# Why the pre-election polls get it so wrong: Is it time to take probability sampling seriously?

**blogs.lse.ac.uk**/politicsandpolicy/why-the-pre-election-polls-get-it-so-wrong-is-it-time-to-take-probability-sampling-seriously-2/

Many of the lessons from the polling debacle of 1992 have been learned, but it may be time to address the underlying causes of the error rather than just treating the symptoms, writes **Oliver Heath**.

The last time the pre-election polls got it this wrong was 1992 – when the polls predicted a slight lead for Labour only for the Tories to win the popular vote by some 7 percentage points. Since then the polling companies have changed the way that they carry out and report their polls, but reflecting on what went wrong in 1992 may shed some light on what happened this time out, and whether the lessons of previous mistakes have truly been learned.

Many of the alleged culprits for the 1992 debacle are similar to those being bandied about to explain the 2010 debacle, and the curious phenomenon of the 'shy tory' has once again reared it's timid head. After 1992, like now, an inquiry was set up to investigate the sources of the error. The report to the Market Research Society (MRS) by the working party on the opinion polls and the 1992 general election was published in 1994, following an earlier interim statement published just a few months after the election. In addition a number of academic studies analysed the performance of the 1992 election polls and examined the sources of error.

The MRS report concluded that three main factors contributed to the error in the 1992 polls: late swing, disproportionate refusals and inadequacies in the quota variables used. All these source of error went in the same direction, resulting in an underestimation of the Conservative vote and an overestimation of the Labour vote.

However, the 1992 report received a somewhat critical reception and a number of later academic studies argued that the report failed to get to the heart of the problem. The report highlighted inadequacies in the implementation of quota sampling – but as Roger Jowell and colleagues argued in a series of articles in the 1990s – did little to address the inadequacy of quota sampling per se. Even back then people within the academic community were urging the pollsters to think more seriously about selection and non-response bias, and to consider the adoption of probability sampling methods in order to reduce these biases.

So what has happened since then? One major change has been to do with the quota variables that pollsters use and the variables they use to weight their data. The MRS report concluded that the 1992 election poll samples were biased primarily because some quota targets were inaccurate (to do with the size of the working class population) and the choice of quota variables was inadequate, in that the variables used were often not very strongly associated with vote choice (such as age and gender). The polling community has taken both of these criticisms on board, and now routinely adjust their data according to political variables, such as recall of past vote or party identification.

Unsurprisingly these variables are very strongly associated with vote intention, some might argue too strongly, and thus in times of political flux they may provide a rather unstable base for adjustment. But perhaps more importantly, these types of quota or weighting variables still focus to a certain extent on addressing the symptoms of the error rather than the cause.

To a certain extent pollsters have had some success at treating the symptoms, in that their attempts to correct for error probably result in more accurate estimates of vote intention than if they simply did nothing. One way to further improve the accuracy of polls would therefore be to further think about how they can better correct for error, perhaps by considering more or better quota and weighting variables. But a more radical solution is to try and think of ways in which there is less error to correct for in the first place.

One way to think about this is to consider what would happen if the pollsters did nothing to adjust their data, and simply selected samples using random probability. The telephone polls have reasonably good coverage of the adult population – though with the decline of landlines this coverage probably isn't as good as it was in 1992 – but they achieve very low response rates, which can lead to biased samples. By contrast internet polls, that rely on a panel of volunteers who are paid to complete surveys, have very poor coverage of the adult population and are not at all representative of the adult population (though the exact composition of these panels are shrouded in secrecy). A purely random sample would therefore also be biased, even before non-response is taken into account.

The accuracy of both types of poll thus depend upon the adjustments that are made to the data, either after it has been collected in the form of weighting or before the sample has been selected in terms of quotas. The more data has to be adjusted, the more scope there is for something to go wrong. The resulting estimates of vote intention are therefore based more on a predictive model of what pollsters think public opinion looks like than they are based on an actual snapshot of what the public really think.

After the 1992 debacle pollsters were urged to think about how they select their samples and to consider methods of random selection. If the polling industry is to regain credibility among the general public, media, politicians and academics it will need to address this issue seriously. With the rise of internet polling there is now scope to do this, but it will cost money. A number of online surveys in Europe have experimented with recruiting panel members through random selection rather than relying on volunteers, and although the start-up costs are high – it may be a price worth paying.

*Note: This article gives the views of the author, and not the position of the General Election blog, nor of the London School of Economics. Please read our comments policy before posting. Featured image credit: Impru20 CC BY-SA 3.0*

**About the Author**

**Oliver Heath** is a Reader in Politics at Royal Holloway, University of London.