

# POLITICAL POLLING – PART I: EVOLUTION OF POLLING

Data Analysis for Journalism and Political Communication  
(Spring 2026)

Prof. Bell



# WHEN POLLING MISSES

# The Literary Digest

NEW YORK OCTOBER 31, 1936

## Topics of the day

### LANDON, 1,293,669; ROOSEVELT, 972,897

Final Returns in The Digest's Poll of Ten Million Voters

Well, the great battle of the ballots in the Poll of ten million voters, scattered throughout the forty-eight States of the Union, is now finished, and in the table below we record the figures received up to the hour of going to press.

These figures are exactly as received from more than one in every five voters polled in our country—they are neither weighted, adjusted nor interpreted.

Never before in an experience covering more than a quarter of a century in taking polls have we received so many different varieties of criticism—praise from many; condemnation from many others—and yet it has been just of the same type that has come to us every time a Poll has been taken in all these years.

A telegram from a newspaper in California asks: "Is it true that Mr. Hearst has purchased THE LITERARY DIGEST?" A telephone message only the day before these lines were written: "Has the Repub-

lican National Committee purchased THE LITERARY DIGEST?" And all types and varieties, including: "Have the Jews purchased THE LITERARY DIGEST?" "Is the Pope of Rome a stockholder of THE LITERARY DIGEST?" And so it goes—all equally absurd and amusing. We could add more to this list, and yet all of these questions in recent days are but repetitions of what we have been experiencing all down the years from the very first Poll.

**Problem**—Now, are the figures in this Poll correct? In answer to this question we will simply refer to a telegram we sent to a young man in Massachusetts the other day in answer to his challenge to us to wager \$100,000 on the accuracy of our Poll. We wired him as follows:

"For nearly a quarter century, we have been taking Polls of the voters in the forty-eight States, and especially in Presidential years, and we have always merely mailed the ballots, counted and recorded those

returned and let the people of the Nation draw their conclusions as to our accuracy. So far, we have been right in every Poll. Will we be right in the current Poll? That, as Mrs. Roosevelt said concerning the President's reelection, is in the 'lap of the gods.'

"We never make any claims before election but we respectfully refer you to the opinion of one of the most quoted citizens to-day, the Hon. James A. Farley, Chairman of the Democratic National Committee. This is what Mr. Farley said October 14, 1932:

"Any sane person can not escape the implication of such a gigantic sampling of popular opinion as is embraced in THE LITERARY DIGEST straw vote. I consider this conclusive evidence as to the desire of the people of this country for a change in the National Government. THE LITERARY DIGEST poll is an achievement of no little magnitude. It is a Poll fairly and correctly conducted."

In studying the table of the voters from

The statistics and the material in this article are the property of Funk & Wagnalls Company and have been copyrighted by it; neither the whole nor any part thereof may be reprinted or published without the special permission of the copyright owner.

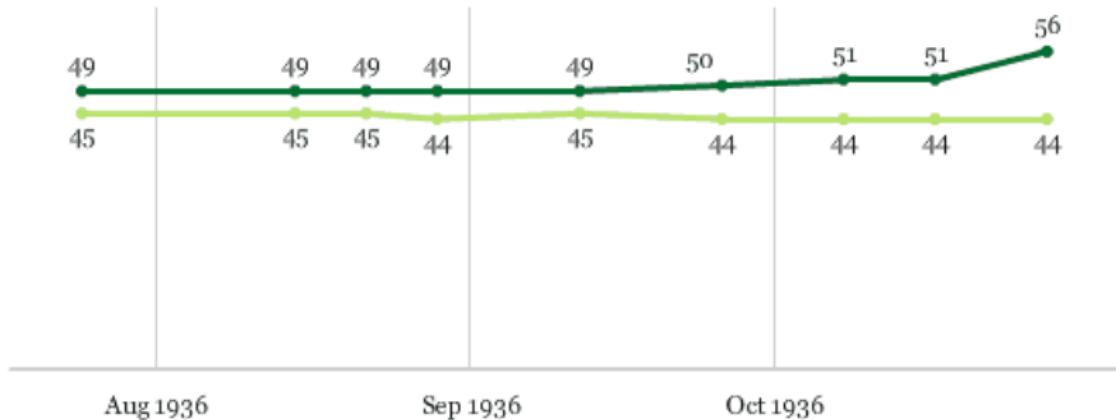


# WHEN POLLING MISSES

1936 Presidential Trial Heats, Based on National Adults

Democrat Franklin D. Roosevelt vs. Republican Alf Landon

■ % Roosevelt      ■ % Landon



Winner: Franklin D. Roosevelt with 61% of the vote

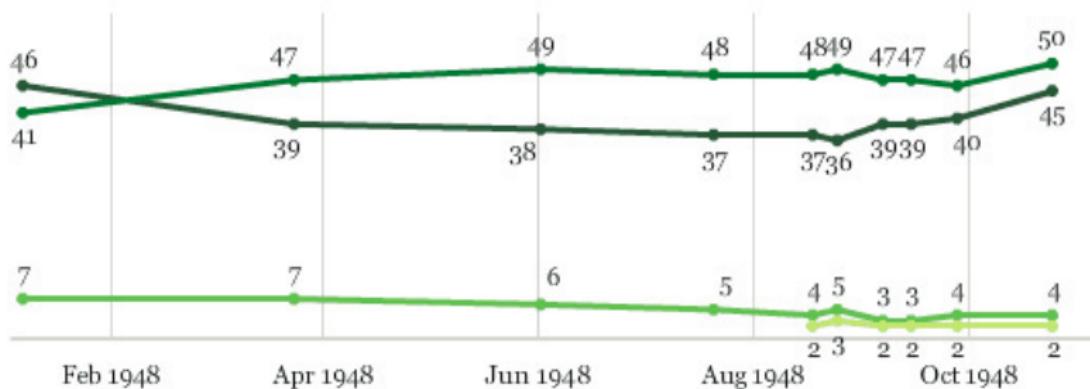
GALLUP<sup>®</sup>

# WHEN POLLING MISSES

1948 Presidential Trial Heats, Based on National Adults

Democrat Harry Truman vs. Republican Thomas Dewey  
vs. Progressive Henry Wallace vs. "Dixiecrat" Strom Thurmond

% Truman    % Dewey    % Wallace    % Thurmond



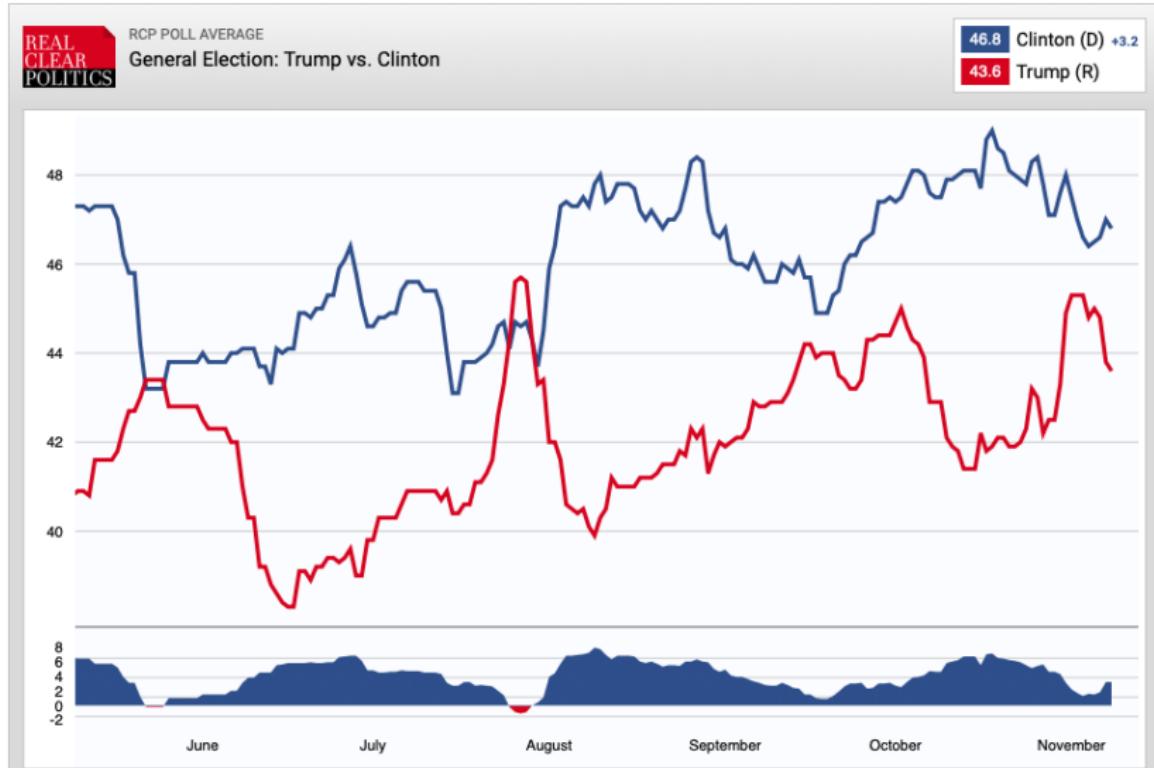
Winner: Harry Truman with 50% of the vote

GALLUP<sup>®</sup>

# WHEN POLLING MISSES



# THE 2016 ELECTION POLLS



# THE 2016 ELECTION POLLS

- ① The polls weren't that wrong

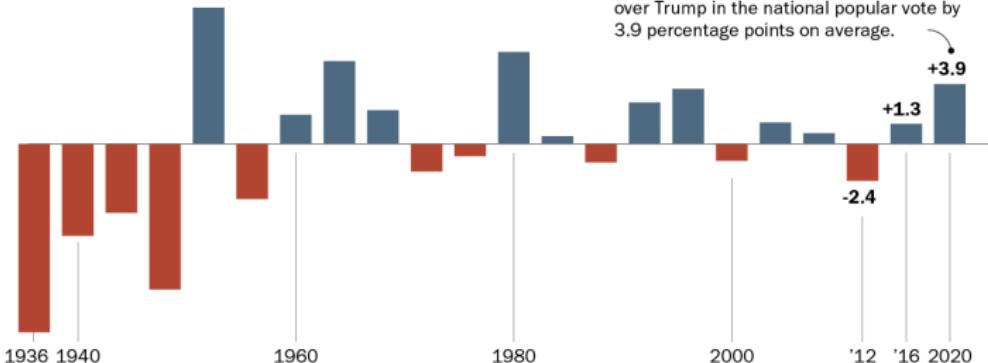
# THE 2016 ELECTION POLLS

## Polling errors in U.S. presidential elections

Bars represent average signed error

Blue bars indicate overestimation of Democratic vote margin.

Red bars indicate overestimation of Republican vote margin.



Note: The average signed error is the difference between the actual margin separating the candidates in the general election and the average margin in the polls. A negative error means that the Republican candidate's margin was overstated and a positive error means the Democratic candidate's margin was overstated.

Source: American Association for Public Opinion Research (AAPOR) Task Force on 2020 Pre-Election Polling: An Evaluation of the 2020 General Election Polls

PEW RESEARCH CENTER

# THE 2016 ELECTION POLLS

- ① The polls weren't that wrong
- ② The polls were wrong due to non-response bias, a form of selection bias

## Sampling error is not the only kind of polling error

Error from ...	Error name	Reflected in margin of error
Excluding parts of the population	Noncoverage	No
Low response rates from certain groups	Nonresponse	No
People misunderstanding the question or misreporting their opinions	Measurement	No
Interviewing a sample rather than entire population	Sampling	Yes

PEW RESEARCH CENTER

## TYPES OF MISSING DATA

- **Missing Completely at Random (MCAR)**: Observations are missing for non-systematic reasons

## TYPES OF MISSING DATA

- **Missing Completely at Random (MCAR)**: Observations are missing for non-systematic reasons
- **Missing at Random (MAR)**: Missing observations are correlated with a factor other than the outcome that is being measured

## TYPES OF MISSING DATA

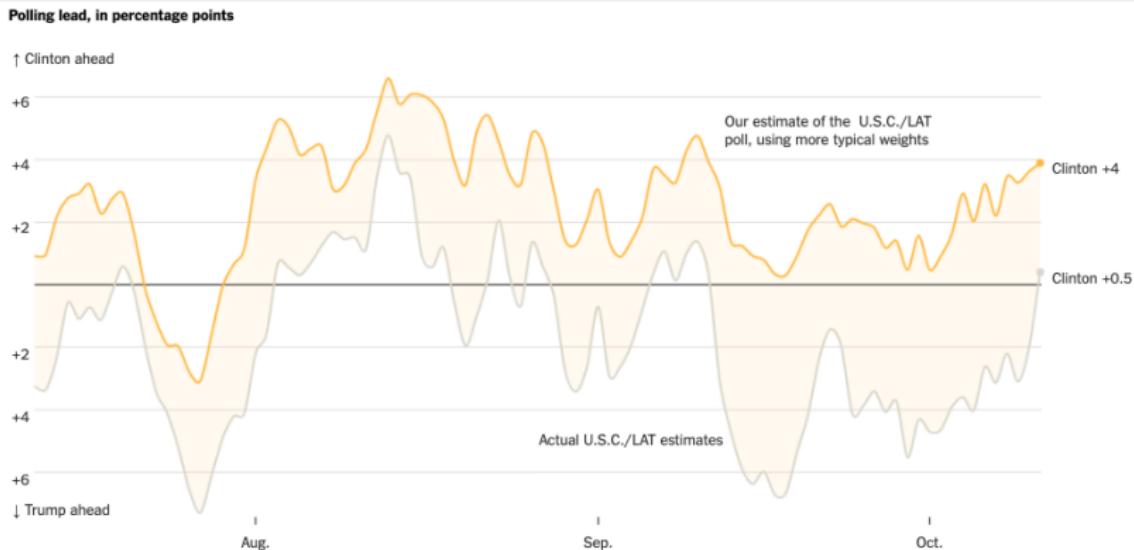
- **Missing Completely at Random (MCAR)**: Observations are missing for non-systematic reasons
- **Missing at Random (MAR)**: Missing observations are correlated with a factor other than the outcome that is being measured
  - ▶ MAR data can be resolved using a statistical technique called “weighting” in which answers from less-responsive groups are counted more, and answers from more-responsive groups are counted less

# TYPES OF MISSING DATA

- **Missing Completely at Random (MCAR)**: Observations are missing for non-systematic reasons
- **Missing at Random (MAR)**: Missing observations are correlated with a factor other than the outcome that is being measured
  - ▶ MAR data can be resolved using a statistical technique called “weighting” in which answers from less-responsive groups are counted more, and answers from more-responsive groups are counted less
  - ▶ Like making a “public opinion soup” but can go wrong (e.g., LA Times)

# TYPES OF MISSING DATA

The New York Times



# TYPES OF MISSING DATA

- **Missing Completely at Random (MCAR)**: Observations are missing for non-systematic reasons
- **Missing at Random (MAR)**: Missing observations are correlated with a factor other than the outcome that is being measured
  - ▶ MAR data can be resolved using a statistical technique called “weighting” in which answers from less-responsive groups are counted more, and answers from more-responsive groups are counted less
  - ▶ Like making a “public opinion soup” but can go wrong (e.g., LA Times)
- **Missing Not at Random (MNAR)**: Missing outcomes are correlated with the outcome that is being measured

# TYPES OF MISSING DATA

- **Missing Completely at Random (MCAR)**: Observations are missing for non-systematic reasons
- **Missing at Random (MAR)**: Missing observations are correlated with a factor other than the outcome that is being measured
  - ▶ MAR data can be resolved using a statistical technique called “weighting” in which answers from less-responsive groups are counted more, and answers from more-responsive groups are counted less
  - ▶ Like making a “public opinion soup” but can go wrong (e.g., LA Times)
- **Missing Not at Random (MNAR)**: Missing outcomes are correlated with the outcome that is being measured
  - ▶ MNAR data is a challenge because the only thing that predicts what data is missing is the fact that it is missing

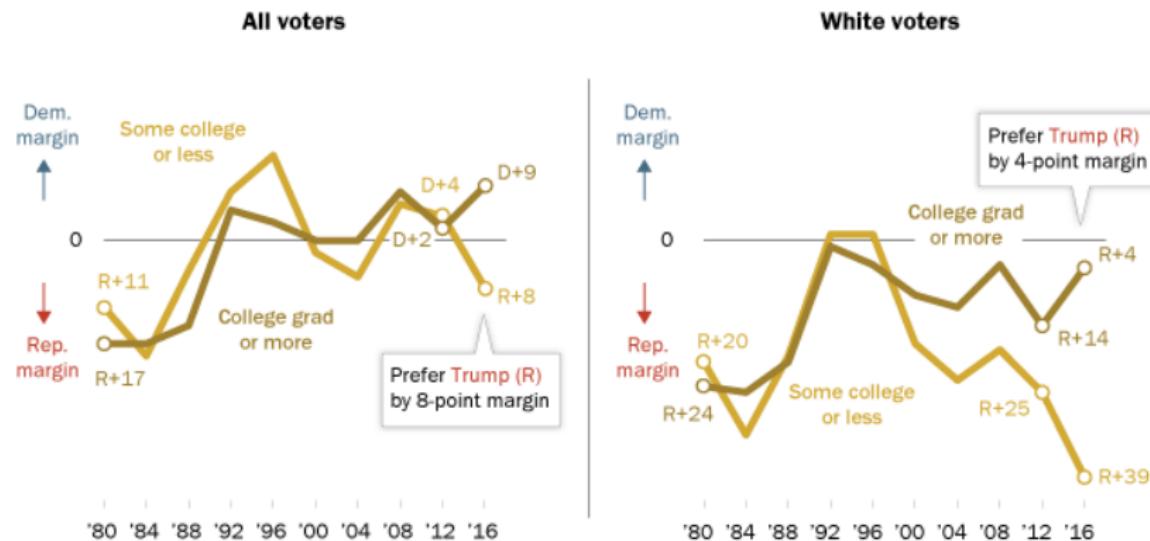
# THE 2016 ELECTION POLLS

- ① The polls weren't that wrong
- ② The polls were wrong due to **missing at random (MAR)** non-response bias by education

# THE 2016 ELECTION POLLS

## Wide education gaps in 2016 preferences, among all voters and among whites

Presidential candidate preference, by educational attainment



Source: Based on exit polls conducted by Edison Research for the National Election Pool, as reported by CNN. Data from prior years from national exit polls. In 1980, race was coded by the interviewer instead of being asked of the respondent.

PEW RESEARCH CENTER

# THE 2016 ELECTION POLLS

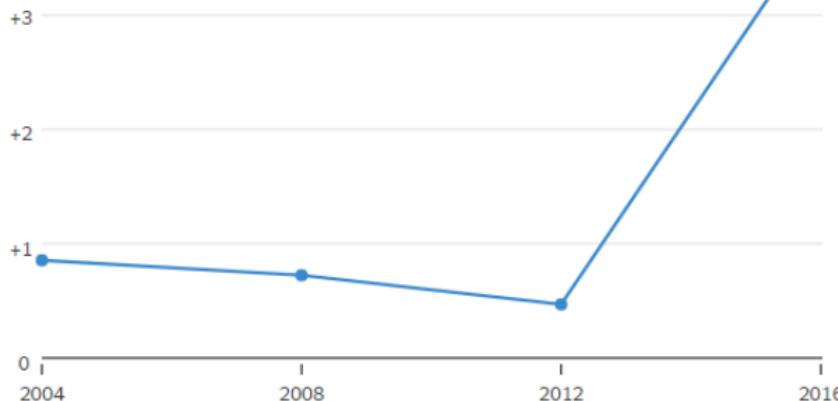
## Education Weighting Vastly More Important in '16

A poll that didn't weight by education might have been imperceptibly more Democratic-leaning in past elections, but was notably biased in 2016.

The effect of neglecting to weight by education in a typical national survey (pct. margin)

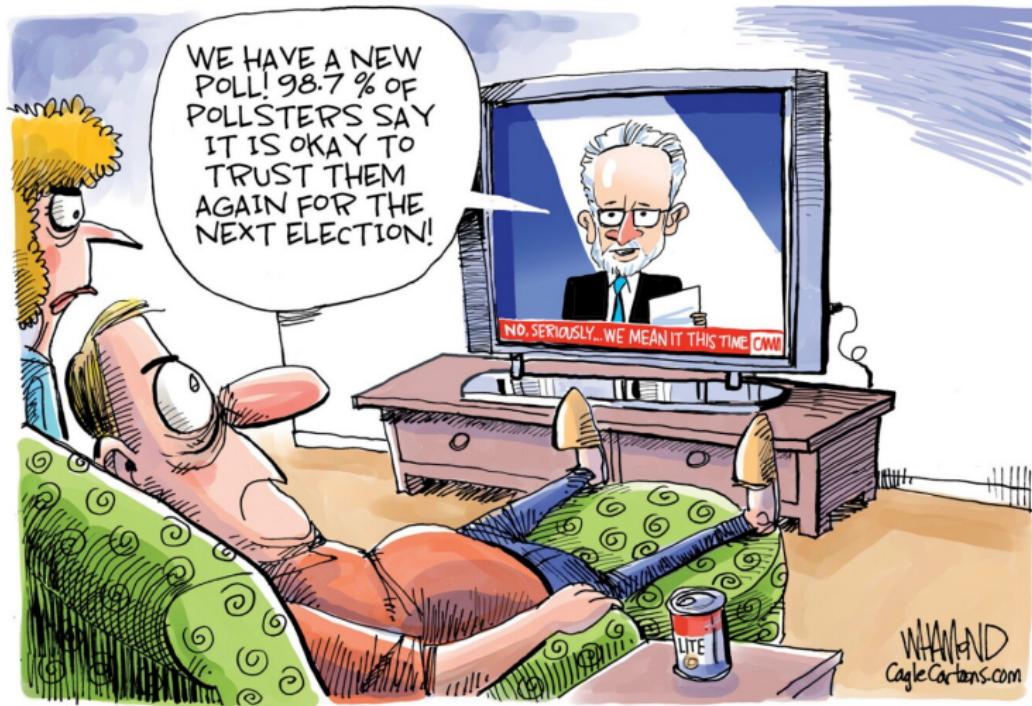
+4 Dem.

 The Upshot

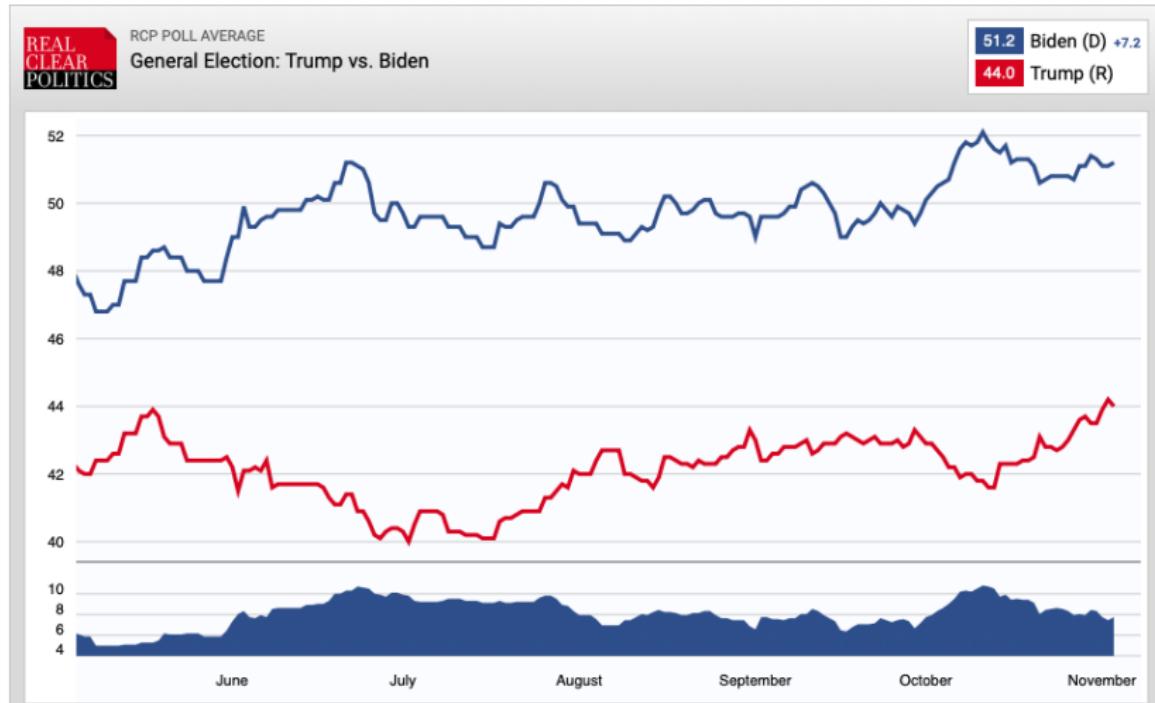


Source: Upshot analysis of surveys from the Roper Center, Pew Research and New York Times/CBS News. 2004 N=12,904, 2008 N=21,387, 2012 N=20,079. "Typical national survey" is an R.D.D. survey of adults with a self-reported likely-voter screen, weighted by race, education, age, sex, region and telephone usage.

# THE 2020 ELECTION POLLS



# THE 2020 ELECTION POLLS



# THE 2020 ELECTION POLLS

- ➊ The polls weren't that wrong (again)

# THE 2020 ELECTION POLLS

- ① The polls weren't that wrong (again)
- ② It is challenging to poll during a pandemic and with record-high turnout (the models were wrong)

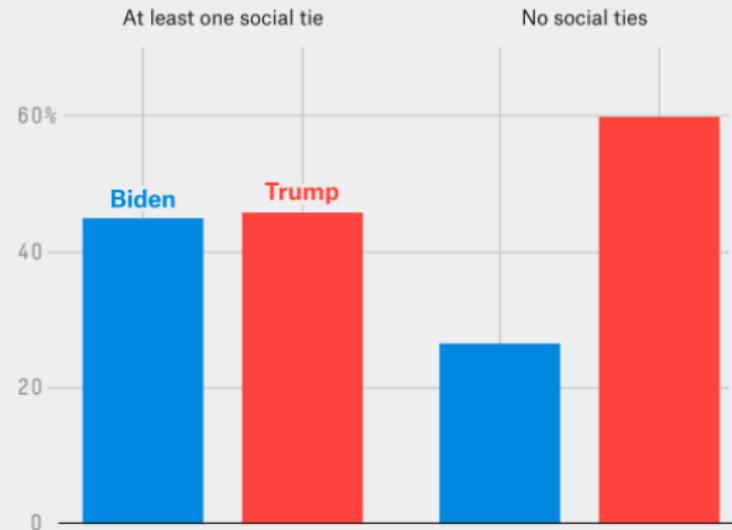
# THE 2020 ELECTION POLLS

- ① The polls weren't that wrong (again)
- ② It is challenging to poll during a pandemic and with record-high turnout (the models were wrong)
- ③ The polls were wrong due to missing at random (MAR) non-response bias (again)

# THE 2020 ELECTION POLLS

## Socially disconnected voters back Trump

Share of white voters with and without social networks that said they supported Donald Trump or Joe Biden in a pre-election survey



To determine one's social network, respondents were asked to name people with whom they had "discussed important personal matters or concerns" in the previous six months.

FiveThirtyEight SOURCE: SURVEY CENTER ON AMERICAN LIFE, AMERICAN NATIONAL SOCIAL NETWORK SURVEY SEPT. 2020

# THE 2020 ELECTION POLLS

- ① The polls weren't that wrong (again)
- ② It is challenging to poll during a pandemic and with record-high turnout (the models were wrong)
- ③ The polls were wrong due to missing at random (MAR) non-response bias (again)
- ④ The polls were wrong due to **missing not at random (MNAR)** non-response bias

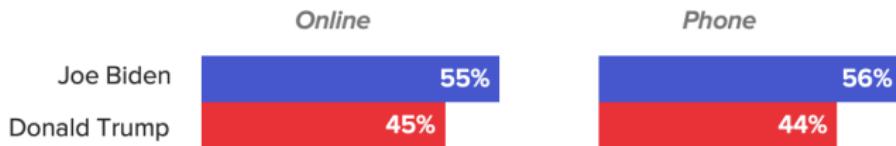
# THE 2020 ELECTION POLLS

- ① The polls weren't that wrong (again)
- ② It is challenging to poll during a pandemic and with record-high turnout (the models were wrong)
- ③ The polls were wrong due to missing at random (MAR) non-response bias (again)
- ④ The polls were wrong due to **missing not at random (MNAR)** non-response bias
  - ▶ "Shy Trump voters" theory

# THE 2020 ELECTION POLLS

---

If the 2020 presidential election were held today, for whom would you vote?



MORNING CONSULT®

Source: Morning Consult

Study conducted Sept. 10-14, 2020, among 1,698 respondents online and 1,244 respondents by live telephone.

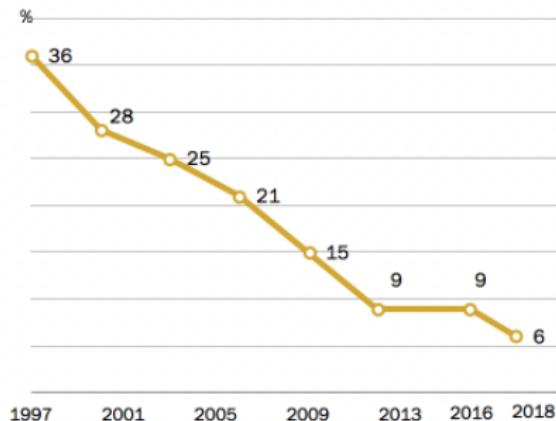
# THE 2020 ELECTION POLLS

- ① The polls weren't that wrong (again)
- ② It is challenging to poll during a pandemic and with record-high turnout (the models were wrong)
- ③ The polls were wrong due to missing at random (MAR) non-response bias (again)
- ④ The polls were wrong due to **missing not at random (MNAR)** non-response bias
  - ▶ "Shy Trump voters" theory
  - ▶ Trump voters are less likely to respond to surveys

# THE 2020 ELECTION POLLS

**After brief plateau, telephone survey response rates have fallen again**

*Response rate by year (%)*



Note: Response rate is AAPOR RR3. Only landlines sampled 1997-2006. Rates are typical for surveys conducted in each year.

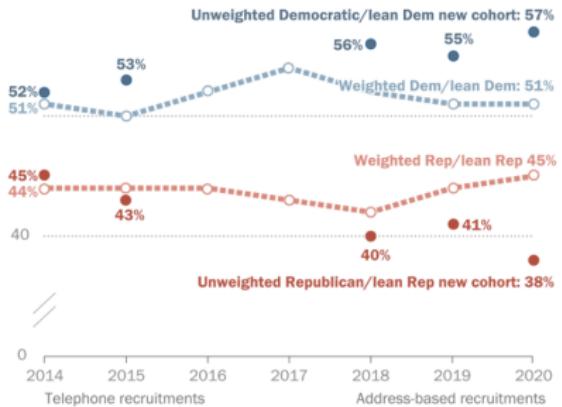
Source: Pew Research Center telephone surveys conducted 1997-2018.

PEW RESEARCH CENTER

# THE 2020 ELECTION POLLS

## More recent recruitment cohorts have been less Republican than early ones

% of newly recruited panelists who are Republican or lean to the Republican Party (unweighted)



Note: All ATP surveys are weighted on partisanship in addition to other variables listed in Appendix A. The 2017 cohort is not shown because it featured an unusual sample design that is not comparable to those displayed. Political party identification was not measured on the panel in 2014, and so the figures for the 2014 cohort come from a survey fielded March 10 to April 6, 2015. All other figures in this analysis are based on surveys fielded in the year shown.

Source: Surveys of U.S. adults conducted March 10-April 6, 2015, Nov. 30-Dec. 21, 2015, Oct. 25-Nov. 8, 2016, Sept. 14-28, 2017, Sept. 24-Oct. 8, 2018, Oct. 29-Nov. 11, 2019, and Aug. 3-Sept. 20, 2020.

\*Confronting 2016 and 2020 Polling Limitations\*

PEW RESEARCH CENTER

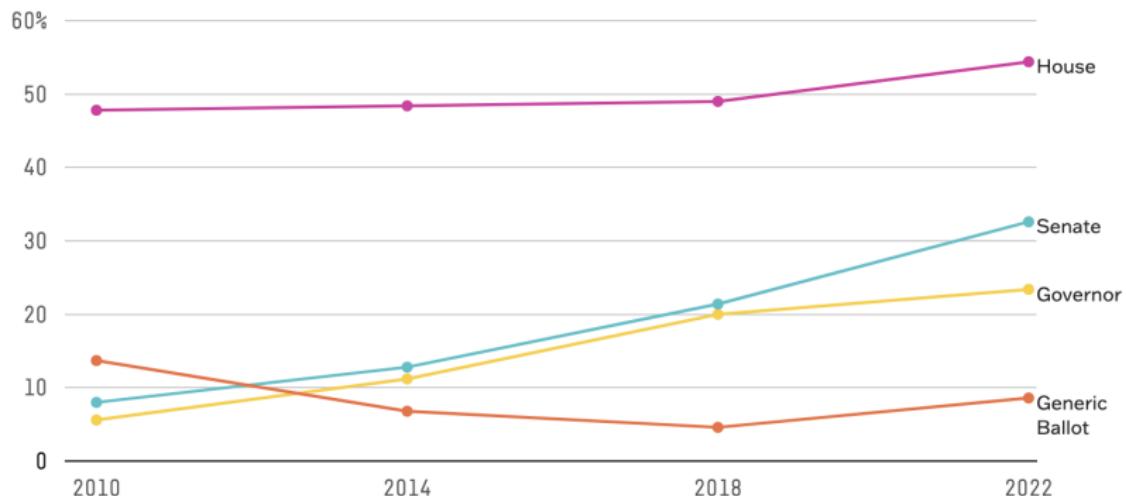
# THE 2022 ELECTION POLLS

- ① Republican-aligned pollsters were over-represented in polling averages, creating an expectation of a “red wave”

# THE 2022 ELECTION POLLS

## A larger share of polls are coming from partisan sources

The share of polls of the generic ballot and U.S. Senate, U.S. House and governor races conducted or sponsored by campaigns or partisan organizations in midterm elections, 2010-2022

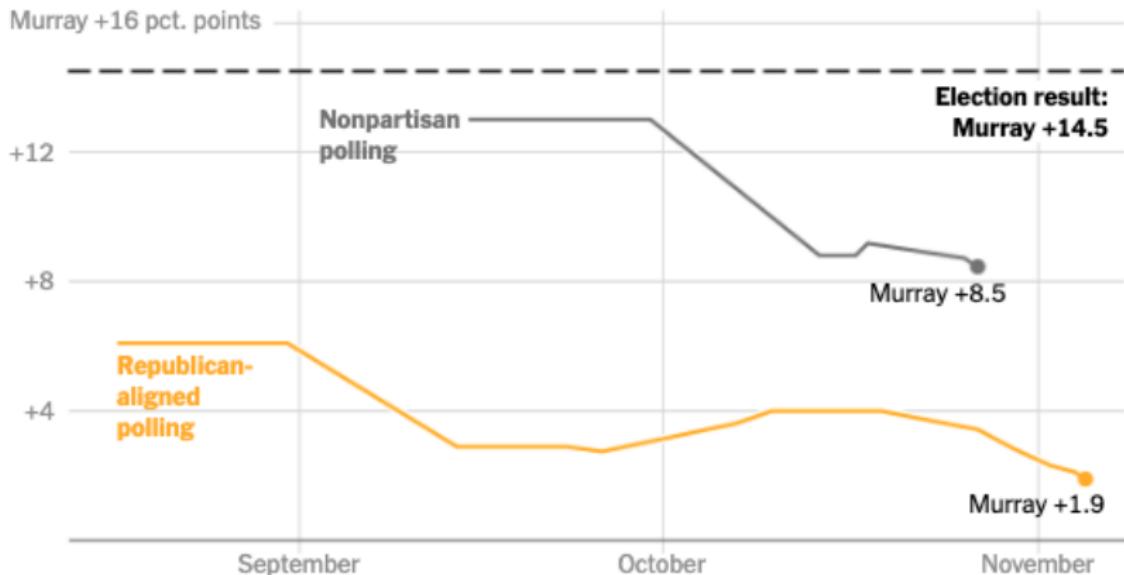


Based on polls whose final date in the field was between 180 and 15 days before the midterm election.

FiveThirtyEight

SOURCE: POLLS

# THE 2022 ELECTION POLLS



Source: New York Times analysis of Washington Senate race polls aggregated by FiveThirtyEight. Notes: Trends are calculated with a 14-day average. Polling groups considered Republican-aligned include those identified by The New York Times and FiveThirtyEight. Polling groups considered nonpartisan are those not known to be aligned with or funded by a political party. By Jason Kao

# THE 2022 ELECTION POLLS

- ① Republican-aligned pollsters were over-represented in polling averages, creating an expectation of a “red wave”
- ② Pollsters engaged in **herding** – adjusting the results of their findings to more closely match the results of other polls (or not releasing outlying polls at all) – to avoid being wrong

# THE 2022 ELECTION POLLS



John Anzalone  
@JohnAnzo

I appreciate this statement acknowledging @MonmouthPoll was an outlier. But I disagree with "In the end we must put out the numbers we have." When our firm comes out of the field and we believe we have an outlier we shit can the numbers and redo the poll at our expense. Period

**M** MonmouthPoll @MonmouthPoll

Statement from poll director @PollsterPatrick regarding this week's poll:  
<https://t.co/n4aZibDUej>

3:10 PM · Aug 28, 2019

---

261 Likes 56 Retweets

# THE 2022 ELECTION POLLS

- ① Republican-aligned pollsters were over-represented in polling averages, creating an expectation of a “red wave”
- ② Pollsters engaged in **herding** – adjusting the results of their findings to more closely match the results of other polls (or not releasing outlying polls at all) – to avoid being wrong
- ③ Something is unique about elections in which Donald Trump is a candidate

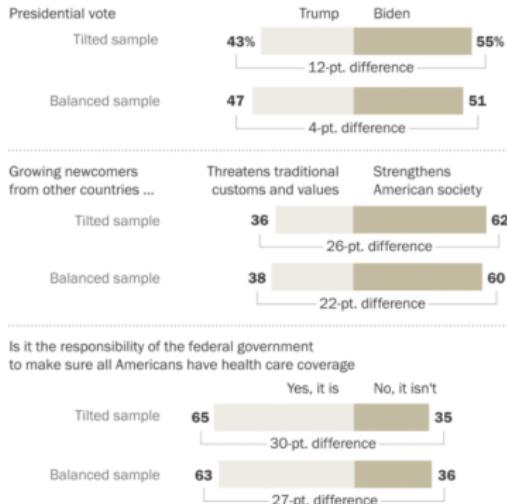
# THE 2022 ELECTION POLLS

- ① Republican-aligned pollsters were over-represented in polling averages, creating an expectation of a “red wave”
- ② Pollsters engaged in **herding** – adjusting the results of their findings to more closely match the results of other polls (or not releasing outlying polls at all) – to avoid being wrong
- ③ Something is unique about elections in which Donald Trump is a candidate
  - ▶ Issue polling seems to be unaffected by biases influencing horse race polling

# THE 2022 ELECTION POLLS

## A large overstatement of voter support for Joe Biden has a minimal effect on measures of opinion about issues

% supporting each option in two different samples, one that is too Democratic and the other correctly balanced



Source: Survey of U.S. adults conducted July 27–Aug. 2, 2020.  
“What 2020’s Election Poll Errors Tell Us About the Accuracy of Issue Polling”

PEW RESEARCH CENTER

# THE 2024 ELECTION POLLS



RealClearPolitics Poll Average

48.7 Harris +0.1

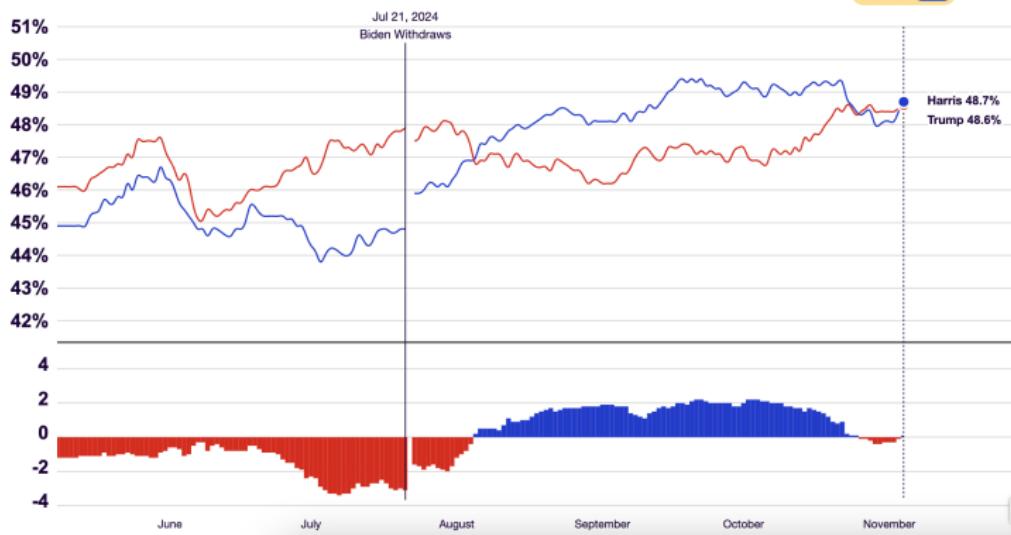
48.6 Trump

## Trump vs. Harris

(Trump vs. Biden Before July 21, 2024)



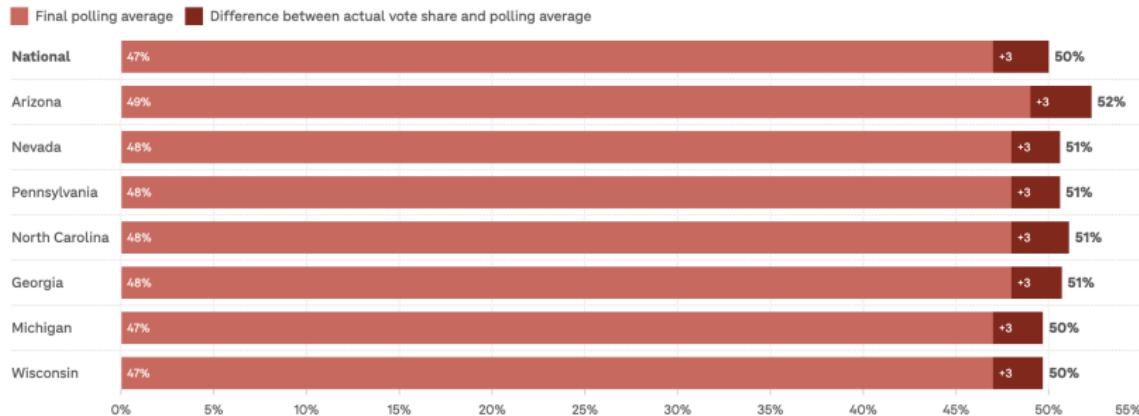
Harris +0.1



# THE 2024 ELECTION POLLS

In every swing state, Trump outperformed his final polling numbers by a few points

President-elect Trump's vote share



Source: Edison Research exit polls for 2024

Credit: Domenico Montanaro and Juweel Adolphe/NPR

# THE 2024 ELECTION POLLS

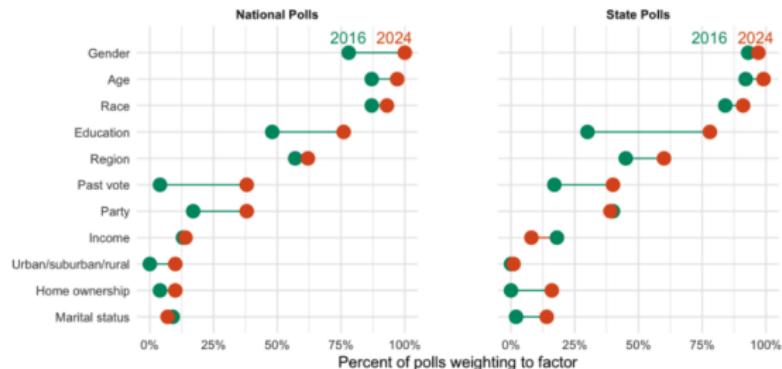
- Pollsters used new techniques to address partisan non-response, like weighting on **recalled vote**

Harris' surge in the polls has people wondering if we might see a **pro-Trump polling error** for the third presidential cycle in a row. An anecdote from Siena College's pollster is revealing:

Levy told me that, in 2020, the people working the phones for Siena frequently reported incidents of being yelled at by mistrustful Trump supporters. "In plain English, it was not uncommon for someone to say, 'I'm voting for Trump—fuck you,' and then hang up before completing the rest of the survey," he said. (So much for the "shy Trump voter" hypothesis.) In 2020, those responses weren't counted. This time around, they are. Levy told me that including these "partials" in 2020 would have erased nearly half of Siena's error rate.

National and state polls weighted more on education and past vote in 2024 compared to 2016

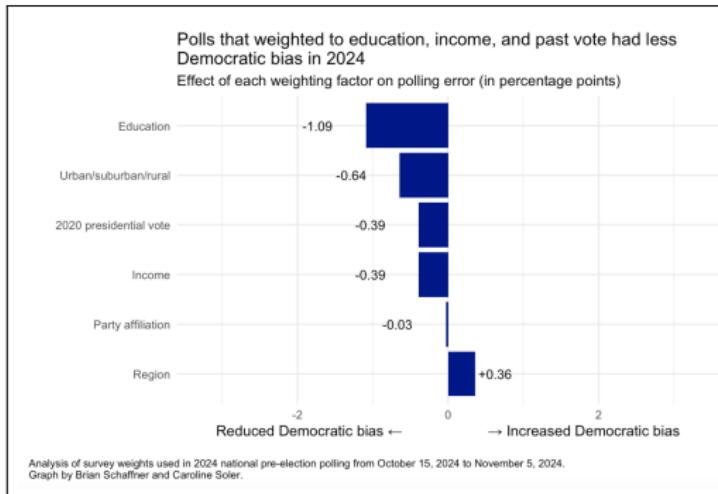
Graph shows the percent of polls weighting to each factor in 2016 and 2024 by poll type



Analysis of survey weights used in 2024 and 2016 pre-election polling from the last three weeks of the election cycle.  
Graph by Brian Schaffner and Caroline Soler.

# THE 2024 ELECTION POLLS

- Pollsters used new techniques to address partisan non-response, like weighting on **recalled vote**



Raphael Nishimura  
@rnishimura ...

Had Selzer weighted by 2020 recall vote, she would have got Trump 50 Harris 44. And I think that could have been further improved if they also weighted non-voters accordingly (she says they assign a weight of 1.0 for those cases). Weighting by recall vote works, folks!

PRESIDENTIAL HORSE RACE, TOP PREFERENCE PLUS LEANING

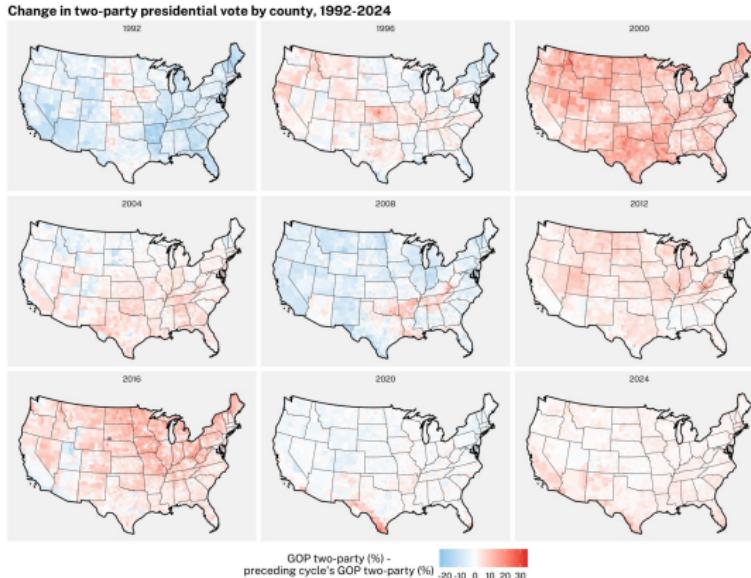
	Count	%
HARRIS	357	44%
TRUMP	401	50%
RFK JR	18	2%
OLIVER	1	0%
SOMEONE ELSE	8	1%
WOULDN'T VOTE	1	0%
NS/DON'T REMEMBER	12	2%
DON'T WANT TO TELL (ALR ONLY)	10	1%

808 WGTTEST2  
656 Total Unweighted

2:59 PM · Nov 17, 2024 · 6,501 Views

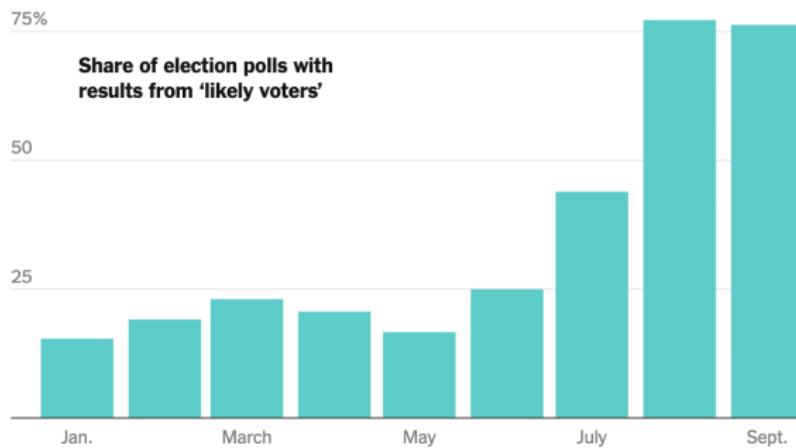
# THE 2024 ELECTION POLLS

- Pollsters used new techniques to address partisan non-response, like weighting on **recalled vote**
- Recalled vote weighting assumes that the electorate looks the same from year-to-year



# THE 2024 ELECTION POLLS

- Pollsters used new techniques to address partisan non-response, like weighting on **recalled vote**
- Recalled vote weighting assumes that the electorate looks the same from year-to-year
- This is especially the case as pollsters begin using “likely voter” models



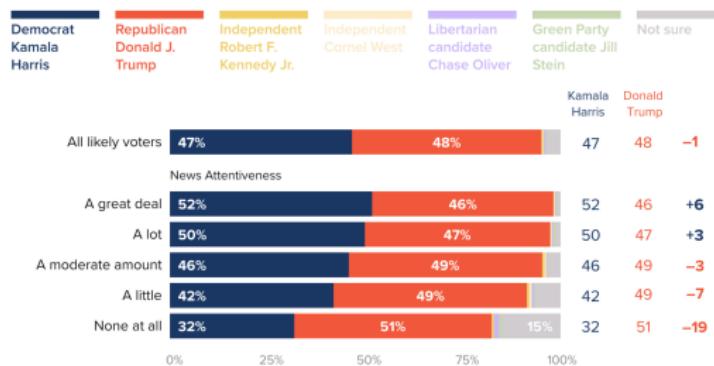
Note: Data through Sept. 19. • Source: Times analysis of national presidential election polls • By The New York Times

# THE 2024 ELECTION POLLS

- Pollsters used new techniques to address partisan non-response, like weighting on **recalled vote**
- Recalled vote weighting assumes that the electorate looks the same from year-to-year
- This is especially the case as pollsters begin using “likely voter” models

## Support for Harris Decreases as News Consumption Decreases

If the November 2024 election for U.S. president was being held tomorrow, and these were the candidates, who would you vote for?

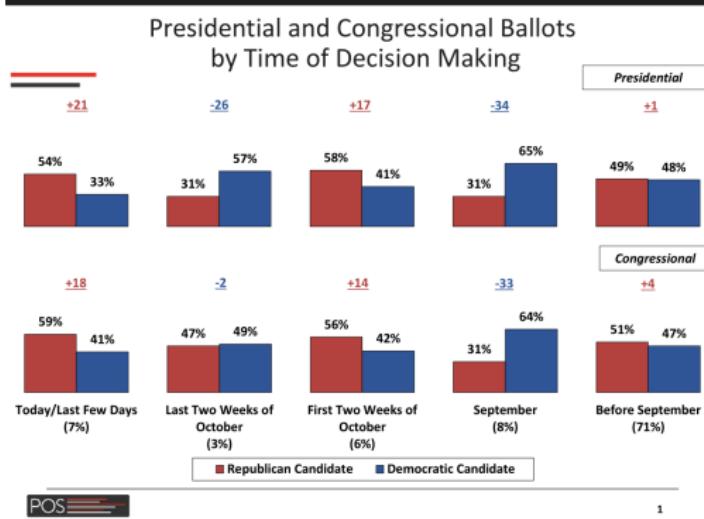


Oct 5–Nov 3, 2024 pooled surveys of 13,404 interviews of likely voters

DATA FOR PROGRESS

# THE 2024 ELECTION POLLS

- Pollsters used new techniques to address partisan non-response, like weighting on **recalled vote**
- Recalled vote weighting assumes that the electorate looks the same from year-to-year
- This is especially the case as pollsters begin using “likely voter” models



# POLLING CHALLENGES

2016 Presidential	Polls wrong due to <b>missing at random (MAR)</b> non-response bias by education
2020 Presidential	Polls wrong due to <b>missing not at random (MNAR)</b> non-response bias
2022 Midterm	The rise of partisan polling and <b>herding</b>
2024 Presidential	Pollsters try weighting on <b>recalled vote</b>