

Ranking pollsters in the UK

Oraclum Intelligence Systems

Public release on April 16, 2018
ORACLUM WHITE PAPER No.08/2017



Ranking pollsters in the UK

by

Oraclum Intelligence Systems

Oraclum White Papers are analytical reports on Oraclum's predictions and prediction methods. They are designed to be informative, provide an in-depth statistical analysis of a given issue, call for a proposal to action, or introduce a unique solution based on one of Oraclum's products.

This white paper was made public on April 16, 2018.

An electronic version is available at <http://www.oraclum.co.uk/about/white-papers/>.

This document was prepared using the TU Delft \LaTeX template.

The cover image is taken from <http://www.meteoearth.com>.



This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.

© Oraclum Intelligence Systems
23 Arnold Close, Hauxton
Cambridge, CB22 5FN
United Kingdom
www.oraclum.co.uk

Executive summary

- This White Paper presents a new method of measuring pollster accuracy for pollsters in British elections
- It represents an effort to vindicate some in the polling business during a time of widespread public anger over their failure to correctly predict recent electoral outcomes in Britain
- This is the first effort of its kind that can be used to rank UK pollsters by accuracy and quality of their methodology

*Oraclum Intelligence Systems
June 2017*

Ranking pollsters in the UK

Opinion pollsters in the UK came under a fierce line of attack from the public and the media following their joint failure to accurately predict the results of the 2015 UK general election and the 2016 Brexit referendum. Months and weeks before the May 2015 election the polls were predicting a hung parliament and a virtual tie between Conservatives and Labour, where the outcome would have been another coalition government (a number of combinations were discussed, including a grand coalition between Labour and Conservatives), or even a minority government.

The results show that the pollsters, on average, missed the difference between the two parties by 6.8%, which translated into about 100 seats. What was supposed to be one of the closest elections in British history turned out to be a landslide victory for the Conservatives.

An error of similar magnitude occurred during the 2016 Brexit referendum. Every pollster was suggesting that Britain will vote Remain. The public outcry against the pollsters was lower this time, partially because people already lost faith in the polls and were expecting them to go wrong, and partially because they were too busy either gloating or sobbing following the referendum results.

Inquiries after both these events were made, accusing pollsters of complacency¹, herding², and deliberate manipulation of their samples³. And while there was certainly something that went wrong in the sampling methods of the pollsters⁴, our effort here is not to further those attacks, but instead to vindicate some in the polling business by offering, **for the first time in the UK, an unbiased ranking of UK pollsters**. Table 1 shows the final ranking.

Our rankings are based on a somewhat technical but still easy to understand methodological approach summarized in detail below. It has its drawbacks, which is why we welcome all comments, suggestions and criticism. We will periodically update our ranking, our method, and hopefully include even more data (local and national), all with the goal of producing a standardized, unbiased overview into the performance of opinion pollsters in the UK.

We hope that our rankings stir a positive discussion on the quality of opinion pollsters in the UK, and we welcome and encourage the usage of our rankings data to other scientists, journalists, forecasters, and forecasting enthusiasts.

Note that in the ranking list we omit the British Election Study (BES), which uses a far better methodology than other pollsters – a face-to-face random sample survey (the gist of it is that they randomly select eligible voters to get a representative sample of the UK population, and then they repeatedly contact those people to do the survey; you can read more about it in the footnote⁵). This has enabled them to give out one of the most precise predictions of the 2015 general election (they gave the Conservatives an 8% margin of victory⁶). However there is a problem – the survey has been (and usually is) done after the elections, meaning that it cannot be used as a *prediction* tool. Because of this, instead of grouping it with the others the BES can only be used as a post-election benchmark.

In addition to tracking pollster accuracy and ranking them, we use the same numbers to construct our adjusted polling average⁷, the benchmark we usually use to compare ourselves

¹<http://www.telegraph.co.uk/news/politics/12108096/The-election-polls-werent-wrong-because-of-mistakes-pollsters-deliberately-cheated.html>

²<http://www.telegraph.co.uk/news/general-election-2015/politics-blog/11701135/Yes-the-pollsters-lied-and-heres-the-proof.html>

³<http://www.southampton.ac.uk/news/2016/01/polling-enquiry.page>

⁴<http://www.britishelectionstudy.com/bes-resources/why-the-polls-got-it-wrong-and-the-british-election-study-face-to-face-survey-got-it-almost-right/>

⁵<http://www.britishelectionstudy.com/bes-impact/f2f-release/>

⁶<http://www.britishelectionstudy.com/bes-resources/why-the-polls-got-it-wrong-and-the-british-election-study-face-to-face-survey-got-it-almost-right/>

⁷We formulate a **joint weight of timing** (the more recent the poll, the greater the weight), **sample size** (the greater the sample size, the greater the weight), whether the poll was done **online or via telephone**, and the **ranking for each poll**, allowing us

Table 1: The unbiased ranking of UK pollsters.

Pollster	Number of polls analyzed	Joint within-between index 2016	Joint within-between index 2015	Joint within-between index 2014	Joint within-between index 2010	Final weighting index (ranking)	Precision index
1. YouGov	188	9.36	2.75	10.45	8.14	7.67	7.67
2. TNS	26	10.70	0.51	10.83	6.58	7.15	7.15
3. ICM	43	8.92	6.58	-0.07	9.71	6.29	6.29
4. Opinium	35	6.48	4.39	2.12	9.92	5.73	5.73
5. Ipsos MORI	18	4.20	4.61	6.92	7.07	5.70	5.70
6. BMG	5	11.60	6.47			4.52	9.03
7. Survation	26	4.22	3.08	7.75		3.76	5.02
8. ComRes	40	-4.72	5.40		9.78	2.61	3.49
9. Angus Reid	16				9.92	2.48	9.92
10. Harris Int.	14				9.92	2.48	9.92
11. BPIX	7				6.94	1.73	6.94
12. Lord Ashcroft	17		6.91			1.73	6.91
13. Populus	44	-4.72	0.34		9.92	1.39	1.85
14. Panelbase	12		-0.54	2.26		0.43	0.86
15. ORB	7	0.92				0.23	0.92

Source of data: UK Polling Report and Wikipedia. All calculations (and mistakes) are our own.

against (e.g. see the benchmark for Brexit⁸, or the same benchmark for the US⁹). The adjusted polling average was wrong in both cases (unlike our predictions), which goes to show that even when adjusting for the usual polling bias, the cases of Brexit and Trump were very specific. There was a systematic bias and underestimation of both Brexit and Trump in the polls. This is why all the poll-based forecasts went wrong.

How do we produce our rankings?

The rankings are based on past performance of pollsters for four earlier elections, the 2016 Brexit referendum, the 2015 general election, the 2014 Scottish referendum, and the 2010 general election. In total we analyzed over 500 polls from 15 pollsters (not all of which participated in all three elections). We realize the sample could have been bigger by including local and previous general elections, however given that many pollsters from 10 years ago do not produce election polls anymore, and given that local elections are quite specific, we focus only on these three national-level one regional-level election. We admit that the sample should be bigger and will think about including the local polling outcomes, adjusted for their type. There is also the issue of methodological standard of each pollster which we don't take into account, as we are only interested in the relative performance each pollster had in the previous elections.

Given that almost all the pollsters failed to predict the outcome of the Brexit referendum and the 2015 general election, we look at the performance between pollsters as well, in order to avoid penalizing them too much for these failures. If no one saw it coming, they are all equally excused, to a certain extent. If however a few did predict correctly, the penalization against all others is more significant. We therefore jointly adjust the **within accuracy** (the accuracy of an individual pollster with respect to the final outcome) and the **between accuracy** (the accuracy of an individual pollster with respect to the accuracy of the group).

to calculate the **final weighted average across all polls** in a given time frame.

⁸<http://oraclum.eu/method-benchmarks-brexit-referendum/>

⁹<http://oraclum.eu/post-election-analysis-polls-underestimated-trump/>

1. Within accuracy

To calculate the precision of pollsters in earlier elections we again have to assign weights for timing and sample size, in the same way as earlier described (older polls are less important, greater sample size is more important). Both of these factors are then summed up into the total weight for a given poll across all pollsters. We then take each individual pollster and calculate its weighted average. This is calculated as the sum of the product of all its polls and their sample and timing weights, divided by the sum of all weights:

$$\sum_i x_i w_i / \sum_i w_i \tag{1}$$

where x_i is an individual poll and w_i the corresponding weight. w_i is calculated as the sum of three weights, for timing (using an adjusted exponential decay formula, decreasing from 4 to 0, where half-life is defined by $t_{1/2} = \tau \ln 2$), for sample size ($N/1000$), and the ranking weight (described in the text).

By doing so we can calculate the average error each pollster made in a given election. This is done for all three elections in our sample allowing us to calculate their within accuracy for each election. We calculate the average error for an individual pollster as the simple deviation between the weighted average polling result and the actual result for the margin between the first two parties in the elections (e.g. Conservatives and Labour). Or in plain English, how well they guessed the difference between the winner and the runner-up.

To clarify, define x_i as the difference between total predicted vote share of party A \bar{v}_A and party B \bar{v}_B for pollster i , and y as the difference between the actual vote share of the two parties. Assume A was the winner, and B was the runner-up. The within accuracy of pollster i (z_i) is then defined simply as $z_i = |x_i - y|$. The closer the value of z_i is to 0, the more accurate the pollster. From this we calculate the within index as:

$$I_w = 10 - z_i \tag{2}$$

2. Between accuracy

After determining our within index, we estimate the accuracy between pollsters (by how much they beat each other) and sum them both into a single accuracy index. To do this we first calculate the average error for all pollsters during a single election. We then simply subtract the average joint error from each individual error (z_i , as defined earlier).

$$I_b = z_i - \frac{1}{n} \sum_{i=1}^n z_i \tag{3}$$

This represents our **between index**: the greater the value, the better the pollster did against all others (note: the value can be negative).

3. Joint within-between ranking

To get our **joint within-between index** we simply sum up the two, thereby lowering the penalization across all pollsters if and when all of them missed. In this case those who missed less than others get a higher value improving their overall performance and ranking them higher on the scale.

$$I_{wb} = I_w + I_b \tag{4}$$

We repeat the same procedure across all three elections and produce two final measures of accuracy. The first is the **final weighting index** (which we use for the ranking itself and whose values we use as an input in the Brexit polls), and the second is the **precision index**. The difference between the two is that the precision index does not factor in the number of elections, whereas the final index does. The precision index is thus the simple average of the within-between indices, while the final index is the sum of all three divided by the total number of elections we observed regardless of how many of them the pollster participated in. The two are the same if a pollster participated in all three elections, but they differ if the pollster participated in less than three elections.

To conclude, the numbers reported under the final weighting index column represent the ranking weight from the beginning of this text. Combined with the timing and sample size weights, it helps us calculate the final weighted average of all polls thereby helping us configure our strongest benchmark, the adjusted polling average.

About Oraclum Intelligence Systems

Oraclum Intelligence Systems Ltd is a company that uses the power of social networks, big data, and machine learning to predict election outcomes and uncover patterns of consumer behaviour. We use our own cutting-edge technology to leverage publicly available data in order to unlock competitive advantages for our clients.

We have developed a unique, science-based forecasting method, the **Bayesian Adjusted Social Network (BASON) Survey**, which successfully solves the biggest issue in polling – it eliminates groupthink bias of respondents. We analyze the social networks of customers/voters in order to draw inferences from these networks and predict outcomes with incredible precision. Our BASON survey is the only poll in the world that **successfully predicted both Brexit and the victory of Donald Trump**. It is the best prediction tool available on the market. Combining it with our machine learning software yields unprecedented precision in figuring out what the customers/voters really want and why.

WWW: www.oraclum.co.uk

BLOG: www.oraclum.eu

 www.facebook.com/oraclumintelligencesystems

 www.twitter.com/@oraclum_UK



SCIENCE BEHIND
PREDICTION



INNOVATION
IN MARKET RESEARCH



MERGING BIG DATA
AND PREDICTIVE POWER