blaming their failure on a late surge, the war and people changing their mind on the day...”

Narrative for a LePen win: “...In retrospect it was obvious. The late surge to LePen overtook Macron and despite it going back-and-forth she could never be ruled out. As the war in Ukraine progressed, the events meant that LePen was seen as the better candidate and French national security voters plumped for her approach to the war. She was seen as having the best answers to the issues of the day (Covid-19, energy bills etc) and secured victory on April 24th. Instead of Pécresse eating LePen's lunch as originally predicted months earlier, the reverse happened. The modellers realised that the Overton Window had changed and that the votes they assigned to Pécresse should have been assigned to LePen instead...”

PART 9: RESULTS OF ROUND ONE ON 2022-04-10

Table: Results of round one candidate per candidate [0529a]

| Candidate | valid votes cast | % |
|------------------|------------------|---------|
| LePen | 8,133,828 | 23.15% |
| Macron | 9,783,058 | 27.85% |
| Mélenchon | 7,712,520 | 21.95% |
| Pécresse | 1,679,001 | 4.78% |
| Zemmour | 2,485,226 | 7.07% |
| others | 5,339,314 | 15.20% |
| Valid votes cast | 35,132,947 | 100.00% |

PART 10: WHAT DID OUR PREDICTORS PREDICT BEFORE ROUND TWO?

We now need to record how well they did prior to the round two election

Probability of winning round two (aka probability of being the next President)

We see below the values over time for the probability of winning round two. Cases where no predictions were made are denoted by a “-“

Table: Average bookmakers probability of winning round two per candidate at datum dates 1-7 [0508g]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|-----------|--------|--------|--------|--------|--------|--------|--------|
| LePen | Bookmaker | 7.06% | 5.27% | 7.88% | 18.84% | 23.32% | 12.21% | 6.43% |
| Macron | Bookmaker | 82.01% | 92.74% | 93.27% | 84.38% | 80.22% | 91.06% | 96.88% |
| Mélenchon | Bookmaker | 2.24% | 1.59% | 1.99% | 1.56% | 1.16% | - | - |
| Pécresse | Bookmaker | 11.32% | 3.14% | 0.88% | 0.72% | 0.20% | - | - |
| Zemmour | Bookmaker | 6.80% | 3.90% | 2.41% | 2.07% | 1.42% | - | - |

Table: Average non-academic probability of winning round two per candidate at datum dates 1-7 [0508h]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|--------------|--------|--------|--------|--------|--------|--------|--------|
| LePen | Non-academic | 6.33% | 4.67% | 7.33% | 17.37% | 22.03% | 9.67% | 4.67% |
| Macron | Non-academic | 84.33% | 90.33% | 90.67% | 77.53% | 74.17% | 90.67% | 96.00% |
| Mélenchon | Non-academic | 1.33% | 1.33% | 2.00% | 1.93% | 1.63% | - | - |
| Pécresse | Non-academic | 5.33% | 3.00% | 1.00% | 0.63% | 0.97% | - | - |
| Zemmour | Non-academic | 5.67% | 2.67% | 1.33% | 1.27% | 1.27% | - | - |

Vote shares in round two per candidate

We see below the values over time for vote shares per candidate in round two. The abnormally large numbers for Pécresse are highlighted in red.

Table: Average pollster round two vote share per candidate at datum dates 1-7 [0326a]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|----------|--------|--------|--------|--------|--------|--------|--------|
| LePen | Pollster | 44.17% | 42.50% | 44.67% | 46.83% | 47.42% | 46.33% | 44.25% |
| Macron | Pollster | 59.84% | 61.33% | 61.67% | 58.75% | 58.83% | 53.67% | 55.75% |
| Mélenchon | Pollster | 37.00% | 35.00% | 37.80% | 38.50% | 42.00% | - | - |
| Pécresse | Pollster | 40.08% | 39.42% | 37.50% | 37.25% | 35.50% | - | - |
| Zemmour | Pollster | 36.75% | 35.33% | 33.25% | 34.00% | 34.00% | - | - |

Table: Average academic round two vote share per candidate at datum dates 1-7 [0508i]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|----------|--------|-------|--------|-------|-------|--------|--------|
| LePen | Academic | 42.90% | ditto | 42.41% | ditto | ditto | 42.41% | 42.41% |
| Macron | Academic | 55.38% | ditto | 56.25% | ditto | ditto | 57.59% | 57.59% |
| Mélenchon | Academic | - | - | - | - | - | - | - |
| Pécresse | Academic | 46.35% | ditto | 45.10% | ditto | ditto | - | - |
| Zemmour | Academic | - | - | - | - | - | - | - |

Table: Average non-academic round two vote share per candidate at datum dates 1-7 [0508j]

| Candidate | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|--------------|--------|--------|--------|--------|--------|--------|--------|
| LePen | Non-academic | 42.20% | 41.90% | 44.80% | 43.50% | 45.20% | 46.20% | 44.70% |
| Macron | Non-academic | 59.58% | 60.75% | 62.25% | 60.10% | 59.55% | 53.80% | 55.30% |
| Mélenchon | Non-academic | 35.60% | 33.70% | 34.70% | 34.70% | 35.00% | - | - |
| Pécresse | Non-academic | 45.70% | 45.10% | 34.80% | 44.50% | 44.60% | - | - |
| Zemmour | Non-academic | 38.20% | 36.30% | 36.70% | 36.90% | 37.00% | - | - |

Vote shares in round two per candidate and opponent

We see below the values over time for vote shares per candidate an opponent in round two. Again the abnormally large numbers for Pécresse are highlighted in red.

Table: Average pollster round two vote share per candidate and opponent at datum dates 1-7 [0326a]

| Candidate | Against | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|-----------|----------|--------|--------|--------|--------|--------|--------|--------|
| LePen | Macron | Pollster | 44.17% | 42.50% | 44.67% | 46.83% | 47.42% | 46.33% | 44.25% |
| Macron | LePen | Pollster | 55.83% | 57.50% | 55.33% | 53.17% | 52.58% | 53.67% | 55.75% |
| Macron | Mélenchon | Pollster | 63.00% | 65.00% | 62.20% | 61.50% | 58.00% | - | - |
| Macron | Pécresse | Pollster | 59.92% | 60.58% | 62.50% | 62.75% | 64.50% | - | - |
| Macron | Zemmour | Pollster | 63.25% | 64.67% | 66.75% | 66.00% | 66.00% | - | - |
| Mélenchon | Macron | Pollster | 37.00% | 35.00% | 37.80% | 38.50% | 42.00% | - | - |
| Pécresse | Macron | Pollster | 40.08% | 39.42% | 37.50% | 37.25% | 35.50% | - | - |
| Zemmour | Macron | Pollster | 36.75% | 35.33% | 33.25% | 34.00% | 34.00% | - | - |

Table: Average academic round two vote share per candidate and opponent at datum dates 1-7 [0508i]

| Candidate | Against | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|----------|----------|--------|-------|--------|-------|-------|--------|--------|
| LePen | Macron | Academic | 42.90% | ditto | 42.41% | ditto | ditto | 42.41% | 42.41% |
| Macron | LePen | Academic | 57.10% | ditto | 57.59% | ditto | ditto | 57.59% | 57.59% |
| Macron | Pécresse | Academic | 53.65% | ditto | 54.90% | ditto | ditto | - | - |
| Pécresse | Macron | Academic | 46.35% | ditto | 45.10% | ditto | ditto | - | - |

Table: Average non-academic round two vote share per candidate and opponent at datum dates 1-7 [0326a]

| Candidate | Against | Type | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|-----------|--------------|--------|--------|--------|--------|--------|--------|--------|
| LePen | Macron | Non-academic | 42.20% | 41.90% | 44.80% | 43.50% | 45.20% | 46.20% | 44.70% |
| Macron | LePen | Non-academic | 57.80% | 58.10% | 55.20% | 56.50% | 54.80% | 53.80% | 55.30% |
| Macron | Mélenchon | Non-academic | 64.40% | 66.30% | 65.30% | 65.30% | 65.00% | - | - |
| Macron | Pécresse | Non-academic | 54.30% | 54.90% | 65.20% | 55.50% | 55.40% | - | - |
| Macron | Zemmour | Non-academic | 61.80% | 63.70% | 63.30% | 63.10% | 63.00% | - | - |
| Mélenchon | Macron | Non-academic | 35.60% | 33.70% | 34.70% | 34.70% | 35.00% | - | - |
| Pécresse | Macron | Non-academic | 45.70% | 45.10% | 34.80% | 44.50% | 44.60% | - | - |
| Zemmour | Macron | Non-academic | 38.20% | 36.30% | 36.70% | 36.90% | 37.00% | - | - |

Observation

On these numbers Macron was predicted to win round two, with a consistent lead throughout the datum dates. Although LePen did eat into Macron's lead up to the round one election, she lost ground during round two. We note that our selected academics did not make round two vote share predictions for Mélenchon and Zemmour, which affected our assessments on a five-person-unforced basis. The abnormally large estimates for Pécresse (labelled in red) were also noted. We will discuss these points later.

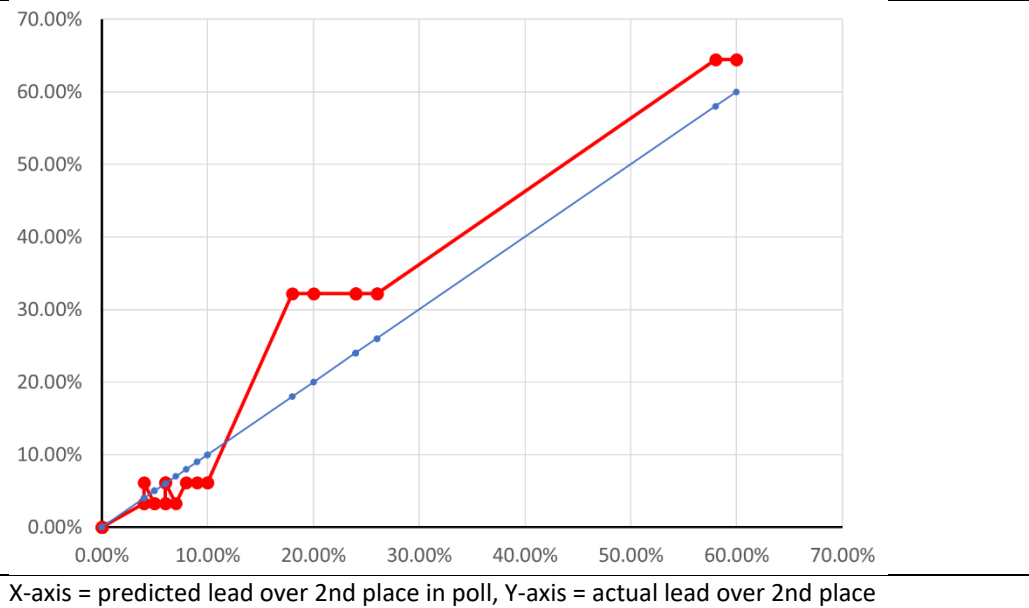
PART 11: INVESTIGATION OF ROUND TWO

Following on from round one, can we use our techniques again. Below we depict the prediction scatter graph and the polling graph

Prediction scatter graph for round two

Again let's take all the polls within a week of the second-round election since 2002 and build a prediction scatter graph. The X axis depicts the polling lead for each candidate over the 2nd-placed person in that opinion poll (so >0 is 1st place, 0 is 2nd place, and <0 is 3rd place or less). The Y axis depicts the achieved lead over second place in the first-round election. If it's a line that goes diagonally thru the axis bottom-left to top-right, then all the predictions for a place in the top two is correct. But if any point touches the axes outside the origin or is in the top left or bottom right quadrant, then something has gone. Here's the graph.

Table: graph of predicted vs actual leads over second place in round two 2002-2017[0528a]

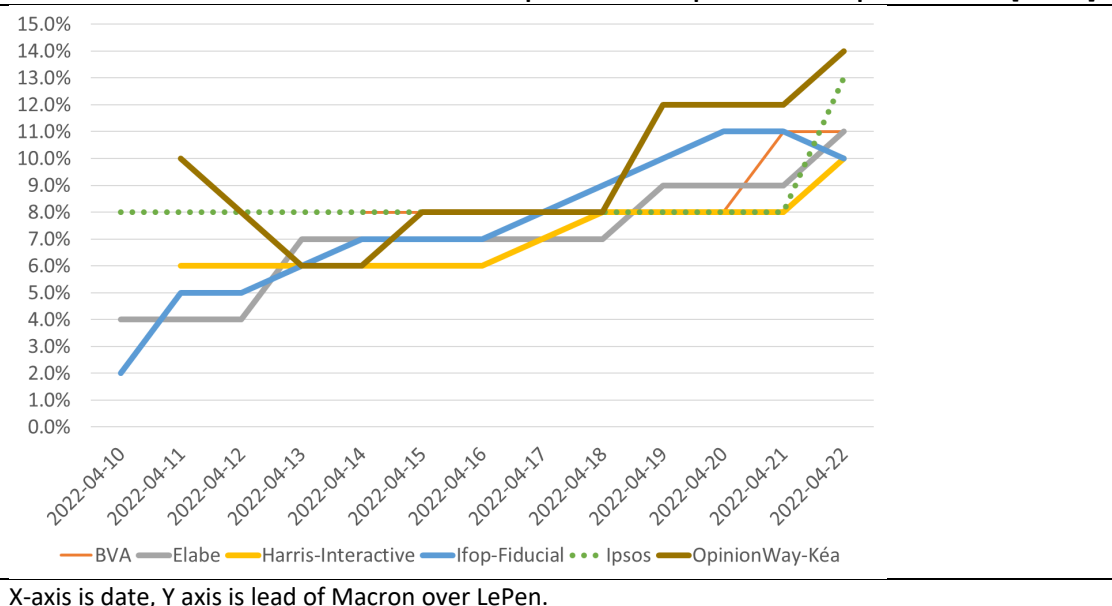


So what does it tell us? Unlike the equivalent graph for round one, all the points are in the top-right quadrant. That means that since 2002 all the polls correctly predicted second place (because all the differences between first and second were positive before the round two election, and were all positive afterwards). Since round two only has two candidates, that meant that all the polls predicted first place too.

The line for a lead of 10% or less is under the diagonal, which means that poll leads under that threshold are clawed back on the day, albeit not enough to overtake the lead. Those leads that are 18% or more are above the diagonal, which means they exceeded expectations on the day: the first place was by a bigger margin than expected.

Within-polls polling graph for round two

Table: the lead of Macron over LePen for each poll between April 10th and April 22th 2022[0528b]



At our seventh datum date (April 23rd - polls stop in France two days before the election) LePen was between 10% and 14% behind in our five polls. This implies that she would still lose on the day, since (as we saw in the prediction scatter graph) that deficit has never been clawed back in the period inspected.

INTERLUDE: SUMMARY AT 2022-04-24

This paper is intended for publication by Significance after the election, but early versions have to be sent to the editor for inspections. Periodically therefore we tidied the working version of the paper and summarised the state of play at a given date. This is the summary for April 24th 2022 at 18:20.

At 18:20 the consensus is that Macron will win. LePen failed to make an impact during the debates and the odds began to move away from her. The feeling on politicalbetting.com is that Macron may in fact do better than the polls say. But we are never confident about predicting the size of the victory, leaving that to more sophisticated observers

*At the moment **the only prediction we can confidently make for round two is that Macron will win against LePen in round two.***

It is now 18:30 BST, 30 minutes before the predictions are issued at 19:00 BST. The only thing to do now is to send off this prediction to the editor to provide a timestamp, take records of the predictors before 7pm, and await the result.

PART 12: RESULTS OF THE SECOND ROUND ON 2022-04-24

Table: Results of round two candidate per candidate [0529b]

| Candidate | Votes | % |
|------------------|------------|---------|
| LePen | 13,288,686 | 41.45% |
| Macron | 18,768,639 | 58.55% |
| Mélenchon | - | - |
| Pécresse | - | - |
| Zemmour | - | - |
| Valid votes cast | 32,057,325 | 100.00% |

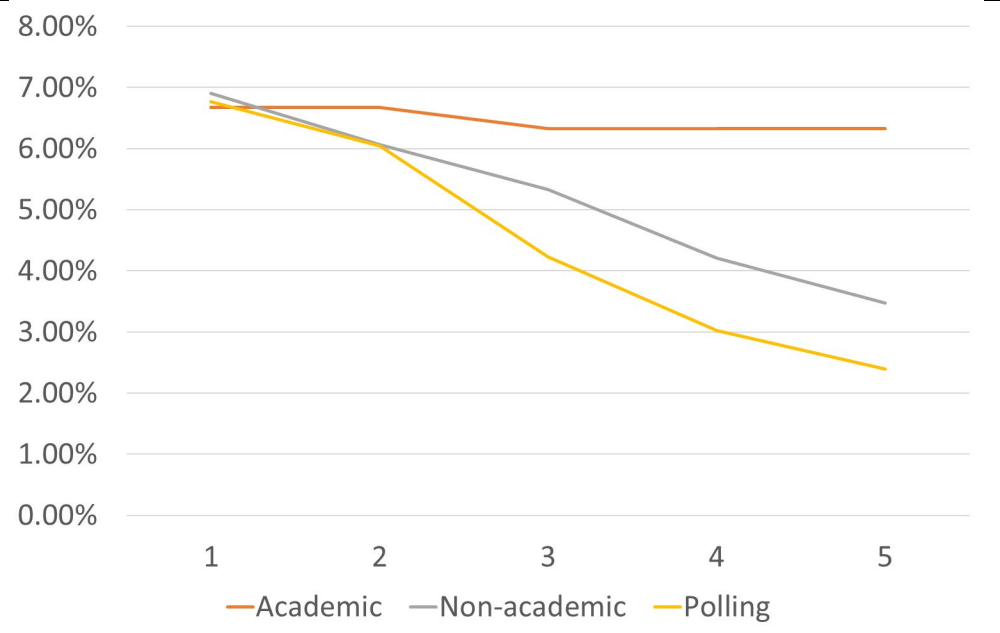
PART 13: HOW WELL DID OUR PREDICTORS DO AFTER ROUND ONE?

The predictors for round one were the predicted probability of a top-two finish, and the predicted vote shares. Here we examine how well they did.

MAE for round one over time

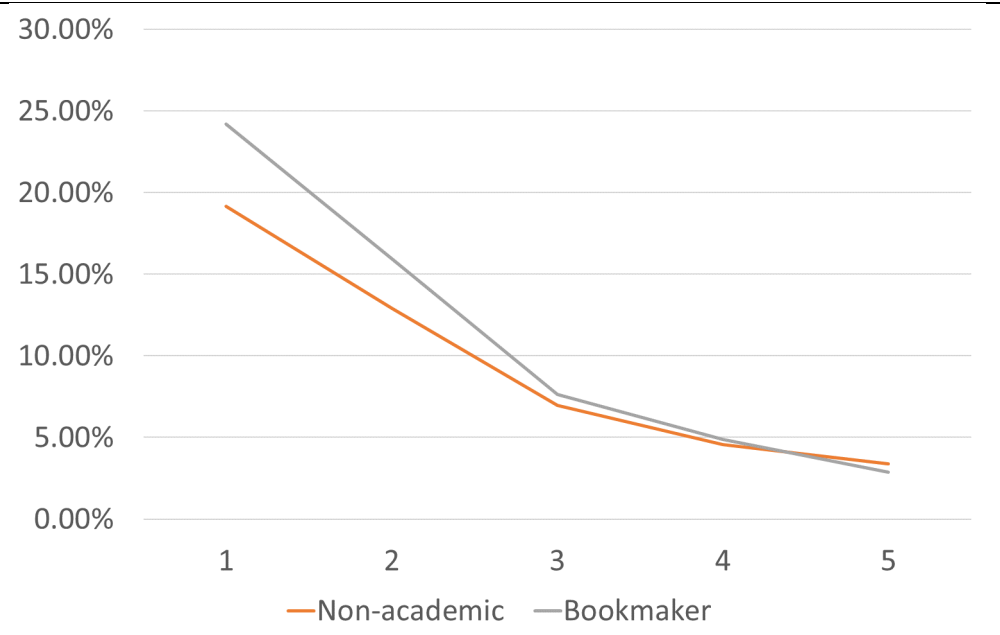
Below we graphically depict the MAE accuracy of the predicted vote shares in round one, and the predicted probability of top two finishes.

Table: MAE for predicted vote share for round one at datum dates 1-5[0603a]



X-axis is datum date, Y axis is mean absolute error

Table: MAE for predicted probability of top two finish at datum dates 1-5[0603b]



X-axis is date, Y axis is mean absolute error

With the exception of the academic predictors who did one-and-done predictions and could not correct as time passed, the accuracy of our vote share predictors increased over time but did not reach the 2% threshold required for a five-person-unforced (5PU) metric.

As noted in our previous work, the accuracy of probabilistic predictions is inherently worse than the accuracy of deterministic predictions (that's why they're usually measured using calibration instead of MAE). In this case they improved dramatically as time wears on but again did not achieve our 2% threshold.

WIN for round one over time

Below we depict in tables the WIN accuracy of the predicted vote shares in round one, and the predicted probability of top two finishes.

Table: Mean WIN indicator for predicted vote share for round one at datum dates 1-5

| Predictor | 1 | 2 | 3 | 4 | 5 |
|--------------|---|---|---|---|---|
| Academic | 1 | 1 | 1 | 1 | 1 |
| Non-academic | 1 | 1 | 1 | 1 | 1 |
| Pollsters | 1 | 1 | 1 | 1 | 1 |

Table: Mean WIN indicator for predicted probability of top two finish at datum dates 1-5

| Predictor | 1 | 2 | 3 | 4 | 5 |
|--------------|---|---|---|---|---|
| Bookmaker | 1 | 1 | 1 | 1 | 1 |
| Non-academic | 1 | 1 | 1 | 1 | 1 |

To put it simply, at no point in the five datum dates prior to the round one election did anybody seriously think that anybody other than Macron and LePen would contest round two.

PART 14: HOW WELL DID OUR PREDICTORS DO AFTER ROUND TWO?

The predictors for round two were predicted probability of winning, and the predicted vote shares. Here we examine how well they did.

MAE for round two over time

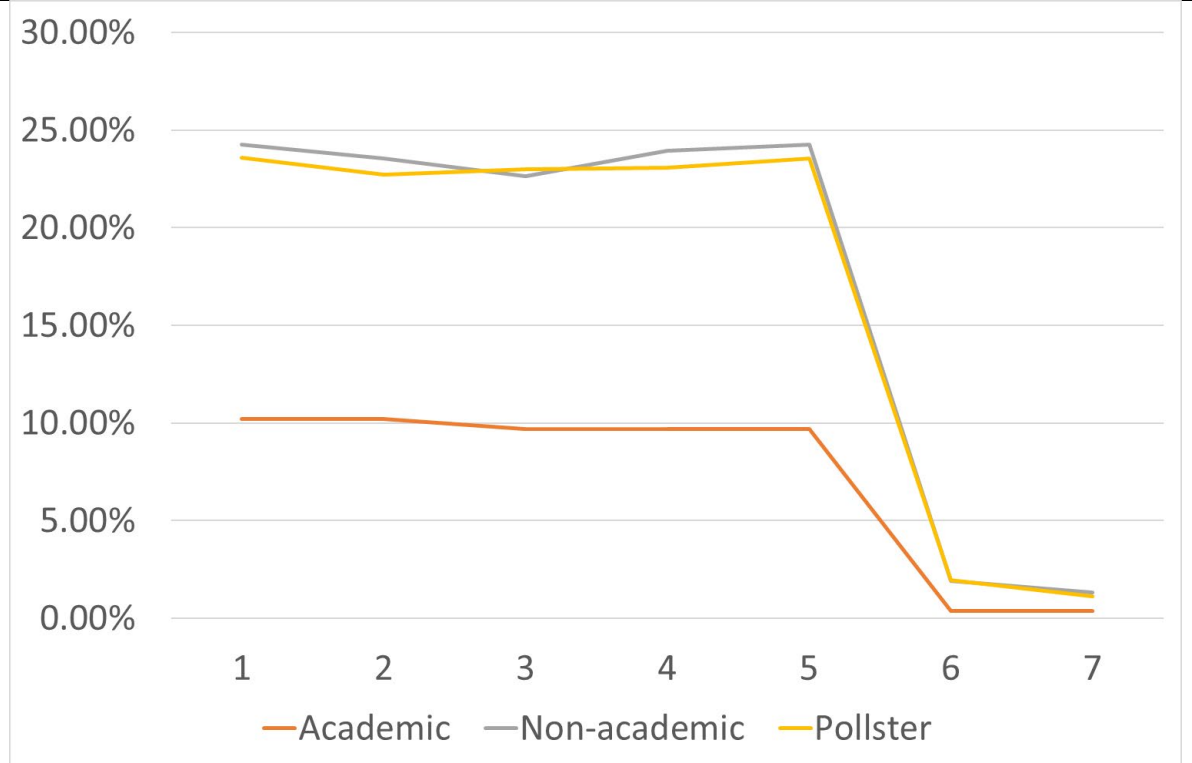
Below we graphically depict the MAE accuracy of the predicted probability of winning round two, and the predicted vote shares in round two.

Table: MAE for predicted probability of winning round two at datum dates 1-7[0603c]



X-axis is datum date, Y axis is mean absolute error

Table: MAE for predicted vote shares in round two at datum dates 1-7[0603d]



X-axis is date, Y axis is mean absolute error

Here we see something interesting. The predictors are unacceptably large, even wildly unacceptable, during round one, but after the round one hurdle had been passed and it was just Macron and LePen left, the errors shrank rapidly and met our 2% threshold in all cases. Why is this?

One possible explanation is format flattery. We are using five-person-unforced 5PU format, so the fact that the second-round votes for Mélenchon, Pécresse and Zemmour defaulted to zero after round one (because they didn't make it) meant that the average accuracy increased: if we had been using 2PU or 2PF things might have been different. Nevertheless the prosaic explanation may be the best one: simply put LePen just performed poorly in the second round and never seriously threatened Macron.

WIN for round two over time

Below we depict in tables the WIN accuracy of the predicted vote shares in round two, and the predicted probability of winning round two.

Table: Mean WIN indicator for probability of winning round two at datum dates 1-7

| Predictor | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|--------------|---|---|---|---|---|---|---|
| Bookmaker | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Non-academic | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Table: Mean WIN indicator for predicted vote shares in round two at datum dates 1-7

| Predictor | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|--------------|---|---|---|---|---|---|---|
| Academic | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Non-academic | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Pollsters | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

To ram the point home, at no point in the seven datum dates prior to the round two election did anybody seriously think that anybody other than Macron would win.

PART 15: CONCLUSION

On an initial reading this was a straightforward election. The multiplicity of candidates meant that prediction was difficult and the errors large, but as time passed and the competitors were winnowed down to two, the predictors stabilized and became useful. So it appears there is nothing interesting here.

But one interesting phenomenon did manifest: the abnormally large estimates for Péresse in round two, which on the face of it was unjustified. It is plausible that Mme. Péresse was actually popular in the early stages and that would explain it, but even up to datum date four (5th April, five days before the round one election) Hypermind was still saying Péresse would do better than LePen in round two, despite the fact that she had no real chance of getting there. Investigating the reasons for this would be fascinating but outside the scope of this article.

PART 16: AFTERWORD

This article is published at a crux point. The publications in Significance began in 2013 under the then-editor Julian Champkin, but the Forecast Error series was shepherded to maturity under present editor Brian Tarran, who has attempted with much patience and some skill to sculpt the author's prose into something readable by a human being. It has been remarkably successful, developing from a minor article looking the 2010 election to its now fully operational state, with simple prediction methods in train and several articles planned for the coming years. The new editor Anna Britten will hopefully afford the author the same leeway and find the strength to not kill the author with hammers, which her predecessors somehow restrained themselves from doing despite the mangled syntax.

As for the future articles, again a crux. Forecast Error articles scheduled for the coming years include "Forecast Error: Money" (late 2023), "Forecast Error: Models" (late 2024?), and the next UK General Election, but the main attraction is the "Forecast Error: 2024 US Presidential Election" article scheduled for just before and after November 2024. If it is the Trump vs Biden rematch that is rumored, it will be a civilisational clash and a fitting peak for the decade's worth of work that the Forecast Error series represents.

And I hope you will continue to watch its career with great interest, and meet me there.

PART 17: DISCLAIMER

The author is a statistician who works in the private sector. To pass the compliance requirements of the author's employer please note that the author's affiliation for this article is "independent" and the views expressed in this article are those of the author and in no way do they reflect the opinions of any entity or organisation with which the author is – or has in the past, or may in the future be – affiliated.