Review of TDMS analysis of South Carolina 2020 Primary

Ray Lutz, CitizensOversight 2020-03-09

On March 2, 2020, Theodore de Macedo Soares (TDMS) published an article entitled "SOUTH CAROLINA 2020 DEMOCRATIC PARTY PRIMARY Exit Poll Versus Reported Vote Count" which can be found at this url: http://tdmsresearch.com/2020/03/02/south-carolina-2020-democratic-party-primary/

This article has been circulated with the notion that "the disparities exceed the exit poll's margin of error" and includes the figure 8.3% "Gain/loss from exit poll." This article has been further circulated in social media with the notion that this analysis proves that the vote count has been modified.

As I have done extensive simulation of risk-limiting audits, my simulator can be easily applied in this case to provide a graphical representation to provide better idea of whether the difference between the exit poll and the results are indeed a concern. This should add more understanding of these results.

Monte Carlo Simulation

The simulator essentially runs many possible random exit polls of actual voters. It models all 528,726 voters in the district, based on how they actually voted in the official results. The simulator runs 10,000 simulated exit polls, assuming an unbiased random sample. In reality, exit polls do not sample randomly and may be concentrated in a few polling places rather than everywhere, and therefore include unintentional bias. In this case, absentee (vote-by-mail) ballots were not included in the audit at all, at least as it is conducted by this auditing firm.

The article by TDMS went into some detail about the complexity of a contest with 8 contestants, Biden, Sanders, Steyer, Buttigieg, Warren, Klobuchar, Gabbard and "others," but the easier way to analyze this is just to consider two contestants at a time, for example Biden vs. the non-Biden candidates (the "pool"). Each voter can vote for only one candidate, and that voter will thus either vote for Biden or not. This same approach can be repeated for each candidate of interest.

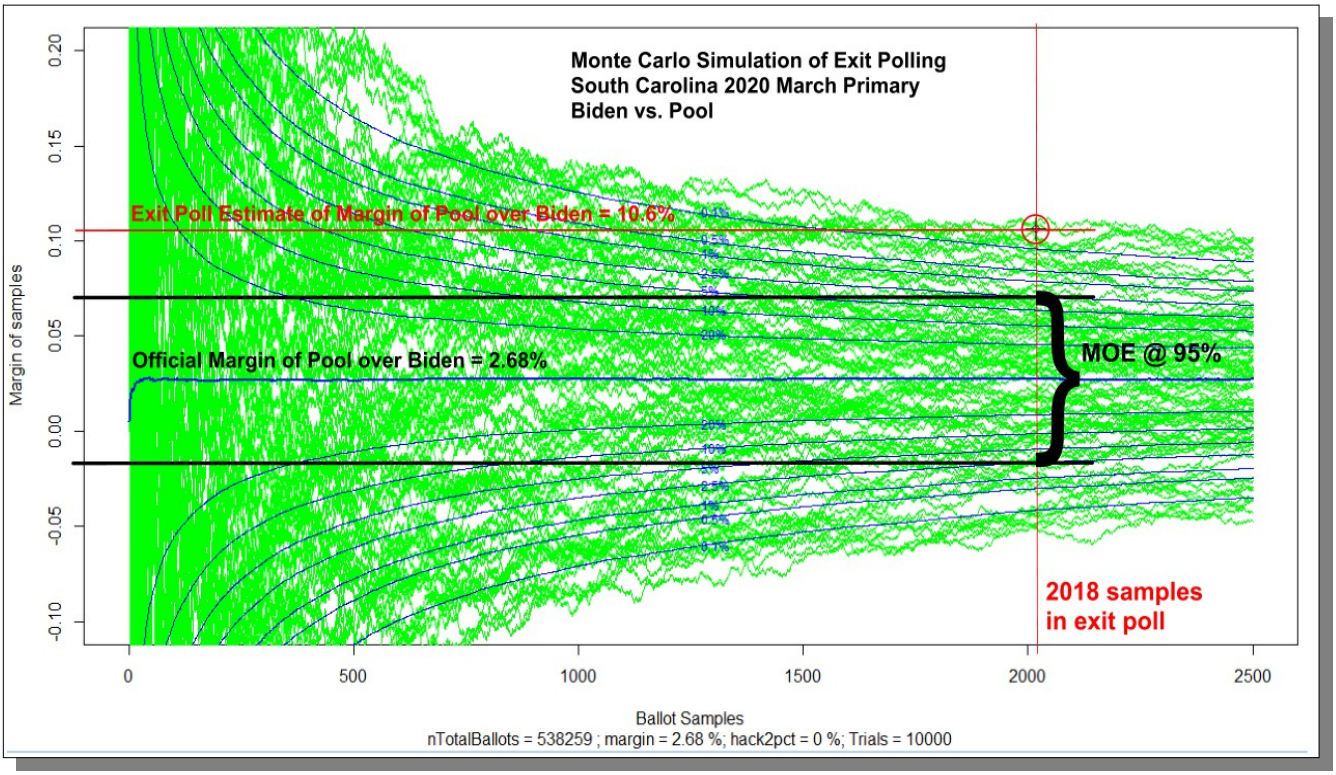
Biden

In the official results, the pool beats Biden by 2.68%. Biden received 261,897 votes and the pool received 276,362 votes, a vote margin of 14,465 votes, which is 2.68% of the total votes. The exit poll had 2018 samples, of which 902 votes were for Biden and the pool received 1116, for a margin of 214 votes, giving the pool a 10.6% lead over Biden. The difference in the sample margin and the actual margin is 10.6 - 2.68 = 7.92% but that does not take into account the margin of error of the exit poll.

The illustration below is the simulation result. The horizontal axis is the ballot samples included in the sample for any given exit poll, and the vertical axis is the margin of the samples at that point in time. Each green line that you see is a simulated exit poll. We know that at first, the first person asked will vote either for Biden or not-Biden, and thus all curves will start at the margin of -1.0 and +1.0. If the second sample is the opposite value, then the next point is at 0, etc. This plot shows what would happen

if we actually sampled 10,000 voters, and we repeated the simulation 10,000 times. As more samples are accumulated, the margin of those samples will converge to the official margin, but at any point in a simulated exit poll, the margin is determined by the random accumulation of voters at that point.

We assume in this plot that the voters actually did vote the way the official results claims, with a margin of 2.68% of the pool over Biden, which is the middle blue line (which is the actual average of all 10,000 exit polls, and that is why it is somewhat erratic at first). This sort of plot is essentially a law of physics, they always look the same. In this case, if you polled all voters, it would converge to the value 2.68% as the final 538,259th voter is averaged in. This is also sometimes referred to as the law of large numbers, because although the dispersion is great at first, it converges relatively quickly and will be very close to the true margin long before we get to the end.



The blue contour curves that you see are a way to show the "height" of the distribution at each point. Between the most-central two blue 20% curves indicates that 20% of the exit polls will be lower or higher than those curves, i.e. 40% of the time, whereas between the two, 60% of all exit polls will be confined to that region. The 2.5% contours provides the point where 95% of all exit polls will exist within those contours, since only 2.5% are higher and 2.5% are lower. The point where these contours cross the samples in the exit poll (2018) establishes the margin of error (MOE), which is traditionally taken as the limit of the distribution of 95% of the possible exit polls, and is with respect to the sampling error only, and not due to the many other inadequacies of exit polls. The MOE at the 2.5% contours is shown. The red circular target is the exit poll result.

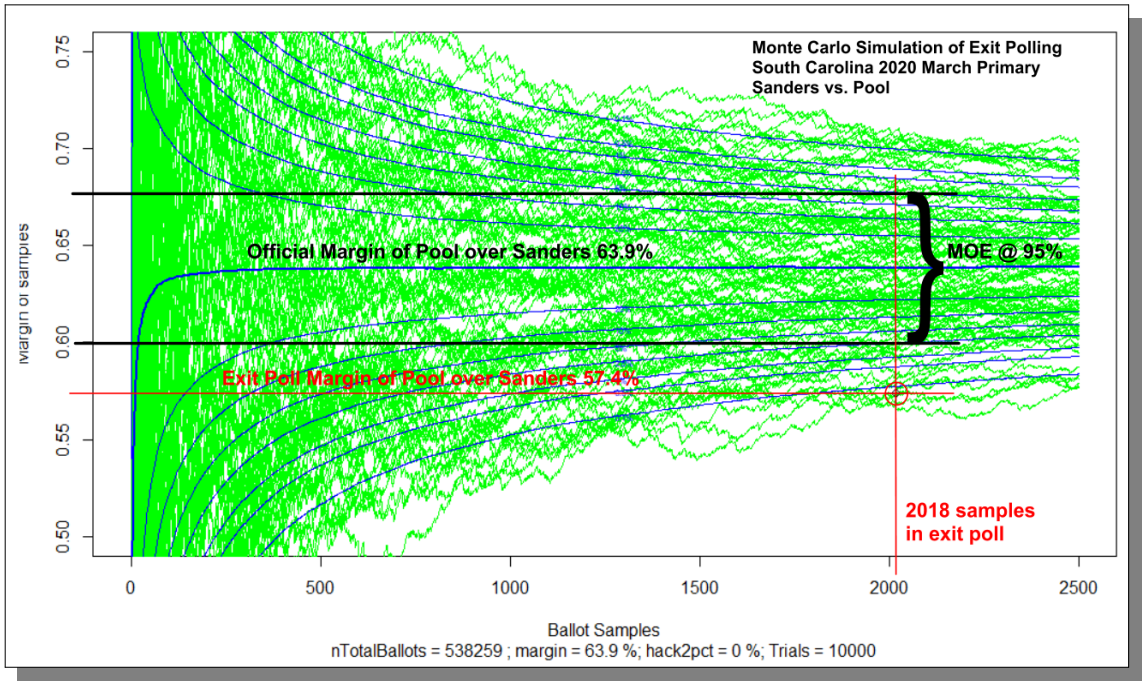
We can see that it is far from impossible for the exit poll to get this result as we see at least a few curves at that point, but indeed, it is not likely from a purely sampling margin of error standpoint.

If the exit poll estimated say 15% margin over Biden, you can see that (absent bias) there is no real likelihood that the poll could get that value, as no curves even when run 10,000 times ever gets that high after 2018 samples. The same would be true if the poll said that Biden would beat the pool by say 10% as it is completely out of range at the -0.10 point (where the text "2018 samples in exit poll" is). What you can see here is that the 95% margin of error is simply the area where we have a 95% expectation of being in that range, and 1 out of 20 times, the exit poll could be outside that range, and you can see here that there are quite a large number of exit polls simulated beyond the traditional MOE limits.

In these diagrams, not all 10,000 trials are shown, only those that reach either a maximum or minimum in the interval modeled. This lightens up the center (which is otherwise solid green) and gives us an idea of how much the sampled margins wander. We can see in this plot that although the prediction made by the exit poll is outside the 95% MOE, it is still a plausible outcome even without any considerations regarding sampling bias.

Sanders

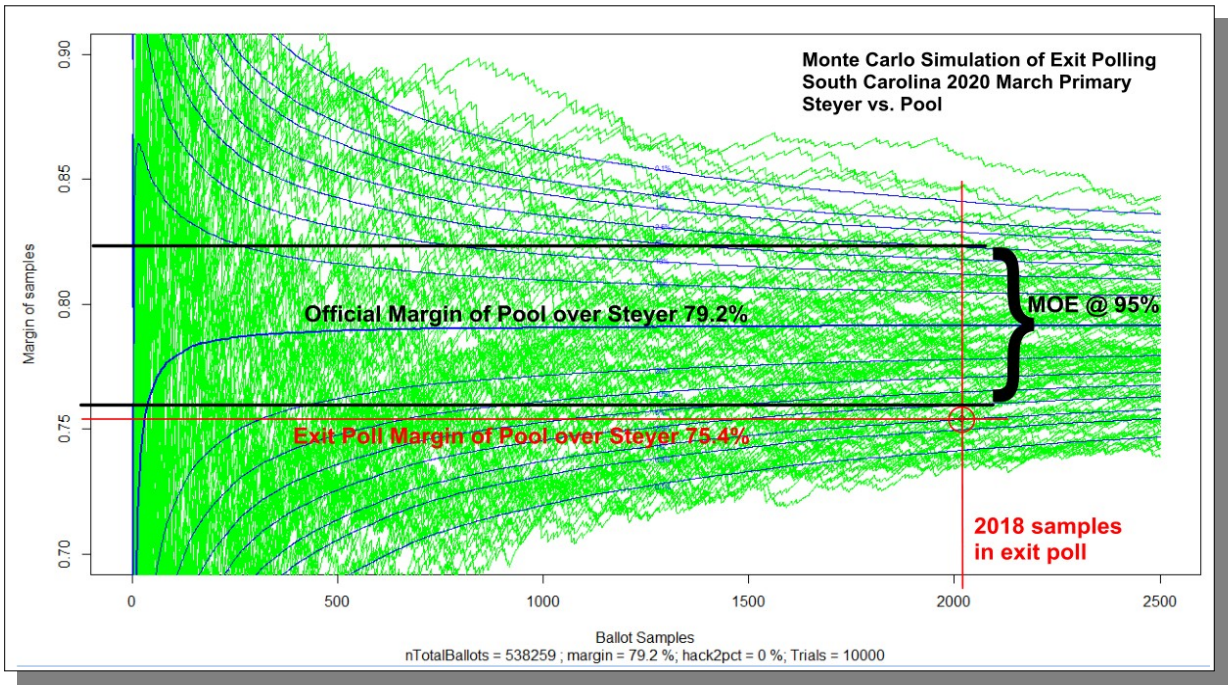
Bernie Sanders received 106,342 votes, 19.8% share, while non-Sanders votes were 482,917. The margin of victory is votes or 63.9%. The exit poll had Sanders at 21.3% out of 2018 votes, or 430 votes, and the vote margin of 1158, and the exit poll margin of victory of pool over Sanders was 57.4%.



Steyer

Steyer, the third highest ranking candidate, went down from the predicted 12.3% share to 11.3%. Steyer received 61,048 votes, while non-Steyer votes totaled 528,211. The margin of victory was 467,163

votes or 79.2%. The exit poll had Steyer at 12.3% out of 2018 votes, or 248 votes, the vote margin was 1522, and the exit poll margin of victory of pool over Steyer was 75.4%.



Conclusion

The exit poll results in the March 2020 South Carolina primary are at the extreme limits of what are possible results from exit polls, assuming there is no sampling bias. However, they are not so extreme as to be implausible, even assuming non-biased sampling. If we include the fact that sampling bias is unavoidable in exit polls, these results suggest that we must be aware that exit polling always has a sampling bias which must be added to the "margin of error." The results should not automatically trigger an assumption that the election has been "fixed" or any election fraud exists. Nevertheless, all elections should be audited.

About Ray Lutz

Ray Lutz holds a Master's degree in electronic and computer engineering and has significant industry and standards experience in document processing equipment, including printers, scanners, facsimile, imaging, etc. Ray founded Citizens Oversight in 2006 and is heavily involved in election integrity oversight, and is developing a cloud-based ballot-image auditing system, "Audit Engine."

Contact Information:
raylutz@citizensoversight.org