

# AAPOR Task Force on 2020 Pre-Election Polls: Performance of the Polls in the Democratic Primaries

Josh Clinton (Chair), Vanderbilt University

Jennifer Agiesta, CNN

Megan Brennan, Gallup

Camille Burge, Villanova University

Marjorie Connelly, AP-NORC Center

Ariel Edwards-Levy, Huffington Post

Bernard Fraga, Indiana University

Emily Guskin, Washington Post

Sunshine Hillygus, Duke University

Chris Jackson, Ipsos

Jeff Jones, Gallup

Scott Keeter, Pew

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 AAPOR Executive Council. The opinions expressed in this report are those of the authors 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.

## Overview & Executive Summary<sup>1</sup>

In June 2019, the Executive Council of the American Association for Public Opinion Research appointed a task force of 19 academic experts, pollsters and statisticians to examine the performance of the 2020 pre-election polls. The task force was charged to examine the scope and performance of pre-election polls in the 2020 Democratic primaries and the general election. This report focuses on the polls conducted for the 2020 Democratic primaries.

To evaluate the performance of the polls we consider two metrics: 1) how well polls did predicting election winners and 2) the absolute error on the margin of victory. Because pre-election polls are a snapshot in time, measuring vote preferences at the time the poll was conducted, it is important to understand that vote preferences can and do change. Even so, poll evaluations often consider how well polls do predicting election winners as a measure of poll performance (e.g. Kennedy et al., 2017, section 2.3). This report extends Kennedy et al.'s (2017) analysis of primary election poll performance, including both the prediction of election winners and average absolute error.

To evaluate the performance of pre-election polls for the primary elections in 2020, the task force collected every pre-election poll that was publicly released -- relying on existing databases (e.g., Real Clear Politics<sup>2</sup>, FiveThirtyEight<sup>3</sup>) and actively monitored media stories to collect polls that were not included in existing databases. Although we did not intentionally exclude any polls, polls that were not released for public consumption (e.g., internal polls for candidates) were excluded from analysis. For every poll identified, we collected all of the available public material -- including the article citing the poll, summary of descriptives, crosstabulations, and the methodology report. Using this information, we classified polls according to their mode and other additional information that may be relevant for polling performance related to the field period, sample (e.g., registered voter, likely voter), and demographics used in weighting procedures. (See Appendix A for a complete listing of the data collected for each poll.)

---

<sup>1</sup> A tremendous thanks to Sarah Lentz (University of Pennsylvania) and Mellissa Meisels (Vanderbilt University) who did an amazing job collecting, organizing, and conducting some preliminary analyses. **None** of this report would be possible without their tireless work and assistance. This report was prepared by Josh Clinton with the invaluable advice and consultation of the Task Force members. Clinton is responsible for any errors or omissions.

<sup>2</sup> <https://www.realclearpolitics.com/elections/2020/>

<sup>3</sup> <https://projects.fivethirtyeight.com/polls/>

Our assessment follows the practices of prior task forces and focus on the performance of pre-election polls for which the final day of the field period was within the final two weeks before each primary contest. To ensure transparency, both the data and the code used in the analysis are available for review and replication.

When assessing the performance of pre-election polls in the 2020 Democratic primaries, it is critical to recall how quickly and dramatically the shape of the race changed between February 26, 2020, and Super Tuesday on March 3, 2020. Three days before the South Carolina Democratic primary, on February 26th, 2020, Rep. James Clyburn (D-SC) endorsed former Vice President Joe Biden. Biden went on to win the South Carolina primary on February 29, 2020, by nearly 20 percentage points over Vermont senator Bernie Sanders. Shortly thereafter -- on the eve of Super Tuesday (March 3rd) -- both former South Bend mayor Pete Buttigieg and Minnesota senator Amy Klobuchar suddenly announced they were dropping out and endorsing Biden. Following Super Tuesday, former New York mayor Michael Bloomberg dropped out of the contest and endorsed Biden on March 4th. It is hard to imagine a primary campaign where the campaign changed -- and consolidated -- so quickly and so dramatically. Not only were some candidates (and therefore poll responses) no longer viable, but the last-minute endorsements may have also shifted voters' opinions. Appendix B summarizes some of the larger events that occurred during the 2020 Democratic primary season to highlight what may have impacted voter opinion.

This background is important because rapidly changing situations create severe challenges for pre-election polls. First, events can change the choices that are viable and available to voters on Election Day. Insofar as pre-election polls ask about candidates who are no longer relevant come Election Day, errors are inevitable. The fact that two competitive candidates announced they were dropping out and endorsing Biden on the eve of Super Tuesday meant that the performance of every pre-election poll in a Super Tuesday state was adversely impacted. Second, events may change voters' minds. For example, insofar as Rep. Clyburn's endorsement mattered, polls conducted prior to his endorsement could not capture the consequences of that endorsement.

To be clear, the potential consequences of events highlight the inherent risk involved in using pre-election polls to predict election outcomes. Late-breaking events can cause the polls to miss the actual vote totals, sometimes badly so.

Analyzing the performance of pre-election polls in the 2020 Democratic primaries reveals a number of high-level conclusions:

- **The ability of pre-election polls overall to correctly predict the winning candidate was similar to pre-election primary polls in recent years -- 81% correctly predicted the winning candidate.** The events of South Carolina and the last-minute consolidation of the primary field on the eve of Super Tuesday had a large impact on the ability of the subset of these pre-election polls that were conducted in Super Tuesday states to predict the winner -- only 61% of these Super Tuesday states predicted the winning candidate.
- **Average absolute polling error was slightly higher for the 2020 primary elections than in recent primary contests. This was largely because of late-**

**breaking events prior to the South Carolina primary and Super Tuesday contests.** Among all polls conducted in the last two weeks of a contest, the overall average absolute polling error was 10 points on the margin of victory. The average absolute polling error was 7 points on the margin of victory among the pre-election polls done for states other than South Carolina and Super Tuesday, which is lower than prior years except 2004 (which was also 7.0)

- **On average, there is no evidence that the accuracy of pre-election polling in the 2020 Democratic primaries depended on whether the poll was done by human interviewers, over the internet, or using multiple modes (e.g., interactive voice recordings and the internet).**

## 2020 Pre-election Primary Polls: Number and Mode

A total of 191 polls were released in the last two weeks before a Democratic primary and 137 of those polls were released within the last seven days of the election. We classified polls according to whether the poll was conducted using human interviewers calling cell phones and landlines (Live Phone); online respondents (Online); a combination of both Interactive Voice Recordings (IVR) and online respondents (IVR/Online); or some other methodology (Other/Misc.).

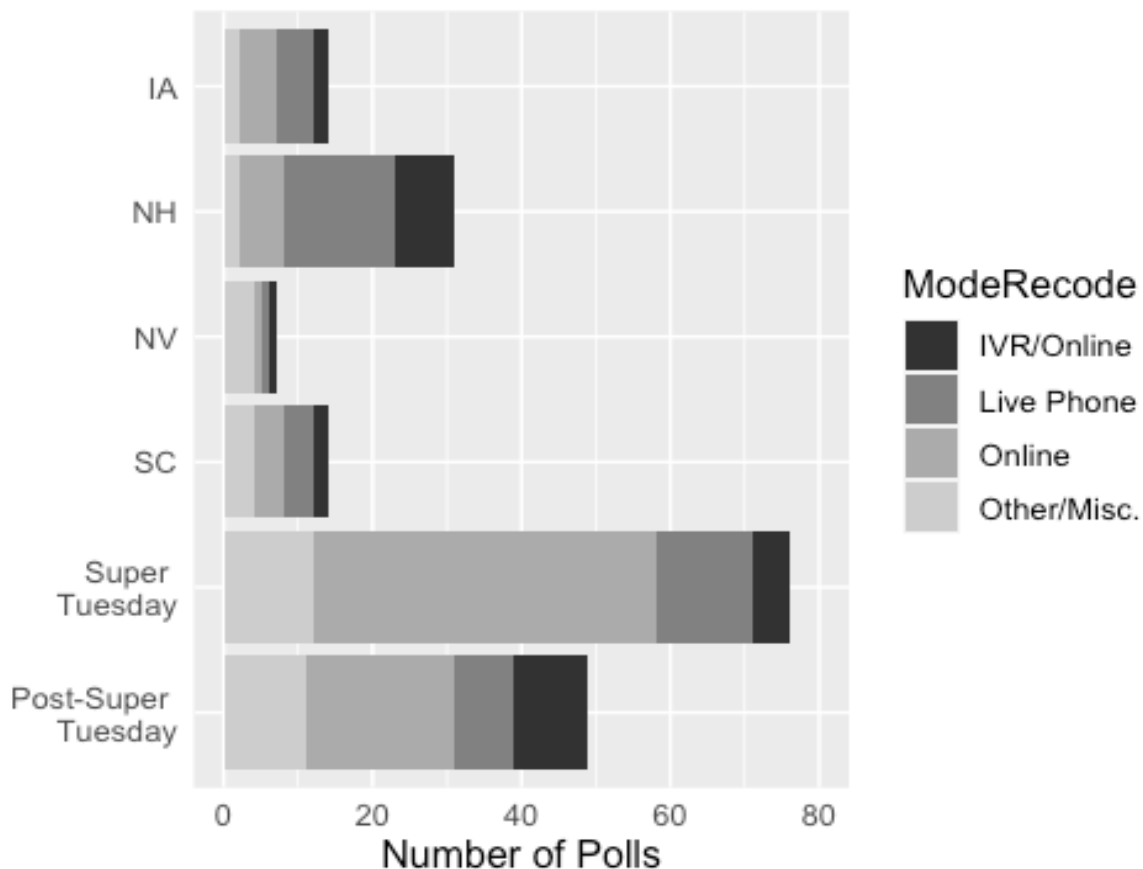
Reflecting trends noted in AAPOR's 2016 Ad Hoc Committee report, *An Evaluation of 2016 Election Polls in the U.S.* (Kennedy, et al., 2017), online polls continue to constitute a majority of the polls conducted before state primary contests. Using the same classification as the 2016 AAPOR Ad Hoc Committee report, Table 1 reports the modes of data collection for all pre-election polls released in the last two-week period, those released in the final week, and for the very last poll prior to Election Day.

		IVR/Online	Live Phone	Online	Other/Misc.	Total
Last two weeks	Number	28	46	82	35	191
	Percent	15%	24%	43%	18%	100%
Last week	Number	20	27	65	25	137
	Percent	15%	20%	47%	18%	100%
Last poll	Number	2	2	21	3	28
	Percent	7%	7%	75%	11%	100%

**Table 1. Mode of Data Collection for Pre-Election Primary Polls by Timing of Poll Relative to Election Day, 2020 US Democratic Primary Elections.**

As shown in Table 1, nearly twice as many pre-election polls were conducted online (43%) as were done with a human interviewer over the phone (24%). Moreover, the percentage of polls using online methods increased as Election Day got closer. In fact, 75% of the final pre-election polls that were released before Election Day were done online. Table 1 makes it clear that the overall performance of pre-election primary polling is increasingly determined by the performance of online polls.

It is also of interest to examine how the composition of pre-election polls varied over the course of the 2020 Democratic primary season. Figure 1 reports the number of polls conducted using each mode for the various primary contests. Polls are grouped for each of the first four state contests (Iowa, New Hampshire, Nevada, South Carolina), followed by the Super Tuesday states and states with contests after Super Tuesday. Although a plurality of polls for the New Hampshire primary were conducted using human interviewers over the phone, polls done for the Super Tuesday contests and beyond were more likely to be done using online methods than any other.



**Figure 1: Number of Polls by Mode and Primary, 2020 Democratic Primaries.** Only polls conducted within the last 14 days of a primary are included. Super Tuesday includes all polls done for Super Tuesday contests and Post-Super Tuesday includes all polls done for primaries after Super Tuesday.

### Polling Error: Overall & Historical

To evaluate the performance of polls conducted during the 2020 Democratic primaries, we rely on two measures: 1) the percentage of polls correctly predicting the winner, and 2) the absolute error on the margin of victory.

In prior evaluations of pre-election polling, AAPOR -- following prior evaluations performed by the National Council on Public Polls<sup>4</sup> -- has largely relied on two measures of pre-election polling error -- the signed error and the absolute error. The 2016 AAPOR Ad Hoc Committee report *An Evaluation of 2016 Election Polls in the U.S.* defines these measures as:<sup>5</sup>

- **Absolute Error:** *"absolute error on the projected vote margin (or "absolute error"), which is computed as the absolute value of the margin (%Clinton-%Trump) in the poll minus the same margin (%Clinton-%Trump) in the certified vote."*
- **Signed Error:** *"signed error on the projected vote margin (or "signed error"), which is computed in the exact same manner as the absolute error but without taking the absolute value. This statistic can be positive or negative..."*

The absolute error examines how well the winning margin in the certified vote compares to the margin between the first- and second-place candidates for each poll. For example, if candidate X beats candidate Y in a primary by a 10-point margin in the certified vote, and if a poll has candidate X leading candidate Y by 7 points, the absolute error for that poll is  $10 - 7 = 3$ . The signed error is the difference between the vote and poll margin without taking the absolute value.<sup>6</sup>

To limit the impact of outliers when summarizing the overall performance of pre-election polls, we use both the mean and median absolute error of pre-election polls in the 2020 Democratic primaries. Although we examine whether the error of pre-election polls varies by mode and contest, we intentionally make no effort to evaluate the performance of individual polls.

The signed error makes sense in a general election contest because the measure reveals whether the polls systematically over- or understate the difference in support between the major-party candidates, but it is not especially meaningful in a primary where the identity of the top two candidates can vary across contests. Because there is no reason to think that pre-election primary polls would systematically over- or underestimate the margin between the winning and second-place candidate, we follow prior evaluations and base our evaluation on the absolute error.

---

<sup>4</sup> The website of the National Council of Public Polls is available [here](http://www.ncpp.org/), although it is no longer being updated: <http://www.ncpp.org/> .

<sup>5</sup> <https://www.aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx>

<sup>6</sup> When there are multiple candidates, it is possible that the polls predict a different first and second place candidate than the actual winner and runner-up. Following past practices, the calculation is always based on the difference between the winner and the second-place candidate in the certified vote compared to the margin between those same candidates in the poll.

	Last Two Weeks	Last Week	Last Poll
% Correctly Predict Winner	81%	80%	82%
Average Absolute Error	9.99	9.23	10.86
Median Absolute Error	8.00	7.00	8.00
# Polls	191	137	28

**Table 2: Overall Polling Error for 2020 Democratic Primary Election Polls.** The absolute error is calculated using the absolute value of the difference between two other differences: the difference in the percentage of certified votes received by the first and second place candidates and the difference in the support each of those candidates receives in the poll. Overall performance is summarized using the average and median errors for: all 191 polls conducted in the last two weeks; the subset of 137 polls conducted in the last week; and the subset of final polls released in the 28 primaries.

As Table 2 shows, among the 191 polls released in the last two weeks of the primary contests, 81 percent of the polls accurately predicted the winning candidate. This percentage was similar to that for polls released in the last week and among the last poll released in each primary contest. The average absolute error was 9.99 among polls released in the last two weeks -- indicating that, on average, the margin of victory in the pre-election poll was off, on average, from the margin in the final certified vote by nearly 10 points across all primary contests. Among the subset of polls conducted in the final week, the average absolute error was slightly lower (9.23). Because some polls had particularly large errors, using the median to summarize the performance results in slightly better performance measures. The median absolute error among polls released in the last two weeks, for example, was 8 points.

### Historical Context

To place the performance of 2020 pre-election primary polls in context, it is useful to compare this performance to prior competitive presidential primaries. To do so we compare the results reported in Table 2 to the performance of pre-election polling from prior primary elections reported in Table 1 of *An Evaluation of 2016 Election Polls in the U.S.* (Kennedy et. al., 2017). These historical comparisons are based on polls conducted in the last two weeks of each contest.

	# Polls	% Correctly Predict Winner	Average Absolute Error
2000	172	99	7.7
2004	129	100	7.0
2008	555	79	7.6
2012	95	64	8.3
2016	457	86	9.3
2020	191	81	10.0

---

**Table 3. U.S. Primary Polling Error Over Time, 2000-2020.** The absolute error is calculated using the absolute value of the difference between two other differences: the difference in the percentage of certified votes received by the first and second-place candidates and the difference in the support each of those candidates received in the poll. Values from prior years taken from Table 1 of *An Evaluation of 2016 Election Polls in the U.S.* (Kennedy et. al., 2017).

As the results reported in Table 3 reveal, in terms of the percentage of polls correctly predicting the winner, the performance of the 2020 pre-election polls was similar or better than prior performance going back to 2008. Specifically, 2020 Democratic primary polls were more likely to predict the winner than in the 2012 pre-election primary polls for the Republican primary, nearly identical to the performance in the Democratic and Republican primaries of 2008, and slightly worse than the performance of polls during the 2016 Democratic and Republican primaries. An important caveat when comparing the performance of pre-election polls across time is that the ability of pre-election polls to predict the winning candidate may depend not only on the polls themselves, but also on the competitiveness of the election contests. For instance, the 2000 and 2004 polls identified the winner in almost every contest, perhaps because of the less competitive nature of those earlier primaries.

In terms of the average absolute error, the performance of the 2020 pre-election polls was slightly worse in 2020 than recent years. The average absolute error of 10 points in 2020 was nearly a point higher than the 9.3 point error in 2016 and nearly 2 points higher than the 8.3 point error in 2012. Highlighting how campaign events can affect the ability of pre-election polls to predict election results, all 76 pre-election polls conducted for Super Tuesday states were affected by the consolidation that occurred on the eve of Super Tuesday, and the 14 pre-election polls in South Carolina prior to February 26th missed the impact of Rep. Clyburn's endorsement of Vice President Biden.

## Polling Error by Primary Contest

To explore how changing political contexts may impact the ability of polls to predict election outcomes, we next grouped the state polls according to when the contest took place to allow us to replicate the analysis separately for pre-election polls in: Iowa, New Hampshire, Nevada, South Carolina, Super Tuesday states, and states with elections following Super Tuesday. If the last-minute events noted earlier (listed in Appendix B) affected the accuracy of the pre-election polls, we should observe differences in how well these groups of polls performed.

Examining the performance of pre-election polls according to this grouping reveals notable variation in the accuracy of pre-election polling during the 2020 Democratic primaries. Table 4 presents the comparison and reveals that the polls had substantially higher absolute error in South Carolina and the Super Tuesday states, whereas the absolute error for pre-election polls in the other 2020 primary contests was much lower. Excluding polls



in South Carolina and Super Tuesday, for example, the average absolute error in the other contests was 7.03, and 93 percent of those polls correctly predicted the winning candidate. In fact, the performance of this subset of 2020 pre-election polls was actually better than the overall performance of pre-election primary polls conducted in 2016, 2012, and 2008 (Table 3).

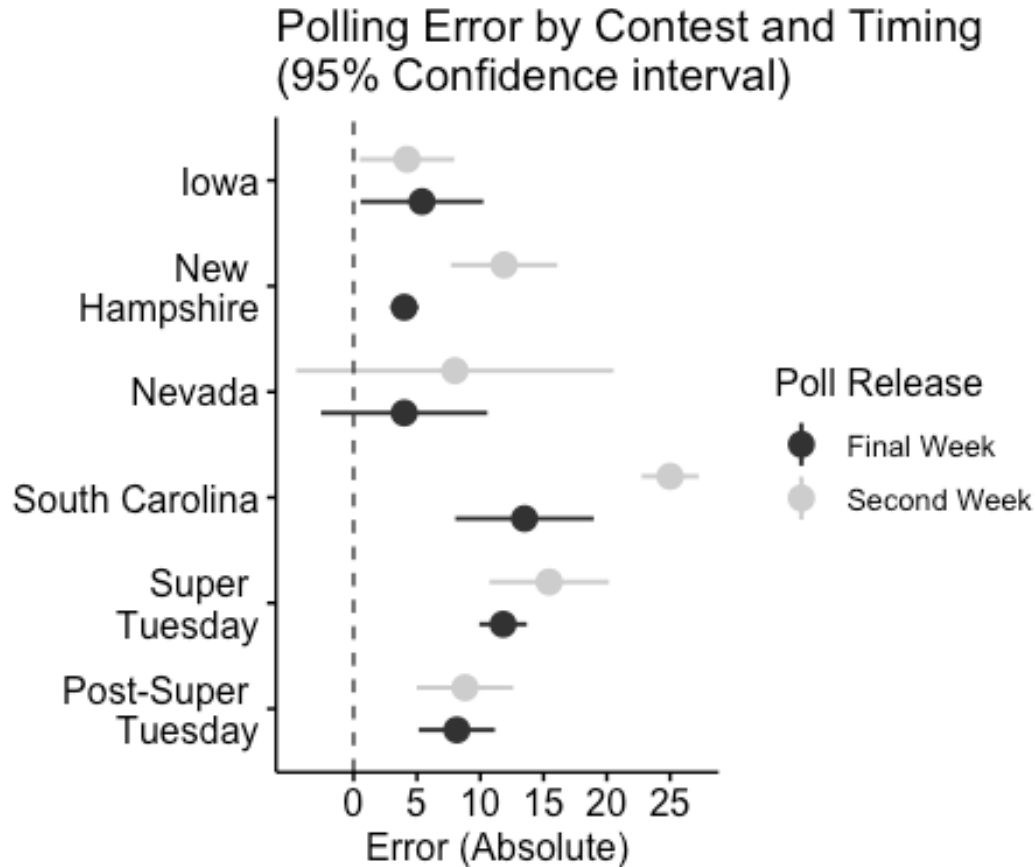
Although it is unsurprising that the performance of pre-election polls improves when contests with the largest polling errors are excluded, it is useful to consider the performance of those polls to highlight how much campaign events can affect the usefulness of pre-election polls for predicting election outcomes. It is also important to note that the average absolute error of pre-election polls -- even when we focus on contests where the polls did best -- was still greater than the margin of error due to sampling variability. It remains critically important for pollsters to communicate that the margin of error, which is a measure of sampling error, is not the same as the total expected polling error. That's because polls are subject to nonsampling error as well as sampling error.

	# Contests	# Polls	% Correctly Predict Winner	Average Absolute Error
Iowa	1	14	64	4.6
New Hampshire	1	31	97	6.3
Nevada	1	7	86	6.3
South Carolina	1	14	100	16.8
Super Tuesday	14	76	61	12.7
Post Super Tuesday	10	49	100	8.3

**Table 4. 2020 Democratic Primary Polling Error by Contest.** Polls conducted in the last two weeks of a primary are included. The absolute error is calculated using the absolute value of the difference between two other differences: the difference in the percentage of certified votes received by the first and second place candidates and the difference in the support each of those candidates received in the poll.

Just as the performance of pre-election polls may vary depending on when the contest occurs in the primary calendar, the performance of pre-election polls may also vary depending on how close the poll's field period is to Election Day. Polls that are in the field closer to Election Day are better able to account for late-breaking changes in the race, and we would consequently expect that their performance is at least as good as polls that were conducted two weeks out from Election Day.

To examine differences in poll performance across time, both within and between primary contests, Figure 2 compares the performance of polls whose field period concluded in the second-to-last week prior to the election (Second Week) with those with field periods ending in the last week of the contest (Final Week).



**Figure 2. Average Absolute Error for 2020 Democratic Primary Polls Conducted Two Weeks Prior to the Election Date (Second Week) versus the Last Week Prior to the Election Date (Final Week), by Contest.** The absolute error is calculated using the absolute value of the difference between two other differences: the difference in the percentage of certified votes received by the first- and second-place candidates and the difference in the support each of those candidates received in the poll.

The relationships depicted in Figure 2 clearly highlight how late-breaking events can sometimes, but not always, impact polling error. Leading up to the Iowa Caucuses, for example, polls conducted two weeks out had about the same amount of absolute error as polls conducted in the final week. In contrast, polls conducted closer to Election Day in the New Hampshire primary were much more accurate than those done two weeks out. In fact, the average absolute error for polls conducted two weeks out in New Hampshire was nearly three times larger than the average absolute error for polls in the last week. Polls conducted in the last week of the Nevada caucuses were slightly more accurate than those conducted two weeks out, but the difference was statistically indistinguishable.

The level of absolute error in pre-election polls in South Carolina depended on how close the poll was done to Election Day. The average absolute error was more than twice as large for polls done two weeks out from Election Day than it was for polls done in the final week. For polls in Super Tuesday states, polls done closer to Election Day had slightly lower absolute error, but the differences were statistically indistinguishable. Appendix C

presents the results separately by primary contest, but the results do not reveal any obvious patterns of performance.

## Polling Error By Mode

Another question of considerable interest is the extent to which the ability of pre-election polls to predict election outcomes varies depending on whether the poll was conducted by human interviewers, over the internet, using a combination of the internet and interactive voice recordings (IVR), or using some other method or combination of methods. To be clear, this analysis is complicated by the considerable methodological variation within and between the types of polls being done. Therefore, it is difficult to isolate the effects of survey mode from the effects of other factors such as differences in the sampling frame, likely voter modeling and the statistical adjustments (e.g., weighting to account for nonresponse). A second complication is that, insofar as the mode of data collection for polls varies across time both within and across contests, it will be impossible to isolate the effect of mode from these other factors. For example, the fact that the last poll released in nearly every contest was conducted online means that we cannot determine whether the performance of those polls is due to their timing or their mode (or, as noted earlier, the statistical adjustments that were used).

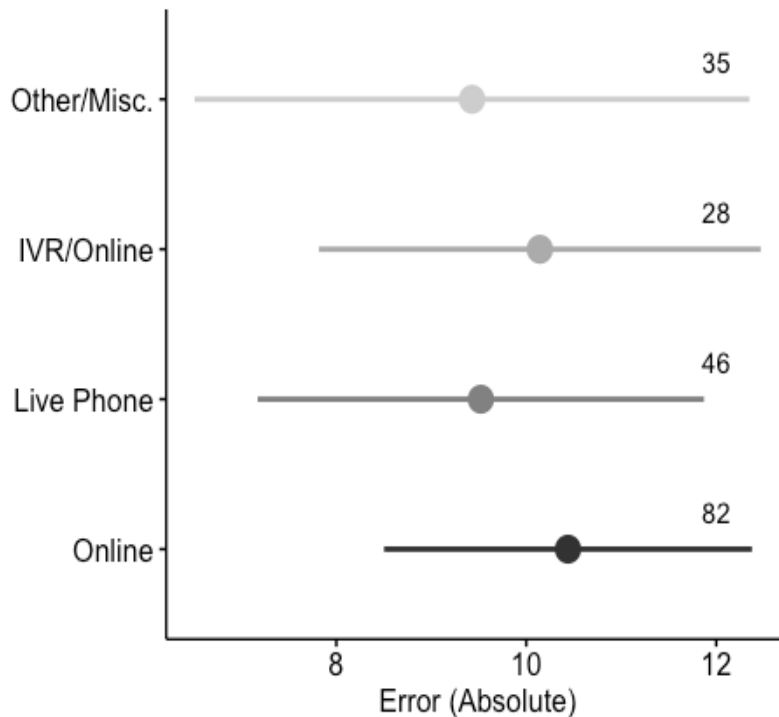
That said, we followed prior analyses of potential mode effects and compared how polling error varied across mode. Table 5 presents the results. In terms of the percentage of polls correctly predicting the winner, IVR/online polls had the highest percentage of polls that correctly predicted the winner (96%), followed by online polls (82%). The lowest rate of correctly predicting the winner occurred among live phone polls (72%), but it is impossible to draw conclusions about the overall accuracy based on this summary given the variation in when the various types of polls were conducted both within and across contests.

Put differently, comparing the performance of pre-election polls by the mode used is difficult because the mode varied over time. Polls done closer to the election – and also those following the consolidation of the candidate field in contests after Super Tuesday -- were more likely to be conducted online or using an IVR/online hybrid than polls conducted in earlier primary contests. If the expected competitiveness of the primary affected pollsters' decisions about whether or not to conduct a poll, then some modes may be used more frequently in easier-to-predict contests (e.g., pollsters might be reluctant to employ higher-cost methodologies for contests that are not expected to be competitive). If so, the ability of polls to predict the winner may be as much a function of where polls are being done as how they are being done.

	# Polls	% Correctly Predict Winner	Average Absolute Error	Median Absolute Error
IVR/Online	28	96	10.1	9.5
Live Phone	46	72	9.5	6.5
Online	82	82	10.4	9.0
Other/Misc.	35	77	9.4	7.0

**Table 5. 2020 Democratic Primary Polling Error by Mode of Data Collection.** The absolute error is calculated using the absolute value of the difference between two other differences: the difference in the percentage of certified votes received by the first- and second-place candidates and the difference in the support each of those candidates received in the poll.

We can also compare how well the various polling modes do at estimating the margin of victory. Because it is not necessarily easier to estimate the margin of victory in an uncompetitive primary than a competitive primary, comparing the average absolute error of various polling modes arguably avoids the problem that can occur from the fact that it is easier to predict the winner of some contests than others. As Table 5 reveals, there are only slight differences in the average absolute error across the various modes of data collection. Polls done using a human interviewer (Live Phone) have slightly lower absolute error than Online and IVR/Online polls, but the differences are statistically indistinguishable. Figure 3 provide a graphic visualization of the average absolute error (and associated 95% confidence interval) for each set of polls.



**Figure 3. 2020 Democratic Primaries Average Absolute Error by Mode of Data Collection.** The absolute error is calculated using the absolute value of the difference between two other differences: the difference in the percentage of certified votes received by the first and second-place candidates and the difference in the support each of those candidates received in the poll. The number of polls in each group is reported above.

## Summary and Conclusion

Pre-election polls provide, at best, a characterization of what voters would do if the election were to be held while the polling data were being collected. As a result, in addition to the usual concerns that can affect the accuracy of public opinion surveys (e.g., nonresponse), trying to predict an election outcome based on the results of a pre-election poll can be difficult if late-breaking events occur. Despite these difficulties, pre-election polls are often used formally or informally to predict election results. Thus, it is important to provide an assessment of how well pre-election polls perform. Not only can such an assessment help diagnose potential issues with current polling practices to help improve survey methodology, but the evaluation can also help highlight the danger in assuming that polling results can always accurately predict election results.

Analyzing the performance of the 197 pre-election polls that were conducted in the last 14 days of a Democratic primary reveals several takeaway messages.

First, 81% of pre-election polls correctly predicted the winner, similar to pre-election primary polls in recent years. The difficulty of predicting outcomes when late-breaking events occur is highlighted by the fact that only 61% of the polls done for Super Tuesday states predicted the winner because of late-breaking events.

Second, the average absolute polling error was slightly larger for the 2020 primary elections than in recent primary contests. This was largely because of late-breaking events prior to the South Carolina primary and Super Tuesday contests. Among all polls, the overall average absolute polling error was 10 points on the margin of victory. The average absolute polling error was 7 points on the margin of victory among the pre-election polls done for states other than South Carolina and Super Tuesday states.

Third, the mode of data collection used by the polls was not related to absolute error. The accuracy of polls was similar regardless of how the respondents were interviewed.

Using pre-election polls to predict election outcomes is inherently risky because late-breaking events can change a race after the polling stops, as we observed in the South Carolina primary and Super Tuesday contests in 2020. Overall, the performance of pre-election polls in the 2020 Democratic primaries was not noticeably worse than the performance in prior years. While absolute error was notably larger in South Carolina and the Super Tuesday states, this was due to the consequences of late-breaking events rather than a failure of the polls' methodology. The 2020 Democratic primaries thus offer both a cautionary tale about the sensitivity of election outcomes and the performance of pre-election polls to late-breaking events, but also the reassuring note that the polls were no less accurate than prior years in the absence of shifting political circumstances.

## References

Kennedy, C., Blumenthal, M., Clement, S., Clinton, J., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., Saad, L., Witt, E., Wlezien, C. (2017). "An Evaluation of 2016 Election Polls in the U.S." A report by AAPOR's Ad Hoc Committee on 2016 Election Polling, accessible here: <https://www.aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx>

## Appendix A: Notes on Data Collection

First, tremendous thanks to Sarah Lentz at the University Pennsylvania and Mellissa Meisels at Vanderbilt University who did an amazing job collecting, organizing, and running some preliminary analyses. **None** of this would be possible without them as they have been simply amazing.

All of the analyses were done using R so that everything could be replicated. Only polls conducted in the last two weeks prior to every Democratic primary contest were analyzed, although for completeness, we collected information on every poll – including polls that ask about primary preference at the national level, even though primaries are held at the state level rather than nationally.

When analyzing poll performance by the mode of data collection, we re-categorized some polls to remove the less frequent modalities. For example, a poll may have been reported as *phone* without distinguishing between a live human interview poll or IVR or if it was reported as using *proprietary technology*, the mode was recoded to "Other/Misc." For transparency, the recoding and the resulting characterization were done as follows:

```
allpolls$ModeRecode <- recode(allpolls$Mode,
  "Live Phone/IVR" = "Other/Misc.",
  "IVR" = "IVR/Online",
  "Phone" = "Other/Misc.",
  "proprietary technology" = "Other/Misc.",
  "Phone/Online" = "Other/Misc.",
  "unsure - see note" = "Other/Misc.")
```

Other variables that were collected include:

Variable	Description
<b>poll.id</b>	unique identifier of each poll
<b>Geography</b>	state or national geography where the poll was fielded
<b>Poll</b>	name of the poll
<b>Funded</b>	name of organization that funded the poll

<b>Conducted</b>	name of organization that conducted the poll
<b>Partisan</b>	whether the poll is a partisan-affiliation poll
<b>538</b>	whether or not FiveThirtyEight tracked the poll; if so, indicate with a 1; if not, indicate with a 0
<b>RealClear</b>	whether or not RealClearPolitics tracked the poll; if so, indicate with a 1; if not, indicate with a 0
<b>StartDate</b>	the start date of the poll
<b>EndDate</b>	the end date of the poll
<b>FieldDates</b>	the range of dates when the poll was in the field
<b>ReleaseDate</b>	the date when the poll results were released to the public
<b>PrimaryDate</b>	the date of the primary or caucus
<b>DaysinField</b>	number of days the poll was in the field
<b>Mode</b>	how the poll was conducted
<b>LandlineOnly</b>	for Live Phone polls, whether the poll used landline only; if so, indicate with a 1; if not, indicate with a 0; NA indicates unknown or not a Live Phone poll
<b>LandlineAndCell</b>	for Live Phone polls, whether the poll used landline and cell phone; if so, indicate with a 1; if not, indicate with a 0; NA indicates unknown or not a Live Phone poll
<b>LandlineCellText</b>	for Live Phone polls, whether the poll used landline, cell phone, and text messaging; if so, indicate with a 1; if not, indicate with a 0; NA indicates unknown or not a Live Phone poll
<b>pctCellPhone</b>	for polls conducted by telephone, the percent of respondents contacted on a cell phone (versus a landline)
<b>SampleSize.demo</b>	sample size of the full sample size of the poll (unless equal to the sample size of the democratic voter subset)
<b>sample_source</b>	source used to find sample
<b>TargetPop</b>	who is specified target population; should be either LV (likely voter) or RV (registered voter)
<b>LVmodel</b>	information on the Likely Voter model, if provided
<b>MoE</b>	margin of error for the sample

<b>Geography</b>	where the poll was fielded
<b>NumCandAsked</b>	number of candidates asked about in the poll
<b>type_simple</b>	which general contest the poll is asking about; should be Pres-P for Presidential Primary
<b>type_detail</b>	which more specific contest the poll is asking about; should be Pres-D for Democratic Primary
<b>Gender</b>	whether or not the poll was weighted to gender; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Age</b>	whether or not the poll was weighted to age; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Race/Ethnicity</b>	whether or not the poll was weighted to race/ethnicity; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Education</b>	whether or not the poll was weighted to education; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>16 Vote</b>	whether or not the poll was weighted to 2016 presidential vote; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Income</b>	whether or not the poll was weighted to income; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Region</b>	whether or not the poll was weighted to region; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Marital Status</b>	whether or not the poll was weighted to marital status; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Household Size</b>	whether or not the poll was weighted to household size; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Employment</b>	whether or not the poll was weighted to employment status; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided



	information provided
<b>Party ID</b>	whether or not the poll was weighted to party identification; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Ideology</b>	whether or not the poll was weighted to political ideology; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Mode.weight</b>	whether or not the poll was weighted to mode; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>TelephoneUsage</b>	whether or not the poll was weighted to telephone usage; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>PartyReg</b>	whether or not the poll was weighted to party registration; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>CD</b>	whether or not the poll was weighted to congressional district; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>SocialMediaUse</b>	whether or not the poll was weighted to social media usage if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>InternetUse</b>	whether or not the poll was weighted to internet use; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>PrimaryVoteHistory</b>	whether or not the poll was weighted to primary vote history; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Religion</b>	whether or not the poll was weighted to religion; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>InternetAccess</b>	whether or not the poll was weighted to internet access; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>VoterRegistration</b>	whether or not the poll was weighted to voter registration; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided

	information provided
<b>Turnout</b>	whether or not the poll was weighted to turnout; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Volunteerism</b>	whether or not the poll was weighted to volunteerism; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>County</b>	whether or not the poll was weighted to county; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>16PrimVote</b>	whether or not the poll was weighted to 2016 primary vote; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>CaucusHistory</b>	whether or not the poll was weighted to past caucus history; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Home-ownership</b>	whether or not the poll was weighted to home ownership; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Metropolitan status</b>	whether or not the poll was weighted to metropolitan status; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Urbanicity</b>	whether or not the poll was weighted to urbanicity; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Media.Market</b>	whether or not the poll was weighted to media market; if so, indicate with a 1; if not, indicate with a 0; NA indicates no weighting information provided
<b>Trump2016</b>	percent of respondents who voted for Trump in 2016
<b>Clinton2016</b>	percent of respondents who voted for Clinton in 2016
<b>PctDemocrats</b>	percent of Democrats in the sample
<b>PctRepublicans</b>	percent of Republicans in the sample
<b>PctIndependents</b>	percent of Independents in the sample
<b>PctStrong Democrat</b>	percent of strong Democrats in the sample

<b>PctLean Democrat</b>	percent of leaning Democrats in the sample
<b>PctLean Republican</b>	percent of leaning Republicans in the sample
<b>PctStrongRepublican</b>	percent of strong Republicans in the sample
<b>PctWhite</b>	percent of Whites in the sample
<b>PctBlack</b>	percent of Blacks in the sample
<b>PctHispanic</b>	percent of Hispanics in the sample
<b>PctAsian</b>	percent of Asians in the sample
<b>PctOther</b>	percent of Other races in the sample

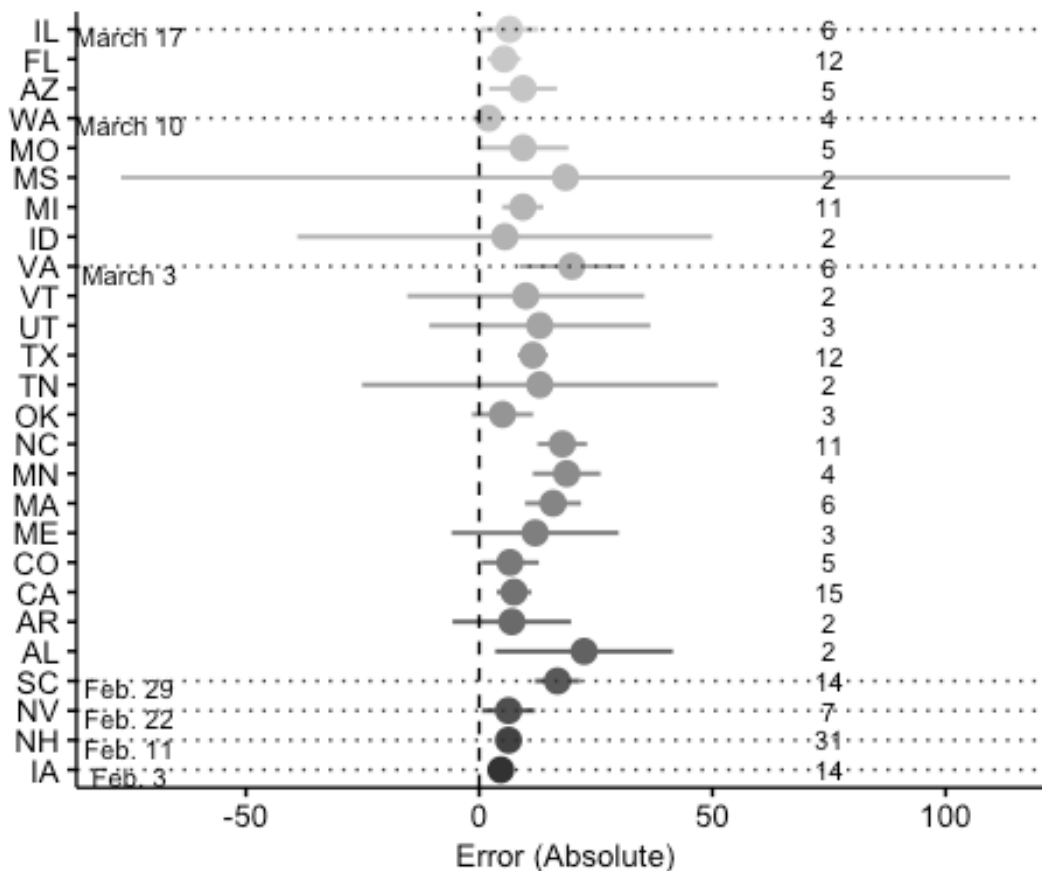
## Appendix B: List of Events Occurring During 2020 Democratic Primary

<b>Date</b>	<b>Event</b>
6/26 – 6/27/2019	First Democratic primary debate
7/9/2019	Tom Steyer entered the race
7/30 – 7/31/2019	Second Democratic primary debate
8/19/2019	John Hickenlooper withdraws from the race
8/21/2019	Jay Inslee withdraws from the race
8/28/2019	Kirsten Gillibrand withdraws from the race
9/12/2019	Third Democratic primary debate
9/20/2019	Bill de Blasio withdraws from the race
10/15/2019	Fourth Democratic primary debate
11/1/2019	Beto O'Rourke withdraws from the race
11/20/2019	Fifth Democratic primary debate
12/2/2019	Steve Bullock withdraws from the race
12/3/2019	Kamala Harris withdraws from the race
12/19/2019	Sixth Democratic primary debate
1/2/2020	Julián Castro withdraws from the race
1/9/2020	WHO announces mysterious coronavirus-related pneumonia in Wuhan, China
1/13/2020	Cory Booker withdraws from the race
1/14/2020	Seventh Democratic primary debate
1/21/2020	CDC confirms first US coronavirus case
2/3/2020	US declares public health emergency
2/3/2020	Iowa caucus
2/7/2020	Eighth Democratic primary debate
2/11/2020	New Hampshire primary
2/11/2020	Michael Bennet and Andrew Yang withdraw from the race
2/19/2020	Ninth Democratic primary debate

2/22/2020	Nevada caucus
2/25/2020	Tenth Democratic primary debate
2/26/2020	James Clyburn endorses Biden
2/29/2020	South Carolina primary
2/29/2020	Tom Steyer withdraws from the race
3/1/2020	Pete Buttigieg withdraws from the race
3/2/2020	Amy Klobuchar withdraws from the race
3/3/2020	Super Tuesday
3/4/2020	Michael Bloomberg withdraws from the race
3/5/2020	Elizabeth Warren withdraws from the race
3/10/2020	Super Tuesday II
3/11/2020	WHO declares COVID-19 a pandemic
3/13/2020	Trump declares COVID-19 a national emergency
3/15/2020	Eleventh Democratic primary debate
3/17/2020	Arizona, Florida, and Illinois primaries
3/18/2020	California becomes first state to issue statewide Stay-at-Home Order
3/19/2020	Tulsi Gabbard withdraws from the race
3/26/2020	Senate passes Coronavirus Aid, Relief, and Economic Security (CARES) Act
4/7/2020	Wisconsin primary
4/8/2020	Bernie Sanders withdraws from the race
4/10/2020	Alaska primary
4/17/2020	Wyoming caucus
4/28/2020	Ohio primary
5/2/2020	Kansas primary
5/12/2020	Nebraska primary
5/19/2020	Oregon primary
5/22/2020	Hawaii primary
5/28/2020	US COVID-19 deaths pass the 100,000 mark
6/2/2020	Washington, DC, Indiana, Maryland, Montana, New Mexico, Pennsylvania, Rhode Island, and South Dakota primaries
6/9/2020	Georgia and West Virginia primary
6/10/2020	US COVID-19 cases reach 2 million
6/23/2020	Kentucky and New York primaries
7/7/2020	Delaware and New Jersey primaries
7/11/2020	Louisiana primary
8/11/2020	Connecticut primary

## Appendix C: Polling Error by State Primary

We can also completely disaggregate the 197 pre-election polls to calculate the polling error for every state primary with at least two pre-election polls. To do so we plot the average absolute polling error – and 95% confidence intervals for that average -- for every contest using all polls released in the final 14 days of the election. There are differences in the number and composition of polls across time, as there are in the nature of the contest (e.g., the number of candidates, open/closed primary, etc.) but the comparison is instructive because it suggests that there are no obvious patterns in the average polling error across primary contests.



**Figure A1. Average Absolute Error by Primary Contest.** The absolute error is calculated using the absolute value of the difference between two other differences: the difference in the percentage of certified votes received by the first and second place candidates and the difference in the support each of those candidates received in the poll. The number of polls conducted in the last two weeks of each primary is reported.