How Google sees the race: Search data sheds light on likely voters

November 6 2012, by Peter Reuell

The challenge for pollsters is that although many people may say they intend to vote, studies suggest that as many as two-thirds of those who do not vote told pollsters that they would. Credit: File photo, Justin Ide/Harvard Staff Photographer

As the dust settles after the presidential election, pundits and politicians from both sides of the aisle will begin the work of analyzing the results, trying to determine which issues tipped the scales and where the candidates went right or wrong.

All the while, another analysis will be under way, behind the scenes. Just as they do following every election, pollsters will take a hard look at how accurate their predictions were. The results of that analysis will then be used to improve the way they model everything from who is likely to cast a ballot to voter turnout, all with the goal of sharpening their predictions for the next election.

Harvard's Seth Stephens-Davidowitz says there's a better way.

Stephens-Davidowitz, a Ph.D. student in economics, uses Google Insights for Search, an online tool for extracting data from the millions of daily Google searches, and then uses statistical tools to analyze the data to gain insights on who is likely to vote and on voter turnout on Election Day.

"There are a number of problems with traditional polling," Stephens-Davidowitz said. "Technically, polls are anonymous, but they don't feel anonymous, so for whatever reason, people tend to say what they think a pollster wants to hear. One of the other issues that isn't talked about much is that many polls are down to only about a 10 percent response rate, so they have had to turn to different weighting models.

"Meanwhile, Google and other digital tools are becoming ubiquitous," he continued. "I think we may have already reached a crossing point—I believe if we put our mind to it, we could do better than traditional polls."

As an example of the potential advantages Google data holds over traditional polls, Stephens-Davidowitz pointed to efforts to identify likely voters.

The challenge for pollsters is that although many people may say they intend to vote, studies suggest that as many as two-thirds of those who do

not vote told pollsters that they would. People who search Google for phrases like "how to vote" or "where to vote," however, are much more likely to actually cast ballots, Stephens-Davidowitz said.

"Telling a pollster you plan to vote is pretty meaningless," he said. "But if you're Googling how to vote, that can be pretty meaningful in terms of who is a likely voter."

There is evidence that the technique works, he said. By analyzing searches conducted in the lead-up to both the 2008 presidential election and the 2010 midterms, Stephens-Davidowitz was able to accurately predict where voter turnout would rise and fall with much greater accuracy than other indicators, such as early voting.

Stephens-Davidowitz warned that Google data shouldn't be seen as a perfect predictor of electoral behavior, but as another tool that can complement traditional polling techniques.

"It's a new technique—we don't have that many elections where this type of material has been available, so we may not have the perfect weighting or the perfect way to analyze the data," he said. "Still, I think it has a lot of potential, especially since turnout is such a big factor in elections.

"The advantages to this technique are that the sample sizes are much larger than traditional polls, you don't have people hanging up on you, and you have the candor of the online world," he added. "When you put those three things together, it's a powerful combination."

This story is published courtesy of the Harvard Gazette, *Harvard University's official newspaper. For additional university news, visit* Harvard.edu.

Provided by Harvard University