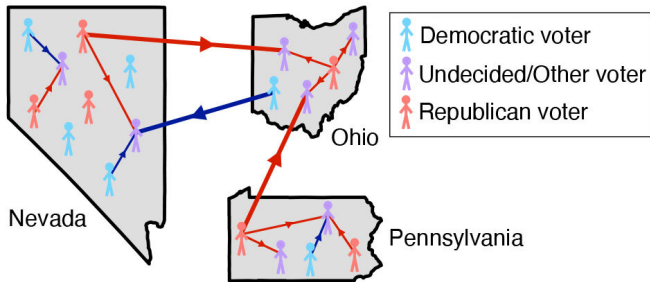


# Forecasting elections with a model of infectious diseases

28 October 2020



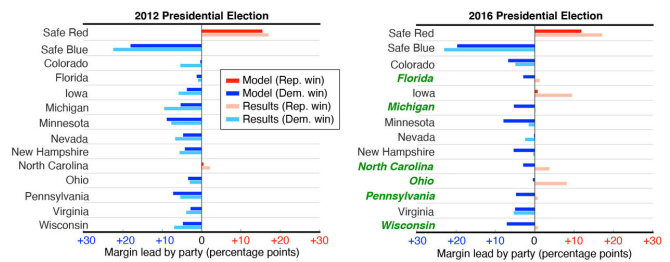
Voters can interact both within and between states, thus potentially influencing each other's political opinions. Credit: Figure courtesy of Alexandria Volkening, Daniel F. Linder, Mason A. Porter, and Grzegorz A. Rempala.

Forecasting elections is a high-stakes problem. Politicians and voters alike are often desperate to know the outcome of a close race, but providing them with incomplete or inaccurate predictions can be misleading. And election forecasting is already an innately challenging endeavor—the modeling process is rife with uncertainty, incomplete information, and subjective choices, all of which must be deftly handled. Political pundits and researchers have implemented a number of successful approaches for forecasting election outcomes, with varying degrees of transparency and complexity. However, election forecasts can be difficult to interpret and may leave many questions unanswered after close races unfold.

These challenges led researchers to wonder if applying a [disease model](#) to elections could widen the community involved in political forecasting. In a paper publishing today in *SIAM Review*, Alexandria Volkening (Northwestern University), Daniel F. Linder (Augusta University), Mason A. Porter (University of California, Los Angeles), and Grzegorz A. Rempala (The Ohio State University) borrowed ideas from epidemiology to develop a new method for forecasting elections. The team

hoped to expand the community that engages with polling data and raise research questions from a new perspective; the multidisciplinary nature of their infectious disease model was a virtue in this regard. "Our work is entirely open-source," Porter said. "Hopefully that will encourage others to further build on our ideas and develop their own methods for forecasting elections."

In their new paper, the authors propose a data-driven mathematical model of the evolution of political opinions during U.S. elections. They found their model's parameters using aggregated polling data, which enabled them to track the percentages of Democratic and Republican voters over time and forecast the vote margins in each state. The authors emphasized simplicity and transparency in their approach and consider these traits to be particular strengths of their model. "Complicated models need to account for uncertainty in many parameters at once," Rempala said.



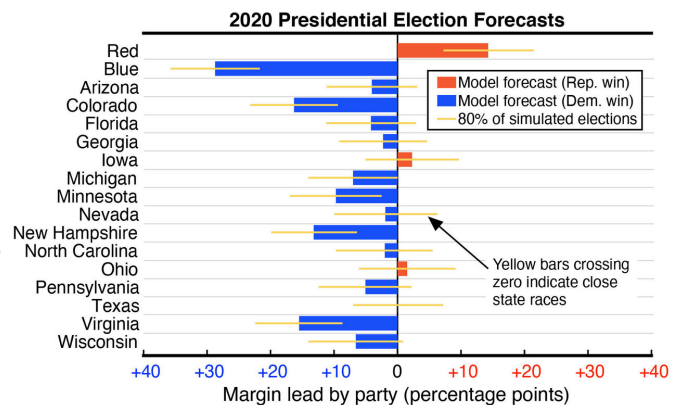
Forecasts of the 2012 and 2016 presidential elections, developed using polling data from up until Election Day. The darker colored bars show the model's forecasts, and the lighter colored bars are the election results. States for which the model produced incorrect predictions are written with green text. Credit: Figure courtesy of Alexandria Volkening, Daniel F. Linder, Mason A. Porter, and Grzegorz A. Rempala.

This study predominantly focused on the influence

that voters in different states may exert on each other, since accurately accounting for interactions between states is crucial for the production of reliable forecasts. The [election outcomes](#) in states with similar demographics are often correlated, and states may also influence each other asymmetrically; for example, the voters in Ohio may more strongly influence the voters in Pennsylvania than the reverse. The strength of a state's influence can depend on a number of factors, including the amount of time that candidates spend campaigning there and the state's coverage in the news. To develop their forecasting approach, the team repurposed ideas from the compartmental modeling of biological diseases. Mathematicians often utilize compartmental models—which categorize individuals into a few distinct types (i.e., compartments)—to examine the spread of infectious diseases like influenza and COVID-19. A widely-studied compartmental model called the susceptible-infected-susceptible (SIS) model divides a population into two groups: those who are susceptible to becoming sick and those who are currently infected. The SIS model then tracks the fractions of susceptible and infected individuals in a community over time, based on the factors of transmission and recovery. When an infected person interacts with a susceptible person, the susceptible individual may become infected. An infected person also has a certain chance of recovering and becoming susceptible again.

Because there are two major political parties in the U.S., the authors employed a modified version of an SIS model with two types of infections. "We used techniques from mathematical epidemiology because they gave us a means of framing relationships between states in a familiar, multidisciplinary way," Volkening said. While elections and disease dynamics are certainly different, the researchers treated Democratic and Republican voting inclinations as two possible kinds of "infections" that can spread between states. Undecided, independent, or minor-party voters all fit under the category of susceptible individuals. "Infection" was interpreted as adopting Democratic or Republican opinions, and "recovery" represented the turnover of committed voters to undecided ones.

In the model, committed voters can transmit their opinions to undecided voters, but the opposite is not true. The researchers took a broad view of transmission, interpreting opinion persuasion as occurring through both direct communication between voters and more indirect methods like campaigning, news coverage, and debates. Individuals can interact and lead to other people changing their opinions both within and between states.



Forecast of the 2020 U.S. presidential election that was made on October 27 using the researchers' model of infectious diseases. Senatorial and gubernatorial forecasts, as well as a link to the relevant code, are available at <https://modelingelectiondynamics.gitlab.io/2020-forecasts>. Credit: Figure courtesy of Samuel Chian, William L. He, Christopher M. Lee, Daniel F. Linder, Mason A. Porter, Grzegorz A. Rempala, and Alexandria Volkening.

To determine the values of their models' mathematical parameters, the authors used polling data on senatorial, gubernatorial, and presidential races from [HuffPost Pollster](#) for 2012 and 2016 and [RealClearPolitics](#) for 2018. They fit the model to the data for each individual race and simulated the evolution of opinions in the year leading up to each election by tracking the fractions of undecided, Democratic, and Republican voters in each state from January until Election Day. The researchers simulated their final forecasts as if they made them on the eve of Election Day, including all of the polling data but omitting the election results.

Despite its basis in an unconventional field for election forecasting—namely, epidemiology—the resulting model performed surprisingly well. It forecast the 2012 and 2016 U.S. races for governor, Senate, and presidential office with a similar success rate as popular analyst sites [FiveThirtyEight](#) and [Sabato's Crystal Ball](#). For example, the authors' success rate for predicting party outcomes at the state level in the 2012 and 2016 presidential elections was 94.1 percent, while FiveThirtyEight had a success rate of 95.1 percent and Sabato's Crystal Ball had a success rate of 93.1 percent. "We were all initially surprised that a disease-transmission model could produce meaningful forecasts of elections," Volkening said.

After establishing their model's capability to [forecast](#) outcomes on the eve of Election Day, the authors sought to determine how early the model could create accurate forecasts. Predictions that are made in the weeks and months before Election Day are particularly meaningful, but producing early forecasts is challenging because fewer polling data are available for model training. By employing polling data from the 2018 senatorial races, the team's model was able to produce stable forecasts from early August onward with the same success rate as FiveThirtyEight's final forecasts for those races.

Despite clear differences between contagion and voting dynamics, this study suggests a valuable approach for describing how political opinions change across states. Volkening is currently applying this model—in collaboration with Northwestern University undergraduate students Samuel Chian, William L. He, and Christopher M. Lee—to [forecast the 2020 U.S. presidential, senatorial, and gubernatorial elections](#). "This project has made me realize that it's challenging to judge forecasts, especially when some elections are decided by a vote margin of less than one percent," Volkening said. "The fact that our [model](#) does well is exciting, since there are many ways to make it more realistic in the future. We hope that our work encourages folks to think more critically about how they judge forecasts and get involved in [election](#) forecasting themselves."

**More information:** Alexandria Volkening et al,

Forecasting Elections Using Compartmental Models of Infection, *SIAM Review* (2020). [DOI: 10.1137/19M1306658](#)

Provided by Society for Industrial and Applied Mathematics

APA citation: Forecasting elections with a model of infectious diseases (2020, October 28) retrieved 27 February 2021 from <https://phys.org/news/2020-10-elections-infectious-diseases.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.*