

# Want the most telling presidential polling data? Professor says turn to Twitter

September 22 2016, by Joe O'Connell



Credit: Younghee Jang/Northeastern University

National polls have traditionally been a go-to barometer to gauge public opinion of presidential candidates and who has the inside track on the race for the White House. But Northeastern assistant professor of

political science Nick Beauchamp says state-level polls provide an even better example of voter intention. The problem, he says, is that state-level polling is rare and often focuses primarily on swing states.

So Beauchamp created an innovative computational model to gauge up-to-date voter intentions in individual states using Twitter. In a paper published on Sept. 13 in the *American Journal of Political Science*, Beauchamp explains that social media is an ideal platform for predictive polling data.

"There is a strong need for temporal data in a short period of time," said Beauchamp, who studies political persuasion and how political opinions are formed and change over time. "State-level polling is expensive and hard to do well. But if we could afford it, we would poll on an even finer level."

Beauchamp sampled about 120 million political tweets from 24 [swing states](#) that were posted during the 2012 presidential election, from Sept. 1 to Election Day, and used machine learning methods to discover which of the thousands of words people used in their tweets were correlated with changes in polls that were already out there. These correlations allowed him to accurately estimate vote intention in other states and on days that were not polled.

Beauchamp found that these Twitter measurements matched dips and rises in the polls during the campaign, predicted all but two of the states correctly in the final election, and could estimate the next day's polling results before those polls were released.

This method worked especially well during fast-changing moments in the campaign, such as during the first 2012 debate where the approach was able to detect the surge towards and then away from Romney. "The whole process seemed to work better during or right after presidential

debates," Beauchamp noted. "People were synchronized by these shared public events."

Beauchamp said he understands some of the skepticism that can arise from using [social media](#) as a presidential predictive model, specifically the lack of substance of tweets and that participants are not representative of the whole voter base. However, he argues the same could be said about traditional polls.

"This is the same sort of problem that polling runs into in all forms," he said. "People who actually answer calls from pollsters are not indicative of the greater voting population. In both cases, we have to adjust the data to make it a more representative reflection of national or state-level opinion. With the Twitter data, though, there's a bit more work to find the words and phrases that specifically correlate with representative measures of vote intention."

In the long term, Beauchamp said he hopes this method can be extended to determine people's opinions of issues and topics that are rarely or never polled, such as local candidates or city services; specific policy approaches to issues like gun control, immigration, or transgender rights; various economic measures; or campaigns in countries less well polled than the US.

**More information:** Nicholas Beauchamp. Predicting and Interpolating State-Level Polls Using Twitter Textual Data, *American Journal of Political Science* (2016). [DOI: 10.1111/ajps.12274](https://doi.org/10.1111/ajps.12274)

Provided by Northeastern University

Citation: Want the most telling presidential polling data? Professor says turn to Twitter (2016,

September 22) retrieved 25 April 2024 from <https://phys.org/news/2016-09-presidential-polling-professor-twitter.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.