Opinion Polls - Myth Versus Reality (2.0)

Earlier this year I published an article in Linkedin where i explained how and why the main opinion polls got it wrong at the federal election.

I recently came across an excellent article from Data Analysis Australia (DAA) that actually backs up what I said. You can find it here …..

Here are my original reasons on what happened and why, plus the analysis from the DAA article which I have *highlighted in italics* for the benefit of comparison:

- Australia is becoming increasingly a politically diverse country, so for example, an issue such as the Adani Coal Mine will play out differently in Queensland than it would here in Victoria.

- The problem is that the main pollsters typically survey the entire country, so unless their sample is large enough, it is possible that the issues such as this one would not have become as obvious as they should have done.

Polls are a form of survey, where a modest number of people - typically from hundreds to several thousand – are asked about their voting intentions. If the sample of voters is properly selected, statistical theory tells us that certain accuracy can be achieved. The key to this is that the sample needs to be representative of the wider population.

- Then we get what is called **non response error**.This occurs when large sections of the population are not, or cannot, be effectively sampled and surveyed. So the representative quality of the sample declines. This has occurred before on recent occasions, think Trump and Brexit. The most likely cause for this problem is the decline in the number of landlines, as these people cannot easily be located. Let alone contacted for a survey.

  Until about ten years ago most polls used telephone samples, based on the assumption that most households had a landline telephone. So called random digit dialling could be used to overcome problems of unlisted numbers and all numbers mapped to geographical areas so a survey could be readily targeted. Today this is no longer a workable assumption, with around 40% of

adults being "mobile only", and many people having multiple phones.

Response rates to surveys have gone down increasing the biases from certain people avoiding surveys. This could be due to a number of reasons including concerns about security and confidentiality and that people are getting inundated with spam and don't want to participate anymore.

- It appears that the major parties were picking up trends in for example Queensland, that the main pollsters missed. Now this may just be the political parties being wise after the event, but let us take them at their word. So they saw something the main pollsters missed. Incidentally the exit polls were also very wrong as well.

- Here we bring **stratified sampling** in to play. This is a sampling method where we drill down into a specific location or strata. For this purpose we can look at individual electorates as strata areas and if these were sampled and surveyed individually, then we get a more detailed picture. As long as the sample was of a reasonable size, so around 500 people, then the chances of accurate polling are very good. So if the major parties did take this approach they would have clearer picture than the mainstream poster who were more focussed on the "bigger picture" across the nation.

Over the years, better methods of sampling and better estimates that have been developed, including Stratified Sampling - Where the population is partitioned into sub-populations and different sampling rates applied to each is a well-known method that can improve estimates. For example, it might be reasonable to assume that party allegiance varies from state to state, so treating the states as strata and then having different sampling rates in each state may improve sampling efficiency.

- Robo polling doesn't help either. As you are not talking to real person you can literally say anything you like when you key in your response. We find it harder to lie to people, so if people said one thing to the robo poll and then voted differently, that should not come as a surprise. I appreciate that survey costs are an issue, but this method of polling (in my opinion) is highly vulnerable to respondent manipulation. For example a person can if they so

choose, deliberately send in an answer designed to influence the results of a survey in thier local area to suit their own political views. If this becomes widespread, the impact on the survey results (especially within marginal electorates) can be profound.

Below is what DDA thinks of this type of technology.

Data Analysis Australia has always taken a conservative approach to surveys, aiming to achieve good sampling and good estimation. This has led to our work focusing on projects where quality is paramount, such as in legal matters. This does not mean we do not use the latest technology, rather we do not let technology dominate to the exclusion of quality.

That sums things up perfectly in my view.