Topline Results

[Rethink Priorities](#) is an independent, non-partisan, non-profit 501(c)3 policy think tank that does polling and policy analysis. Rethink Priorities is not funded by any candidate or political party committee and does not poll on behalf of any political candidate or party. This poll was conducted out of general interest, to test Rethink Priorities's capabilities to accurately poll and forecast policies of interest.

Trump vs. Biden (Biden +9.8)

A national poll of likely voters and found that **Joe Biden (D) has a +9.8 point lead over Donald Trump (R).** The poll sampled 4933 Americans and was adjusted to match a US nationally representative likely voter electorate by weighing on race, age, gender, education, income, socioeconomic status, region, 2016 Presidential vote, and religious attitudes. The raw margin of error is +/-2 points with 95% confidence.

The poll found Biden has 51.3% support and Donald Trump has 41.5% support among likely voters, with 3.6% backing another candidate, and 3.6% remaining undecided.

Using this polling data alone and historical information about polling accuracy, but no other information, we naively expect an 80% chance that Biden's actual margin on election day will be between +4.8 and +14.9.

Generic Congressional Ballot (Dems +5.3)

In the poll, **Democrats have a +5.3 point lead nationally over Republicans in the Generic Congressional Ballot.** In this poll, 47.7% of likely voters in the poll support the Democratic candidate, 42.4% support the Republican candidate, 2.5% support another candidate, and 7.4% are undecided.

Using this polling data alone and historical information about polling accuracy, but no other information, we naively expect an 80% chance that the actual national popular vote for Democrats in Congress (Generic Congressional Ballot) on election day will be between Republicans +2.6 and Democrats +13.1.

Rethink Priorities National Poll

Report by **Peter Hurford**, Co-Executive Director of Rethink Priorities
peter@rethinkpriorities.org

Poll conducted: 20 October 2020
Analysis published: 28 October 2020

Table of Contents

Methods

Survey Design

This poll was designed on Surveymonkey. The survey was 45 questions long and took an average of 9 minutes to complete. Questions consisted of matters of policy and relevant demographics.

Survey Deployment

This poll was conducted entirely on Prolific, an online platform where people are recruited and paid to complete surveys. The platform is non-political and non-partisan.

The survey was live on 20 October 2020 between 3pm and 10pm Central time.

Our survey was advertised to participants on the platform as "A Survey about Attitudes" with the description "In this survey you will be asked a number of questions regarding your attitudes to certain policy proposals. You will also be asked some basic demographic information." The nature of the survey was not disclosed any further, so we would not expect any additional selection bias in who takes the poll, beyond the bias already present in using an online platform like Prolific.

Only Americans were allowed to take our survey, and Prolific had about 42,394 eligible participants at the time. We sampled 4600 of them. We paid $1.47 for participants to complete the survey, which worked out to an average hourly rate of $9.80. This is normal for Prolific.

We further used Prolific surveyed 900 California residents, to allow for more California-specific analysis not included in this write-up. These participants were sampled and compensated similarly. These results were combined with our other results to create the full dataset.

Quality Filtering

Online surveys do not always produce accurate information - sometimes participants could be deliberately dishonest or otherwise low quality to the extent where it is best to remove them when conducting with analysis. In this survey, we removed 567 such people.

We started with 5500 possible responses. 396 responses were removed for being duplicates as determined by Prolific's internal ID system.

93 responses were removed because when asked "How honestly have you answered these questions?" at the very end of the survey, they replied "Not honestly at all" or "Somewhat honestly" instead of "Very honestly" or "Completely honestly" (see Robinson-Cimpian, 2014).

67 responses were removed because they failed an attention check - when asked "Which of these, if any, do you read, listen, or watch for news?", they indicated that they watched a news program that did not exist (there was one such program, "The Current Show with Al Franken" on the list - the list contained 12 real programs).

11 responses were removed due to failing a multiple low incidence check (see Lopez and Hillygus, 2018) which uses probability methods to screen for respondents dishonestly entering in unlikely information. This was done using the survey_dud_detector Python package developed by Peter Hurford at Rethink Priorities.

Results were also checked for straightlining using the survey_dud_detector Python package developed by Peter Hurford at Rethink Priorities, but this was not found to be an issue above and beyond the other quality checks and no respondents were removed for this issue.

After all of this quality filtering, there were 4933 remaining responses.

Demographic Weighting

Surveys only capture a sample of the population, so we know that the result probably won't exactly match the "true" result that we would get if we surveyed everyone in the population or that we would expect to see on election day.

The margin of sampling error describes how close we can reasonably expect a survey result to fall relative to the true population value. A margin of error of plus or minus 3 percentage points at the 95% confidence level means that if we fielded the same survey 100 times, we would expect the result to be within 3 percentage points of the true population value 95 of those times.

Without adjustment, surveys tend to overrepresent people who are easier to reach and underrepresent those types of people who are harder to reach. In order to make the results more representative we weight the data so that it matches the population – based on a number of demographic measures. Weighting is a crucial step for avoiding biased results, but it also has the effect of making the margin of error larger. Using US Census data, we can get a rough sense of the proportions of gender, race, and age we would expect to see in our sample.

We used the surveyweights Python package developed by Peter Hurford at Rethink Priorities to create weights to adjust for race, age, gender, education, income, socioeconomic status, region, 2016 Presidential vote, and religious attitudes. These weights were used to upsample and downsample responses accordingly to produce results that would end up matching the US Census data. All data to form weights, and sourcing for that information, is contained within the publicly available source code for the package.

Likely Voter Weighting

To create a view of what will happen on election day, it is important to only sample those who vote. However, many people ineligible to vote and eligible voters who end up deciding not to vote still show up in our survey sample. To resolve this, we create a probabilistic likely voter model and weigh our results accordingly.

Our likely voter model is based on the Perry-Gallup index outlined in Pew 2016:

- Respondents were given 1 point for thinking about the election "Quite a lot", 0.7 points for thinking about the election "Some", and no points for thinking about the election "Only a little".
- Respondents were given 1 point for saying they planned to vote.
- Respondents were given 1 point for saying they were "Very likely" to vote, 0.7 points for saying they were "Likely" to vote, 0.4 points for saying they were "Neither likely nor

unlikely" to vote, 0.1 points for saying they were "Somewhat unlikely" to vote, and 0.05 points for saying they were "Unlikely" to vote.
- Respondents were given 1 point for either saying they voted in 2016 or for not being old enough to vote in 2016.
- Respondents were given 1 point for saying they either plan to vote in person or that they have already received a mail-in ballot. If they requested but have not received their mail-in ballot, they were given 0.8 points. If they said they planned to vote by mail but have not yet requested their mail-in ballot, they were given 0.4 points.
- This produces a scale from 0.05 to 5.
- If a respondent indicates they were already voted, they were given a score of 5 instead of whatever score is calculated from above.
- If a respondent indicated they were not registered to vote, they were given a score of 0 instead of whatever score is calculated from above.
- Scores were mapped to probabilities of voting by transforming the decimal score to a whole number score from 0-7 by multiplying by 2.86 and rounding to the nearest whole number. These scores were then mapped 0 -> 0.11, 1 -> 0.13, 2 -> 0.23, 3 -> 0.34, 4 -> 0.4, 5 -> 0.59, 6 -> 0.63, 7 -> 0.83.
- Probabilities were then converted to likely voter weights by normalizing all the probabilities to sum to 1 across the dataset.
- These likely voter weights were then combined with demographic weights to adjust the sample to a likely voter electorate.

Electoral Modeling

Weighted results from the topline figures were used to construct *80% naive modeled confidence intervals* (NMCIs). NMCIs are intended to use polling results to predict actual election results using data from the poll alone and historical information about polling accuracy, but not other polls or other non-polling information like fundamentals. NMCI are not just the CIs implied by the raw polling margin of error. An 80% NMCI from *x* to *y* is meant to imply that there is an 80% chance that the true value as observed on election day will fall between *x* and *y*.

Biden vs. Trump

80% NMCIs for the Biden vs. Trump race were constructed by looking at the poll's raw weighted margin of error as calculated on the weighted N after applying likely voter weights (weighted N = 1867, raw weighted margin of error = +/- 2.9).

We allocate undecideds evenly between the two candidates, with some uncertainty and margin of error, allowing for a 20% chance that up to 75% of undecideds break for one particular candidate.

The final NMCIs were constructed with an additional 4pts of margin to account for historical accuracy of Presidential polling within three weeks of the election (see [Silver 2019](#)).

Generic Congressional Ballot

80% NMCIs for the Generic Congressional Ballot were constructed by looking at the poll's raw weighted margin of error as calculated on the weighted N after applying likely voter weights (weighted N = 1806, raw weighted margin of error = +/- 3.0).

We allocate undecideds evenly between the two parties, with some uncertainty and margin of error, allowing for a 20% chance that up to 75% of undecideds break for one particular party.

The final NMCIs were constructed with an additional 6pts of margin to account for historical accuracy of Congressional polling within three weeks of the election (see [Silver 2019](#)).

Appendix 1: Estimating Turnout

In the sample poll, 68.3% of participants indicated they would be likely voters. Using this polling data alone and historical information about polling accuracy, but no other information, we model that the actual turnout will be 62.7% of the voting eligible population.

We further naively expect an 80% chance that turnout will be between 57.4% and 67.0%.

We think that there is a risk we oversampled voters even after applying demographic weights, so we adjust our turnout model for this fact. Our survey found that 37% of registered voters say they voted (after demographic weighting), while we assume that the actual rate is 34% (given that ~52M people already voted as of the survey date on 20 October, out of estimated ~153M registered voters). This creates an adjustment of 34%/37%, or 0.918.

Our observed turnout rate of 68.3% was then adjusted to 68.3% * 0.918, or 62.7%. We then include the raw weighted margin of error plus an additional margin of 4pts to construct our 80% NMCI.

Appendix 2: Code

Rethink Priorities values transparency and invites scrutiny of its methods. The code and data for all our data quality filtering, demographic weighting, likely voter weighting, and electoral modeling is [available publicly on GitHub](#) under an MIT license.