

Election polls are 95% confident but only 60% accurate, study finds

October 27 2020



Credit: Unsplash/CC0 Public Domain

How confident should you be in election polls? Not nearly as confident as the pollsters claim, according to a new Berkeley Haas study.

Most election polls report a 95% [confidence](#) level. Yet an analysis of 1,400 polls from 11 election cycles found that the outcome lands within the [poll](#)'s result just 60% of the time. And that's for polls just one week before an election—accuracy drops even more further out.

"If you're confident, based on polling, about how the 2020 election will come out, think again," said Berkeley Haas Prof. Don Moore, who conducted the analysis with former student Aditya Kotak, BA 20. "There are a lot of reasons why the actual outcome could be different from the poll, and the way pollsters compute confidence intervals does not take those issues into account."

Many people were surprised when President Donald Trump beat Hillary Clinton in 2016 after trailing her in the polls, and speculated that polls are getting less accurate or that the election was so unusual it threw them off. But Moore and Kotak found no evidence of declining accuracy in their sample of polls back to 2008—rather, they found consistently overconfident claims on the part of pollsters.

"Perhaps the way we interpret polls as a whole needs to be adjusted, to account for the uncertainty that comes with them," Kotak said. In fact, to be 95% confident, polls would need to double the margins of error they report even a week from election day, the analysis concluded.

As a statistics and computer science student on an undergraduate research apprenticeship in Moore's Accuracy Lab during the 2019 [presidential primary](#), Kotak grew curious about the [confidence intervals](#) included with polls. He noticed that polls' margin of error was frequently mentioned as a footnote in news articles and election forecast methodologies, and he wondered whether they were as accurate as their margins of error implied they should be.

Kotak brought the idea to Moore, who studies overconfidence from both

a psychological and statistical perspective. Much of the research on polling accuracy considers only whether the poll correctly called the winner. To gage poll confidence, they decided to take a retroactive look at polls based on how long before an election they were conducted, and consider not whether a candidate won or lost, but whether the actual share of the vote fell within the margin of error the poll had reported. For example, if a poll showed that 54% of voters favored a candidate, and it had a 5% margin of error, it would be accurate if the candidate garnered 49% to 59% of the vote, but would be a miss if the candidate won with more than 59% of the vote (or less than 49%).

Moore and Kotak obtained 1,400 polls conducted ahead of the general elections of 2008, 2012, and 2016, as well as the Democratic presidential primaries in Iowa and New Hampshire from 2008 and 2016 and the Republican primaries in the same states from 2012 and 2016. Because some polls asked about multiple candidates, the sample included results of over 5,000 surveys of how people said they'd vote on particular candidates, as well as the accompanying margins of error.

Analyzing the polls in seven-day batches, they found a steady decline in accuracy the farther from an election the poll was conducted, with only about half proving to be accurate 10 weeks before an election. This makes sense, since unforeseen events occur—such as former FBI director James Comey announcing an investigation into Clinton's emails just a week before the 2016 presidential [election](#). Yet most polls, even weeks out, reported the industry standard 95% confidence interval.

Sampling error and confidence intervals

The confidence interval quantifies how sure one can be that the sample of people surveyed reflects the whole voter population. A 95% confidence interval, for example, means that if the same sampling procedure were followed 100 times, 95 of those samples would contain

the true voter population. Therein lies the problem, however.

The confidence level takes into account "sampling error," a statistical term that quantifies how likely it is that by pure chance, the sample varies from the larger population of voters from which the sample was drawn. For example, not surveying a large enough group of voters would increase the sampling error. But sampling error does not include any other kinds of errors—such as surveying the wrong set of people to begin with.

"People often forget that margins of error for polls only capture the statistical sources of error," said David Broockman, an associate professor in Berkeley's Department of Political Science. "This analysis shows just how large the remaining non-statistical sources of error are in practice."

Added Prof. Gabriel Lenz, also of Berkeley Political Science, "This is a fascinating analysis, and future work could sort out the sources of the inaccuracy, such as low-quality pollsters, difficulty screening likely voters, last-minute changes in voter intentions, and more."

It's easy to take sampling [error](#) into account in polling statistics, but much harder to account for all the other unknowns, Moore said. It's a lesson that goes far beyond polling.

"Because we base our beliefs on imperfect and biased samples of information, sometimes we will be wrong for reasons that we did not anticipate," he said.

More information: Public Election Polls are 95% Confident but only 60% Accurate, psyarxiv.com/rj643

Provided by University of California - Berkeley

Citation: Election polls are 95% confident but only 60% accurate, study finds (2020, October 27)
retrieved 26 June 2024 from <https://phys.org/news/2020-10-election-polls-confident-accurate.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.