



AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Task Force on 2020 Pre-Election Polling:

An Evaluation of the 2020 General Election Polls

Josh Clinton (Chair), Vanderbilt University¹

Jennifer Agiesta, CNN

Megan Brenan, Gallup

Camille Burge, Villanova University

Marjorie Connelly, AP-NORC Center

Ariel Edwards-Levy, CNN

Bernard Fraga, Emory University

Emily Guskin, Washington Post

D. Sunshine Hillygus, Duke University

Chris Jackson, Ipsos

Jeff Jones, Gallup

Scott Keeter, Pew Research Center

Kabir Khanna, CBS News

John Lapinski, University of Pennsylvania

Lydia Saad, Gallup

Daron Shaw, University of Texas

Andrew Smith, University of New Hampshire

David Wilson, University of Delaware

Christopher Wlezien, University of Texas

This report was commissioned by the AAPOR Executive Council as a service to the profession. The report was reviewed and accepted by the AAPOR Executive Council. The opinions expressed in this report are those of the author(s) and do not necessarily reflect the views of the AAPOR Executive Council. The authors, who retain the copyright to this report, grant AAPOR a non-exclusive perpetual license to the version on the AAPOR website and the right to link to any published versions.

¹ The Task Force thanks Sarah Lentz (University of Pennsylvania), Rezwana Uddin (NBC News), Mellissa Meisels (Vanderbilt University), Sara Kirshbaum (Vanderbilt University), Ami Ikuenobe (University of Pennsylvania), and Grayson Peters (University of Pennsylvania) for outstanding research assistance. The data collection and analysis would not have been possible without their careful and exhaustive work on behalf of the Task Force. In addition to the Task Force members, we would also thank Eunji Kim (Vanderbilt University), Steve Rogers (St. Louis University), John Sides (Vanderbilt University), Alan Wiseman (Vanderbilt University), and seminar participants at the Center for the Study of Democratic Institutions at Vanderbilt University for thoughtful and constructive feedback on an earlier version.

EXECUTIVE SUMMARY

The November 3, 2020, presidential election was historic by many standards, most notably because it was conducted during a global pandemic that resulted in a record high proportion of voters casting their ballots early and by mail. The election also featured the highest rate of voter turnout in decades.²

Most national polls accurately estimated that President Joe Biden would get more votes than President Donald Trump nationally, but Biden’s certified margin of victory fell short of the average margin in the polls at both the national and state levels. Polling overstated support for Biden relative to Trump, and Biden’s 306-232 victory in the Electoral College was narrower than predicted by many election forecasters.³

The shadow of 2016 hung over the 2020 election. Presidential election polls were widely criticized for “getting it wrong” in 2016. Many believed the polls predicted the wrong winner although the 2016 national polls were remarkably accurate in estimating former Secretary of State Hillary Clinton’s national popular vote margin. Still, polls in a handful of key states overstated Clinton’s lead or underestimated Trump’s lead, leading to widespread discounting of Trump’s Electoral College chances by election observers and rendering his victory a shock to many. The state-level polling error became the focus of a task force the American Association for Public Opinion Research (AAPOR) convened in spring 2016 to evaluate the accuracy of that year’s polls, which resulted in very specific methodological recommendations for political polling (Kennedy et al. 2016).⁴

In October 2019, the Executive Council of AAPOR proactively convened a task force to examine the performance of pre-election polls in the 2020 elections. The Executive Council appointed 19 members to the Task Force on 2020 Pre-Election Polling from industry, nonprofit organizations, media, and academia to ensure a diversity of opinions, approaches, and expertise.

The Task Force collected all publicly available poll results at the national and state levels for the purpose of evaluating 2020 polling error and the extent that polls overstated the Democratic-Republican margin in the 2020 general election. The main findings of that review are as follows.

² <https://www.pewresearch.org/fact-tank/2021/01/28/turnout-soared-in-2020-as-nearly-two-thirds-of-eligible-u-s-voters-cast-ballots-for-president/>

³ Initial returns suggested that polls did even worse than they ended up doing after all of the votes were counted. The outcome was not decided on election night: During the ongoing pandemic, the process of counting early votes and mail-in votes after Election Day resulted in the initial election returns being more favorable for Trump than the final certified vote.

⁴ <https://www.aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx>

The 2020 Pre-Election Polling Error

- The 2020 polls featured polling error of an unusual magnitude: It was the highest in 40 years for the national popular vote and the highest in at least 20 years for state-level estimates of the vote in presidential, senatorial, and gubernatorial contests.⁵ Among polls conducted in the final two weeks, the average error on the margin in either direction was 4.5 points for national popular vote polls and 5.1 points for state-level presidential polls.
- The polling error was much more likely to favor Biden over Trump. Among polls conducted in the last two weeks before the election, the average signed error on the vote margin was too favorable for Biden by 3.9 percentage points in the national polls and by 4.3 percentage points in statewide presidential polls.
- The polling error for the presidential election was stable throughout the campaign. The average error matched closely for polls conducted in the last two weeks, in the final week, and even in the final three days. The challenges polls faced in 2020 did not diminish as Election Day approached.
- Beyond the margin, the average topline support for Trump in the polls understated Trump's share in the certified vote by 3.3 percentage points and overstated Biden's share in the certified vote by 1.0 percentage point.⁶ When undecided voters are excluded from the base, the two-candidate support in the polls understated Trump's certified vote share by 1.4 percentage points and overstated Biden's vote share by 3.1 percentage points.
- The overstatement of the Democratic-Republican margin in polls was larger on average in senatorial and gubernatorial races compared to the presidential contest. For senatorial and gubernatorial races combined, polls on average were 6.0 percentage points too favorable for Democratic candidates relative to the certified vote margin. Within the same state, polling error was often larger in senatorial contests than the presidential contest.
- Whether the candidates were running for president, senator, or governor, poll margins overall suggested that Democratic candidates would do better and Republican candidates would do worse relative to the final certified vote.
- No mode of interviewing was unambiguously more accurate. Every mode of interviewing and every mode of sampling overstated the Democratic-Republican margin relative to the final certified vote margin. There were only minor differences in the polling error depending on how surveys sampled or interviewed respondents. Regardless of whether respondents were sampled using random-digit dialing, voter registration lists, or online recruiting, polling margins on average were too favorable to Democratic candidates.
- On average, polls overstated the Democratic-Republican margin in states more supportive of Trump in 2016. In states Trump won by more than five points in 2016, the average signed error on the margin was 5.3 percentage points too favorable for Biden; on the other hand, in states Clinton won by more than five percentage points in 2016, the average signed error on the margin was 3.5 percentage points too favorable for Biden. Even after controlling for state-level differences in demographics and voting administration, the average signed error was larger in states that favored Trump in 2016.

⁵ It was the largest in 40 years for the national popular vote, but the performance of the state-level presidential polls has only been tracked since 2000.

⁶ This calculation includes undecided voters in the base.

Factors That Do Not Explain the Polling Error

Several proposed explanations can be ruled out as primary sources of polling error in 2020. Our analyses suggest the following.

- **Polling error was not caused by late-deciding voters voting for Republican candidates.** More voters voted prior to Election Day in 2020 than ever before and the number of undecided voters was relatively small. Only 4% of poll respondents, on average, gave a response other than “Biden” or “Trump” when asked by state-level presidential polls conducted in the final two weeks. Unlike in 2016, respondents deciding in the last week were as likely to support Biden as Trump, according to the National Election Pool exit polls.
- **Polling error was not caused by a failure to weight by education.** A suspected factor in 2016 polling error was the failure to weight by education (Kennedy et al. 2016). In the final two weeks of the 2020 election, 317 state-level presidential polls (representing 72% of all polls conducted during this period) provided information on the statistical adjustments accounting for coverage and nonresponse issues; of these 317 polls, 92% accounted for education level in the final results.
- **Polling error was not primarily caused by incorrect assumptions about the composition of the electorate in terms of age, race, ethnicity, gender, or education level.** There is no evidence that polling error was caused by the underrepresentation or overrepresentation of particular demographics. Reweighting survey data to match the actual outcome reveals only minor changes to demographic-based weights.
- **Polling error was not primarily caused by respondents’ reluctance to tell interviewers they supported Trump.**⁷ The overstatement of Democratic support occurred regardless of mode and the overstatement of Democratic support was larger in races that did not involve Trump (i.e., senatorial and gubernatorial contests).
- **Polling error cannot be explained by error in estimating whether Democratic and Republican respondents voted.** Trump supporters and Biden supporters were equally likely to vote after saying they would. This conclusion is based on validating the vote of registration-based samples shared with the Task Force by some AAPOR Transparency Initiative members.
- **Polling error was not caused by the polls having too few Election Day voters or too many early voters.** Among the 23 state-level presidential polls conducted in the final two weeks that reported how respondents said they would vote, the proportion of Election Day voters closely matched the percentage of certified votes cast on Election Day.

⁷ It is plausible that Trump supporters were less likely to participate in polls overall. Nonetheless, *among those who chose to respond to polls* there is no evidence that respondents were lying. A separate and likely problem is that some people chose not to respond to polls at all.

Factors That May Explain the Polling Error

Some explanations of polling error can be ruled out according to the patterns found in the polls, but identifying conclusively why polls overstated the Democratic-Republican margin relative to the certified vote appears to be impossible with the available data. Reliable information is lacking on the demographics, opinions, and vote choice of those not included in polls (either because they were excluded from the sampling frame or else because they chose not to participate), making it impossible to compare voters who responded to polls with voters who did not. Some educated guesses are possible based on patterns emerging from available data but conclusive statements are impossible. It cannot be ruled out that there is a multitude of overlapping explanations for the pattern of polling error.

Voter file information and certified vote information were compared to poll results but the most relevant information is unavailable; for example, it is unknown if Republicans who responded to polls voted differently than those who did not respond.⁸ If the voters most supportive of Trump were least likely to participate in polls then the polling error may be explained as follows: Self-identified Republicans who choose to respond to polls are more likely to support Democrats and those who choose not to respond to polls are more likely to support Republicans. Even if the correct percentage of self-identified Republicans were polled, differences in the Republicans who did and did not respond could produce the observed polling error.

This hypothesis is not unreasonable, considering the decreasing trust in institutions and polls especially among Republicans (e.g., Cramer 2016). Trump provided explicit cues to his supporters that polls were “fake” and intended to suppress votes (e.g., Haberman 2020). These statements by Trump could have transformed survey participation into a political act whereby his strongest supporters chose not to respond to polls.⁹ If so, self-identified Republican voters who participated in polls may have been more likely to support Democrats than those who chose not to participate in polls. Unfortunately, this hypothesis cannot be directly evaluated without knowing how nonresponders voted. Not only is the percentage of voters who self-identify as Republicans unknown but so too is the vote choice of the self-identified Republicans who chose not to participate.¹⁰

Many potential explanations for the polling error cannot be evaluated without knowing how respondents and nonrespondents compare. The polls may have differed relative to the 2020 electorate as follows: too many Democrats or too few Republicans, too many new voters or too few, or the wrong percentage of unaffiliated voters.¹¹ Or perhaps the polling error was caused by differences in the vote choice of the voters who were and were not included in polls (perhaps because the voters refused to participate). Any or all of these possibilities could produce an overestimate of the Democratic-Republican margin, but it is impossible to identify the precise cause(s) of the polling error documented here without knowing the opinions and demographics of voters who were and were not included in polls.

⁸ An analogous argument extends to unaffiliated voters and voters who were new or newly energized to participate in the 2020 election.

⁹ On the importance of elite cues for public opinion and behavior see, for example: Barber and Pope 2018; Bullock 2011; Broockman and Butler 2017; Lenz 2012.

¹⁰ Voter files can be used to estimate partisanship but estimates of individual-level partisanship in the absence of party registration data (or participation in a party primary) are often based on either precinct-level data or an imputation based on demographics (along with correlations between demographics and partisanship among past survey respondents). The former raises questions about the validation of ecological inferences and the latter must assume that the relationship among survey respondents can be used to impute partisanship among survey nonrespondents. But that is precisely the problem of concern.

¹¹ While we have estimates of these percentages based on voter file records, the characteristics that are of most interest are often estimated.

Even so, the present analyses help quantify the nature of the polling error and suggest what may have happened.

- **At least some of the polling error in 2020 was caused by unit nonresponse.** The overstatement of Democratic support could be attributed to unit nonresponse in several ways: between-party nonresponse, that is, too many Democrats and too few Republicans responding to the polls; within-party nonresponse, that is, differences in the Republicans and Democrats who did and did not respond to polls; or issues related to new voters and unaffiliated voters in terms of size (too many or too few) or representativeness (for example, were the new voters who responded to polls more likely to support Biden than new voters who did not respond to the polls?). Any of these unit nonresponse factors could have contributed to the observed polling error. Without knowing how nonrespondents compare to respondents we cannot conclusively identify the primary source of polling error.
- **Factors that worked well in correcting for nonresponse in previous elections (including demographic composition, partisanship, or 2016 vote) did not render accurate vote estimates for the 2020 election.** Poll data provided by some AAPOR Transparency Initiative members were reweighted to match the 2020 certified outcome. It was necessary to increase the percentage of Republicans (or 2016 Trump voters) and decrease the percentage of Democrats (or 2016 Clinton supporters) in the outcome-reweighted sample. In contrast, there are only slight differences between the originally weighted poll data and the outcome-reweighted data in terms of standard demographic categories.
- **Weighting to a reasonable target for partisanship and past 2016 vote does not fully correct the polling error.** Reweighting the polls to reproduce the 2020 outcome requires a much larger margin for Trump in 2016 than actually occurred among respondents who report voting in 2016. The larger 2016 margin for Trump among those who reported voting for Trump in 2016 could be caused by the following: an issue with the weighting targets, i.e., the implied vote share among 2016 voters who voted in 2020 was different from the 2016 actual outcome; or differences in opinion within groups that responded, e.g., the 2016 Trump supporters who responded to polls were more likely to vote for Democrats than those who did not. It is impossible to know which caused the larger 2016 margin.
- **It is possible that 2020 pre-election polls were not successful in correctly accounting for new voters who participated in the 2020 election.** There were many new voters in 2020 and it is unclear whether the proportion of new voters in the polls matched the proportion of actual new voters. It is also unclear whether the new voters who responded to polls had similar opinions to those who did not respond. Given the relative proportion and self-reported voting behavior of these new voters in the data available to the Task Force, this group of voters pushed the overall polling margins in the Democratic direction. Error in polling this group could have produced the observed polling error.