

# **Why Did the Polls get it wrong in the 2015 General Election?**

## **Evaluating the Inquiry into Pre-Election Polls**

**By**

**Paul Whiteley**

**(University of Essex)**

### **Introduction**

The opinion polls got it wrong in the 2015 General Election. The Conservatives won a small overall majority in the House of Commons, contrary to almost all predictions except for the exit poll on the night itself. Most previous polls had predicted a hung parliament, with many indicating that Labour would be the largest party. The average Conservative and Labour vote shares predicted by eight different survey agencies were identical at 33.6 per cent which understated the actual Conservative vote by 4.2 per cent and over-stated the Labour vote by 2.4 per cent in the final pre-election polls. We have to go back to the general election of 1992 to find a comparable set of inaccurate pre-election polls. In contrast the exit poll was spot on, but that had the advantage of interviewing people who actually voted, since the respondents were approached outside the polling stations. The same was clearly not true for participants in the pre-election polls.

This outcome prompted the British Polling Council to set up an inquiry into what went wrong and their preliminary report was published in January of this year<sup>1</sup>. To put this in context there were a total of 1,942 polls conducted in Britain between 2010 and 2015. This compares with the approximately 3,500 polls in the field between 1945 and 2010. So polling exploded in the years of the Coalition government. The report provides a careful analysis of the various factors which might explain what happened and it was written by an experienced team of researchers. They draw attention to a number of factors including unrepresentative samples, problems of weighting the data, biased responses and interviewees who misled

pollsters about their voting intentions. Perhaps not surprisingly they ended up concluding that there is '*no silver bullet, the risk of polling misses in the future can be reduced not removed*' (p.54).

This article critically evaluates the findings of the report, which by and large are sensible and judicious. But it draws attention to two problems which might give a misleading picture of what went wrong, and so cause misunderstandings about what should be done in the future. The first problem relates to mode comparisons: that is, whether a survey was conducted on the internet, by telephone or by interviewers talking to people in their homes. The second relates to respondents misreporting their voting intentions in these polls, something which is rather neglected in the report.

### **How Do Pollsters Make Forecasts?**

Before looking at the problems in the polls it is helpful to explain how pollsters go about making their forecasts of the vote. Firstly, they collect a sample of respondents typically using quota samples in the case of internet surveys and random samples in the case of telephone surveys. A quota sample tries to replicate the characteristics of the electorate in the sample, using information from a variety of sources such as the Census. For example, young males of working age make up a given percentage of the electorate and the polling company will aim to replicate this proportion in the sample. Internet quota samples are drawn from a big panel of potential respondents compiled by the polling companies from many sources, and they are paid a modest amount to participate in the survey. In contrast telephone surveys use a computer to dial numbers at random in order to select the respondents. This type of random sampling faces considerable difficulties in persuading individuals to be interviewed, and the pollsters have to call many telephone numbers before they get a respondent.

Once the data are collected then they are weighted to look like the known demographic profile of the electorate. Weighting means that not all respondents are counted equally. To take an illustrative example, if a survey were to pick up only half of the young people needed to be representative of the electorate, then those who actually do appear in the sample are counted twice. This makes opinions in the survey as a whole more likely to reflect the electorate in general. In relation to turnout, additional weighting often takes place in order to adjust for the probability that an individual will actually vote, using information acquired from questions in the survey. Weighting is a useful device but it cannot make a really bad sample representative of the wider population. When this is all done the forecast of how people will vote on the day can be made.

### **What Went Wrong in 2015?**

The authors of the report stressed that they looked for general explanations of what went wrong rather than specific issues linked to a particular polling agency or method of collecting the data. They dismissed a number of possible explanations such as voter registration problems, question wording biases, distortions introduced by postal voting and survey modes. The problem with many of these explanations is that if they were relevant they would show up in successive general elections, producing problems for all pre-election polls. In fact the polls have been fairly accurate in predicting outcomes, subject to the usual margins of error, in the great majority of elections since the Second World War. This was certainly true for elections between 1997 and 2010.

The report goes on to concentrate on four possible explanations for the failure: late swing, sampling issues, herding behaviour and misreporting. Late swing refers to the possibility that some voters opted for the Conservatives at the expense of Labour and other parties at the last minute, and this was not captured in the polls because they were in the field before it happened. Sampling issues relate to whether or not the surveys were actually

representative of the wider electorate. Herding behaviour refers to the fact that polling agencies assiduously follow each other and if one agency appears to be out of line with its competitors, it may be tempted to readjust its weighting schemes in order to bring them back into line. It should be said that pollsters have a strong incentive to get things right and they do not do this kind of thing in order to deliberately bias the results; but they are often motivated by legitimate concerns that their own surveys might have methodological weaknesses if they are out of line with their competitors. Finally, there is ‘over-reporting’, that is, respondents saying they are going to vote or that they actually voted when they did not, something to be discussed more fully below.

On the question of a late swing, pollsters try to identify this by re-interviewing some of the individuals from their pre-election surveys at the last minute, to see if there is any evidence of a change. They also interview people before and after the election to see if there is a change in their reported behaviour. The conclusion of the report was that there was some evidence of a last-minute swing towards the Conservatives, although it was not very strong and its contribution to the overall error in the polls was fairly modest.

The issue of sampling bias is more complicated and it comes down to the difference between random samples and quota samples. Traditionally the ‘gold standard’ in survey research has been random samples in which respondents are pre-selected using probability theory and the interviewers talk to people in their homes – so-called ‘face-to-face’ surveys. The basic idea of a random sample is that everyone in the electorate has a chance, albeit very small, of being chosen for interviewing. This means that people in hard to reach groups such as old people not connected to the internet can still end up in the sample, making it representative of the electorate as a whole. This is not necessarily true for quota samples.

That said, the report rightly does not recommend that pollsters go to random sample face to face interviewing designs in order to forecast elections. This is because such designs

face formidable problems; they cost a great deal of money and it can take several months to collect the data. The biggest problem, however, is that response rates for these surveys have been declining over time. The first British Election Study conducted by David Butler and Donald Stokes in 1963 had a response rate of 79 per cent, so that almost 8 out of 10 of the people chosen to be interviewed actually participated in the study. In contrast the response rate for the 2015 British Election Study face-to-face survey was 55.9 per cent. This raises the obvious problem that a well-designed probability sample can cease to be representative of the electorate if nearly half of the people chosen to be interviewed don't participate.

The report makes a number of comparisons between the 2015 British Election Study and the British Social Attitudes face-to-face surveys and various internet and telephone polls. These suggest that the former were more accurate than the latter when it came to forecasting the vote. But this is misleading. Firstly, the internet and telephone polls were done before the election whereas the face-to-face surveys were done afterwards. This is problematic because the election itself changes opinions and behaviour. There is a tendency for some respondents to claim that they voted for winners after an election even when they did not. In the 2015 context, this makes support for the Conservatives appear higher in a post-election survey.

A fairer comparison would be to examine the party vote shares in the British Election Study face-to-face survey in 2015 with a post-election internet survey. This can be done using the Essex internet panel survey of 2015 which interviewed the same people before and after the election<sup>2</sup>. The Conservative vote share in this post-election internet survey was 35.7 per cent compared with the actual election vote share of 37.7 per cent, and the Labour vote share was 29.8 per cent compared with the actual result of 31.2 per cent. These discrepancies are all within the margin of error, indicating that a post-election internet survey can be quite accurate.

An additional problem for these comparisons is that the British Election Study face-to-face survey was itself fairly unrepresentative of the electorate. We know this by looking at the percentage of respondents in the survey who actually voted. It is possible to check the turnout records of survey respondents at the local level since these are held in case there is a legal challenge to an election in a particular constituency. These records show whether respondents voted, but they do not of course show how they voted. This exercise showed that 76 per cent of the BES survey respondents actually voted, in an election which had an overall turnout of 66 per cent. As the report points out: *'Having too many or too few of some types of people in sample translates into error in predicted vote if those different types of people also vote differently'* (p.39). This means that comparisons made between the BES survey and pre-election polls are problematic. All of the surveys appear to have too many politically engaged respondents relative to the electorate, because disengaged people are increasingly reluctant to participate in any kind of survey.

On the issue of herding, one way of identifying this is to look at how variable the different polls were during early parts of the campaign compared with the final pre-election polls. The argument is that the range of results across polling companies should be roughly the same over the entire campaign, but if this narrows significantly as the campaign comes to an end that could be evidence of herding. One objection to this is that voters might be reaching a consensus as the campaign progresses which would itself produce a narrowing of the range across polls. But this can be addressed by applying a common weighting scheme to the polls instead of the different weighting schemes that the pollsters actually use.

Herding means pollsters adjusting the weighting schemes, and so if convergence between polls is driven by the voters rather than by differences in the weighting schemes it will look exactly the same in the common scheme as in the separate schemes. However, if this comparison throws up differences, it suggests that weighting schemes are being adjusted

in a way consistent with herding behaviour. Of course it is possible to ask pollsters if they did actually adjust their weightings in light of what their competitors were doing after the event, but it is not clear that such a question will get a straight answer. There is in fact evidence to support herding in the data, but it is modest and certainly not enough to explain why the polls got it wrong.

The final issue of over-reporting was referred to in the report but not analysed at all. This is quite an important issue, and is examined next.

### **Do People Lie to Pollsters?**

It has long been recognized that some people lie to interviewers about their turnout. An extensive report on ‘over-reporting’ as it is called was published in 1968 in the United States. This used data from the American National Election Studies, which at the time validated the electoral participation of their respondents. Over-reporting means people claiming to have voted when they did not, and there is a second group of under-reporters, or those who claimed not to have voted when the records show that they actually did (a much smaller group). The former arises primarily from people misleading interviews and the latter more from people forgetting what they did on polling day, which is quite possible when face-to-face surveys take months to complete. The overall conclusion of the report was that:

*‘Estimates of turnout have consistently exceeded the population figures by about 12 to 13 per cent’.*<sup>3</sup>

Over-reporting of turnout occurs for a number of reasons. Firstly, there is a ‘social desirability bias’ causing respondents to lie about their turnout because they want to appear to be good citizens in the eyes of the interviewer. Secondly, there is a ‘spiral of silence’ argument which suggests that voters will mislead interviewers about their turnout if they think that public acknowledgements of voting for a party will be frowned on by others if that party is unpopular at the time. This idea gave rise to the concept of ‘Shy Tories’ in the 1992

general election, in which the polls underestimated support for the Conservatives in the pre-election surveys. These Tories would tell the interviewers that they were not going to vote, and so avoided answering the follow-up question about which party they would support in the election. Thirdly, there is a more recent concept of ‘Lazy Labour’ supporters in which respondents said they were going to support the party in pre-election polls but then did not vote on the day. This had the effect of over-estimating support for Labour prior to the election. These are interesting ideas but a multivariate analysis of survey data from the 2010 election showed little support for them.<sup>4</sup>

A fourth explanation suggests that people who are under pressure to vote from friends, family, political parties and others are more likely to lie to interviewers about their participation if they fail to vote. This is prompted by a sense of guilt at not responding to the pressure to vote in the first place. In effect, failing to live up to one’s normative beliefs creates pressures for dishonesty. In the United States some early findings showed that black respondents were more likely to over-report than whites because they felt guilty about letting down their fellow ethnics by failing to vote. A similar point can be made about people who perceive that big differences exist between the political parties. If they feel this is true but end up not voting then they are more likely to over-report. Both of these happened in the 2010 general election.

All of the British Election study post-election face-to-face surveys conducted since 1992 have validated the voting behaviour of their respondents. This makes it possible to identify any trends in over-reporting which might be occurring during a period of almost a quarter of a century. We will focus on the respondents who claim to have voted but did not do so, rather than those who forgot that they had participated in the election. Figure 1 shows trends in the percentage of respondents who over-reported in successive British Election Study surveys from 1992 to 2015. To reiterate, this refers to the difference between the



validated turnout that is, the percentage of respondents in the surveys who actually voted, and the percentage who claimed to have voted when they were asked about it. A second measure in the Figure looks at the difference between the percentage of respondents in the surveys who actually voted and the national turnouts in these elections. Clearly if the surveys are accurate then the percentage turnout in the survey should be approximately the same as the turnout in the general election. So looking at validated turnout is a way of measuring the accuracy of the surveys in replicating the turnouts in these elections.

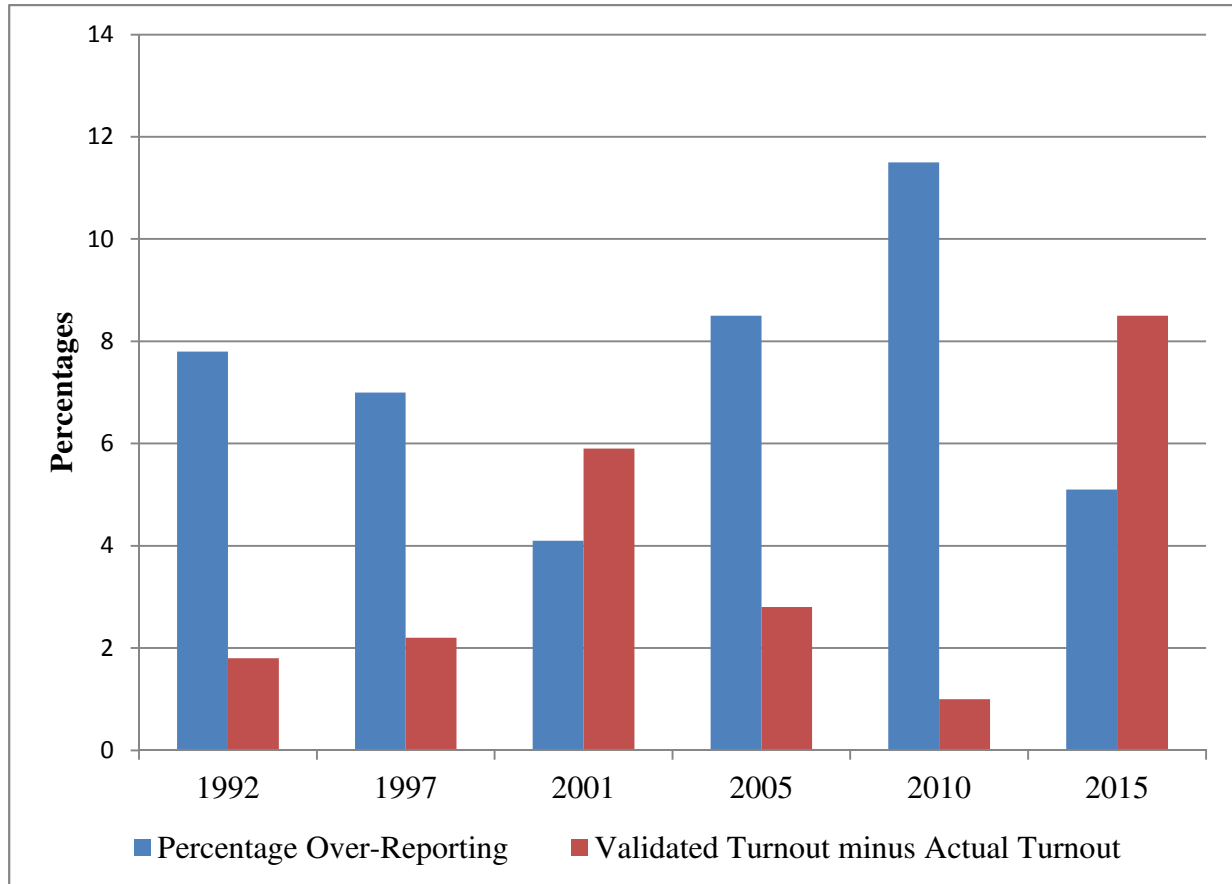
■ Figure 1 about here --

The figure suggests that between the 2001 and 2010 elections there was a trend increase in over-reporting in the surveys. In 2001 some 4.1 per cent of respondents were over-reporters and this rose to 11.5 per cent by 2010. However, in the 2015 survey the trend appeared to reverse itself and over-reporting fell back 5.1 per cent. Yet this conclusion is misleading, since we need to take into account how accurate the surveys were in measuring actual turnout in the election. This is where the discrepancy between the actual turnout of survey respondents and the national turnout in the election comes in. As already indicated, the 2015 survey was much less representative of actual levels of turnout than the earlier surveys. By definition over-reporters are non-voters and so clearly if a survey contains too many validated voters this will reduce the number of over-reporters – because it means that there are fewer non-voters than there should be in the sample. If we take this into account then it is not clear that over-reporting declined significantly in 2015. Overall the data tends to support the idea that over-reporting is growing in Britain over time. People are more likely to lie to interviewers about their participation in the election than used to be the case in the past.

## **Discussion and Conclusions**

It is important to get the performance of the polls in the 2015 general election in perspective. Internet and telephone polls have been pretty successful in forecasting the results of the great majority of general elections since the war, and they have been equally successful in predicting the results of mayoral elections in London and elections for the devolved assemblies in Wales and Scotland. Statistical theory predicts that every so often we are going to get a rogue result just by chance and it certainly looks like the general election of 2015 was one of these. That said, if there is a trend increase in people refusing to participate in surveys which to some extent mirrors the decline in electoral turnout that has occurred since the 1990s, and the people who do participate are more likely to lie to the pollsters, then it is going to get more difficult to predict the outcome of elections in the future.

**Figure 1 Trends in Over-Reporting and in the Validated Turnout of Respondents in British Election Surveys 1992 to 2015**



<sup>1</sup> <http://www.britishpollingcouncil.org/general-election-7-may-2015/>

<sup>2</sup> See Harold Clarke, Peter Kellner, Marianne Stewart, Joe Twyman and Paul Whiteley, *Austerity and Political Choice in Britain*, 2015 London: Palgrave for a full analysis of this data.

<sup>3</sup> Clausen, Aage R. 1968-1969. Response validity: vote report. *Public Opinion Quarterly* 32 (Winter 1968-1969): 588-606, p. 589.

<sup>4</sup> See Paul Whiteley, 'Why Do People Lie to the Pollsters?' *Political Insight*, April 2016: 16-19.