A 2016 Review: Why Key State Polls Were Wrong About Trump

Cohn, Nate . New York Times (Online) , New York: New York Times Company. May 31, 2017.

ABSTRACT (ENGLISH)

Pollsters are leaning toward optimistic explanations that their problems of last year can be fixed.

FULL TEXT

Nearly seven months after the presidential election, pollsters are still trying to answer a question that has rattled trust in their profession: Why did pre-election polls show Hillary Clinton leading Donald J. Trump in the battleground states that decided the presidency? Is political polling fundamentally broken? Or were the errors understandable and correctable?

At their annual conference in New Orleans this month, polling experts were inching toward the latter, more optimistic explanation. And there is mounting evidence to support their view.

At least three key types of error have emerged as likely contributors to the pro-Clinton bias in pre-election surveys. Undecided voters broke for Mr. Trump in the final days of the race, or in the voting booth. Turnout among Mr. Trump's supporters was somewhat higher than expected. And state polls, in particular, understated Mr. Trump's support in the decisive Rust Belt region, in part because those surveys did not adjust for the educational composition of the electorate —a key to the 2016 race.

Some of these errors will be easier to fix than others. But all of them seem to be good news for pollsters and others who depend on political surveys.

It might seem strange to argue that the polls could miss the result of an election and could still be trusted in the future. But there are some kinds of polling errors that pollsters can accept, even if the public never will.

There's nothing pollsters can do, for example, if undecided voters break for one candidate in the final hours. Even an error that puts more blame on the pollsters might be acceptable, provided it can be fixed.

Errors have happened enough in past elections to know that an upset was well within the realm of possibility in 2016. The Upshot model estimated that a polling misfire was about as likely as a baseball strikeout or a missed midrange field goal in football. It's not pretty, but it happens and will happen again, and a team wouldn't release a batter or a kicker because of a strikeout or a missed kick.

At the annual conference of the American Association of Public Opinion Research (AAPOR), as well as at a number of other meetings held earlier this year, evidence pointed toward an explanation in one of these categories:

Undecided voters turned to Trump

A postelection survey by Pew Research, and another by Global Strategy Group, a Democratic firm, re-contacted people who had taken their polls before the election. They found that undecided and minor-party voters broke for Mr. Trump by a considerable margin —far more than usual. Similarly, the exit polls found that late-deciding voters supported Mr. Trump by a considerable margin in several critical states. These three results imply that late movement boosted Mr. Trump by a modest margin, perhaps around two points.

But there are at least some reasons to question whether undecided voters really broke that decisively for Mr. Trump. For one, these measures might be biased by the tendency for respondents to over-report voting for the winner. And, at the time, few pollsters believed that undecided voters were poised to break overwhelmingly for Mr. Trump. These voters were a fairly representative mix of Republicans and Democrats, young and old, the well

educated and the less well educated, and whites and nonwhites.

The postelection shift is also potentially consistent with a subtly different explanation: the so-called "shy Trump" effect. The idea is that Trump supporters took telephone surveys but were embarrassed to divulge their support for an unpopular candidate. If true, the "undecided" voters were really Trump voters all along; they just didn't want to admit it to pollsters until after their candidate won.

This hypothesis has received a lot of attention, but there's not much evidence to back it up. Mr. Trump did not fare noticeably better in polls conducted online, where voters wouldn't have to admit socially undesirable views to a person on the other end of the telephone. Several studies reached similar conclusions.

Another issue is that the polls underestimated Republicans up and down the ticket.

"I find the idea that voters were embarrassed to support Ron Johnson in Wisconsin to be kind of absurd," argued Nick Gourevitch, a pollster at Global Strategy Group. "Instead, it's likely that the polls underestimated Republicans in ways that were not unique to Trump."

That doesn't mean there couldn't have been any "shy Trump" effect. A Morning Consult study, for instance , found some evidence of a modest effect. It just means it wasn't large enough to be clearly detected, let alone large enough to explain the preponderance of the polling error in key states.

Likely-voter screens may have tilted polls in Clinton's direction

It's hard to make a general statement about whether pollsters' likely-voter screens, which try to determine which voters are likely to turn out, were biased toward Mrs. Clinton, since there are many different kinds of such screens . But so far, there have been at least three publicly presented analyses of pre-election polls conducted via voter files: from two Democratic firms, TargetSmart and Global Strategy Group, and from the New York Times Upshot/Siena survey. All three of these analyses found that turnout was modestly more favorable to Mr. Trump than pre-election polls predicted.

These polls are not especially representative of the average public poll. Most public surveys simply ask respondents whether they'll vote, while these voter file-based surveys relied on past vote history. Our Upshot/Siena polls also asked people about their intention to vote, and Mrs. Clinton's supporters were likelier than Mr. Trump's supporters to stay home after indicating their intention to vote.

If other surveys received similarly biased responses, it could have tilted their results by a couple of percentage points in Mrs. Clinton's direction.

Education weighting seems to explain a lot

Education was a huge driver of presidential vote preference in the 2016 election, but many pollsters did not adjust their samples —a process known as weighting —to make sure they had the right number of well-educated or less educated respondents.

It's no small matter, since well-educated voters are much likelier to take surveys than less educated ones. About 45 percent of respondents in a typical national poll of adults will have a bachelor's degree or higher, even though the census says that only 28 percent of adults (those 18 and over) have a degree. Similarly, a bit more than 50 percent of respondents who say they're likely to vote have a degree, compared with 40 percent of voters in newly released 2016 census voting data.

This was a big deal in 2016, since Mrs. Clinton fared very well among well-educated voters. Her lead might have increased by around four percentage points in a typical national survey that wasn't weighted by education. The effect shrinks in polls weighted more heavily, including by party registration or past turnout, but there were virtually no public polls that were weighted this way.

The tendency for better-educated voters to respond to surveys in greater numbers has been true for a long time. What's new is the importance of education to presidential vote choice. Mrs. Clinton led Mr. Trump by 25 points among college-educated voters in pre-election national polls, up from President Obama's four-point edge in 2012. This made it a lot more important to weight by education. In the past, it barely mattered whether a political poll was weighted by education —which is probably part of why so many didn't do so.

The education issue doesn't just explain why polls were tilted toward Mrs. Clinton —it also helps explain why the

state polls fared so much worse than national polls. Most national polls were weighted by education, even as most state polls were not.

For all of the talk about last year's polling failure, the national surveys were quite accurate by any standard. Mrs. Clinton ultimately won the national popular vote by two points, just short of her three- or four-point lead in the final national surveys. The performance of national surveys has been one of the better reasons to assume that last year's misfire wasn't a broad indictment of public opinion polls.

This argument has become even more credible as pollsters have released their pre-election survey data. The preliminary data —so far from four pollsters in a dozen surveys —indicates that national surveys might have done pretty well in the Rust Belt states where the state polls were off by the most.

The education issue won't be easy for all state pollsters to fix. Most state polls contact registered voters, not all adults the way national surveys do, and there isn't an authoritative and up-to-date measure of the educational composition of registered or actual voters. Weighting by education can create additional complications. The effect of education can vary by race, so it can be a problem, for instance, if a pollster has the right number of well-educated voters but the wrong number of well-educated white voters. Several high-quality state pollsters did not weight by education, in part for these reasons.

But many lower-quality state pollsters did not even ask about education at all, suggesting that it wasn't on their radar as a potential issue in 2016. That's surprising. The potential for bias should have been fairly obvious, given the media coverage of Mr. Trump's strength among less educated voters and the well-established difference in response rates along educational lines.

The More Pessimistic Case

In theory, all of this could add up to something near a complete explanation of the polling error in 2016. It would explain why the national polls did well, while state polls did not. It would explain why the state polls performed most poorly in states with large numbers of white working-class voters. It wouldn't be easy to fix, but it certainly could be done. The AAPOR postelection report reflects this optimistic view of the state of survey research.

But there are a few loose ends. And those loose ends keep the possibility of a more pessimistic explanation alive. The loosest end: The state polls weighted by education didn't fare as well as the national polls. In fact, it's not clear whether a state poll weighted by education and with the benefit of a perfect likely-voter screen would have shown Mr. Trump ahead in the key states.

CNN/ORC, Quinnipiac and Marquette Law School all conducted live-interview polls weighted by education in Wisconsin and Pennsylvania over the final stretch. Methodologically, they were fairly similar to the best national polls. But Mrs. Clinton still had a considerable lead of four to six percentage points.

The same story seems to hold up among campaign pollsters. The public's window into private polling isn't very clear, but presentations by analysts from the Democratic firms Civis Analytics and Global Strategy Group indicated that Mrs. Clinton would have still led in their final surveys in Midwestern states or Pennsylvania, even after weighting by education and correcting the likely electorate.

Perhaps undecided voters could explain the remaining error. After all, these same states had the largest number of voters who switched to Mr. Trump after voting for President Obama; it makes sense that they would have been relatively likely to be undecided.

But there's a more pessimistic possibility, one that has been mainly promulgated by analysts at Civis Analytics. Their theory, advanced in postelection interviews and presentations at AAPOR, is that Trump voters weren't responding to telephone surveys because they had lower levels of civic engagement.

The notion that polls would miss disengaged voters is not new. It's probably the best-known response bias in polling: Respondents are likelier than people who don't respond to surveys to be voters, to be trusting of others, and to be volunteers or engaged in their communities. If these voters were also inclined toward Mr. Trump, it would help explain the bias toward Mrs. Clinton in pre-election polls.

This would pose a more serious challenge to survey research. It would suggest that declining response rates have finally taken a toll on the accuracy of political polling, and that would be hard to fix. The 2016 election wouldn't be

like a typical strikeout or a missed field goal —it would be more like an aging player who could no longer muster the leg strength or bat speed he had a decade ago.

It's certainly possible that Mr. Trump's white working-class supporters were less likely to respond to telephone surveys. But the data, at least in the public realm, is not very clear.

In Upshot/Siena surveys, registered Democrats were indeed likelier than registered Republicans to respond. This wasn't an issue for our polls —we adjusted our sample by party registration —but most public polls don't have a mechanism to correct for partisan nonresponse. If this were true more generally, and there's no way to know whether it was, that would have biased most other surveys toward the Democrats by a modest amount. A Civis analyst presented similar findings at AAPOR.

Partisan nonresponse, however, is a fairly well-established phenomenon, especially in lopsided media environments. Two of our three final surveys, for instance, were conducted after the final presidential debate —a fairly positive media environment for Mrs. Clinton. So the Upshot/Siena poll data does not necessarily indicate that there's a long-term bias in polling toward the Democrats, and it doesn't address whether voters with high civic engagement were likelier to support one candidate over the other.

The most relevant public data is a recent report by the Pew Research Center. It found, perhaps surprisingly, that Mr. Trump was stronger among civically engaged white voters, suggesting that nonresponse among socially isolated voters wouldn't have biased surveys toward Mrs. Clinton. SurveyMonkey conducted a similar analysis after the election and also did not find that Mr. Trump's support was correlated with civic engagement.

"Weighting by volunteerism didn't make a meaningful difference," SurveyMonkey's Jon Cohen said. "Basically, Clinton and Trump voters said they volunteer at similar levels."

Mr. Trump's supporters might well have been less likely to respond to telephone surveys. But, at least in publicly available data, this case remains underdeveloped.

What's very clear is that several typical sources of polling error, in addition to the education issue in lower-quality state polls, contributed to a pro-Clinton bias in pre-election polls. It may or may not explain all of the error, but it probably explains most of it.

That seems to be positive news for pollsters. Without a good explanation for the misfire, it would be understandable to wonder whether lower response rates had degraded the political survey research to the point where our metaphorical field goal kicker ought to consider retirement.

But the education gap among supporters of Mr. Trump and the eventual Democratic nominee will probably persist into 2020, and it will be especially challenging to pollsters in the Northern states that had a relatively high percentage of working-class whites and that may again play an outsize role in the Electoral College.

The lack of high-quality state polling will also be tough to fix, especially if the cost of polling rises further or if local newspaper budgets continue to shrink or vanish. The failure of many state pollsters to even ask respondents about education does not inspire much confidence in their ability to stave off less predictable sources of bias. Many of the challenges that pollsters faced in 2016 aren't going away. Next time, the challenges could easily be greater.

The Upshot provides news, analysis and graphics about politics, policy and everyday life. Follow us on Facebook and Twitter . Sign up for our newsletter .

Source URL: https://www.nytimes.com/2017/05/31/upshot/a-2016-review-why-key-state-polls-were-wrong-about-trump.html?partner=bloomberg

Credit: NATE COHN

DETAILS

Subject:	Polls &surveys; Election results; Education; Voter behavior; Bias; Political parties

People:	Trump, Donald J
Identifier / keyword:	Presidential Election of 2016; Polls and Public Opinion; Research; Clinton, Hillary Rodham; Trump, Donald J
URL:	https://www.nytimes.com/2017/05/31/upshot/a-2016-review-why-key-state-polls-were-wrong-about-trump.html?partner=bloomberg
Publication title:	New York Times (Online); New York
Publication year:	2017
Publication date:	May 31, 2017
Section:	upshot
Publisher:	New York Times Company
Place of publication:	New York
Country of publication:	United States, New York
Publication subject:	General Interest Periodicals--United States
Source type:	Blog, Podcast, or Website
Language of publication:	English
Document type:	News
ProQuest document ID:	1903876868
Document URL:	https://www.proquest.com/blogs-podcasts-websites/2016-review-why-key-state-polls-were-wrong-about/docview/1903876868/se-2?accountid=9840
Copyright:	Copyright 2017 The New York Times Company
Last updated:	2018-08-02
Database:	Global Newsstream

LINKS

Find It@ CSUF