Before Citing a Poll, Read the Fine Print

By NATE SILVER

On Saturday, a survey came out showing Mitt Romney with a large, 21-point lead in South Carolina. The poll is something of an outlier relative to other recent polls of the state, all of which show Mr. Romney ahead, but by margins ranging from 2 to 9 points.

The poll, conducted by Ipsos for Reuters, has already attracted more than 200 citations in the mainstream media. Most of these articles, however, neglected to mention a key detail: in a break with Ipsos’ typical methodology, the survey was conducted online.

Reuters did disclose this in its write-up of the poll, but it wasn’t mentioned until the 17th paragraph:

The Reuters/Ipsos poll was conducted online from January 10-13 with a sample of 995 South Carolina registered voters. It included 398 Republicans and 380 Democrats.

There are a couple of other important details here as well, none of which necessarily speak favorably to the poll’s potential accuracy. The poll was conducted among registered rather than likely voters, something which is almost certainly
a mistake so close to a primary since turnout in primaries is normally quite low. And it contained a relatively small sample size: 398 Republicans, about half the average of other recent surveys of the state.

Now it becomes easier to understand why the poll showed such distinct results from others conducted at the same time: it used a very different, and possibly rather dubious, methodology.

Internet-based polls are very likely to be a part of polling’s future, and my view is not necessarily that they should be dismissed out of hand. However, they need to be approached with caution.

The central challenge that Internet polls face is in collecting a random sample, which is the sine qua non of a scientific survey. There is no centralized database of e-mail addresses, nor any other method to “ping” someone at random to invite them to participate in an online poll. Many people have several e-mail addresses, while about 20 percent of Americans still do not go online at all.

The situation can be contrasted with the platonic ideal of a telephone poll, in which everybody has a phone number, and they each have an equal chance of being reached through a random digit dial method.

In reality, telephone polling falls short of the platonic ideal, while the best online polls take steps to make their samples effectively random. Some telephone polls, especially those that are conducted through automated scripts, do not call
cellphone numbers, even though more than a quarter of American households do not have landline telephones at all, with the fraction increasing by several percentage points every year. Meanwhile, households often share a single number between a family, or they have multiple telephone lines; careful pollsters take steps to ensure that their samples are not biased by these problems, but others apply a blitzkrieg approach to polling and do not.

Practices for conducting online polls vary significantly from survey firm to survey firm. At the favorable extreme is the company Knowledge Networks, which goes so far as to provide Internet service to people who do not already have it.

The company YouGov, meanwhile, which has produced reasonably accurate results with online polls, recruits participants by advertising at nonpolitical websites, and relies on databases of e-mail addresses that they purchase from marketers. While not strictly random, this method has some promise of developing a relatively unbiased sample, as it is not correlated with political ideology in any obvious way. My view is that a thoughtfully-conducted online poll like YouGov is probably no worse than a telephone poll that does not call cellphone numbers, although neither is ideal.

At the other end of the spectrum are online polls that are blatantly unscientific. The worst example is the company IBOPE Zogby, which has had extremely inaccurate results in past elections. Rather than making any effort to recruit a random sample, the company instead relies on people who sign up for the survey voluntarily. What’s worse, Zogby
encourages people who are in their database to invite their friends to join the panel as well. Since most of us have friends and acquaintances who share similar political beliefs and similar demographic characteristics, this potentially biases the sample even more.

Companies like Zogby sometimes claim that they create a random sample by inviting only a subset of their database to participate in any given survey, but this argument is dubious. If the initial method for recruiting the sample is highly nonrandom, then picking a random sample from among the nonrandom pool will not solve the problem, any more than taking a test tube from a tainted barrel of wine will restore it to being a fine Bordeaux.

My view is that online polls should be regarded as “guilty until proven innocent.” Because a company like YouGov has a reasonably robust track record of results, for instance, and those results testify to a good (although not great) level of reliability, I am happy to include their surveys in our polling-based forecasts. On the other hand, we do not include Zogby’s online polls in our forecasts at all. Finally, when a pollster like Ipsos begins conducting online surveys for the first time, we do include its results but weight them much less than polls that use a more traditional methodology.

In most contexts, news organizations do not have the option of hedging their bets in this way: they either deem a poll to be credible and publish its results, or they do not. It is not my intention to play the role of “poll czar” and tell news organizations what they should and should not do.
My predisposition, however, is that providing more information to the reader is usually the right default. In this case, that might imply reporting the results from some of the better online polls, but also highlighting very clearly that they have used an unconventional methodology that may violate the assumption of a truly random sample in some important ways.

It would help, of course, if the polling organizations were more upfront about disclosing their methodology rather than burying the information in the fine print. Ipsos is not the only offender here. The company Rasmussen Reports, for instance, recently adopted a hybrid approach in which it uses an online panel to supplement results from its automated telephone polling, but I rarely see this mentioned when Rasmussen Reports polls are cited.

The other details of a poll can matter as well. Sometimes, a poll will be released to the public several days after its last interviews were conducted. News accounts sometimes report on the poll as though it is newer than other surveys, and therefore says something important about the momentum in the race, when an examination of the fine print may reveal that the poll is already outdated. In the context of a general election, where polling results are fairly stable, this may not matter so much. But voter preferences can significantly realign themselves on a daily basis in advance of primaries and caucuses, making the timing of a poll exceptionally important.

As landline penetration decreases, meanwhile, it is probably
increasingly important to monitor which polls include cellphone numbers in their samples and which do not. Most major news organizations, like The New York Times, now include cellphones in their sample as a matter of course, but others are inconsistent about doing so.

Finally, there are cases when even strong pollsters can make an oversight in how they develop their samples — for instance, by failing to include independent voters in a primary or caucus poll of a state where they are welcome to participate.

Let me acknowledge some hypocrisy here. For instance, I certainly do not mention that a poll has failed to include cellphones every time that I report on its results, nor do I mention every time when a pollster has a dubious track record. Some of this is for practical reasons: if every 400-word blog post were accompanied by 600 words of caveats and qualifications, it would hard to get the main thrust of the analysis across. Also, because our forecast methods rely on averaging different polls together, and because our averaging method puts more weight on surveys that have a better track record and which apply more careful methodological practices, the hope is that a methodological quirk here and there will not have that much impact on the big picture. If you think I am striking the wrong balance, please let me know in the comment section.

Nevertheless, particularly when a poll produces what appears to be an outlying result, as the recent Ipsos poll of South Carolina did, reporters and analysts should treat it with
suspicion and consider whether the discrepancies are explained by questionable methodology. Too often these outlier polls receive more attention precisely because their results are surprising or unexpected, but they usually deserve less.