SIAM REVIEW
Vol. 62, No. 4, pp. 837–865

Forecasting Elections Using Compartmental Models of Infection*

Alexandria Volkening†
Daniel F. Linder‡
Mason A. Porter§
Grzegorz A. Rempala¶

Abstract. Forecasting elections—a challenging, high-stakes problem—is the subject of much uncertainty, subjectivity, and media scrutiny. To shed light on this process, we develop a method for forecasting elections from the perspective of dynamical systems. Our model borrows ideas from epidemiology, and we use polling data from United States elections to determine its parameters. Surprisingly, our model performs as well as popular forecasters for the 2012 and 2016 U.S. presidential, senatorial, and gubernatorial races. Although contagion and voting dynamics differ, our work suggests a valuable approach for elucidating how elections are related across states. It also illustrates the effect of accounting for uncertainty in different ways, provides an example of data-driven forecasting using dynamical systems, and suggests avenues for future research on political elections. We conclude with our forecasts for the senatorial and gubernatorial races on 6 November 2018 (which we posted on 5 November 2018).

Key words. elections, compartmental modeling, polling data, forecasting, complex systems

AMS subject classifications. 34F05, 37N99, 60G10, 91D10

DOI. 10.1137/19M1306658

Contents

*Received by the editors December 16, 2019; accepted for publication (in revised form) August 24, 2020; published electronically November 3, 2020.
https://doi.org/10.1137/19M1306658

Funding: The work of the first, second, and fourth authors was partially supported by the Mathematical Biosciences Institute and the National Science Foundation (NSF) under grant DMS-1440386. The work of the fourth author was also supported by the NSF under grant DMS-1853587. The work of the first author was also supported by the NSF under grant DMS-1764421 and by the Simons Foundation/SFARI under grant 597491-RWC.

†NSF–Simons Center for Quantitative Biology, and Department of Engineering Sciences and Applied Mathematics, Northwestern University, Evanston, IL 60208 USA (alexandria.volkening@northwestern.edu, https://www.alexandriavolkening.com).

‡Medical College of Georgia, Division of Biostatistics and Data Science, Augusta University, Augusta, GA 30912 USA (dlinder@augusta.edu, https://www.augusta.edu/mcg/dphs/bds/people/daniel_linder.php).

§Department of Mathematics, University of California, Los Angeles, Los Angeles, CA 90095 USA (mason@math.ucla.edu, https://www.math.ucla.edu/~mason/).

¶Division of Biostatistics, College of Public Health, The Ohio State University, Columbus, OH 43210 USA (rempala.3@osu.edu, https://neyman.mbi.ohio-state.edu).

1. Introduction. Despite what was largely viewed as an unexpected outcome in the 2016 United States presidential election, recent work [44] suggests that national polling data are not becoming less accurate. Election forecasting is a complicated, multistep process that often comes across as a black box. It involves polling members of the public, identifying likely voters, adjusting poll results to incorporate demographics, and accounting for other data (such as historical trends). The result is a high-stakes, high-interest problem that is rife with uncertainty, incomplete information, and subjective choices [50]. In this paper, we develop a new forecasting method that is based on dynamical systems and compartmental modeling, and we use it to help examine U.S. election forecasting.

People typically use two primary types of data to forecast elections: polls and "fundamental data." Fundamental data consist of different factors on which voters may base their decisions [37]; such data include economic data, party membership, and various qualitative measurements (e.g., how well candidates speak) [41, 73]. Mainstream forecasting sources, such as newsletters and major media websites, offer varying levels of detail about their techniques and often rely on a combination of polls and fundamental data. Some analysts forecast vote margins at the state or national level (e.g., [16, 31, 43, 77]), while others (e.g., [6, 11, 38, 67]) call outcomes by party without giving margins of victory. We will refer to the former approaches as "quantitative" and the latter as "qualitative," though both settings typically employ quantitative data. Among quantitative forecasters, it is important to distinguish between those that aggregate publicly available polls from

a range of sources and those that gather their own in-house polls (e.g., *The Los Angeles Times* [31]). For example, FiveThirtyEight [77] is a poll aggregator that is known for its pollster ratings; they weight polls more heavily from sources that they judge to be more accurate [74]. After adjusting polls to account for factors such as recency, poll sample, convention bounce,[1] and polling source, FiveThirtyEight uses state demographics to correlate random outcomes, such that similar states are more likely to behave similarly [74]. For instance, in one of FiveThirtyEight's simulations, they may adjust the projected vote among Mormons by a few percentage points in favor of the Democratic candidate; in another, they may adjust the vote among Hispanic individuals toward the Republican candidate. By replicating this process many times for a wide range of characteristics, FiveThirtyEight produces outcomes that tend to be correlated in states with similar demographics [74].

In the academic literature, many statistical models (see, e.g., [41, 49, 50]) combine a variety of parameters—including state-level economic indicators, approval ratings, and incumbency—to forecast elections. See [54] for a review. Although some of these methods [24, 49] blend polls and fundamental data, Abramowitz's Time for Change model [14] and the work of Hummel and Rothschild [41] rely on fundamental data without using any polls. Models that are based on fundamental data alone can provide early forecasts, as they do not need to wait for polling data to become available. However, these forecasts are not dynamic and do not measure current opinion. To provide both election-day forecasts and estimates of current opinion, Linzer [55] augmented fundamental data with recent polls using a Bayesian approach. Although the media often stresses daily variance in polls as election campaigns unfold, the political-science community has cautioned that such fluctuations are typically insignificant and may represent differences in technique between polling sources, rather than true shifts in opinion [37, 42, 87]. Therefore, to account for nonrepresentative poll samples or "house effects" (i.e., bias that is introduced by the specific methods that each polling organization, which is often called a "house," uses to collect and weight raw poll responses [42]), some statistical models [42, 85] adjust and weight polling data in different ways (in a similar vein as FiveThirtyEight [74]). One can also consider simpler approaches, such as poll aggregation, for reducing error and improving the accuracy of forecasts [84].

Although there is extensive work on mathematical modeling of political behavior (see, e.g., [19, 20, 33, 34]) and opinion models more generally [25, 62], most of these studies focus on phenomena such as opinion dynamics or on questions that are related only tangentially to elections, rather than on engaging with data-driven forecasting. For example, Braha and de Aguiar [20] and Fernández-Gracia et al. [33] combined voter models with data on election results to comment on vote-share distributions and correlations across U.S. counties. In a series of papers (e.g., [35, 36]), Galam used a "sociophysics" approach (without reliance on polls or fundamental data) to suggest race outcomes and shed light on the dynamics that may underlie various election results. Very recently, Topîrceanu [80] developed a temporal attenuation model for U.S. elections with a basis in national polling data. Reference [80] includes measurements of each candidate's momentum in time and is able to produce forecasts at the national level.

[1] Candidates often receive a brief increase (i.e., a "bounce") in support after their party's convention. Because of this, some analysts adjust polling data shortly after the Republican and Democratic conventions take place [74].

Accounting for interactions between states is crucial for producing reliable forecasts, and FiveThirtyEight's Nate Silver [74] has stressed the importance of correlating polling errors by state demographics. Such correlations play an important role in the forecasts of FiveThirtyEight. Their approach [74], which one can view as indirectly incorporating relationships between states through noise, relies on geographic closeness or demographic similarity between states; these are inherently undirected quantities. For example, if Ohio and Pennsylvania are viewed as similar by FiveThirtyEight, so are Pennsylvania and Ohio. However, it is possible that states influence each other in directional ways. For example, voters in Ohio may more strongly influence the population in Pennsylvania than vice versa. The strength with which states influence each other can depend on where candidates are campaigning, the people with whom voters in different states interact, which distant states are featured prominently in the news, which states most resonate with local voters, and other factors. Linzer [55] estimated national-level influences on state voters on a daily basis using a statistical-modeling approach, but we are not aware of prior work that has estimated directed, asymmetric state–state relationships. We are also not aware of poll-based forecasting approaches that take a mathematical-modeling perspective.

To make election forecasting more transparent, broaden the community that engages with polling data, and raise research questions from a dynamical-systems perspective, we propose a data-driven mathematical model of the temporal evolution of political opinions during U.S. elections. We use a poll-based, poll-aggregating approach to specify model parameters, allowing us to provide quantitative forecasts of the vote margin by state. We consider simplicity in the election-specific components of our model a strength and thus do not weight or adjust the polling data in any way. Following Wang [84], we strive to be fully transparent; we provide all of our code, data, and detailed reproducibility instructions in a GitLab repository [82]. We have a special interest in exploring how states influence each other, and (because it provides a well-established way to frame such asymmetric relationships) we borrow techniques from the field of disease modeling. Using a compartmental model of disease dynamics, we treat Democratic and Republican voting intentions[2] as contagions that spread between states. Our model performs well at forecasting the 2012 and 2016 races, and we used it to forecast the 6 November 2018 U.S. gubernatorial and senatorial elections before they took place. We posted our forecasts [81] on arXiv on 5 November 2018. For the 2018 senatorial races, we also explore how early we can make accurate forecasts, and we find that our model is able to produce stable forecasts from early August onward.

Admittedly, our forecasting method involves many simplifications. Our goal is to apply a data-driven, dynamical-systems approach to elections that we hope leads to more such studies in the future. Most importantly, our model demonstrates how one can employ mathematical tools (e.g., dynamical systems, uncertainty quantification, and network analysis) to help demystify forecasting, explore how subjectivity and uncertainty impact forecasting, and suggest future research directions in the study of political elections. See [26] for our model's forecasts for the 2020 U.S. elections, which are forthcoming at the time of this writing.

[2]Throughout our paper, we use the term "Democratic voter" (respectively, "Republican voter") to indicate an individual who is inclined to vote for a Democratic candidate (respectively, Republican candidate). We thereby focus on an individual's current voting inclination, rather than on any possible affiliation that they may have with a political party.

2. Background: Compartmental Modeling of Infections. Because we repurpose our forecasting approach from compartmental modeling of disease dynamics [21, 30, 39], we begin with an introduction to these techniques. Compartmental models [21, 30, 39, 45, 46, 47] are a standard mathematical approach for studying biological contagions, including the current COVID-19 pandemic [83]. Developed initially for analyzing the spread of diseases such as influenza [27], compartmental models (which are often combined with network structure to incorporate social connectivity [48, 60]) have also been used to study phenomena such as social contagions [17, 18]. Compartmental models are built on the idea that one can categorize individuals into a few distinct types (i.e., "compartments"). One then describes contagion dynamics using flux terms between the various compartments. For example, a susceptible–infected–susceptible (SIS) model divides a population into two compartments. At a given time, an individual can be either susceptible (S) or infected (I). As we illustrate in Figure 1(a), the fraction of susceptible and infected individuals in a community depends on two factors:

- *transmission*: when an infected individual interacts with a susceptible individual, the susceptible person has some chance of becoming infected;
- *recovery*: an infected person has some chance of recovering and becoming susceptible again.

Suppose that $S(t)$ and $I(t)$ are the fractions of susceptible and infected individuals, respectively, in a well-mixed population at time t. One then describes infection spread using the following set of ordinary differential equations (ODEs):

$$\frac{dS}{dt} = \underbrace{\gamma I}_{\text{recovery with rate } \gamma} - \underbrace{\beta SI}_{\text{infection with rate } \beta}, \tag{2.1}$$

$$\frac{dI}{dt} = -\gamma I + \beta SI\,, \tag{2.2}$$

where $S(t)+I(t)=1$ and β and γ correspond to the rates of disease transmission and recovery, respectively. We provide additional background on compartmental models in Appendix A, but it may be helpful to think of the equations $\frac{dS}{dt} = f(S,I)$ and $\frac{dI}{dt} = g(S,I)$, where f and g are unknown functions. If we Taylor expand the functions f and g, then (2.1)–(2.2) are the lowest-order terms in this expansion that make sense for our application. In particular, the term for transmission must depend on both S and I, because both susceptible and infected individuals need to be present for a new infection to occur. SIS models have been extended to account for more realistic details, such as multiple contagions, communities, and contact structure between individuals or subpopulations [15, 48, 57, 60].

3. Model and Methods. We now construct a model of party choices in elections in the form of a compartmental model, describe how we specify its parameters, and overview how we incorporate uncertainty into our forecasts. See section 3.4 and Figure 2 for a summary of the steps that we follow to generate a forecast. We also summarize some of the main simplifications that our method involves in section 3.4. Our code and the data sets that we use in our model are available on GitLab [82].

3.1. Our Model of Election Dynamics. Our model of election dynamics is a two-pronged SIS compartmental model (see Figure 1(a)). First, we reinterpret "susceptible" individuals as undecided (or independent or minor-party) voters. Because most U.S. elections are dominated by two parties, we consider two contagions: Democratic and Republican voting intentions. We track these quantities within each state

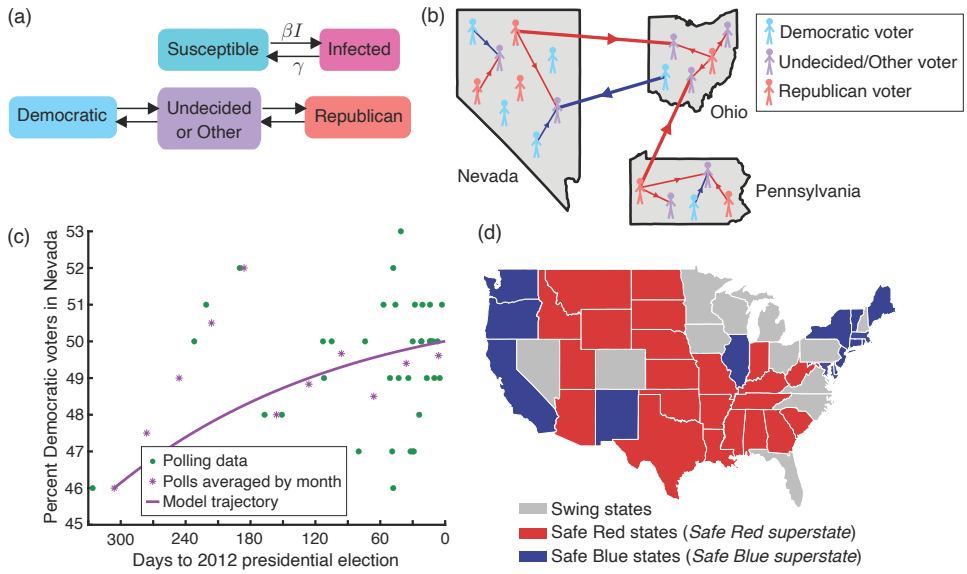

Fig. 1 *Overview of our modeling approach.* (a) *A susceptible–infected–susceptible (SIS) compartmental model tracks the fraction of susceptible ($S(t)$) and infected ($I(t)$) individuals in a community in time; these quantities evolve according to infection and recovery. We repurpose this approach by including two types of infections (Democratic and Republican voting inclinations), interpreting infection as opinion adoption, and replacing recovery with turnover of committed voters to undecided ones.* (b) *We assume that individuals interact within and between states. Red (respectively, blue) lines indicate interactions between undecided voters and Republican voters (respectively, Democratic voters), and thick and thin lines correspond to interactions between voters in different states and in the same state, respectively. In this cartoon example, Pennsylvania influences Ohio, but voters in Nevada and Pennsylvania do not interact. We show individual voters and their interactions in this figure, but our model works at a population level and tracks voter percentages. Additionally, this image is a schematic and is not to scale.* (c) *Example model dynamics of a Democratic voting inclination in Ohio leading up to the* 2012 *presidential election. We take the mean of the polls by month to obtain the data points, which we indicate using purple asterisks. We specify parameters by minimizing the difference between our model* (3.1)–(3.3) *and these monthly data points. We simulate the temporal evolution of opinions in the year leading up to each election, but we focus on the result when there are* 0 *days until the election.* (d) *For elections with many state races, we combine all reliable Republican and Democratic regions into two "superstates" (in Red and Blue, as we use traditional party colors). We show the superstates for presidential elections. See Table* SM1 *for the superstates that we use in other elections.*

and make the assumption that populations are well-mixed within each state. Let

$$S^i = \text{fraction of undecided voters in state } i\,,$$
$$I_{\mathrm{D}}^i = \text{fraction of Democratic voters in state } i\,,$$
$$I_{\mathrm{R}}^i = \text{fraction of Republican voters in state } i\,,$$

where $I_{\mathrm{D}}^i + I_{\mathrm{R}}^i + S^i = 1$. We account for four behaviors:

- *Democratic transmission*: An undecided voter can decide to vote for a Democrat due to interactions with Democratic voters. As we discuss below, we interpret "interactions" and "transmission" broadly.
- *Republican transmission*: An undecided voter can decide to vote for a Republican due to interactions with Republican voters.

- *Democratic turnover*: An infected person has some chance of changing their mind to undecided (this amounts to "recovering").
- *Republican turnover*: An infected person has some chance of becoming undecided (i.e., recovering).

We assume that committed voters can "transmit" their opinions to undecided voters but that undecided voters do not sway Republican or Democratic voters. Marvel et al. used a similar assumption in a model of more general ideological conflict [56]. Although the language of contagions does not necessarily apply to social dynamics [52], we find it useful for our work. Our use of terminology from contagion dynamics highlights that our model is not a specialized election model, as part of our goal is to show how a general framework can give meaningful forecasts in high-dimensional systems. We thus expect similar ideas to provide insight into the forecasting of many complex systems.

By extending the traditional SIS model (2.1)–(2.2) to account for two contagions and M states or "superstates" (see Figure 1(d)), we obtain the following ODEs:

$$\frac{dI_{\mathrm{D}}^i}{dt}(t) = \underbrace{-\gamma_{\mathrm{D}}^i I_{\mathrm{D}}^i}_{\text{Dem. loss}} + \underbrace{\sum_{j=1}^{M} \beta_{\mathrm{D}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{D}}^j}_{\text{Dem. infection}}, \tag{3.1}$$

$$\frac{dI_{\mathrm{R}}^i}{dt}(t) = \underbrace{-\gamma_{\mathrm{R}}^i I_{\mathrm{R}}^i}_{\text{Rep. loss}} + \underbrace{\sum_{j=1}^{M} \beta_{\mathrm{R}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{R}}^j}_{\text{Rep. infection}}, \tag{3.2}$$

$$\frac{dS^i}{dt}(t) = \gamma_{\mathrm{D}}^i I_{\mathrm{D}}^i + \gamma_{\mathrm{R}}^i I_{\mathrm{R}}^i - \sum_{j=1}^{M} \beta_{\mathrm{D}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{D}}^j - \sum_{j=1}^{M} \beta_{\mathrm{R}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{R}}^j, \tag{3.3}$$

where N is the total number of voting-age individuals in the U.S.; N^j is the number of voting-age individuals in state j [2, 4, 7]; and γ_{D}^i and γ_{R}^i describe the rates at which committed Democratic and Republican voters, respectively, convert to being undecided. Similarly, β_{D}^{ij} and β_{R}^{ij} correspond, respectively, to the transmission (i.e., influence) rates from Democratic and Republican voters in state j to undecided individuals in state i. In addition to the presence of state-level variables and two contagions in (3.1)–(3.3), the other difference between our election-dynamics model and the traditional SIS system in (2.1)–(2.2) is the new terms that include the factor N^j/N. These terms appear in (3.1)–(3.3) because of how we choose to approximate the mean number of interactions between undecided voters in state i and committed voters in state j. See Appendix A for details.

The β parameters, which pertain to transmission dynamics, allow us to model directed relationships as a form of network between states. We take a broad interpretation of "transmission." Although opinion persuasion (i.e., transmission) can occur through communication between undecided and committed voters [61], we expect that it can also occur through campaigning, news coverage, and televised debates. We hypothesize that these venues are an indirect means for voters in one state to influence voters in another state. For example, if news coverage of Republican campaigning in Pennsylvania resonates with undecided voters in Ohio, there may be an associated indirect route of opinion transmission from Pennsylvania to Ohio. Therefore, we consider large β_{R}^{ij} to signify that Republican voters in state j strongly influence undecided voters in state i, and such "strong influence" may be due either to conversations

(or other direct interactions) between voters or due to indirect effects like state–state affinities that are influenced or activated by media.

3.2. Parameter Fitting. Broadly, we obtain our parameters in (3.1)–(3.3) by fitting to about a year (or less, in the case of our early forecasts) of state polls [8, 10]. We gather these polls from HuffPost Pollster [8] for our 2012 and 2016 forecasts and from RealClearPolitics [10] for our 2018 forecasts. Our parameters are different for each election and year, as we use the data that are specific to each race for fitting. To fit model parameters for a given election and year (e.g., the 2012 senatorial races), we begin by formatting its associated polling data. First, we assign each poll a time point by taking the mean of its start and end dates. Because some states are polled more frequently than others, we then adjust our data so that each state or superstate has an equal number of data points that we can use to fit our model. We do this by binning the polls for each state in 30-day increments that extend backward from the appropriate election day (6 November 2012, 8 November 2016, or 6 November 2018). Counting backward from election day, the bin that is closest to the election is $(0, 30]$, the next bin is $(30, 60]$, and so on. Our earliest bin includes polls from between 330 and 300 days before an election,[3] so the maximum number of bins that we consider is 11. Most of our forecasts are in November; for these, we use all $T = 11$ bins. For our earlier forecasts in Figure 6, we use a smaller subset of the polling data. As an example, to forecast the 2018 senatorial elections on 7 August, we bin the polling data from between 330 and 90 days of the election in 30-day increments to obtain $T = 8$ bins.

Within each 30-day bin, we take the mean of the polling data to help remove small-scale fluctuations in the polls. If a given state has no polls within a bin, we approximate the associated data point using linear interpolation. In many cases, there are no polls for a state early in the year, so we set all missing early data points for that state to its earliest data point. To arrive at T data points each for the Safe Red and Safe Blue superstates, for each of these T points, we take a weighted average of the individual data points of each of the states within these conglomerates. To determine the weightings, we use the states' voting-age population sizes. The result of this process is T data points per state or superstate that we forecast; see Figure 1(c) for an example.

Let $\{R^j(t_i), D^j(t_i), U^j(t_i)\}_{i=1,\dots,T}$ denote the T data points for state (or superstate) j that we obtain through the above process. The variables $R^j(t_i)$, $D^j(t_i)$, and $U^j(t_i)$ are the fractions of people who are inclined to vote for a Republican, are inclined to vote for a Democrat, and are undecided (or have other plans) in state j at time point t_i. To describe our fitting procedure, we define a "concentration" vector

$$\mathbf{C}(t_i) = [R^1(t_i), \dots, R^M(t_i), D^1(t_i), \dots, D^M(t_i), U^1(t_i), \dots, U^M(i)]\,,$$

where M is the number of states and superstates that we forecast for a given election (see Table SM1). For a candidate parameter set, $(\beta, \gamma) = \{\beta_{\mathrm{R}}^{jk}, \beta_{\mathrm{D}}^{jk}, \gamma_{\mathrm{R}}^{j}, \gamma_{\mathrm{D}}^{j}\}_{j,k=1,\dots,M}$, we define $\mathbf{c}^{\beta,\gamma}$ to be the solution of (3.1)–(3.3) using these parameters:

$$\mathbf{c}^{\beta,\gamma}(t) = [I_{\mathrm{R}}^1(t), \dots, I_{\mathrm{R}}^M(t), I_{\mathrm{D}}^1(t), \dots, I_{\mathrm{D}}^M(t), S^1(t). \dots, S^M(t)]\,.$$

[3] We made one adjustment to this rule for the 2012 presidential race. For this race, when a state either has polls between 330 and 300 days before the election or has no polls within 330 days of the election, our earliest bin for that state includes polls from between 400 and 300 days until the election.

We obtain our parameters $(\hat{\beta}, \hat{\gamma})$ for a given election by minimizing the least-squares deviation between the averaged polling data and the solutions of (3.1)–(3.3) at the T time points. That is,

$$(\hat{\beta}, \hat{\gamma}) = \underset{(\beta,\gamma)}{\operatorname{argmin}} \sum_{i=1}^{T} \|\mathbf{C}(t_i) - \mathbf{c}^{\beta,\gamma}(t_i)\|_2^2 .$$

These parameter estimates are consistent and converge weakly to a Gaussian distribution if the data are from a density-dependent Markov jump process [65].

We base our parameters for a given election on $2 \times T \times M$ data points that represent the fractions of Republican and Democratic voters at T time points in each of M states or superstates. (We also know the fractions of other voters, but these data points are correlated with the above data, because $R^j(t_i) + D^j(t_i) + U^j(t_i) = 1$.) As a comparison, there are $2 \times M$ turnover parameters and $2 \times M^2$ transmission parameters in (3.1)–(3.3). For example, we use 308 data points to specify 420 parameters in our November forecasts of presidential elections, which include $M = 14$ states and superstates.

Our GitLab repository [82] includes all of the model parameters that we generate using the procedure that we just described. Across all election years (2012, 2016, and 2018) and election types (gubernatorial, senatorial, and presidential) that we consider, the mean minimum transmission parameter is 0.000000 and the mean maximum transmission parameter is 0.555818. Our recovery parameters $\{\gamma_{\mathrm{R}}^i, \gamma_{\mathrm{D}}^i\}$ vary from a mean minimum of 0.000000 to a mean maximum of 0.047622. As an example, we illustrate the parameters for the 2018 senatorial races in Figures SM4–SM6 in our supplementary material.

3.3. Uncertainty. For some of our forecasts (specifically, for our 2018 forecasts and for our case study of the 2016 presidential race in section 4.2), we generalize our model (3.1)–(3.3) to a system of stochastic differential equations (SDEs):

$$dI_{\mathrm{D}}^i(t) = \underbrace{\left(-\gamma_{\mathrm{D}}^i I_{\mathrm{D}}^i + \sum_{j=1}^{M} \beta_{\mathrm{D}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{D}}^j\right) dt}_{\text{deterministic dynamics in (3.1)}} + \underbrace{\sigma dW_{\mathrm{D}}^i(t)}_{\text{uncertainty}}, \tag{3.4}$$

$$dI_{\mathrm{R}}^i(t) = \left(-\gamma_{\mathrm{R}}^i I_{\mathrm{R}}^i + \sum_{j=1}^{M} \beta_{\mathrm{R}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{R}}^j\right) dt + \sigma dW_{\mathrm{R}}^i(t), \tag{3.5}$$

$$dS^i(t) = \left(\gamma_{\mathrm{D}}^i I_{\mathrm{D}}^i + \gamma_{\mathrm{R}}^i I_{\mathrm{R}}^i - \sum_{j=1}^{M} \beta_{\mathrm{D}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{D}}^j - \sum_{j=1}^{M} \beta_{\mathrm{R}}^{ij} \frac{N^j}{N} S^i I_{\mathrm{R}}^j\right) dt + \sigma dW_{\mathrm{S}}^i(t), \tag{3.6}$$

where we now consider I_{D}^i, I_{R}^i, and S^i to be stochastic processes. We let W_{D}^i, W_{R}^i, and W_{S}^i be Wiener processes; these are the components of $\mathbf{W}_{\mathrm{D}}$, $\mathbf{W}_{\mathrm{R}}$, and $\mathbf{W}_{\mathrm{S}}$, respectively. The parameters in this system have the same values as those that we fit for the corresponding deterministic model (3.1)–(3.3). By simulating many (e.g., we use 10,000) elections using (3.4)–(3.6), we obtain a distribution of possible outcomes; this allows us to quantify uncertainty in a given race. We explore the effects of uncorrelated and correlated noise on our forecasts in section 4.2.

In graphics that describe its 2016 presidential forecasts, FiveThirtyEight [79] included both its expected vote margin for each state and a confidence interval that

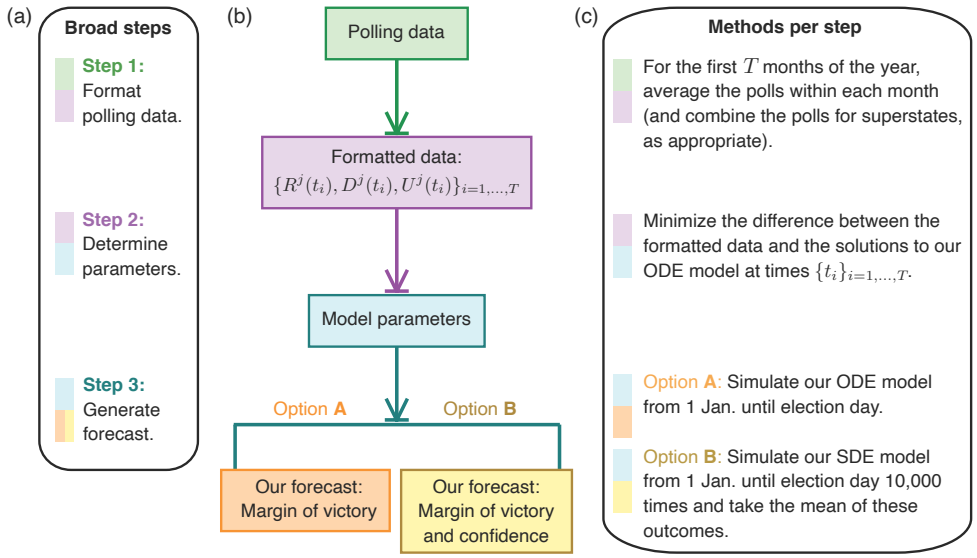

Fig. 2 *Summary of our approach.* (a) *To forecast a given election, we follow three steps. We start by formatting polling data from* [8, 10]. (b) *Once we fit our model parameters, we have two options for forecasting elections. When using Option A, we generate forecasts of the vote margin by simulating our ODE model* (3.1)–(3.3) *from* 1 *January through election day. When using Option B, we instead simulate many realizations of our SDE model* (3.4)–(3.6) *to produce a distribution of vote margins, and we take the mean of this distribution to be our forecast vote margin.* (c) *Our final forecasts in November have* $T = 11$ *months, but earlier forecasts for the* 2018 *races (see Figure* 6*) have* $T < 11$ *months.*

indicates the range that encompassed the middle 80% of its model outcomes for that state. To determine our noise strength σ in (3.4)–(3.6), we measure the length of these confidence intervals in FiveThirtyEight's final forecasts. Based on our estimates, the intervals range in length from about 13 to about 19 percentage points for swing states. We tested a few values of σ and chose $\sigma = 0.0015$ to roughly match the length of our 80% confidence intervals[4] for the vote margin to these measurements for our 2016 presidential forecasts. For example, across the 10,000 simulations of (3.4)–(3.6) that we show in Figure 4(b), the mean length of our 80% confidence intervals for swing states is about 15 percentage points.

3.4. Summary of Our Approach and Important Simplifications. Our forecasting process consists of three steps, which we illustrate in Figure 2. Our first step is to take the mean of the polling data [8, 10] within each month (specifically, within each 30-day increment extending backward from election day) to generate T data points in time for each state or superstate. Because we focus on November forecasts, there are $T = 11$ months in most of our simulations. This corresponds to one data point in each month from January to November of an election year. For our earlier forecasts for the 2018 races in Figure 6, there are $T < 11$ months. As an example, our forecasts in August have $T = 8$. Our second step is to fit our parameters to our trajectory of polling data in the first T months of a year using our deterministic model (3.1)–(3.3). Third, after we fit our parameters to polls for a given election, we

[4]The middle 80% confidence interval is the range in which a set of outcomes lies after we have removed the bottom 10% and top 10% of the outcomes from the set. We determine these cutoffs using the built-in PRCTILE function in MATLAB (version 9.3).

use one of our models to simulate the daily evolution of political opinions from the preceding January through election day. In such simulations, we use the polling data point at $T = 1$ to specify our initial conditions (see Appendix B.3). Because we have both a deterministic model (3.1)–(3.3) and a stochastic model (3.4)–(3.6), we have two options for simulating elections. For our initial study in section 4.1, we use our ODE model (3.1)–(3.3) to simulate elections. (We label this approach as "Option A" in Figure 2.) Following this initial study, we generate forecasts by simulating many realizations of our SDE model (3.4)–(3.6). This alternative approach (which we label as "Option B" in Figure 2) provides a distribution of the vote margin in each state, allowing us to forecast the margin of victory (namely, the mean of this distribution) in each state and our confidence in this margin. Except for Figures SM1(b,c), we use a time step of $\Delta t = 3$ days for parameter fitting, specify a time step of $\Delta t = 0.1$ days for simulating our models once we have determined their parameters, and simulate 10,000 realizations of (3.4)–(3.6) to generate the forecasts that are based on our SDE model. We refer to these values as our "typical" simulation parameters.[5] In each figure caption, we indicate whether we use our ODE model or our SDE model for the simulations in that figure.

We do not claim that our approach is the most accurate method of forecasting elections. Instead, we propose it as a data-driven model that admittedly involves many simplifications, some of which are instructive to mention before we discuss our simulation results. Important simplifications include the following:

- Although it is generally not realistic, we assume that voters mix uniformly (e.g., everyone has the same influence on everyone else), aside from the state structure, which is analogous to patches in epidemiology. Accounting for additional network structure may improve forecasts [22, 58, 86].
- We combine all sources of opinion adoption into time-independent transmission parameters β_{R}^{ij} and β_{D}^{ij}. Including time-dependent transmission parameters may lead to richer model behavior, such as oscillations over time in a state's vote margin.
- If undecided voters remain at the end of our simulation, we assume that they vote for minor-party or other candidates.
- We assume that all polls are equally accurate. Unlike FiveThirtyEight [74], we do not weight polls more strongly based on recency or make any distinction between partisan and nonpartisan polls (or polls of likely voters, registered voters, or all adults). Notably, Wang [84] has illustrated that, when aggregated, polling data can be accurate without needing to be weighted or adjusted to account for polling source.
- Because of our approach to fitting parameters, we use the earliest available formatted polling data to initialize our models. Using data-assimilation techniques [51] to determine initial conditions may improve the accuracy of our forecasts.

Despite these simplifications, our forecasting method performs as well as popular analysts. We discuss our results in section 4.

4. Results. We now use our ODE model (3.1)–(3.3) to simulate the races for governor, senator, and president in 2012 and 2016. Because realistic forecasts should

[5] Our forecasts in Figures SM1(b,c) are the only situations in which we do not use our typical simulation parameters. We determined the model parameters for our simulations in Figures SM1(b,c) using a time step of $\Delta t = 15$ days, and we then produced our forecasts based on 4,000 simulations of (3.4)–(3.6) with a time step of $\Delta t = 0.1$ days.

incorporate uncertainty, we follow this exploration of past races with a short study of the impact of noise on our 2016 presidential forecast. To do this, we compare simulations of our SDE model (3.4)–(3.6) with uncorrelated and correlated noise. We conclude by using our SDE model to forecast the gubernatorial and senatorial midterms on 6 November 2018.

4.1. 2012 and 2016 Election Forecasts. By fitting our parameters to polling data for senatorial, gubernatorial, and presidential races in 2012 and 2016 without incorporating the final election results, we can simulate forecasts as if we made them on the eve of the respective election days. In Figure 3, we summarize our forecasts for these races. To measure the accuracy of our forecasts, we compute a success rate for predicting ("calling") party outcomes at the state level:

(4.1)
$$\text{success rate} = 100 \times \frac{\text{number of state or district races that are called correctly}}{\text{total number of state or district races that are forecast}},$$

where we consider only state and district races for which our model provides forecasts.[6] As we show in Table 1, our model has a success rate that is similar to those of popular forecasters FiveThirtyEight [79] and Sabato's Crystal Ball [67]. For example, our success rate across all 102 of the (state or Washington, D.C.) forecasts that we made for the presidential elections in 2012 and 2016 is 94.1%, whereas FiveThirtyEight and Sabato's Crystal Ball achieved success rates of 95.1% and 93.1%, respectively.

Figure 3 and Table 1 highlight two forecasting goals: (1) estimating the vote share by state (e.g., the percentages of the state vote that are received by the Democratic and Republican candidates) and (2) calling the winning party by state (i.e., which party's candidate wins the election in a given state). Many qualitative forecasters (e.g., [11, 67]) focus on the second goal, whereas our model and FiveThirtyEight [77] pursue both goals.

4.2. Accounting for and Interpreting Uncertainty. Sources of uncertainty and error in election forecasting include sampling error and systematic bias from the specific methods of different polling sources [42, 55, 74, 84]. Consequently, election forecasting involves not only calling a race for a specific party and estimating vote shares, but also specifying the likelihood of different outcomes. This raises a third goal of forecasters: (3) quantifying uncertainty, such as by estimating a given candidate's chance of winning an election or by providing a confidence interval for their vote margin. We suggest that this is one of the key places where mathematical techniques can contribute to election forecasting. As a case study, we investigate two methods of accounting for uncertainty. Specifically, we compare the forecasts that result from simulating 10,000 realizations of our SDE model (3.4)–(3.6) (see section 3.3) for the 2016 presidential race with uncorrelated noise with those that arise from following the same process with correlated noise.

Although U.S. elections are decided at the level of states, polling errors are correlated in regions with similar populations [74]. Therefore, if a pollster is wrong in Minnesota, they may also be off in states (such as Wisconsin) with shared features [74]. This type of error makes it possible for polls of a bloc of states to all

[6] As we discuss in Appendix B.1.1, there are a few special cases in which we do not forecast a given state (see Table SM1). For example, we do not forecast single-party state races or 2012 gubernatorial races in states for which we have no polling data. See section SM2 for a discussion of alternative choices that we could have made when calculating accuracy in the face of these special cases.

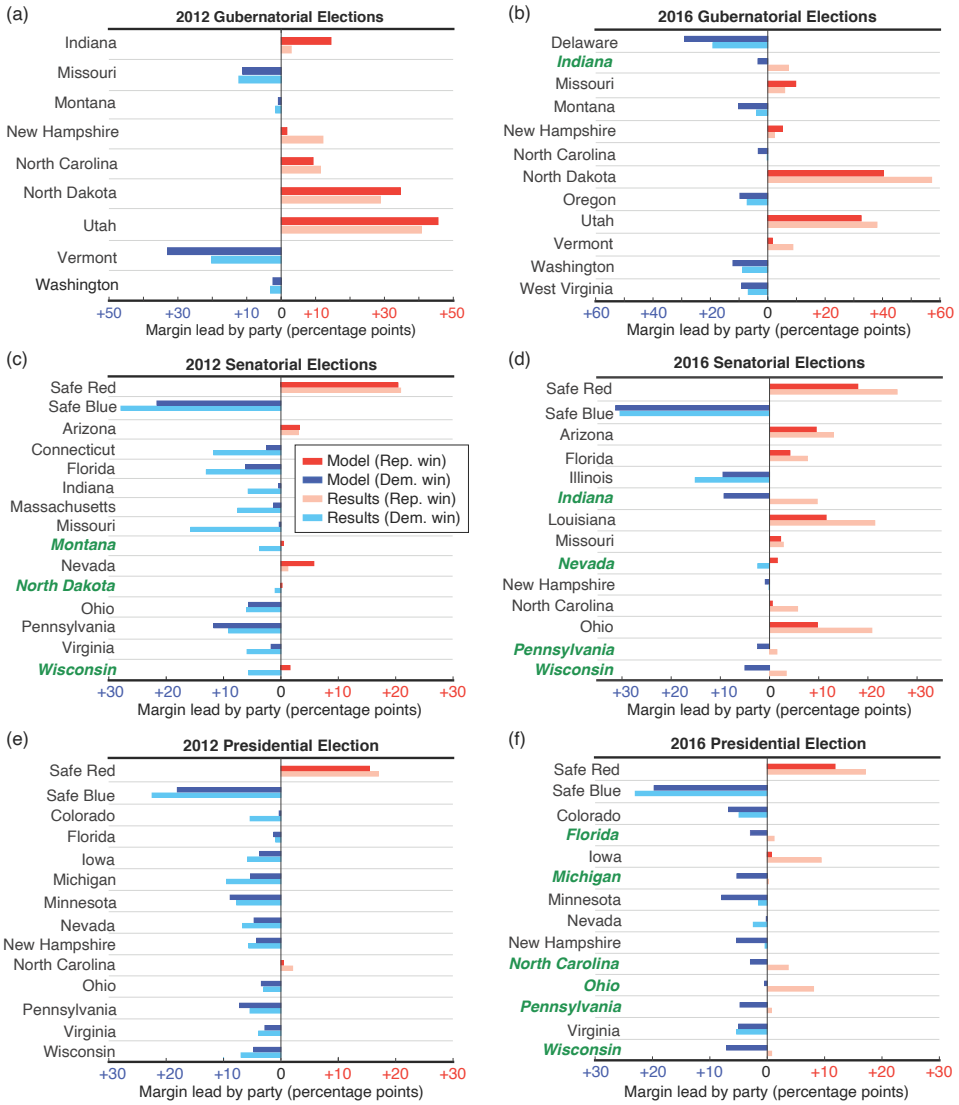

Fig. 3 *Simulations of* (3.1)–(3.3). *We calculate our* 2012 *and* 2016 *forecasts using polling data up until election day. Our forecasts do not include any election results, and they should be interpreted as forecasts from the night before an election. Comparison of our forecasts for* (a, b) *gubernatorial,* (c, d) *senatorial, and* (e, f) *presidential elections with results from* [12, 53]. *The horizontal axis shows the percentage-point leads by the Democratic (blue) or Republican (red) candidates. Shorter bars represent closer elections, and bars that extend to the right (respectively, left) correspond to Republican (respectively, Democratic) leads. The states that we forecast incorrectly are in a bold, italic, green font. "Safe Red" and "Safe Blue" refer to superstates that are composed, respectively, of reliably Republican and reliably Democratic states. We assemble the superstates based on forecaster opinions and historical data (see Appendix* B.2*).*

Table I *Comparison of the success rates for our model* (3.1)–(3.3) *and two popular sources. We measure a forecaster's "success rate" in* (4.1) *as the percent of (state or Washington, D.C.) races that they correctly forecast, in the sense that they identified the true winner by party, among the races that we forecast using our model (see Table* SM1*). Importantly, we leave races that we do not forecast (e.g., single-party races) out of these computations. See Appendix* B.1.1 *and section* SM2 *for more details. For FiveThirtyEight* [79]*, we use the* 2016 *polls-only forecast.*

Election	FiveThirtyEight [79]	Our model	Sabato [67]
2016 presidential	90.2%	88.2%	90.2%
2016 senatorial	90.9%	87.9%	93.9%
2016 gubernatorial	NA	91.7%	83.3%
2012 presidential	100%	100%	96.1%
2012 senatorial	NA	90.3%	93.5%
2012 gubernatorial	NA	100%	77.8%

be wrong together, leading to an unforeseen upset. To explore these dynamics, we compare the impact of uncorrelated noise with the effect of additive noise that is correlated on a few sample demographics. Specifically, we consider the fractions of Black, Hispanic, and adult college-educated individuals in a population. We correlate on these demographics because these data are readily available; future work should incorporate additional data.

To correlate noise in (3.4)–(3.6), we first quantify the similarity of two states, i and j, using the Jaccard index $J^{i,j} = \min\{X^i, X^j\}/\max\{X^i, X^j\}$, where X^i is the fraction of a given demographic in state i and X^j is the fraction of that demographic in state j. The Jaccard index indicates the covariance for our increments of $\mathbf{W}_{\mathrm{R}}$ and $\mathbf{W}_{\mathrm{S}}$ in (3.4)–(3.6). We define $\mathbf{J}_{\mathrm{B}}, \mathbf{J}_{\mathrm{E}}$, and $\mathbf{J}_{\mathrm{H}}$ to be the Jaccard indices that we find using the fractions of non-Hispanic Black individuals, adults without a college education, and Hispanic individuals, respectively.[7] We calculate $\mathbf{J}_{\mathrm{B}}$ and $\mathbf{J}_{\mathrm{H}}$ using 2016 U.S. Census Bureau data [3], and we base $\mathbf{J}_{\mathrm{E}}$ on data from 247WallSt.com [29]. For our forecasts with correlated noise, each time that we simulate an election, we select one Jaccard index uniformly at random among $\mathbf{J}_{\mathrm{B}}$, $\mathbf{J}_{\mathrm{E}}$, and $\mathbf{J}_{\mathrm{H}}$ to use as our covariance. For example, if we select $\mathbf{J}_{\mathrm{H}}$, then the increment $d\mathbf{W}_{\mathrm{R}}(t_n)$ has a multivariate normal distribution with mean 0 and covariance $\mathbf{J}_{\mathrm{H}}$ at each time step in that simulation. Consequently, for each such simulated election, at any given time, we are more likely to adjust the vote in a set of states with a similar feature (e.g., a high Hispanic population) in the same direction (e.g., in favor of the Democratic candidate) than we are to adjust the vote in these states in opposite directions.

Our case study of the 2016 presidential race illustrates how accounting for uncertainty in different ways influences forecasts, echoing points that have been raised by Nate Silver and his collaborators [74]. In Figure 4, we demonstrate that uncorrelated noise, which can model uncertainty in a single state or a single poll without assuming a larger systematic (e.g., country-wide) polling error, results in a low likelihood of a win by Donald Trump (the Republican candidate) in the 2016 presidential election. By contrast, correlating outcomes by demographics, which can model systematic polling

[7] We compute the fraction of a given demographic in the Safe Red (respectively, Safe Blue) superstate by calculating the mean of the fractions in all of the states inside the Safe Red (respectively, Safe Blue) superstate. We do not weight these averages by state population. Additionally, for presidential elections, we do not include Washington, D.C., in our estimate of education level in the Safe Blue superstate, because data on education levels in Washington, D.C., are not listed on 247WallSt.com [29].

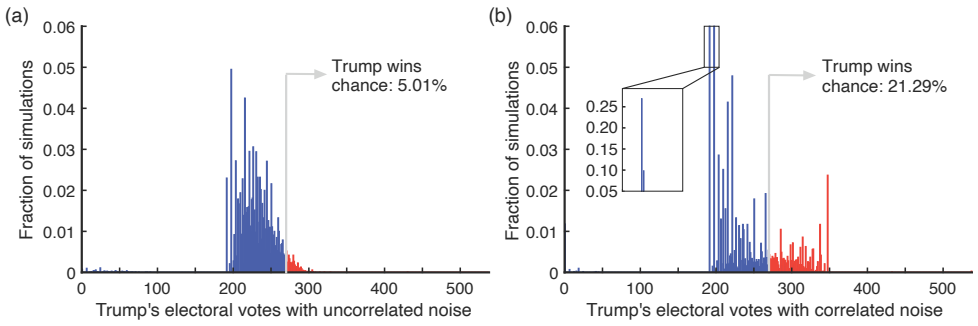

Fig. 4 *Impact of incorporating uncertainty in different ways, as demonstrated by simulations of our SDE model* (3.4)–(3.6) *for the* 2016 *U.S. presidential election.* (a) *Uncorrelated additive white noise gives Donald Trump an approximately* 5% *chance of winning the electoral college, whereas* (b) *correlating noise on state demographics increases his chance of winning to about* 21%. *The tall bar in the magnified image in panel* (b) *illustrates that many of our simulations forecast that only the Red superstate votes Republican, leading to* 191 *electoral votes for Trump. The smaller bar in the magnified image corresponds to model outcomes in which only the Red superstate and either Iowa or Nevada vote Republican. We generate distributions by simulating* 10,000 *elections using* (3.4)–(3.6) *with* $\sigma = 0.0015$.

errors (e.g., due to misidentifying likely voters) in similar states, increases Donald Trump's chances of winning the election by a factor of about four. This agrees with Nate Silver's comment [74] that failing to account for correlated errors tends to result in underestimations of a trailing candidate's chances. As we discuss in section SM1 of our supplementary material, because of an indexing error in one of our files, in an earlier version of our model, we correlated state outcomes on the demographics of the wrong states. After correcting this error, we obtained similar results, suggesting that it is the mere presence of correlated noise that improves Donald Trump's chances in these forecasts and that the noise does not need to be correlated on the specific state demographics that we used. FiveThirtyEight [74], for example, correlates state outcomes on party, region, religion, race, ethnicity, and education.

In our computations, we do not attempt to account directly for errors in polls. Instead, we take the simple approach of assuming that we can incorporate all sources of uncertainty as an additive noise term in our SDE model (3.4)–(3.6). There has been extensive work on quantifying uncertainty (see [51]); exploring alternative ways of accounting for uncertainty is an important future direction for research on forecasting complex systems.

4.3. 2018 Senatorial and Gubernatorial Forecasts. The 2018 midterm elections provided a fantastic opportunity for us to test our model. Our final forecasts, which we posted on 5 November 2018 (the night before the 2018 elections) [81], rely on polls that we gathered from RealClearPolitics [10] through 3 November[8] and are based on our SDE model (3.4)–(3.6). We account for uncertainty by correlating noise on education, ethnicity, and race (as in Figure 4(b)). Because of computational time

[8]There is a time delay between when polls are completed and when the data become available from RealClearPolitics [10]. The latest polls in our gubernatorial and senatorial data sets were completed on 1 November and 2 November, respectively. Polls do not always become available in the temporal order of polling day, so this does not imply that our data include all of the polls that occurred before these dates. For example, RealClearPolitics [10] occasionally updates its website with additional early polls, despite its prior posting of more recent polls.

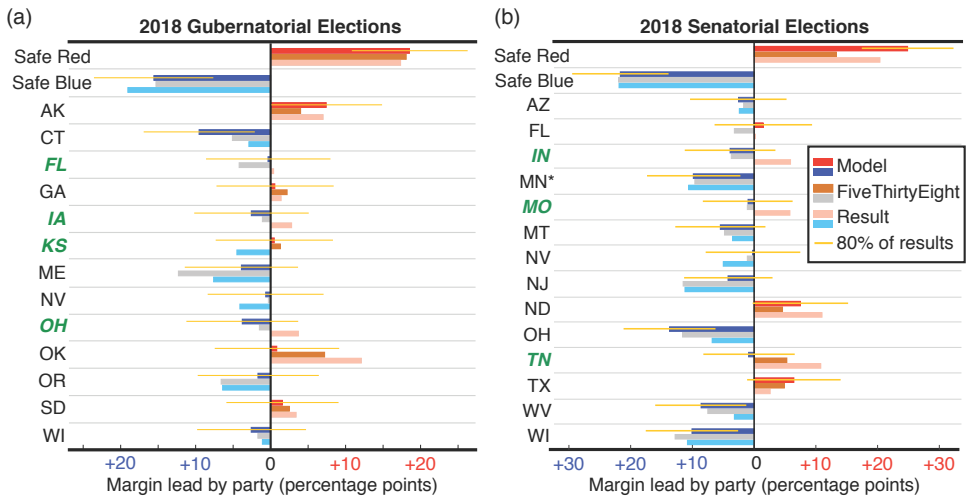

Fig. 5 *Forecasts of vote margins for the* 2018 (a) *gubernatorial and* (b) *senatorial races. We base our forecasts on* 10,000 *elections that we simulate using* (3.4)–(3.6) *with noise that we correlate on state demographics. We base our forecasts on poll data that we collected from RealClearPolitics* [10] *through* 3 *November* 2018. *(The election took place on* 6 *November.) We compare them to FiveThirtyEight's* 6 *November forecasts (according to the "classic" version of the FiveThirtyEight algorithm)* [77] *and the election results* [5]. *(We use an asterisk to mark the MN special election.) The bold, italic, green font indicates state races that we called incorrectly. The length of the bars that extend to the right (in red) and left (in blue) indicate the mean percentage leads by the Republican and Democratic candidates, respectively. (A value of* 0 *represents a tie.) The narrow orange bars indicate the regions that encompass the middle* 80% *of our simulated election results.*

constraints, we based our original senatorial forecast from 5 November on 4,000 simulations of (3.4)–(3.6) and used a larger time step ($\Delta t = 15$ days) than usual for parameter fitting. Additionally, after checking our polling data without the election-day rush, we found several typos that we corrected for the forecasts in the main text. See section SM1 in our supplementary material for details. We include our original forecasts [81] from 5 November 2018 in Figures SM1–SM3. In the main text, we present the forecasts that we obtain using our typical simulation parameters.[9] Both forecasts project the same candidate to win in each state.

In Figures 5 and 6, we compare our gubernatorial and senatorial forecasts for swing states and superstates with those of several popular sources. For the gubernatorial races, our Safe Red superstate consists of Alabama (AL), Arizona (AZ), Arkansas (AR), Idaho (ID), Maryland (MD), Massachusetts (MA), Nebraska (NE), New Hampshire (NH), South Carolina (SC), Tennessee (TN), Texas (TX), Vermont (VT), and Wyoming (WY); and our Safe Blue superstate consists of California (CA), Colorado (CO), Hawaii (HI), Illinois (IL), Michigan (MI), Minnesota (MN), New Mexico (NM), New York (NY), Pennsylvania (PA), and Rhode Island (RI). For the senatorial races, our Safe Red superstate consists of Mississippi (MS), Mississippi special (MS*), NE, Utah (UT), and WY; and our Safe Blue superstate consists of Connecticut (CT), Delaware (DE), HI, Maine (ME), MD, MA, MI, MN, NM, NY,

[9]For our typical simulation parameters, we use $\Delta t = 3$ days for parameter fitting and simulate 10,000 realizations of (3.4)–(3.6). See section 3.4 and Appendix B.3 for details.

(a)

2018 Gubernatorial Elections

	Model 21 Oct.	Cook 26 Oct.	IE 1 Nov.	**Model** 3 Nov.	538 4 Nov.	Sabato 5 Nov.	RCP 5 Nov.	538 6 Nov.
Alaska	91.2%		Tilt	90.4%	70.2%			68.9%
Connecticut	95.4%			95.0%	78.6%			79.0%
Florida	52.6%		Tilt	52.6%	75.7%			77.2%
Georgia	53.2%		Tilt	53.1%	59.2%			67.8%
Iowa	72.9%		Tilt	66.9%	52.1%			57.3%
Kansas	52.5%			53.0%	58.2%			57.2%
Maine	76.6%		Tilt	74.6%	94.4%			94.7%
Nevada	50.5%		Tilt	54.1%	55.0%			51.5%
Ohio	73.7%			74.6%	55.2%			59.5%
Oklahoma	52.8%			55.1%	86.2%			85.7%
Oregon	62.3%		Tilt	60.6%	81.4%			82.3%
South Dakota	50.4%		Tilt	60.7%	83.1%			63.1%
Wisconsin	71.6%			68.1%	60.6%			59.7%

Solid Rep. (≥95%)
Likely Rep. (≥75%)
Lean Rep. (≥60%)
Toss-Up (<60%)
Lean Dem. (≥60%)
Likely Dem. (≥75%)
Solid Dem. (≥95%)

(b)

2018 Senatorial Elections

	Model 8 July	**Model** 7 Aug.	**Model** 6 Sept.	**Model** 6 Oct.	Sabato 11 Oct.	IE 12 Oct.	538 30 Oct.	IE 1 Nov.	**Model** 3 Nov.	Sabato 5 Nov.	538 6 Nov.
Arizona	90.0%	84.6%	71.8%	70.0%			64.5%	Tilt	65.3%		61.4%
Florida	79.3%	78.2%	75.1%	66.1%			70.6%	Tilt	59.4%		70.4%
Indiana	55.8%	88.5%	88.9%	80.8%			67.2%		75.1%		71.8%
Minnesota*		99.1%	97.9%	97.3%			89.7%		95.3%		92.4%
Missouri	65.8%	61.4%	67.9%	63.4%			60.1%	Tilt	57.1%		56.9%
Montana	91.3%	83.3%	79.9%	77.0%		Tilt	85.3%	Tilt	83.3%		76.0%
Nevada	66.7%	58.5%	54.2%	56.6%			59.2%	Tilt	51.8%		57.0%
New Jersey	67.2%	51.6%	71.0%	78.4%			89.9%		77.5%		94.6%
North Dakota	72.5%	76.1%	77.1%	88.0%		Tilt	73.6%		89.3%		73.2%
Ohio	99.8%	99.7%	99.6%	99.5%			95.7%		99.1%		96.7%
Tennessee	87.8%	91.9%	82.0%	69.0%			76.3%		56.4%		80.4%
Texas	92.0%	90.4%	84.7%	85.9%			80.6%		86.1%		78.8%
West Virginia	90.8%	93.1%	93.0%	93.4%		Tilt	89.4%	Tilt	93.4%		87.9%
Wisconsin	94.0%	97.9%	94.6%	96.2%			98%		95.6%		97.7%

Fig. 6 *Ratings for the* 2018 (a) *gubernatorial and* (b) *senatorial races. There is considerable variability in these ratings across forecasters. One forecaster may identify a given race as a toss-up, and another forecaster may identify that race as solidly partisan. As we show in panel* (b)*, our forecasts are consistent (with respect to which party we project to win a given race) after July. We base our forecasts on* 10,000 *simulations of our SDE model* (3.4)–(3.6)*. We show ratings from our model, the Cook Political Report* [11]*, Inside Elections (IE)* [38]*, Sabato's Crystal Ball* [68]*, RealClearPolitics (RCP)* [10]*, and FiveThirtyEight* [77]*. IE* [38] *breaks down its ratings to include a "Tilt" category. Our SDE model and FiveThirtyEight* [77] *provide numbers to quantify uncertainty. For toss-ups, the number or text is red (respectively, blue) if it corresponds to a Republican's (respectively, Democrat's) chance of winning. We indicate the actual election outcome by the coloring of the state name. In panel* (b)*, we do not forecast the MN special (MN*) election in July, because this race had no polls before August.*

Table 2 *Final forecast performance for the* 2018 *gubernatorial and senatorial races. We measure performance by calculating the mean MOV error, the number of state races that are missed or not called, and log-loss error. We use final forecasts, and we note that lower numbers indicate better performance. Across these diagnostics, the forecasters perform similarly, but determining who is the most successful depends on what one values in a forecast. For example, the Cook Political Report* [11] *identified the winning gubernatorial candidate in all of the states for which they provided a forecast, but they left* 12 *states as too close to call. Our SDE model* (3.4)–(3.6) *provides forecasts for all of these states, but it was incorrect for* 4 *of them. The measurements in the table are for the races in Figure* 6*, and they do not include the states that we combined into the Safe Red and Safe Blue superstates. (Additionally, see section* SM2 *and Tables* SM2 *and* SM3 *for alternative ways of measuring forecast accuracy.)*

Forecaster	2018 gubernatorial races				2018 senatorial races			
	Mean MOV error	Num. races missed	Num. races not called	Log-loss error	Mean MOV error	Num. races missed	Num. races not called	Log-loss error
Our model	4.1%	4	0	0.589	4.6%	3	0	0.396
538 [77]	3.1%	4	0	0.548	3.7%	3	0	0.410
Sabato [68, 69]	NA	3	1	0.585	NA	1	0	0.379
Cook [11]	NA	0	12	0.670	NA	0	9	0.553
IE [38]	NA	2	3	0.619	NA	1	1	0.415
RCP [10]	NA	0	12	0.647	NA	0	8	0.565

PA, RI, VT, Virginia (VA), and Washington (WA). Our results in Table 2 are based only on the remaining states, which we treat individually in our model.

Given the probabilistic nature of forecasts, it is not straightforward to evaluate their accuracy [63], and we use the 2018 races to discuss a few ways of quantifying forecast performance. For quantitative forecasters, one natural way of evaluating performance is by computing the error in their forecast margin of victory (MOV) by state. We denote the percentages of people who voted Republican and Democratic in a given state by R^{result} and D^{result}, respectively, and we denote the corresponding forecast percentages by R^{forecast} and D^{forecast}. The MOV error in that state is

$$\text{MOV error} = \left|\left(R^{\text{result}} - D^{\text{result}}\right) - \left(R^{\text{forecast}} - D^{\text{forecast}}\right)\right|.$$

For example, as we show in Figure 5(a), we forecast that the Democratic candidate (Gillum) would win the Florida gubernatorial race by 0.4 percentage points over the Republican nominee (DeSantis). DeSantis edged out Gillum by 0.4 points, so our MOV error is 0.8 points for this specific race. For close races, it is worth comparing vote margins and MOV errors to the margins of error in the polling data. The mean margins of error that were reported in the polls [10] on which we based our parameters were 4.1 and 4.0 for the gubernatorial and senatorial data, respectively. Critically, this reported error is sampling error only; it does not account for other sources of error, such as ones from unrepresentative polling samples, which can result in error that is correlated by demographics [63]. Although it is straightforward to measure accuracy using MOV error, one drawback of doing so is that one can use this measure only for quantitative forecasters (see Table 2).

The baseline measure of how well forecasters do at calling race outcomes—specifically, of whether a state will elect a Republican or a Democrat—often attracts media attention. As we illustrate in Table 2, 2018 was a good year for a couple of the well-known pundits who use qualitative approaches, and Sabato's Crystal Ball [69] was the most successful of these at calling outcomes. The quantitative forecasts of

our SDE model (3.4)–(3.6) and FiveThirtyEight tied for second place (along with Inside Elections [38]). Our SDE model and FiveThirtyEight forecast the incorrect winner in the same 4 races for governor and in 2 of the same races for senator. We also were wrong about Tennessee, and FiveThirtyEight missed Florida [77]. The Cook Political Report [11] and RealClearPolitics [10] performed worse based on this baseline measure, because they left many races as toss-ups.

As we show in Figure 6, there is a lot of variability in how strongly different sources forecast the races. A helpful measure to evaluate classification models is logarithmic loss [28], which rewards confident forecasts that identify the winning candidate and penalizes confident forecasts that do not identify the winner. It is given by

$$\text{log loss} = -\frac{1}{E}\sum_{j=1}^{E}\left(y_i \log p_i + (1-y_i)\log\left(1-p_i\right)\right),$$

where "log" is the natural logarithm, E is the number of states that we treat individually (so $E = 13$ and $E = 14$ for the 2018 gubernatorial and senatorial races, respectively), $y_i = 1$ if the projected candidate wins in state i and $y_i = 0$ otherwise, and p_i is the probability (see the percentages in Figure 6) that we assign to the projected winning candidate in state i. To calculate the log-loss error for qualitative forecasts, we specify $p_i = 0.5$ for "Toss-up," $p_i = 0.55$ for "Tilt," $p_i = 0.675$ for "Lean," $p_i = 0.85$ for "Likely," and $p_i = 0.975$ for "Solid." As we show in Table 2, the forecasts from our SDE model (3.4)–(3.6) rank second and third according to log-loss error for the races for senator and governor, respectively, among our example popular forecasters. In comparison to the log-loss errors in Table 2, a log-loss error of about 0.7 corresponds to a hypothetical forecast that assigns a 50% chance to each of the two candidates in a race.

We have focused predominantly on producing final forecasts (i.e., those that are available right before an election takes place), in part because public attention often centers on how forecasts from the eve of an election compare to race outcomes and in part because our work is a first step toward data-driven election forecasting from a dynamical-systems perspective. However, the most meaningful forecasts are those in the weeks and months before an election day, and there is particular value in forecasts that remain stable across time [54, 55, 84]. Producing an early forecast is very challenging, and it provides a better view than late forecasts of a model's worth [55]. To begin to address these ideas, we show earlier forecasts for the 2018 races in Figure 6. We base these forecasts on less polling data; for example, our 8 July forecast uses polling data up to and including 8 July. We use the same superstate categorizations in these forecasts as we do in our final forecasts, which rely on the ratings of popular forecasters in August and November (see Appendix B.2). Notably, our July, August, September, and October forecasts for swing states in the senatorial races are as accurate as our final forecasts at calling race outcomes (see Figure 6(b)). As the election nears and we incorporate more polling data into our model, our performance (as measured by log-loss error and MOV error) improves. This supports observations [24, 87] that polling data become more reliable over time. In comparison to Table 2, we miss the true vote margins by 6.7, 5.8, and 5.0 percentage points on average in August, September, and October, respectively. Similarly, our log-loss error decays in time; it is 0.56 in August, 0.53 in September, and 0.44 in October. For the gubernatorial elections, our SDE model calls the same state outcomes roughly two weeks before the election as it does in November.

5. Conclusions. We developed a method for forecasting elections by adapting ideas from compartmental modeling and epidemiology, and we illustrated the utility of such a dynamical-systems approach by applying it to the U.S. races for president, senator, and governor in 2012, 2016, and 2018. When making our modeling choices, we tried to limit the number of election-specific details in our methodology. Despite our approach of using poll data without any weighting adjustments, as well as clear differences between voting dynamics and the spread of infectious diseases, we performed similarly to popular forecasters in calling the final outcomes of the races. Moreover, we were able to forecast the outcomes of the senatorial elections in 2018 using polling data prior to August 2018 with the same success rate as FiveThirtyEight's final forecast [77].

We consider our model's basis in the well-studied, multidisciplinary field of mathematical epidemiology a virtue in this initial dynamical-systems effort, as part of our goal is to help demystify election prediction, highlight future research directions in the forecasting of elections (and other complex systems), and motivate a broader research community to engage actively with pollster interpretations and polling data. There are many ways to build on our basic modeling approach and account more realistically for voter interactions.

To give one example of a viable research direction, it will be useful to be more nuanced about how to handle undecided and minor-party voters. FiveThirtyEight [74] assigns a voting opinion (mostly to one of the major parties) to any undecided voters who remain on election day, and *The Huffington Post* factors undecided voters into an election's uncertainty [43]. By contrast, we assumed that any undecided voters at the end of our simulations are minor-party voters. Using the fraction of undecided voters to inform one's choice of noise strength is an interesting direction to pursue. Moreover, because we compare our model with some popular qualitative forecasters, we showed our forecasts as projected vote margins rather than as absolute Republican and Democratic percentages. It may be desirable for future studies to look more closely at how the fractions of Republican, Democratic, and undecided voters evolve in time (see, e.g., [55, 85]).

We assumed that all polls are equally accurate (e.g., we did not consider the time to election and pollster-reported error), and we did not distinguish between partisan and nonpartisan polls or between polls of likely voters, registered voters, and adults. This minimal, poll-aggregating approach echoes the work of Wang [84]. By contrast, FiveThirtyEight [74] relies on measures [78] of polling-firm accuracy to weight polls, and its analysts adjust polls of registered voters and adults to frame all of their data in terms of likely voters. Using FiveThirtyEight's pollster ratings [78] to weight polls in our model would allow us to explore how the various subjective choices of forecasters determine their predictions. Similarly, future work can compare the influence of noise that is correlated based on demographics with the effects of noise that is correlated based on roughly the last 80 years of state voting history. *The Huffington Post* [43] uses the latter (and it does not use demographics).

In our study of election dynamics, we took a macroscopic, simplified view of state and voter interactions. We based our approach on compartmental modeling of contagions because it gives a well-established, multidisciplinary way to include asymmetric state–state relationships in a model. However, when a social behavior or opinion appears to spread in a community, it is often difficult to determine whether transmission is actually occurring. In particular, the appearance of "spreading" may emerge because social contagions are truly spreading between individuals (that is,

individuals are influencing each other), because people form relationships with others who are similar to them and behave in a similar way (e.g., adopting the same opinion) due to their shared characteristics [72], because of some external factors, or because of a combination of such processes. By building more detailed mathematical models of voter behavior in the future, one can help elucidate what role influence plays in political opinion dynamics.

Our models assume that every voting-age individual is equally likely to interact with any other voter in the U.S. Although this mean-field approach fits within the theme of simplicity that we embraced throughout our modeling process, the assumption of uniform mixing is not particularly realistic [22, 58, 86]. For example, several celebrities were heavily involved in encouraging individuals to vote in the 2018 elections, and they have more prominent platforms than a typical voter. Accounting for realistic network structures and exploring frameworks—such as voter models [20, 33], local majority-rule models [34, 35], and threshold models [52]—other than compartmental models may be helpful for capturing relationships between voters. Network models may also allow future studies of how different methods, such as "big nudging" [32, 88], may influence voter turnout and behavior at an individual level.

In future modeling efforts, it will also be useful to incorporate additional types of data (e.g., measurements of partisan prejudice by county [66]), as they become available, into election models to improve both the detail and the quality of forecasts. Our models work on the level of states because state polling data are available. By contrast, House elections are polled less regularly, making fundamental data more important for these races [76]. Although precinct-level data are available for presidential election results (see, e.g., [71]), we are not aware of polls at the precinct level. Because many states have regions with different voting behavior, such as urban versus rural areas, incorporating precinct-level polling data has the potential to lead to significant improvements in forecasts.

When fitting our parameters, we averaged polls [8, 10] by month (see section 3.2). This technique smooths out daily fluctuations, which may be more representative of sampling error than of real shifts in opinion [42], so it may throw out certain interesting dynamics, such as those that occur around party conventions [55]. (As described by FiveThirtyEight [74], candidates often receive a spike in support after their party's convention.) The impact of campaigns and media coverage on opinion dynamics is debated in the political-science community [24, 37, 87]. With a finer view on polls, one can explore the possible effects of time-specific events, such as a large rally or a story about a candidate in the media, using our modeling framework. Similarly, one can build feedback mechanisms into a model to test how the perception of future election results influences an individual's likelihood of voting. The framework of dynamical systems provides a valuable approach for exploring the temporal evolution of opinions and their interplay with external forces (such as the media, rallies, and conventions).

Our compartmental-model approach allowed us to obtain parameters that are related to the strengths of interactions between states and measurements of voter turnover by state for each election year and race (see Figures SM4–SM6). In the future, it will be useful to investigate the stability of our forecasts and parameters using alternative methods (e.g., data-assimilation methods [51]) for fitting these parameters. By comparing our parameters across years, types of elections, and different approaches for fitting, one can help identify blocs of states that are related persistently, analyze which states have the most plastic voter populations, and suggest differences in the political dynamics in presidential, senatorial, and gubernatorial races. One can also

use our parameters from previous elections to provide early forecasts for upcoming races before large-scale polling data become available. These and other future research directions may provide insight into how state relationships evolve across years, allowing researchers to identify ways that the U.S. electorate may be changing in time, which may in turn suggest ideas to incorporate into future forecasts.

Appendix A. Additional Background on Compartmental Models. The dynamics of transmission and recovery in a susceptible–infected–susceptible (SIS) model are described by the following coupled ODEs:

$$\frac{d\tilde{S}}{dt} = \gamma\tilde{I} - \tilde{\beta}[\tilde{S}\tilde{I}]\,,$$
$$\frac{d\tilde{I}}{dt} = -\gamma\tilde{I} + \tilde{\beta}[\tilde{S}\tilde{I}]\,,$$

where $\tilde{S}(t)$ and $\tilde{I}(t)$, respectively, are the mean numbers of susceptible and infected individuals in a population at time t and $[\tilde{S}\tilde{I}]$ is the mean number of contacts between infected and susceptible individuals. To close the system [48], we use the approximation $[\tilde{S}\tilde{I}] \approx \tilde{S} \cdot n \cdot \tilde{I}/N$, where N is the total number of individuals in a population and n is the mean number of contacts per person. We define the notation $\beta := \tilde{\beta}n$ and obtain the following system:

$$\frac{d\tilde{S}}{dt} = \gamma\tilde{I} - \beta\tilde{S}\tilde{I}/N\,,$$
$$\frac{d\tilde{I}}{dt} = -\gamma\tilde{I} + \beta\tilde{S}\tilde{I}/N\,.$$

Writing this system in terms of the population fractions, $S(t) = \tilde{S}(t)/N$ and $I(t) = \tilde{I}(t)/N$, and dividing by N yields (2.1)–(2.2).

A similar process yields our two-pronged deterministic SIS model (3.1)–(3.3) for election forecasting. The key difference is how we calculate $[S^i I_{\mathrm{D}}^j]$ and $[S^i I_{\mathrm{R}}^j]$. Specifically, we use the approximation

$$[\tilde{S}^i \tilde{I}_{\mathrm{D}}^j] \approx \tilde{S}^i \cdot n \cdot \left(N^j/N\right) \cdot \left(\tilde{I}_{\mathrm{D}}^j/N^j\right)\,,$$
$$[\tilde{S}^i \tilde{I}_{\mathrm{R}}^j] \approx \tilde{S}^i \cdot n \cdot \left(N^j/N\right) \cdot \left(\tilde{I}_{\mathrm{R}}^j/N^j\right)\,.$$

We thereby estimate, for example, the number of interactions between undecided voters in state i and Democratic voters in state j as the mean number of interactions that involve an undecided voter in state i multiplied by the probability that the interaction is with someone in state j multiplied by the probability that someone in state j is a Democratic voter. In making these approximations, we are assuming that an undecided voter is equally likely to interact with a Republican voter or a Democratic voter across the U.S. In particular, we do not assume that interactions are more likely to occur between individuals in the same state or between those in neighboring states. We also ignore any effects of homophily (even though people are more likely to interact with others who are similar in some way, such as their political outlook [59]), such that the number of interactions between individuals in different states depends only on the voting-age populations [2, 4, 7] of the states and on the numbers of Republican, Democratic, and undecided voters that are currently in them.

Appendix B. Election-Modeling Details. We now provide additional details about our election-forecasting process. We overview the data that we use and describe

several special cases in Appendix B.1. We then discuss how we select superstates (see Appendix B.2) and how we numerically implement our model (see Appendix B.3).

B.1. Data. We obtained publicly available state polling data for 2012 and 2016 from HuffPost Pollster [8] using the Pollster API v2 [9]. State polling data for 2018 were not available from HuffPost Pollster [8], so we collected 2018 data by hand from RealClearPolitics [10]. See our GitLab repository [82] for the 2012, 2016, and 2018 polling data. We use 2012, 2016, and 2017 estimates of voting-age population sizes from the Federal Register [2, 4, 7] to specify N and N^i in (3.1)–(3.6). We use 2017 data for 2018 because 2018 measurements were not yet available at the time of our analysis.

B.1.1. Special Cases and Notes. Working with election data is often messy, and we comment on a few special cases in our efforts.

(1) Different election days: Unlike the other 2012 races, the Wisconsin gubernatorial election took place in June 2012, so we do not forecast this race.
(2) Single-party races: California had two Democrats running for senator in 2018 and 2016. Because our model assumes a race of a Democrat facing a Republican, we do not use polling data from California when it has a single-party race. Naturally, we also do not forecast these races.
(3) Independent candidates: The Vermont senatorial races featured an independent candidate in place of a Democrat in 2012 and 2018. Consequently, for the 2012 election, we do not forecast Vermont. For the 2018 race, we treat the independent candidate as a Democrat in our models so that we can still provide a forecast for Vermont.
(4) Third-candidate polling data: We focus on polling data that compare two candidates. In particular, we do not include polls that report data for races in which there are three or more candidates who each get reasonably large shares of the vote (with the exception of the Louisiana Senate race in 2016). The polls that we found from RealClearPolitics [10] for New Mexico's 2018 senatorial race were for three candidates, so we do not include New Mexico's polls in our averaged data points for the Safe Blue superstate for this election. We also do not forecast the Maine 2012 senatorial race because it included three popular candidates. For the 2016 Senate race in Louisiana, which uses the so-called "jungle primary" system, there were more than two candidates. No candidate received a majority of the vote on election day, so there was a runoff election between the top Republican candidate and the top Democratic candidate on 10 December 2016. We treat these two candidates as the main candidates in our forecasts of the 2016 Senate election in Louisiana, and we forecast the percentages of the vote that they each received on 8 November 2016. However, in Figure 3(d), we compare our forecast vote margin to the final vote margin in the runoff election.
(5) No polls: In some elections, one or more states have no polls. If these states lie in our Safe Red or Safe Blue superstates, this is not an issue, as we simply assign the vote margin of the appropriate superstate to them. However, in elections for which we forecast each state individually, we cannot forecast states without polling data. Therefore, because polling data from HuffPost [8, 9] were not available for the 2012 gubernatorial races in Delaware and West Virginia, we do not provide forecasts for these races.
(6) Early forecasts: For our 8 July forecasts for the 2018 senatorial races in Figure 6(b), the Minnesota special election had no polls prior to 8 July, so we remove this race from our models for this forecast only.

(7) Demographic estimates: To correlate noise in our model (3.4)–(3.6), we use estimates of the numbers of Hispanic individuals and non-Hispanic Black individuals in each state in 2016. We gathered these estimates from the U.S. Census Bureau through American FactFinder [3]. Later, after we developed our model, American FactFinder was decommissioned, so the website [3] is no longer available. Demographic estimates are now available at [13], but these data are slightly different from the estimates that we used, because the U.S. Census Bureau revises their past estimates when they make new estimates. See [82] for the data that we used.

B.2. Selecting Superstates. We focus on forecasting elections in swing states and treat all reliably Red and Blue states together as two "superstate" conglomerates. (We do not specify that these superstates actually vote Republican and Democratic, respectively; such voting results are outputs of our models.) This raises the question of how to identify states as "safe" or "swing," and we do this differently for different elections. For presidential races, we define our swing states as the ones that FiveThirtyEight has identified as "traditional swing states" [75]. These states are Colorado (CO), Florida (FL), Iowa (IA), Michigan (MI), Minnesota (MN), Nevada (NV), New Hampshire (NH), North Carolina (NC), Ohio (OH), Pennsylvania (PA), Virginia (VA), and Wisconsin (WI) (see Figure 1(d)). Therefore, for the presidential elections, $M = 14$ in (3.1)–(3.6). In our notation, $\{S^1, I_{\mathrm{D}}^1, I_{\mathrm{R}}^1\}$ and $\{S^2, I_{\mathrm{D}}^2, I_{\mathrm{R}}^2\}$ refer to the voter fractions in the Red and Blue superstates, respectively, and $\{S^i, I_{\mathrm{D}}^i, I_{\mathrm{R}}^i\}$ for $i \in \{3, 4, \ldots, 14\}$ are the voter fractions in the 12 swing states.

To define superstates in the senatorial and gubernatorial races, we use the race ratings of popular forecasters. For the 2018 senatorial races, we combine the August 2018 ratings of Sabato's Crystal Ball [69], 270toWin (the consensus version) [6], and *The New York Times* [16]. We determine which states to include in our superstates for the 2018 gubernatorial races based on the ratings of FiveThirtyEight [77], the Cook Political Report [11], Sabato's Crystal Ball [68], and Inside Elections [38] (all accessed on 1 November 2018). We define our Safe Red and Safe Blue superstates for the 2012 senatorial races based on Sabato's Crystal Ball [70] and *The New York Times* [1]. We base our superstates for the 2016 senatorial races on 270toWin [6], Sabato's Crystal Ball [67], and *The Huffington Post* [43]. We treat each state separately for the 2012 and 2016 gubernatorial races. In Table SM1, we give a summary (for each election) of the states that we forecast individually and those that we combine into the Safe Red and Safe Blue superstates.

B.3. Numerical Implementation. For our parameter fitting, we use the OPTIM routine in R (version 3.4.2) [64] to perform constrained optimization of the least-squares objective function (see section 3.2), subject to nonnegative rate constraints [23], with a time step of $\Delta t = 3$ days over T months. (As a simplification, we assume that each month is 30 days long.) We use this time step in all cases, except for the forecasts in Figures SM1(b,c) (for which we use $\Delta t = 15$ days). We simulate our models in MATLAB (version 9.3). For each state (or superstate), we set its initial condition to the earliest of its T data points that we use for parameter fitting. We solve our ODE model (3.1)–(3.3) using a forward Euler scheme and our SDE model (3.4)–(3.6) using the Euler–Maruyama method [40]. We use the constraint that $S^i + I_{\mathrm{R}}^i + I_{\mathrm{D}}^i = 1$ to reduce our system to 2 equations per state (or superstate) for our simulations. For both of our models, we use a time step of $\Delta t = 0.1$ day, and we simulate them from 1 January until an election day. In these simulations, we assume that each month has a length of 30 days. Specifically, we simulate for 306 days for

the 2012 and 2018 races and for 308 days for the 2016 races. The noise strength in our SDE system (3.4)–(3.6) is $\sigma = 0.0015$ in all of our simulations.

Appendix C. Supplementary Materials. We include the following in our supplementary materials:

- The original forecasts for the 2018 elections that we posted on 5 November 2018 [81]; see section SM1 and Figures SM1–SM3.
- A summary of the races that we forecast individually, the races that we include in our Safe Red and Safe Blue superstates, and the races that we do not forecast (see Appendix B.1.1) by year; see Table SM1.
- Alternative ways of measuring forecast accuracy; see section SM2 and Tables SM2 and SM3.
- A summary of the simulation code, formatted polling data, and model parameters that we include in our GitLab repository [82]; see section SM3 and Figures SM4–SM6.

Acknowledgments. We thank Heather Zinn Brooks, Catherine Calder, Samuel Chian, Eli Fenichel, Serge Galam, William He, Brian Hsu, Kyle Kondik, Christopher Lee, Niall Mangan, Subhadeep Paul, Laura Pomeroy, and Samuel Scarpino for helpful comments and pointers to references. A.V. thanks Maciej Pietrzak for his advice on APIs.

REFERENCES

[1] *Election* 2012 *Senate Map*, The New York Times, 2012, https://www.nytimes.com/elections/2012/results/senate.html (accessed 03-11-2018). (Cited on p. 860)

[2] *Estimates of the Voting Age Population for* 2012. *A Notice by the Commerce Department on* 1/30/2013. *Agency: Office of the Secretary, Commerce*, Report 78 FR 6289, Federal Register Notices, 2013. (Cited on pp. 843, 858, 859)

[3] *Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States, States, and Counties: April* 1, 2010 *to July* 1, 2016, U.S. Census Bureau, Population Division, 2017, https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=PEP_2015_PEPSR6H&prodType=table (accessed 22-01-2018). (Cited on pp. 850, 860)

[4] *Estimates of the Voting Age Population for* 2016. *A Notice by the Commerce Department on* 1/30/2017. *Agency: Office of the Secretary, Commerce*, Report 82 FR 8720, Federal Register Notices, 2017. (Cited on pp. 843, 858, 859)

[5] 2018 *Midterm Election Results*, The New York Times, 2018, https://www.nytimes.com/interactive/2018/us/elections/calendar-primary-results.html (last accessed 26-02-2019). (Cited on p. 852)

[6] 270*toWin*, https://www.270towin.com (last accessed 03-11-2018). (Cited on pp. 838, 860)

[7] *Estimates of the Voting Age Population for* 2017. *A Notice by the Commerce Department on* 2/20/2018. *Agency: Office of the Secretary, Commerce*, Report 83 FR 7142, Federal Register Notices, 2018. (Cited on pp. 843, 858, 859)

[8] *HuffPost Pollster*, The Huffington Post, https://elections.huffingtonpost.com/pollster (last accessed 02-11-2018). (Cited on pp. 844, 846, 857, 859)

[9] *HuffPost Pollster API v*2, The Huffington Post, https://elections.huffingtonpost.com/pollster/api/v2; data provided under a CC BY-NC-SA 3.0 license (https://creativecommons.org/licenses/by-nc-sa/3.0/deed.en_US) (last accessed 02-11-2018). (Cited on p. 859)

[10] *RealClearPolitics: Polls*, https://www.realclearpolitics.com/epolls/latest_polls/elections/ (last accessed 22-12-2018). (Cited on pp. 844, 846, 851, 852, 853, 854, 855, 857, 859)

[11] *The Cook Political Report: Ratings*, https://www.cookpolitical.com/ratings (accessed 03-11-2018). (Cited on pp. 838, 848, 853, 854, 855, 860)

[12] *Ballotpedia*, https://ballotpedia.org/Main_Page (accessed 26-02-2019). (Cited on p. 849)

[13] *Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin: April* 1, 2010 *to July* 1, 2019, U.S. Census Bureau, Population Division, 2020, https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-detail.html#par_textimage_673542126 (accessed 09-08-2020). (Cited on p. 860)

[14] A. I. Abramowitz, *Will time for change mean time for Trump?*, PS: Political Sci. Politics, 49 (2016), pp. 659–660. (Cited on p. 839)

[15] L. J. S. Allen, M. Langlais, and C. J. Phillips, *The dynamics of two viral infections in a single host population with applications to hantavirus*, Math. Biosci., 186 (2003), pp. 191–217. (Cited on p. 841)

[16] S. Almukhtar, M. Andre, W. Andrews, M. Bloch, J. Bowers, L. Buchanan, N. Cohn, A. Coote, A. Daniel, T. Fehr, S. Jacoby, J. Katz, J. Keller, A. Krolik, J. C. Lee, R. Lieberman, B. Migliozzi, P. Murray, K. Quealy, J. Patel, A. Pearce, R. Shorey, M. Strickland, R. Taylor, I. White, M. Whitely, and J. Williams, *Live Forecast: Who Will Win the Senate?*, The New York Times, 2018, https://www.nytimes.com/interactive/2018/11/06/us/elections/results-senate-forecast.html (last accessed 06-10-2018). (Cited on pp. 838, 860)

[17] L. M. A. Bettencourt, A. Cintrón-Arias, D. I. Kaiser, and C. Castillo-Chavéz, *The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models*, Phys. A, 364 (2006), pp. 513–536. (Cited on p. 841)

[18] L. Bonnasse-Gahot, H. Berestycki, M.-A. Depuiset, M. B. Gordon, S. Roché, N. Rodriguez, and J.-P. Nadal, *Epidemiological modelling of the* 2005 *French riots: A spreading wave and the role of contagion*, Sci. Rep., 8 (2018), art. 107. (Cited on p. 841)

[19] L. Bottcher, H. J. Herrmann, and H. Gersbach, *Clout, activists and budget: The road to presidency*, PLoS ONE, 13 (2018), art. e0193199. (Cited on p. 839)

[20] D. Braha and M. A. M. de Aguiar, *Voting contagion: Modeling and analysis of a century of U.S. presidential elections*, PLoS ONE, 12 (2017), art. e0177970. (Cited on pp. 839, 857)

[21] R. Brauer and C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*, 2nd ed., Texts Appl. Math., Springer, Heidelberg, Germany, 2012. (Cited on p. 841)

[22] S. Busenberg and C. Castillo-Chavez, *A general solution of the problem of mixing of subpopulations and its application to risk- and age-structured epidemic models for the spread of AIDS*, Math. Med. Biol., 8 (1991), pp. 1–29. (Cited on pp. 847, 857)

[23] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, *A limited memory algorithm for bound constrained optimization*, SIAM J. Sci. Comput., 16 (1995), pp. 1190–1208, https://doi.org/10.1137/0916069. (Cited on p. 860)

[24] J. E. Campbell, *Polls and votes: The trial-heat presidential election forecasting model, certainty, and political campaigns*, Amer. Politics Res., 24 (1996), pp. 408–433. (Cited on pp. 839, 855, 857)

[25] C. Castellano, S. Fortunato, and V. Loreto, *Statistical physics of social dynamics*, Rev. Modern Phys., 81 (2009), pp. 591–646. (Cited on p. 839)

[26] S. Chian, W. L. He, C. M. Lee, D. F. Linder, M. A. Porter, G. A. Rempala, and A. Volkening, 2020 *U.S. Election Forecasts with a Compartmental Model*, https://modelingelectiondynamics.gitlab.io/2020-forecasts/. (Cited on p. 840)

[27] B. J. Coburn, B. G. Wagner, and S. Blower, *Modeling influenza epidemics and pandemics: Insights into the future of swine flu (H1N1)*, BMC Med., 7 (2009), art. 30. (Cited on p. 841)

[28] N. Cohen, *Volume* 24: *The* 2018 *Election, Who Projected It Best?*, https://www.lobbyseven.com/single-post/2018/12/17/Volume-24-The-2018-Election-Who-Projected-It-Best (accessed 22-12-2018). (Cited on p. 855)

[29] E. Comen, T. C. Frohlich, and M. B. Sauter, *America's Most and Least Educated States: A Survey of All* 50, https://247wallst.com/special-report/2016/09/16/americas-most-and-least-educated-states-a-survey-of-all-50/2/ (last accessed 02-11-2018). (Cited on p. 850)

[30] O. Diekmann and J. A. P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, John Wiley & Sons, New York, 2000. (Cited on p. 841)

[31] A. Emamadjomeh and D. Lauter, *Where the Presidential Race Stands Today*, USC Dornsife/Los Angeles Times Daybreak Poll, 2016, http://graphics.latimes.com/usc-presidential-poll-dashboard/ (accessed 19-09-2018). (Cited on pp. 838, 839)

[32] R. Epstein and R. E. Robertson, *The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections*, Proc. Natl. Acad. Sci. USA, 112 (2015), art. e4512. (Cited on p. 857)

[33] J. Fernández-Gracia, K. Suchecki, J. J. Ramasco, M. San Miguel, and V. M. Eguiluz, *Is the voter model a model for voters?*, Phys. Rev. Lett., 112 (2014), art. 158701. (Cited on pp. 839, 857)

[34] S. Galam, *The dynamics of minority opinions in democratic debate*, Phys. A, 336 (2004), pp. 56–62. (Cited on pp. 839, 857)

[35] S. Galam, *The Trump phenomenon: An explanation from sociophysics*, Internat. J. Modern Phys. B, 31 (2017), art. 1742015. (Cited on pp. 839, 857)
[36] S. Galam, *Unavowed abstention can overturn poll predictions*, Frontiers Phys., 6 (2018), art. 24. (Cited on p. 839)
[37] A. Gelman and G. King, *Why are American presidential election campaign polls so variable when votes are so predictable?*, British J. Political Sci., 23 (1993), pp. 409–451. (Cited on pp. 838, 839, 857)
[38] N. L. Gonzales, L. Askarinam, R. Matsumoto, R. Yoon, and S. Rothenberg, *Inside Elections with Nathan L. Gonzales, Nonpartisan Analysis: Year Archive:* 2018, https://insideelections.com/archive/year/2018 (last accessed 27-12-2018). (Cited on pp. 838, 853, 854, 855, 860)
[39] H. W. Hethcote, *The mathematics of infectious diseases*, SIAM Rev., 42 (2000), pp. 599–653, https://doi.org/10.1137/S0036144500371907. (Cited on p. 841)
[40] D. J. Higham, *An algorithmic introduction to numerical simulation of stochastic differential equations*, SIAM Rev., 43 (2001), pp. 525–546, https://doi.org/10.1137/S0036144500378302. (Cited on p. 860)
[41] P. Hummel and D. Rothschild, *Fundamental models for forecasting elections at the state level*, Elect. Stud., 35 (2014), pp. 123–139. (Cited on pp. 838, 839)
[42] S. Jackman, *Pooling the polls over an election campaign*, Aust. J. Political Sci., 40 (2005), pp. 499–517. (Cited on pp. 839, 848, 857)
[43] N. Jackson and A. Hooper, *Huffington Post Election* 2016 *Forecast: President*, The Huffington Post, 2016, http://elections.huffingtonpost.com/2016/forecast/president (accessed 31-10-2018). (Cited on pp. 838, 856, 860)
[44] W. Jennings and C. Wlezien, *Election polling errors across time and space*, Nat. Hum. Behav., 2 (2018), pp. 276–283. (Cited on p. 838)
[45] W. O. Kermack and A. G. McKendrick, *A contribution to the mathematical theory of epidemics*, Proc. R. Soc. London, 115 (1927), pp. 700–721. (Cited on p. 841)
[46] W. O. Kermack and A. G. McKendrick, *Contributions to the mathematical theory of epidemics.* II. *The problem of endemicity*, Proc. R. Soc. London, 138 (1932), pp. 55–83. (Cited on p. 841)
[47] W. O. Kermack and A. G. McKendrick, *Contributions to the mathematical theory of epidemics.* III. *Further studies of the problem of endemicity*, Proc. R. Soc. London, 141 (1933), pp. 94–122. (Cited on p. 841)
[48] I. Z. Kiss, J. C. Miller, and P. L. Simon, *Mathematics of Epidemics on Networks: From Exact to Approximate Models*, Springer, Cham, Switzerland, 2017. (Cited on pp. 841, 858)
[49] C. Klarner, *Forecasting the* 2008 *U.S. House, Senate and presidential elections at the district and state level*, PS: Political Sci. Politics, 41 (2008), pp. 723–728. (Cited on p. 839)
[50] B. E. Lauderdale and D. Linzer, *Under-performing, over-performing, or just performing? The limitations of fundamentals-based presidential election forecasting*, Int. J. Forecast, 31 (2015), pp. 965–979. (Cited on pp. 838, 839)
[51] K. Law, A. Stuart, and K. Zygalakis, *Data Assimilation: A Mathematical Introduction*, Texts Appl. Math. 63, Springer, Cham, Switzerland, 2015. (Cited on pp. 847, 851, 857)
[52] S. Lehmann and Y.-Y. Ahn, *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks*, Springer, Cham, Switzerland, 2018. (Cited on pp. 843, 857)
[53] D. Leip, *Dave Leip's Atlas of U.S. Presidential Elections*, http://uselectionatlas.org (accessed 26-02-2019). (Cited on p. 849)
[54] M. S. Lewis-Beck, *Election forecasting: Principles and practice*, British J. Politics Int. Relat., 7 (2005), pp. 145–164. (Cited on pp. 839, 855)
[55] D. A. Linzer, *Dynamic Bayesian forecasting of presidential elections in the States*, J. Amer. Stat. Assoc., 108 (2013), pp. 124–134. (Cited on pp. 839, 840, 848, 855, 856, 857)
[56] S. A. Marvel, H. Hong, A. Papush, and S. H. Strogatz, *Encouraging moderation: Clues from a simple model of ideological conflict*, Phys. Rev. Lett., 109 (2012), art. 118702. (Cited on p. 843)
[57] R. K. McCormack and L. J. S. Allen, *Multi-patch deterministic and stochastic models for wildlife diseases*, J. Biol. Dyn., 1 (2007), pp. 63–85. (Cited on p. 841)
[58] L. A. Meyers, B. Pourbohloul, M. E. J. Newman, D. M. Skowronski, and R. C. Brunham, *Network theory and SARS: Predicting outbreak diversity*, J. Theoret. Biol., 232 (2005), pp. 71–81. (Cited on pp. 847, 857)
[59] M. E. J. Newman, *Networks*, 2nd ed., Oxford University Press, Oxford, UK, 2018. (Cited on p. 858)

[60] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Epidemic processes in complex networks*, Rev. Modern Phys., 87 (2015), pp. 925–979. (Cited on p. 841)
[61] V. Pons, *Will a five-minute discussion change your mind? A countrywide experiment on voter choice in France*, Amer. Econ. Rev., 108 (2018), pp. 1322–1363. (Cited on p. 843)
[62] M. A. Porter and J. P. Gleeson, *Dynamical Systems on Networks: A Tutorial*, Front. Appl. Dyn. Syst. Rev. Tutor. 4, Springer, Cham, Switzerland, 2016. (Cited on p. 839)
[63] C. Prosser and J. Mellon, *The twilight of the polls? A review of trends in polling accuracy and the causes of polling misses*, Gov. Oppos., 53 (2018), pp. 757–709. (Cited on p. 854)
[64] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018. (Cited on p. 860)
[65] G. A. Rempala, *Least squares estimation in stochastic biochemical networks*, Bull. Math. Biol., 74 (2012), pp. 1938–1955. (Cited on p. 845)
[66] A. Ripley, R. Tenjarla, and A. Y. He, *The Geography of Partisan Prejudice*, The Atlantic, 2019, https://www.theatlantic.com/politics/archive/2019/03/us-counties-vary-their-degree-partisan-prejudice/583072/ (accessed 14-03-2019). (Cited on p. 857)
[67] L. J. Sabato and K. Kondik, *Sabato's Crystal Ball*, http://crystalball.centerforpolitics.org/crystalball/ (accessed 19-09-2018). (Cited on pp. 838, 848, 850, 860)
[68] L. J. Sabato and K. Kondik, *Sabato's Crystal Ball:* 2018 *Governor*, http://www.centerforpolitics.org/crystalball/2018-governor/ (last accessed 22-12-2018). (Cited on pp. 853, 854, 860)
[69] L. J. Sabato and K. Kondik, *Sabato's Crystal Ball:* 2018 *Senate*, http://www.centerforpolitics.org/crystalball/2018-senate/ (last accessed 04-11-2018). (Cited on pp. 854, 860)
[70] L. J. Sabato, K. Kondik, and G. Skelley, *Sabato's Crystal Ball: Projection: Obama Will Likely Win Second Term*, http://crystalball.centerforpolitics.org/crystalball/articles/projection-obama-will-likely-win-second-term/ (last accessed 03-11-2018). (Cited on p. 860)
[71] J. Schleuss, J. Fox, and P. Krishnakumar, *California* 2016 *Election Precinct Maps*, https://github.com/datadesk/california-2016-election-precinct-maps (accessed 14-03-2019). (Cited on p. 857)
[72] C. R. Shalizi and A. C. Thomas, *Homophily and contagion are generically confounded in observational social network studies*, Sociol. Methods Res., 40 (2011), pp. 211–239. (Cited on p. 857)
[73] N. Silver, *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't*, Penguin Press, New York, 2012. (Cited on p. 838)
[74] N. Silver, *FiveThirtyEight: A User's Guide to FiveThirtyEight's* 2016 *General Election Forecast*, https://fivethirtyeight.com/features/a-users-guide-to-fivethirtyeights-2016-general-election-forecast/ (accessed 31-10-2018). (Cited on pp. 839, 840, 847, 848, 850, 851, 856, 857)
[75] N. Silver, *FiveThirtyEight: Politics. The Odds of an Electoral College-Popular Vote Split Are Increasing*, https://fivethirtyeight.com/features/the-odds-of-an-electoral-college-popular-vote-split-are-increasing/ (accessed 02-11-2018). (Cited on p. 860)
[76] N. Silver, *FiveThirtyEight: How FiveThirtyEight's House, Senate and Governor Models Work*, https://fivethirtyeight.com/methodology/how-fivethirtyeights-house-and-senate-models-work/ (last accessed 26-02-2019). (Cited on p. 857)
[77] N. Silver, J. Boice, E. Brillhart, A. Bycoffe, R. Dottle, L. Eastridge, R. King, E. Koeze, A. Scheinkman, G. Wezerek, J. Wolfe, D. Dienhart, A. Jones-Rooy, D. Mehta, M. Nguyen, N. Rakich, D. Shan, and G. Skelley, *FiveThirtyEight: Election* 2018, https://projects.fivethirtyeight.com/2018-midterm-election-forecast/ (last accessed 20-12-2018). (Cited on pp. 838, 839, 848, 852, 853, 854, 855, 856, 860)
[78] N. Silver, A. Jones-Rooy, and D. Mehta, *FiveThirtyEight: FiveThirtyEight's Pollster Ratings, Based on the Historical Accuracy and Methodology of Each Firm's Polls*, https://projects.fivethirtyeight.com/pollster-ratings/. (Cited on p. 856)
[79] N. Silver, J. Kanjana, D. Mehta, J. Boice, A Bycoffe, M. Conlen, R. Fischer-Baum, R. King, E. Koeze, A. McCann, A. Scheinkman, and G. Wezerek, *FiveThirtyEight:* 2016 *Election Forecast. Who Will Win the Presidency?*, https://projects.fivethirtyeight.com/2016-election-forecast/ (last accessed 02-11-2018). (Cited on pp. 845, 848, 850)
[80] A. Topîrceanu, *Electoral Forecasting Using a Novel Temporal Attenuation Model: Predicting the US Presidential Elections*, preprint, https://arxiv.org/abs/2005.01799, 2020. (Cited on p. 839)

[81] A. Volkening, D. F. Linder, M. A. Porter, and G. A. Rempala, *Forecasting Elections Using Compartmental Models of Infections*, preprint, https://arxiv.org/abs/1811.01831v1, 2018. (Cited on pp. 840, 851, 852, 861)

[82] A. Volkening, D. F. Linder, M. A. Porter, and G. A. Rempala, *Forecasting Elections Using Compartmental Models*, https://gitlab.com/alexandriavolkening/forecasting-elections-using-compartmental-models/-/tree/master/Original%20Programs. (Cited on pp. 840, 841, 845, 859, 860, 861)

[83] N. Wang, Y. Fu, H. Zhang, and H. Shi, *An evaluation of mathematical models for the outbreak of COVID*-19, Precis. Clin. Med., 3 (2020), pp. 85–93. (Cited on p. 841)

[84] S. S.-H. Wang, *Origins of Presidential poll aggregation: A perspective from* 2004 *to* 2012, Int. J. Forecast., 31 (2015), pp. 898–909. (Cited on pp. 839, 840, 847, 848, 855, 856)

[85] W. Wang, D. Rothschild, S. Goel, and A. Gelman, *Forecasting elections with non-representative polls*, Int. J. Forecast., 31 (2015), pp. 980–991. (Cited on pp. 839, 856)

[86] D. J. Watts, R. Muhamad, D. C. Medina, and P. S. Dodds, *Multiscale, resurgent epidemics in a hierarchical metapopulation model*, Proc. Natl. Acad. Sci. USA, 102 (2005), pp. 11157–11162. (Cited on pp. 847, 857)

[87] C. Wlezien and R. S. Erikson, *The timeline of presidential election campaigns*, J. Politics, 64 (2002), pp. 969–993. (Cited on pp. 839, 855, 857)

[88] J. Zittrain, *Engineering an election. Digital gerrymandering poses a threat to democracy*, Harvard Law Rev. Forum, 8 (2014), pp. 335–341. (Cited on p. 857)