

Why the Polls Got It Wrong.

Originally posted on OurRVAdventures.com 11/13/2016 by Greg White

Now I'm about to step in it, or as Jan says, "ride that ragged edge."

I've always been crazy about math, pretty much any math, but especially odds, statistics, and sampling. So during election cycles, I'm in hog heaven.

I saw a lot of references to people being 'surprised' at the recent election results. And if you only paid attention to the major news outlets, you certainly were.

On Monday night the –

NBC/WSJ poll showed Clinton up by 4 points.

NBC/Survey Monkey poll showed Clinton up by 6 points.

Bloomberg poll showed Clinton up by 3 points.

Fox News poll showed Clinton up by 4 points.

CBS News poll showed Clinton up by 4 points.

LA Times poll showed Clinton winning the election with 352 electoral votes.

So where did it all go wrong?

The main reason is over-sampling, with a modicum of over-confidence, and maybe a dash of wishful thinking..

Polling companies don't just go out and ask 1000 people who they're going to vote for. That would be too easy. And probably not as quite as accurate.

So instead of just asking 1000 random people, they are selected by 'weighting', i.e., filtering for age, sex, income, race, etc. And it's not always 1000 people. It may be as many as 2000 and as few as 800.

But one of the biggest determiners of polling results coming from three things. Whether the person is listed as a voter, a registered voter, or a likely voter. Each level of voter is more accurate, but more expensive to find.

A voter's choice of candidates is not as accurate as a registered voter's which is not as accurate as a likely voter's.

So why don't all the polling companies use only likely voters and be done with it?

One word. Money.

Every year it costs more and more to select the pollee's that meet all the needed categories. And the biggest reasons for that problem is cellphones.

Twenty years everybody had a landline and the vast majority were in the phone book. So it was quick and easy to reach them and filter for your requirements. But now landlines are on the wane, along with phonebooks.

Instead, everyone has a cellphone and screens their calls. In our case, if a call doesn't come up in my Contacts, I don't answer it. If they want me, they need to leave a voicemail.

So the polling companies have to go through many more people before they find their sample.

So where does 'over-sampling' come into all this. Glad you asked.

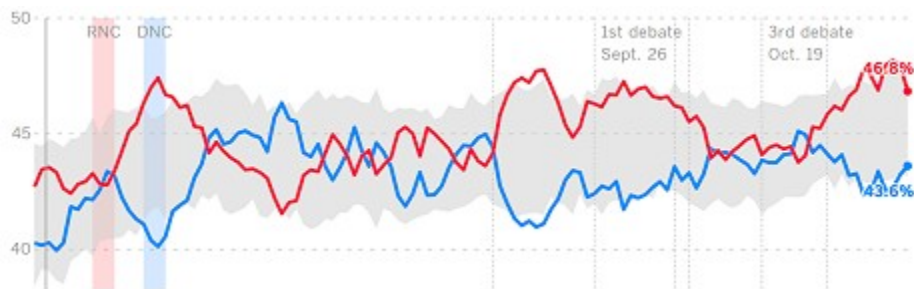
The other category they track is whether or not the person is a declared Democrat, Republican, or Independent. And because Democrats traditional vote in higher percentages than Republicans, the polling companies include more Democrats than Republicans in their sample.

How many you might ask? Aye, there's the rub.

Obviously the more Democrats in your sample, the further ahead the Democrat candidate will seem. And there's no standard amount of extra Democrats. Each polling company has their own ideas about turnout and that varies from election year to election year.

So if a polling company has reason to want one candidate to appear to be ahead of another, they just add more of that candidate's voters in their numbers. Supposedly some companies were polling Democrats +20 over Republicans.

But there were a few polls that turned out to be very accurate this year. The main one was the Dornsife/USC/LA Times Daybreak poll.



Note that this poll shows Trump ahead since the middle of September, except for a few days on and off in October. And a number of other 'outlier' polls showed similar results.

So how can do you find out whether or now they're sampling voters, registered voters, or likely voters? Or how much they're over or under sampling one party or the other?

Well, it's all there, buried in the fine print, the pages and pages of details that accompany the poll information. But you've really got to like math to dig it out.

But if you want the most accurate polls, check with the campaigns themselves. They do their own internal polling, trying to be as accurate as possible, with no fudging, hedging, or fibbing.

And this is why the Trump campaign seemed to be more and more confident in the last week before the election.

And why at the same time, the Clinton campaign, who as of Oct. 31st had a fireworks show scheduled for 9:30pm election eve on a barge in New York Harbor, abruptly canceled the show the morning of Thursday, Nov. 3rd.

So what did they know and when did they know it?

But don't think all these wrong polls are anything new. In the days before the election in 1980, some polls showed Carter ahead of Reagan by as many as 8 points.

Reagan won by 10 points.