Explaining the Convergence Mystery: A Tautology Problem?

David Moore*

Tags: survey practice

Survey Practice

Vol. 1, Issue 5, 2008

Explaining the Convergence Mystery: A Tautology Problem?

The two major types of explanations offered by our experts for the convergence of polls right before the election are based on 1) changes in the pollsters' methodology and 2) changes in the certainty of the vote choice.

The first explanation suggests that in the final weeks of the campaign, many pollsters adjust their likely voter models (mentioned by Lavrakas and Blum) or they increase their sample sizes (Dresser).

Lavrakas argues that the adjustment of the voter models, even when done explicitly to make their outcomes more consistent with other polls, should be seen as a positive action rather than as a “suspicious” activity. However viewed, such last minute changes could account for some of the convergence.

Dresser mentions the tendency of many pollsters to substantially increase their sample sizes for their final pre-election polls, to insure as small a margin of sampling error as they can reasonably afford. He specifically mentions Pew, Harris and ABC/ *Washington Post*.

Blum suggests that the outlier polls during the campaign were either a) less likely to poll in the final three days, or b) more likely to allocate their undecided voters, thus bringing them closer to the mean.

Zukin speculates that the convergence has less to do with pollsters' methods and more to do with the “phenomenon we are measuring.” Specifically, as voters become more certain about their choices, polls will tend to converge toward each other. That sentiment is also found in the explanations by Lavrakas and Blum. This is the same reasoning that Pew's Andrew Kohut made on [NPR](#) on Nov. 2, 2008, when he told the host, Andrea Seabrook, that “the closer we get to the election, the more crystallized public opinion is” and thus “it's pretty typical that the polls – rigorous polls – all come together in the final weeks.”

Mike Traugott's analysis is different from all the rest, because he focuses on

* **Institution:** University of New Hampshire

the historical trends and their relation to a “normal” vote. My sense is that this approach, while important for understanding the 2008 election in the context of American elections generally, does not get at the heart of the polling issue being debated here. The question is not so much why Obama’s victory margin was about seven percentage points, but why the polls in October showed wildly varying results during October and then converged to a small variance right before election day. As Mark Blumenthal showed [in his commentary about the “convergence mystery,”](#) the same pattern of convergence was found in the state polls as well as in the national polls – relatively high variance during October, but dramatically lower variance in the final election estimates (even though Obama’s average lead among those polls varied by less than two percentage points).

WHY ARE POLL RESULTS RELATED TO SOLIDITY OF VOTE?

Virtually all of the factors offered by our experts seem plausible partial explanations for the convergence. The *implications* of these arguments need to be analyzed further, especially as they relate to the confidence we can put in pre-election polls during the campaign, as well as the confidence we can place in polls that deal with public policy issues when the public does not seem to have crystallized opinions, when sample sizes are much smaller, and where there is no “screening” for likely voters or other indicators of citizen engagement in the issues.

For the time being, however, it may be useful to ask the question – how can we accept the notion that the more “uncertain” public opinion is, the greater the disparity we are likely to find in the polls?

Since posting the original article on the convergence mystery, based on *weekly* poll averages, I posted a further analysis on [pollster.com](#) which shows the results of eight *daily* tracking polls for the last four weeks of the campaign. That graph is reproduced below:

Figure 1

The next graph shows the variance in the polls over this same time period. Note that there is a persistent decline over the month of October, interrupted by three spikes – each one about five days after a vice presidential or presidential debate. Given that the polls are mostly three-day averages, that means the spikes all became evident about two days after the debate – which suggests poll variance probably followed commentary about the debate, rather than the debate itself.

Figure 2

These results tend to confirm the experts' consensus that as election day approaches, (when, presumably, voter certainty becomes greater), the more likely the polls are to converge.

But these results also suggest that the debates had the effect of causing short-term increased variance in the polls. Was that because voter uncertainty increased after the debate (and the commentary surrounding the debate)? Is there any evidence for this?

Most important, what is the theory that suggests why polls become more unreliable (inconsistent) just because of voter uncertainty?

The margin of sampling error we calculate has nothing in its formula that relates to the character of what is being measured, only the sample size. Typically, pollsters will add that results can also be affected by nonresponse, question wording, and timing of the interviews – but in the case of the daily tracking polls, none of these additional factors seems to offer a plausible explanation for why results vary among the polls.

All of the daily tracking polls were, of course, conducted “daily,” thus eliminating timing as an explanation for their differences. The vote choice questions are almost all identical, unlike, say, questions on a bailout or other policy matters, when question wording is an accepted explanation for variations in poll results. And nonresponse was not suggested by any of the experts as a factor in the polls' variability.

A TAUTOLOGICAL OR FIRST-LEVEL EXPLANATION FOR CONVERGENCE?

Thus, the issue comes back to crystallized opinions. The experts and the data all

seem to agree that the more crystallized are voter choices, the more “accurate” (consistent) are the individual polls (compared with each other). But this seems less an explanation than a different way of expressing an empirical observation. It’s at best, I would think, a first-level explanation. We know that polls converge, and we have reason to believe that more people have made up their minds the closer polls are to election day. But the explanation also seems dangerously close to a tautology – the same as saying that polls converge toward election time, because we’ve observed that polls converge toward election time.

The question is: Why are polls – conducted at the same time, using virtually the same wording – supposedly more accurate (reliable) when opinion is more crystallized?

And if this relationship between poll variance and uncrystallized opinion is significant, do we need to design a new margin of “crystallization” error, to warn poll users of variation due to voter/citizen uncertainty about the issue being measured?

Your comments are welcome.