Poll Aggregators

At the simplest level, a poll aggregator is an individual or organization that brings together many poll results from a number of different sources. Generally, poll aggregators compile results of polls through a certain time period, using national or state polls. But not all aggregators are alike. Some aggregators report a simple rolling average of recent poll results, removing or giving less weight to older polls from the average. Other aggregators use regression-based analyses — a method for adjusting data to account for unusual results ("outliers"). Other aggregators combine additional data like historical election results or economic data with current polling data through statistical methods – these are often called modelers.

The most well-known aggregators track the support for candidates during a race, although a few also track other common polling questions, such as presidential approval and the direction of the country.

Advantages and Disadvantages of Aggregators

There are advantages and disadvantages to aggregation, and journalists need to employ the same attention to detail as they do when reporting about single polls.

The benefit of a poll aggregator is that an average of polling results should give a reporter a more reliable and wider perspective on a race, rather than relying on just one poll's result.

For example, because of variation among the polls, looking at the average across a number of polls to characterize the presidential race during the summer of 2015 would provide a more complete and more stable picture than looking at the result of just one poll.

During the election cycle, individual media or academic polls are still crucial for reporters. Most of these polls go well beyond the horserace question, providing a snapshot of the electorate and insights as to what voters care about and are thinking at that given time.

A major disadvantage is that some aggregators include all pollsters' results regardless of methodological quality. These aggregators thus include poorly conducted polls and if such pollsters do a lot of polls, their polls will dominate the aggregate.

Further, poll aggregators are only as good as the polls themselves. If pollsters are systematically wrong in a certain direction, when a poll aggregator brings all these polls together, the errors may, in fact, get multiplied. A major concern for aggregators is herding, or the tendency of some pollsters to adjust their results to mirror the findings of other surveys. As Nate Silver has argued, "herding" can make the average poll more accurate even as it makes the polling average less accurate.

Finally, there has been a fear expressed that as the poll aggregators rise in popularity, media organizations will invest in poll aggregators or modelers and fewer resources will go to conducting the polls themselves. High-quality, probability surveys of the American public are expensive and labor intensive. But with fewer high-quality polls, the aggregators may see their accuracy decline with insufficient data to analyze.

Some Aggregators Are Better Than Others

All poll aggregators make their own choices as to which polls to include and the level of importance or weight given to different polls. Some will incorporate a poll's track record of matching past election results. Some aggregators may include partisan pollsters, and others may not, or not allow certain methodologies, such as IVR or robopolls. Some will include other factors, such as whether or not a polling organization is a member of AAPOR's Transparency Initiative. In addition, most of the modelers or forecasters will calculate the probabilities of their predictions -- basically giving the odds that their forecast will be right.

A number of modelers supplement poll data with other information, such as the unemployment rate, with the goal of predicting the outcome of the election.

Following every election, there is a lot of post-game analysis by pundits and political scientists about which pollsters, aggregators, and modelers came closest to calling the election.

In the days leading up to the 2012 election, individual pollsters generally showed a close race, with a few showing Mitt Romney ahead of Barack Obama by a percentage point or in a tie. The aggregators generally came close in 2012 to predicting the final outcome, with some within a few points of Obama's final 3.9 point victory.

But as many a pollster will say, past results do not guarantee future performance. In 2014, a number of aggregators and modelers predicted closer Senate races than what actually occurred at the voting booths, when Republicans dominated in a number of state races.

What You Need to Know about Poll Aggregations

When citing any numbers based on an aggregator or modeler, it is essential for reporters to indicate whether the figure reported is simply an average of recent polls, or whether it is an estimate attempting to predict an election's outcome that incorporates other data into its forecast.

Do you need to report exactly what type of data is included in a model? Not if your story isn't about the model itself. However, linking to an aggregator's or modeler's methodology would provide transparency to your readers.

Reporters should also exercise caution when citing state-level poll aggregators and models. Iowa, New Hampshire, and other battleground states are heavily polled leading up to a primary or election, and there are generally enough polls in a relatively short period of time for aggregators to do a robust analysis of state poll results. But many states are notoriously difficult to survey, due to low-density populations and varying degrees of landline or cell-phone penetration.