

NATIONAL POLITICS | By Charlie Cook, December 30, 2016

## How Analytical Models Failed Clinton



*This story was originally published on [nationaljournal.com](http://nationaljournal.com) on December 27, 2016*

The November elections pitted Democrats against Republicans, conservatives against liberals, Trump-style populists and tea partiers against the establishment and conventional politicians. Another contest, followed mainly by political aficionados, matched traditional pollsters against newly fashionable analytics wizards, some of whom—pretentiously in my opinion—called themselves “data scientists.”

It was well known that traditional polling was having problems. The numbing effect of billions of telemarketing calls and the advent of caller ID and voice mail had reduced response rates (the percentage of completed interviews for every hundred attempts) from the 40s a couple of decades ago to the high single digits. As they struggled to get truly representative samples, pollsters “weighted” their data more than ever before, making assumptions of what the electorate would look like on election days that were weeks, months, or even a year or more away.

Problems with traditional, live-telephone polling led to experimentation and more recently a growing acceptance of new methods like Interactive Voice Response, popularly known as robo-polls, and on-line polling. Each new method brings both good and bad attributes. As a traditionalist, I see the new techniques as bad ideas whose time is regrettably coming.

The other trend is “analytics,” which incorporates information from a variety of sources—Census Bureau studies, commercially available market data combined with past election results, and conclusions gleaned from polling, voter canvassing, and economic measures such as the unemployment rate. This “big data” enable campaigns to model the anticipated electorate, identify voters most likely to be sympathetic to their candidates, and shape their messages accordingly.

The roots of campaign analytics go back to the 1970’s when Democratic campaign consultant Matt Reese and Republican consultant Eddie Mahe promoted a new technology branded Claritas, a geo-demographic targeting system centered on lifestyles and neighborhoods based on a market-segmentation platform developed by computer scientist Jonathan Robbin (Claritas is now owned by Nielsen). It was an idea ahead of its time, too expensive for most campaigns, and it eventually left the political theater altogether.

In 2004 the Howard Dean, George W. Bush-Dick Cheney, and John Kerry-John Edwards presidential campaigns advanced the uses of data to contact voters, but it was the 2008 campaign of Barack Obama that took analytics to a whole new level. The infatuation with analytics after Obama’s reelection in 2012 prompted some of his operatives to say they didn’t need traditional polling anymore.

When Hillary Clinton began putting together her 2016 campaign, she brought on board many Obama veterans, go

ing all in for the new technology. Donald Trump's general-election campaign also employed analytics, though how sophisticated and important it was in his victory is a matter of considerable debate. House and Senate campaign committees and super-PACs also used analytics to varying degrees.

The reliance, or perhaps overreliance on analytics, may be one of the factors contributing to Clinton's surprise defeat. The Clinton team was so confident in its analytical models that it opted not to conduct tracking polls in a number of states during the last month of the campaign. As a consequence, deteriorating support in states such as Michigan and Wisconsin fell below the radar screen, slippage that that traditional tracking polls would have certainly caught.

According to Kantar Media/CMAG data, the Clinton campaign did not go on the air with television ads in Wisconsin until the weeks of Oct. 25 and Nov. 1, spending in the end just \$2.6 million. Super PACs backing Clinton didn't air ads in Wisconsin until the last week of the campaign. In Michigan, aside from a tiny \$16,000 buy by the campaign and a party committee the week of Oct. 25, the Clinton campaign and its allied groups didn't conduct a concerted advertising effort until a week before the election.

In fact, the Clinton campaign spent more money on television advertising in Arizona, Georgia, and the Omaha, Nebraska markets than in Michigan and Wisconsin combined. It was Michigan and Wisconsin, along with Pennsylvania (the Clinton campaign and allied groups did spend \$42 million on television in the Keystone State), that effectively cost Democrats the presidency.

In the end, the national polls fared better than commonly thought. The RealClearPolitics average of national polls showed Clinton ahead by 3.2 percentage points going into Election Day, and the final ABC News/Washington Post, CBS News, NBC News/Wall Street Journal, and Fox News polls each had Clinton ahead by 4 points (the last CNN national poll was taken two weeks before the election and had Clinton ahead by 5 points). She ended up winning the national popular vote by 2.1 percentage points, 48.2 to 46.1. Thus the RCP average was off by 1.1 percentage points, the network polls were off by 1.9 percentage points. They were off by far more in 2012, but nobody noticed because the popular vote and Electoral College tally went the same direction. If one buys the argument that the race changed considerably in the last week, for whatever reason, then some of these polls may not have been off by much if at all.

Like so many other aspects of this election, a lot of small misses added up to one giant error on the outcome of the election. In 54 out of our 58 presidential elections, the winner of the popular vote also prevailed in the electoral vote. A good rule of thumb is that if a candidate wins the popular vote by at least 2 percentage points, he or she will almost certainly capture the Electoral College. So in an election when one candidate is thought to have a comfortable lead of more than 2 percentage points, there is a reasonable expectation that the electoral vote will go in the same direction. But if the final result is hovering at the 2-point threshold, that's a wrinkle that can create an unexpected outcome, as the Clinton team learned to its dismay.

It was the individual state polling that badly missed the mark. In Wisconsin, Clinton led in each of the 32 public polls from mid-August on. The final Marquette University Law School, generally considered to be the most respected in the state, had the Democrat up by 6 points. She lost by eight-tenths of a point.

In Pennsylvania, Clinton led in 37 out of 38 polls beginning in early August. CNN's last poll had Clinton up by 4 points, the final Quinnipiac poll had her up by 5 points, and the RealClearPolitics average had her up by 1.9 percentage points. She lost by eight-tenths of a point.

In Michigan, Clinton was ahead in 25 out of 26 polls taken from the beginning of August on. The Detroit Free Press's last poll had her up by four points, and the RealClearPolitics average had her up by 3.6 points. She lost by two-tenths of a point.

It's worth noting that state polls conducted by news organizations and universities vary enormously in quality and sophistication. Few state-based news organizations spend the kind of money on polling that many once did. Much of the state-level polling is of a dime-store quality, conducted by polling firms that are even unfamiliar to political pros.

Experienced journalists might argue that the overreliance by reporters on both polls and analytics has led to a decrease in shoe-leather, on-the-ground reporting that might have picked up movements in the electorate that the polls missed. As the Michigan results came in on election night, I vividly recalled that two congressmen from Michigan—one a Democrat, the other a Republican—had been warning me for months that Michigan was more competitive than publicly thought. I wished I had listened.

The analytical models for both sides pointed to a Clinton victory, albeit not a runaway. The Clinton campaign and super PACs had several of the most highly regarded polling firms in the Democratic Party, yet in the places that ended up mattering, very little if any polling was done. So while 2016 wasn't a victory for traditional polling, it certainly took a lot of the luster from analytics. In the end, big data mattered very little.

*CLARIFICATION: According to Kantar Media/CMAG, a firm that monitors political advertising, the Clinton campaign's advertising started the week of Nov. 1 in Michigan and Oct. 25 in Wisconsin. The campaign also made a \$70 million national ad buy, \$59 million of which would have been prior to Oct. 25, and some of that would have gone into Michigan and Wisconsin. The campaign also had field organizations in both states.*