GLOBAL 19.09.2017 | Marcus Roberts

Why pollsters normally get it right (more or less)

Marcus Roberts explains why there will always be a need for pollsters

Opinion polls have been getting a bad press recently. Remember when three separate polls put UK Labour leader Ed Miliband ahead of the Tories in the run-up to the 2015 general election, only for Labour to lose a whopping 48 seats, including almost all their traditional Scottish strongholds? Or the Brexit vote, when Nigel Farage predicted sadly that 'Remain will edge it', only to taunt 'who's laughing now?' to MEPs a week later?

As professional pollster Marcus Roberts explains, it's easy to remember the few polls that got it wrong, and forget the hundreds that predict elections with remarkable accuracy.

Why have people lost confidence in pollsters?

Truth be told, a big part of this is actually bad luck! Pollsters have not got any more inaccurate recently but there have been a couple of high profile examples of very close elections where pollsters have fallen the wrong side of the line. The key case is Brexit where throughout the campaign we (and other online pollsters) were showing the results as being too close to call. Our polling done the weekend before the vote had the Leave side ahead, but our final poll before the vote had Remain on 51 percent and Leave on 49 percent. Historically speaking, this would be seen as reasonably accurate, but that doesn't really help if you aren't getting the result correct.

The other thing to bear in mind is that people remember when pollsters get it wrong but forget about the times we get it right. In the French presidential election, and in the Norwegian elections just last week, polls were broadly accurate. And in the British general election this year two online pollsters, YouGov and Survation, both produced final numbers that were pretty much spot on – yet because many other pollsters got it wrong, the perception was that 'polling had failed again'. Sigh.

What factors lead to pollsters overestimating the turnout of certain groups and underestimating others?

Even in a perfect world, all polls come with a margin of error. That is usually around three per cent either way, and it's a statistical inevitability of research. But there are additional areas that increase the error, such as sampling or weighting. The main reason we think pollsters have missed the mark in recent years is because of problems with the representativeness of the sample.

After the 2015 election we at YouGov ran a big review of our methodology and found we needed to make two changes to our sample. First, we found we were surveying too many graduates, so we adjusted to weight by educational qualifications. Second, the people we were surveying were paying too much attention to politics. We have now put measures in place to account for that, effectively reducing the proportion of high political information respondents to our surveys whilst increasing low political attention respondents (also known as normal people).

When it comes to the impact of specific demographic groups' turnout rates things get even more complicated. Take the case of Jeremy Corbyn's youth surge at the British general election. Young voters clearly did have an influence on the result in that:

1. The biggest swings to Labour were amongst the youngest voters, and
2. Early indications show that youth turnout was up.

However, Labour won more votes than the Conservatives amongst all age groups up to around 50, and the Tories faced problems amongst the entire working age population.

In terms of the US, it seems like youth turnout didn't really drop in the US presidential election by as much as it is perceived to have done so. However YouGov's chief scientist Doug Rivers's research has concluded that there was a more general turnout problem, implying that Democrats in 2016 were less likely to vote across all age groups.

Lastly, in the case of the recent showdown between Emmanuel Macron and Marine Le Pen in the French presidential elections, the age divide didn't seem to work the same way in France as it did elsewhere. Those least likely to vote for Le Pen were actually the very old. So your propensity to vote for Macron didn't depend on being young, but on being 'not middle-aged'.

Has there been an improvement in polling techniques over the last

decade?

The main change over the past ten years is that polling has moved from telephone to online. YouGov was a pioneer of online research in the UK, which has now been replicated by many other organisations. The big turning point came during the 2016 EU referendum, when telephone polls were showing a big lead for Remain whilst online polls were showing a close race. Ultimately it was the online polls that were more accurate.

A good example of a major improvement in the industry is YouGov's recent success with a technique called Multilevel Regression and Post-stratification, or 'MRP' for short. This model correctly forecast over 95 per cent of parliamentary seats in June's UK election. It was updated every day until the vote on the 8th of June. This is something we publicly tested during last year's EU campaign and it always had 'Leave' ahead.

As our CEO Stephan Shakespeare explained, MRP 'works by modelling every constituency and key voter types in Britain based on analysis of key demographics as well as voting behaviour in the 2015 general election and the 2016 EU referendum. Turnout is assessed on voters' demographics and is based on analysis from 2010 and 2015 British Election Study data.

Every day in the run-up to the election, YouGov conducted approximately 7,000 interviews with registered voters about how they planned to vote in their constituency. This data was used to assess how each type of voter was shaping the race in every type of constituency in Britain. From this, the model calculated daily voting intention and seat estimates.

Every day in the run-up to the [UK] election, YouGov conducted approximately 7,000 interviews with registered voters about how they planned to vote in their constituency.

Shakespeare went on to explain: 'The model is based on the fact that people with similar characteristics tend to vote similarly – but not identically – regardless of where they live. Given it is a national model, though, it does not account for specific local factors that may shape the vote in some seats.'

One of the most striking predictions of our model was that Labour would gain the constituency of Canterbury, in the South-East of England. Conservatives have won every election in this constituency since its creation in 1918. In 2015, the Conservative Julian Brazier won 43 percent of the vote, versus 25 per cent for Labour, 14 percent for UKIP and 12 per cent for the Lib Dems. This was far from an

obvious opportunity for a Labour gain, but the model put Labour ahead of Conservatives by 45 to 43 percent. In fact, Labour's Rosie Duffield gained the seat on 45.0 of the vote to 44.7 for Julian Brazier. This prediction came from a combination of Canterbury being a relatively urban and Remain-leaning constituency within its region, and the presence of a large number of students, both of which were associated with Labour gains. We did not have enough survey responses in any single constituency to find constituency-specific swings, but Canterbury was part of a larger pattern across constituencies which these shared characteristics, and which helped our model capture this striking result.

The MRP showed the value of 'big data' and sophisticated modelling to allow truly massive surveys of tens of thousands of voters to create an electoral estimate of far greater accuracy than had ever previously been achieved. For more on how it worked do read my colleagues Doug Rivers and Ben Lauderdale here.

The value of 'big data' and sophisticated modelling to allow truly massive surveys of tens of thousands of voters to create an electoral estimate of far greater accuracy than had ever previously been achieved.

Despite their purported inaccuracy, why are polls still important?

Hopefully by now readers won't feel that polls are that inaccurate after all! But regardless, they are still important because they shed light on what would otherwise be a very closed conversation amongst political insiders and journalists that would cut out the public even more than is currently the case. Polls at their best help decision makers make better decisions and they help the public to express their hopes, fears and indifference. And they help us know who does or doesn't have a Zombie Plan too. That's all the importance anyone should need.



Marcus Roberts
London

Marcus is Director of International Projects at YouGov. He previously worked for the Labour Party in the UK and the Democratic Party in the US as well as for a number of think tanks and campaigning groups across Europe and America. He tweets as @marcusaroberts.