Swing State Project

Methodology Resource: September 2024

In collaboration with Cook Political Report and GS Strategy Group, BSG conducted polling in Arizona, Georgia, Michigan, Nevada, North Carolina, Pennsylvania, and Wisconsin among likely 2024 voters from September 19 – September 25, 2024. Surveys were conducted in English using SMS-to-Web and online panel methodologies.

In the aggregate, 2,941 likely voters across the 7 states completed the survey, for a margin of sampling error of ±1.8 at the 95% confidence level. The full dataset included:
- 409 likely 2024 voters in Arizona (±4.9)
- 411 likely 2024 voters in Georgia (±4.8)
- 416 likely 2024 voters in Michigan (±4.8)
- 409 likely 2024 voters in Nevada (±4.9)
- 411 likely 2024 voters in North Carolina (±4.8)
- 474 likely 2024 voters in Pennsylvania (±4.5)
- 411 likely 2024 voters in Wisconsin (±4.8)

For the purposes of this research, likely voters were defined as anyone who is currently registered to vote and has voted in at least 1 of the last 4 presidential or midterm elections, or registered to vote after the 2020 general election for president.

Respondents were allowed to participate in the survey regardless of their self-reported likelihood to vote, as many respondents to indicate in survey that they "possibly will vote," "don't know" if they will vote, or even say they "absolutely will not vote" do ultimately participate in the election they were asked about. In this survey:
- 83% indicated they are "absolutely certain" to vote
- 10% indicated they are "very likely" to vote
- 5% indicated they "possibly will vote"
- 0% indicated they "absolutely will not vote"
- 0% indicated they "already voted" (only asked in states where ballots may have hit inboxes prior to the end of fielding)
- 1% indicated they don't know if they will vote or not

The results were weighted in each state to reflect the likely 2024 voter universe. In the combined data, states were weighted based on electoral votes.

What share of interviews were SMS vs. opt-in panel vs. probability-based panel?
57% were conducted on panel, 43% via SMS.

The panels are opt-in panels. These panels are not matched to the voter file (matched email panels are limited and tend to underrepresent lower-engagement voters, so we prefer to include panels made up of folks who aren't necessarily excited about politics/political surveys). Respondents were allowed to complete the survey via mobile browsers. About 81% took the survey on a smartphone, another 2% on a tablet, and 16% on a desktop/laptop.

How was sampling conducted for the SMS sample?

Our SMS-to-web sample was drawn from the registered voter file.

Were quotas used in drawing the sample? If yes, what characteristics were used and what was the source of targets?

Quotas were not applied to the SMS. Quotas were applied to Panel respondents, and those quotas were based on TargetSmart voter file counts reporting of our likely voter universe for each state.

To what characteristics was the sample weighted to (e.g., what is the "likely 2024 voter universe")? and where did the benchmarks come for that? Was the sample weighted by education? If not, can you provide weighted and unweighted breakdowns of respondents' educational attainment?

The likely voter universe included anyone who voted in at least 1 of the last 4 federal general elections. The data was weighted after fielding to be representative of voters in each of the states individually (per data from the voter file and past elections), and then the states weighted proportionally together according to electoral college representation. Education was weighted.

Is the margin of error adjusted for design effect?
No.

While part of our sample comes from non-probability panels, we provide what the margin of error would be for a probability-based sample with the same base size for context (because it is widely expected that some number be provided).

If you have one, what is the minimum unweighted sample size that the poll requires to report subgroup findings?

Our general rule is not to report anything below n=100 unweighted. In this case, the smallest demographic group included in the article is n=103.

For weighting to past elections, does this mean weighting to the exit poll? Was that just for education? Which past elections were included?

Regarding weighting, we look at a combination of voter file counts from past elections (2016, 2018, 2020, 2022), counts from the voter file on our 'likely voter' universe (voters who voted in 1 of those 4 elections, along with voters registered after the Nov. 2020 election), as well as exit polls. The variables weighted varied by state, depending on how the results came in (certainly every variable was not weighted in every state), but include: age, ethnicity, education levels, gender breakdowns within party identification, urban/suburban/rural classified zip codes, gender within ethnicity, age within ethnicity, DMA, zip code by median income, and whether and how voters report voting in the 2020 election.

Describe the sampling frame for the SMS-to-Web (i.e., the list used to select the sample, for example, RDD, voter list, recruited panel)

Our SMS-to-web sample was drawn from the registered voter file.

Describe the sampling design SMS-to-Web (i.e., how was the sample selected?)

The voter file list contained voters who have voted in at least 1 of the last 4 federal general elections, or are newly registered to vote since the 2020 November election.

Describe the sampling frame for the online surveys (i.e., the list used to select the sample, for example, RDD, voter list, recruited panel)

The online panels are opt-in panels (the PureSpectrum online platform which aggregates across multiple online panels).

Describe the sampling design for the online surveys (i.e., how was the sample selected?)

The online panel is randomly selected, not targeted, and individuals are screened to be voters.

For the online, was it based on an opt-in sample?

(yes)

What proportion of the cell phone numbers called were not working?

1.27% of the numbers were rejected for being invalid or disconnected or misclassified landlines.

What do you do to ensure your respondents are real people and are paying attention to the survey?

We have several ways of ensuring that our surveys are completed by real people (and not bots). We have several "clicker traps" throughout the instrument, where respondents must respond to a question "right" to continue. We have a "straight line" flag, if a respondent is answering all questions in a series with the same response. We also have a time flag, if someone has taken the survey faster than is possible (or likely). Additionally, we have open-ended questions, whose answers we review for anything that is out of the ordinary.

How were the SMS-to-web and online surveys combined?

The data is combined to be representative of voters in each of these states.

Was any weighting done, and if so, what variables were used and what is the source of the population parameters?

The data was weighted after fielding to be representative of voters in each of the states individually (per data from the voter file and past elections), and then the states weighted proportionally together according to electoral college representation.

The variables weighted varied by state, depending on how the results came in (not every variable was weighted in every state), but include: age, ethnicity, education levels, party identification or registration, gender, urban/suburban/rural classified zip codes, gender within ethnicity, age within ethnicity, DMA, zip code by median income, how many of the past four general elections they have voted in, and whether and how voters report voting in the 2020 election.

If a list of registered voters was used for the SMS-to-Web sampling, for each state's sample: what proportion of names in the registered voter database have cell phone numbers?

36,451,804 reg voters with a cell phone / 43,886,605 reg voters

Were any quotas used? If so, at what stage were they applied, what variables were used, and what is the source of your target quotas?

Quotas were not applied to the SMS. Quotas were applied to Panel respondents, and those quotas were based on TargetSmart voter file counts reporting of our likely voter universe for each state.

Who paid for the poll?
BSG and GS Strategy Group both paid for the survey fielding, and analysis was conducted by Cook Political, BSG, and GS Strategies, for purposes of public release.