Estimating the Probability of Events That Have Never Occurred: When Is Your Vote Decisive?

Andrew GELMAN, Gary KING, and W. John BOSCARDIN

Researchers sometimes argue that statisticians have little to contribute when few realizations of the process being estimated are observed. We show that this argument is incorrect even in the extreme situation of estimating the probabilities of events so rare that they have never occurred. We show how statistical forecasting models allow us to use empirical data to improve inferences about the probabilities of these events. Our application is estimating the probability that your vote will be decisive in a U.S. presidential election, a problem that has been studied by political scientists for more than two decades. The exact value of this probability is of only minor interest, but the number has important implications for understanding the optimal allocation of campaign resources, whether states and voter groups receive their fair share of attention from prospective presidents, and how formal "rational choice" models of voter behavior might be able to explain why people vote at all. We show how the probability of a decisive vote can be estimated empirically from state-level forecasts of the presidential election and illustrate with the example of 1992. Based on generalizations of standard political science forecasting models, we estimate the (prospective) probability of a single vote being decisive as about 1 in 10 million for close national elections such as 1992, varying by about a factor of 10 among states. Our results support the argument that subjective probabilities of many types are best obtained through empirically based statistical prediction models rather than solely through mathematical reasoning. We discuss the implications of our findings for the types of decision analyses used in public choice studies.

KEY WORDS: Conditional probability; Decision analysis; Election; Electoral campaigning; Forecasting; Political science; Presidential election; Rare event; Rational choice; Subjective probability; Voting power.

1. INTRODUCTION

When an event is so rare that it has never occurred, despite many trials, estimates of its probability would seem to be a theoretical matter about which statisticians have little to contribute. Through a political science election analysis example, we demonstrate that statistical models can be used to extract information from related data to yield better estimates of the probabilities of even extraordinarily rare events.

Our application is a more extreme example of decision analyses that require the assessment of subjective probabilities. Our results are related to examples from space shuttle safety (Martz and Zimmer 1992), record linkage (Belin and Rubin 1995), and DNA matching (Belin, Gjertson, and Hu 1995; Sudbury, Marinopoulos, and Gunn 1993), where scholars have also found that probabilities estimated using data-based statistical methods are much better calibrated than probabilities assigned by theoretical mathematical models. These fields are like election analysis in that data-free models, typically based on independence assumptions, have led to mistaken conclusions. Our work is also related to analyses that seek to improve estimation by supplementing datasets that have a small (but nonzero) number of rare events with precursor data (Bier 1993) or additional carefully selected observations (Sanchez and Higle 1992).

We estimate the probability that an individual's vote is decisive in U.S. presidential elections. Given the size of the electorate, an election where one vote is decisive (equivalent to a tie in your state and in the electoral college) will almost certainly never occur. Nevertheless, political scientists have sought an estimate of this event through systematic theoretical analyses for over two decades and through informed speculation for much longer (Beck 1975). Although the exact value of the probability of one vote being decisive is a minor issue in and of itself, it turns out to lie at the heart of several important lines of inquiry.

For one, the perception of the probability that a single vote, or block of votes, will be decisive, governs the optimal allocation of campaign resources. Understanding political campaigns and the behavior of political candidates thus involves estimating the probability that a vote will be decisive in each state or region. Candidates for office are also obviously interested in these estimates to maximize their chances of winning. States and various voter groups trying to ensure that they get a fair share of attention from prospective office holders are also interested, because attention during the campaign relates to how they will be treated by the occupant of the White House after the election.

Normatively, many political scientists find electoral systems undesirable if some voter groups are more likely than others to influence an election outcome. For example, the variation from state to state in the probability that your vote is decisive in a U.S. presidential election is often addressed in terms of whether the electoral college favors voters in large or small states (Banzhaf 1968; Brams and Davis 1974; Merrill 1978) or whether the electoral college as a whole treats the political parties equally (Abbott and Levine 1991). (The winner of the presidential election is the candidate who receives a majority of votes from the 538 electoral college delegates. The plurality winner of each state chooses all the electoral college delegates assigned to that state. The

number of delegates is determined by the number of senators plus the number of congressional representatives from each state. We ignore minor exceptions such as Maine's rule that allows their electoral college delegates to be split, delegates deciding to vote for candidates other than for whom they were chosen, and complications caused by third-party candidates who get large fractions of the vote.) From the standpoint of normative philosophy, an election in a democratic system should allow the *possibility* that a single vote can matter; some believe it is desirable to design electoral systems so that the probability of this outcome is relatively high.

Finally, from a "rational choice" perspective, voting becomes more desirable as the probability of it making a difference increases. Different researchers give different numbers for the probability that one vote is decisive, but all agree it is low enough so that considering only the immediate personal costs of voting and direct personal gains from influencing the outcome, it is "irrational" for most Americans to vote in the presidential election (Aldrich 1993; Barzel and Silberberg 1973; Ferejohn and Fiorina 1974; Green and Shapiro 1993; Jackman 1987; Riker and Ordeshook 1968). Scholars in this subfield have at times been consumed with trying to explain, through mathematical models of voter choice, why "so many" people bother to vote (given rational choice assumptions). Thus estimates of the probability of tied elections play a role in understanding or resolving this central puzzle. (Many reasons for voting other than direct personal gain from the election outcome can also be given, but scholars in this subfield have sought to build mathematical models that explain why people vote with only minimal modifications of their parsimonious behavioral assumptions.)

In this article we use a standard model for forecasting presidential elections to estimate the probability that a single vote will be decisive, for voters in each state in every postwar election. We also perform some more approximate calculations to estimate the average probability that one vote will be decisive in a U.S. congressional election and in elections in general. Sections 2 and 3 lay out the theoretical framework for estimating the probability of a decisive vote in the electoral college. Section 4 gives numerical details of our implementation with historical data. Section 5 discusses other elections, and Section 6 discusses the implications of our methods for studies of voting in particular and decision theory in general.

2. THE FACTORS THAT DETERMINE THE PROBABILITY THAT YOUR VOTE IS DECISIVE

2.1 Interpretation in Terms of Forecasts

The question, "What is the probability that your vote will be decisive?" is inherently about uncertainty in the outcome of the election, given the information available to you *before* the election. Thus, to answer this question in even an approximately calibrated way, one must model the uncertainty in the pre-election period. The uncertainty could be measured in many ways, depending on what information is available at the time of the forecast. For this article we

use presidential election forecasting methods based on national and state economic and political variables available a few months before the election, following Rosenstone (1984) and Campbell (1992). As discussed by Gelman and King (1993), these forecasts predict the election about as accurately as polls taken a few days before the election. In fact, no method—including the predictions of informed observers, political insiders, media pundits, sample surveys, or other types of expert analysis or nonstatistical predictions—has been shown to outperform these forecasts. As such, although our model is conditional on the information publically available prior to the general election campaign that we included in our model, this is nearly equivalent to conditioning on all the information available that an individual voter would have just prior to election day. (The puzzling implication, that the campaign has little net effect despite huge sums spent and wide fluctuations in voter preference polls over time, was studied by Gelman and King [1993].)

In any case, the following factors will almost necessarily be involved in the probability that your vote will be decisive in a presidential election. First, the probability that your state election is tied depends on (a) the forecast vote share for the two candidates, (b) the uncertainty in that forecast (a Democratic vote share forecast at $.51 \pm .02$ is more likely to be a tie than one forecast at $.51 \pm .10$), and (c) the number of voters in your state (to yield the probability of an exact tie). Next, the probability that your state, if tied, will be decisive in the national electoral vote total depends on (d) the number of electoral voters assigned to your state and (e) the approximate proximity of the state to the national median vote.

2.2 Comparison to Theoretical Models of Voting

Most of the literature to date on the effects of individual votes has focused on formal probabilistic models of voting, generally based on a model of binomial (Beck 1975; Margolis 1977) or at best compound binomial variation (Chamberlain and Rothchild 1981) of the votes within each state. The probabilities produced by such models do not correspond, even approximately, to the state of uncertainty of participants in the political process before the election. In particular, the formal mathematical models typically assume that the probability of a vote being decisive depends only on the electoral vote and turnout (or, worse, population) in each state—thus ignoring factors (a) and (e)—and further assume that the variability (b) is determined by binomial variation (an assumption not warranted by the data, as we discuss at the end of Section 4.1 and in Section 6.2). Merrill (1978) allowed the probabilities to vary by state, but did not allow the parameters for states to vary over time or with the closeness of the election.

Other work in this field, based on game-theoretic ideas [such as that of Banzhaf (1968) and Brams and Davis (1974)], avoids explicit probability models but can be seen implicitly to assume that votes are assigned at random. (For example, a voting power measure based on counting the number of winning coalitions for which your vote would be decisive is equivalent to determine the probability that

your vote will be decisive, under the assumption that all the other actors vote by flipping coins.) More recent game-theoretic analyses (such as that of Feddersen 1992), which allow votes to depend on additional information, also implicitly assume that the variability (b) approaches 0 if the number of voters is large.

3. USING ELECTION FORECASTS TO COMPUTE THE PROBABILITY THAT A VOTE WILL BE DECISIVE

3.1 Expression in Terms of Conditional and Marginal Probability

Your presidential vote will be decisive if two conditions are satisfied. First, without your vote, your state's election outcome must be exactly tied or one vote away from a tie. (We consider the case of a person who will either vote for one candidate not vote. A voter who is considering switching from the Democratic to Republican candidate will, of course, have a higher probability of being decisive.) Second, your state must be decisive in the national election; given that it is tied, neither party must have an electoral vote majority. We introduce the following mathematical notation for the following known constants:

$$e_i = \text{number of electoral votes assigned to state } i$$

and

$$e_{\text{total}} = \sum_i e_i = 538 = \text{total number of electoral votes;}$$

and the following election outcomes that need to be modeled:

n_i = number of voters in state i (excluding yourself),
v_i = Democratic share of the two-party vote in state i (excluding your vote),
$V_i = \begin{cases} 1 & \text{if } v_i > .5 \\ 0 & \text{otherwise,} \end{cases}$

and

$E_{-i} = \sum_{j \neq i} e_j V_j + 3 =$ Democratic electoral vote in the 49 states excluding i, plus the District of Columbia.

Because the District of Columbia is an unambiguous outlier, and easily predictable, we assume that its 3 electoral votes are certain to go to the Democrats. This is not a controversial coding decision (see, e.g., Rosenstone 1983). One can crudely account for minor parties by separately estimating which states will be won by minor parties and setting E to the total number of electoral votes in the states contested by the Democrats and the Republicans.

Then, if you live in state i,

$$\Pr(\text{your vote matters})$$
$$= \Pr(\text{your vote is decisive in your state})$$
$$\times \Pr(\text{your state will be decisive}$$
$$|\text{your vote is decisive in your state}). \quad (1)$$

The second factor on the right side of (1) is a conditional probability: the probability that state i will be decisive, given a popular vote tie in that state.

We now describe how to evaluate (1) given any state-by-state forecast of the presidential election (i.e., the values n_i and v_i for all 50 states). In practice, n_i can be fairly accurately and uncontroversially estimated from previous elections, even ignoring the slightly higher turnout that many accompany closer contests. Because electoral votes also are known, we require a forecast of the vector of vote shares, $(v_1, \ldots, v_{50})$, representing some state of knowledge before the presidential election in question. Such a forecast is an input to our method and, like all statistical forecasts, must include uncertainty as well as a point estimate. In addition, separate forecasts for all the states are not enough; it must be a joint forecast so that the conditional probability in (1) can be determined (e.g., to find the probability that Utah—a strongly Republican state—will be decisive in the unlikely event that it is tied). We compute the two factors of (1) in turn. First, given the large number of voters in any state, one can with negligible error model the Democratic vote shares, v_i, as continuous variables. If n_i is even, then the first factor in (1) is

$$\Pr(\text{your vote is decisive in your state})$$
$$= \Pr(n_i v_i = .5 n_i) = \Pr(v_i = .5) \approx f_{v_i}(.5)/n_i, \quad (2)$$

using the discrete approximation to the continuous distribution and the notation f_{v_i} for the probability density function for the continuous variable v_i under the forecasting model. Similarly, if n_i is odd, then the first factor in (1) is $\Pr(n_i v_i = .5(n_i - 1)) = \Pr(v_i = .5 - .5/n_i) \approx f_{v_i}(.5)/n_i$ also, assuming that n_i is reasonably large.

We can compute the second factor on the right side of (1) by using the forecasting model to determine the conditional forecast of the other 49 states, given the condition $v_i = .5$ (for all practical purposes, $v_i = .5$ and $v_i = .5 - .5/n_i$ are identical conditions):

$$\Pr(\text{your state is decisive}$$
$$|\text{your vote is decisive in your state})$$
$$= \Pr(E_{-i} \in (.5e_{\text{total}} - e_i, .5e_{\text{total}})|v_i = .5)$$
$$+ \frac{1}{2}\Pr(E_{-i} = .5e_{\text{total}} - e_i|v_i = .5)$$
$$+ \frac{1}{2}\Pr(E_{-1} = .5e_{\text{total}}|v_i = .5). \quad (3)$$

(The factors of 1/2 arise because a vote that causes the national election to be tied is only half as decisive as a vote that changes the election outcome.) The conditional forecasts of the other 49 states must then be combined into a forecast of the national electoral vote.

3.2 Computation

The two factors on the right side of (1), which are given by (2) and (3), can be computed using posterior simulations, using analytic expressions where necessary to avoid having to estimate very low probabilities directly by simulation. The first step is to estimate all the parameters in the model or, in a Bayesian context, obtain a large number of

simulations (e.g., 1,000) of the vector of model parameters. For each state i and each draw of the simulated parameter vector, one can then compute $f_{v_i}(.5)/n_i$ conditional on the model parameters. The probability (2) can then be estimated as the average of these 1,000 values.

The next step, computing (3) for each state, is more complicated, because it is conditional on the event $v_i = .5$. For each state i, one must re-estimate the model, conditional on $v_i = .5$ (i.e., including the event $v_i = .5$ as additional "data" when fitting the model), and obtain a new set of 1,000 simulations of the model parameters. For each of these 1,000 draws, one must then simulate a draw from the predictive distribution of the vector of outcomes v_j in the other 49 states, and from those compute the value of E_{-i}. Expression (3) can then be estimated for state i using the empirical probabilities from the 1,000 simulated values of E_{-i}.

When the probability (3) is very low, perhaps even less than 1/1,000 (for example, for a small state that is much more Republican than the national average), the estimate obtained above may be unacceptably variable. In this case, if one desires a less variable estimate without having to draw many more simulations, one can analytically approximate the distribution of E_{-i} and use that to compute (3). We have found in our simulations that a beta distribution on $(E_{-i} - 3)/(e_{\text{total}} - 3 - e_i)$, fit to the first two moments of the drawn E_{-i} values, works quite well, as we discuss at the end of Section 4.2. We chose this particular distribution because E_{-i} is restricted to the range $[3, e_{\text{total}} - e_i]$.

4. RESULTS UNDER A PARTICULAR FORECASTING MODEL

4.1 The Forecasting Model

For this article we use a method of forecasting presidential elections based on a hierarchical linear regression model described by Boscardin and Gelman (1996), which adds a heteroscedastic specification to the model developed by Gelman and King (1993). These models are generalizations of standard methods in political science for forecasting based on past election results, economic data, poll results, and other political information. (See Campbell 1992 and Gelman and King 1993, Sec. 1, for more references and discussion of the political context. Similar models have been effective for forecasting election results in other countries as well; see, e.g., Bernardo 1986 and Bernardo and Giron 1992.) We estimate probabilities for the 1992 election, based on a forecast using information available before November 1992.

The model has the form

$$v_{it} = (X\beta)_{it} + \delta_t + \gamma_{r_i t} + \varepsilon_{it}, \qquad (4)$$

where i indexes states, t indexes election years, δ_t is a national error term, $\gamma_{r_i t}$ are independent regional errors ($r_i = 1, 2, 3$, or 4, depending on whether state i is in the Northeast, Midwest, West, or South), and ε_{it} are independent state-level errors.

The term $X\beta$ is a regression predictor, based on national, state, and regional variables all measured before the elec-

tion. The national variables—which are constant in each election year—are the Democratic candidate's share of the trial heat polls 2 months before the election; incumbency (0, 1, or -1, depending on the party); the President's approval rating, included as an interaction with the national presidential incumbency variable; and the change in gross national product (GNP) in the preceding year (counted positively or negatively, depending on whether the Democrats or the Republicans are the incumbent party). The statewide variables are the state's vote in the last two presidential elections (relative to the nationwide vote in each case), a presidential and vice-presidential home-state advantage (0, 1, or -1), the change in the state's economic growth in the past year (counted positively or negatively depending on the incumbent party), the partisanship of the state (measured by the proportion of Democrats in the state legislature), the state's ideology (as measured based on the political ideologies of its congressional representatives in 1988), the absolute difference between state and candidate ideologies as used by Rosenstone (1984), and the percent of the state's population that was Catholic in the election (1960) in which one of the candidates was Catholic. We also included an indicator variable for the south in elections in which one of the candidates was a southerner.

Gelman and King (1993) discussed the choice of these predictor variables, estimated the additional uncertainty due to the choice of specification, and provided evidence about the fit of the model. Boscardin and Gelman (1996) provided further tests, including the regional random-effects terms and a heteroscedastic variance function. We omit the details of model choice and model fitting here, because the approach presented in this article is designed to apply to any probabilistic forecast.

The regression coefficients and the variances of the error terms are estimated using data from the 1948 through 1988 elections. All error terms are assumed normally distributed (an assumption not contradicted by our data—i.e., there was no noticeable skewness or outliers). The regional and national error terms are vital for our purposes because they can affect the second factor in equation (1). Improvements in the forecast—for example, by modeling correlations between states or across election years, or including information from other sources such as state opinion polls—could be incorporated by altering the covariance structure or adding more explanatory variables, without altering the essential form of the model.

In addition, our forecasting model allows for unequal variances in the error terms ε_{it}, a fairly minor point for forecasting but potentially crucial for estimating the state-to-state variation in the probability that your vote will be decisive. State-by-state presidential election forecasting methods in the political science literature (e.g., Campbell 1992 and Rosenstone 1984) generally assume equal variances. In contrast, the theoretical models of voting generally assume binomial variation; that is, $\text{var}(\varepsilon_{it})$ proportional to n_{it}^{-1}. To include both possibilities in the same model, we fit a model in which $\text{var}(\varepsilon_{it}) = n_{it}^{-\theta}\sigma^2$, where θ is a parameter that varies between 0 (as in the forecasting models) and 1 (as in the theoretical models) and is estimated from the data.

For our regression model, we obtained a 95% interval for θ of [.09, .36], with estimated standard deviations of about 2.5%, 2.5%, 5.0%, and 3.5%, for the national, regional (outside the South), regional (the South), and state-level error terms.

4.2 The Probability That Your Vote Will Be Decisive

We estimate the parameters in our model using Bayesian simulation (Boscardin and Gelman 1996); our estimates yield a matrix of 1,250 simulations of the parameter vector $(\beta, \sigma^2, \theta, \gamma_1, \ldots, \gamma_4, \delta)$ from their posterior distribution. (We also obtain simulations of the parameters γ and δ for the election years 1948–1988 and their model variances, parameters not required for computing the probability that a single vote is decisive in 1992.) For convenience, we suppress the subscript t, as we are forecasting only one election at a time. For each simulated parameter vector, the probability that state i is tied or one vote less than tied—that is, expression (2)—is

if n_i is even, $\Pr(n_i v_i = .5 n_i | \beta, \gamma, \delta, \theta, \sigma^2)$

if n_i is odd, $\Pr(n_i v_i = .5(n_i - 1) | \beta, \gamma, \delta, \theta, \sigma^2)$

$$\approx \frac{1}{n_i} \, \mathrm{N}(.5 | (X\beta)_i + \gamma_{r_i} + \delta, n_i^{-\theta} \sigma^2), \quad (5)$$

where $\mathrm{N}(x | \mu, \tau^2)$ is the normal density function. Expression (5) depends on n_i, the voter turnout in state i, which is unknown before the election. To compute (5), we use an estimate of n_i obtained by scaling the turnout from the previous presidential election by the increase in the voting age population in the state in the previous 4 years. This correction is not precise, but errors in the turnout have simple and relatively minor effects on the estimated probabilities. For each state i, we then estimate (2) by averaging the probabilities (5) over the 1,250 simulations of the parameter vector; this is the correct estimated probability of a tie or near-tie given the Bayesian simulations.

For each state i, we now compute the conditional probability that the state is decisive, given that it is tied, in two steps. First, we assume the state is tied ($v_i = .5$) and use this as additional information in estimating the model parameters—most importantly, γ_{r_i} and δ. We condition on the information $v_i = .5$ by simply adding another row to the data matrix in the regression (4), corresponding to the "observation" $v_{it} = .5$, then repeating the Bayesian computations to produce 1,250 simulations of the vector of model parameters. For each of the simulations, we then simulate the outcomes v_j for the remaining 49 states using the forecast model: each v_j drawn from a normal distribution with mean $(X\beta)_j + \gamma_{r_j} + \delta$ and variance $n_j^{-\theta} \sigma^2$. We then compute E_{-i} for each simulation and use the results from the 1,250 simulations to estimate the factor (3) for each state. Figure 1 plots the estimate of (3) based on the empirical frequencies versus the estimate based on fitting beta distributions, as described at the end of Section 3. The estimates are quite similar, and so we use the estimates based on the beta approximation so that the estimates of the low probabilities will be more stable using this moderate number of simulation draws.
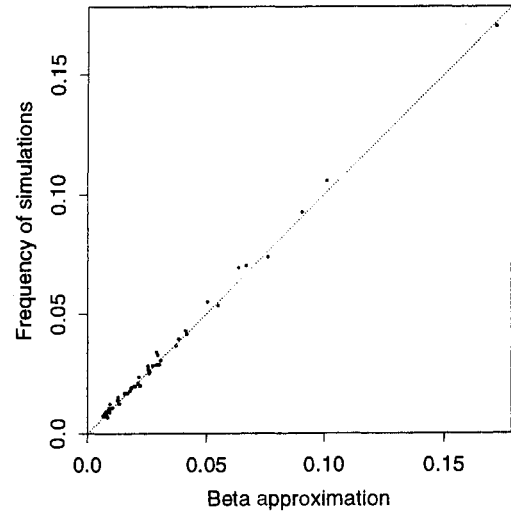

Figure 1. Estimated Probability That a State is Decisive Given Tied, Computed Based on Frequency of Simulations Versus an Estimate From a Fitted Beta Distribution. $\cdots$, equality of probabilities.

4.3 Numerical Results

We used these simulations to compute the probability that a single vote would decide the election in each state for each presidential election year from 1952 to 1992, excluding 1968, when a third-party candidate won several states.

Figure 2 displays for the 1992 election the probability that a single vote is decisive versus the number of electoral votes in each state. The probability is about 1 in 10 million for all states. Voters in some of the smaller politically moderate states have a greater chance (e.g., 1 in 3.5 million in Vermont), whereas those in more extreme states (e.g., Utah and Nebraska) have a lesser chance. If the vote in a politically extreme state is tied, then the probability of a close election at the national level is very low.

Figure 2b displays a summary of the results for the 1952–1988 elections. For six of the elections, the probability is fairly independent of state size (slightly higher for the smallest states) and is near 1 in 10 million. For the other three elections (1964, 1972, and 1984, corresponding to the landslide victories of Johnson, Nixon, and Reagan), the probability is much smaller, on the order of 1 in hundreds of millions for all of the states. This strong dependence of the estimated probability on the size of the victory margin invalidates most of the existing theoretical models. Of course, the probabilities of decisive votes in the landslide elections are sensitive to the tail behavior of our forecasting model; we trust the qualitative findings, but would rely less strongly on the exact numerical results.

For comparison, we estimate the chance that a single vote would be decisive if the popular vote decided the election. The posterior predictive distribution for popular vote in 1992 is easily estimated by the simulations; it is roughly normal with mean 51.5% and standard deviation 5.6%. With about 92 million people predicted to vote, the chance that it would have been an exact tie is approximately 1 in 13.3 million. The electoral college system places a slightly greater importance on the individual votes from all but eight of the states in 1992.
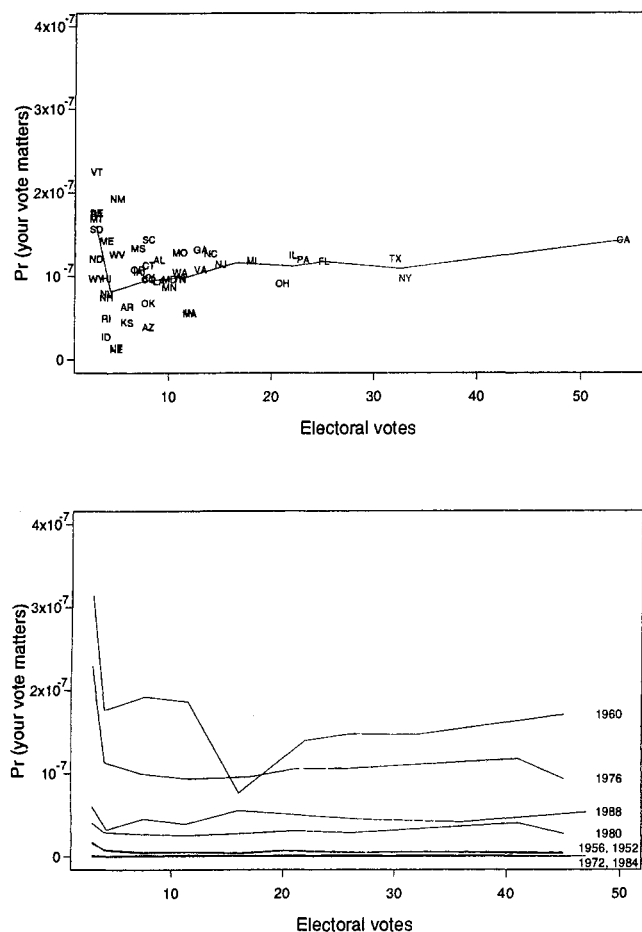
Figure 2. Probability That One Vote Decides the Election, by State, Versus Electoral Votes in the State for (a) 1992 and (b) 1952–1988 (Excluding 1968). In both figures the solid lines were obtained by binning according to electoral votes and then averaging.

5. APPROXIMATE RESULTS FOR U.S. CONGRESSIONAL AND OTHER ELECTIONS

As an external check on our model, we estimate the probability that any generic election is tied using equation (2). Suppose that n people vote in the election, and that the forecast is a normal distribution with mean μ and standard deviation τ; then the probability that a single vote will be decisive is approximately $(\sqrt{2\pi}\tau n)^{-1}\exp(-(\mu - .5)^2/(2\tau^2))$, as discussed by Margolis (1977). One way to interpret this result is in terms of upper bounds. The probability of a tie is clearly maximized at $\mu = .5$. As for σ, it is hard to imagine a real election that could be forecast to within a standard error of less than, say, 2% of the vote. This yields $20/n$ as an *upper bound* on the probability that your vote is decisive in a close election.

A typical value of n for an election to the U.S. Congress is 200,000, which gives an upper bound of 1 in 10,000 of your vote making a difference. Another way to look at this is that even in the closest elections, it is not in practice possible to forecast the outcome to within less than about 10,000 votes. Of course, most Congressional elections are not forecast to be so close, and so the probability of a tie is usually much lower.

Another way to attack the problem is empirically, by averaging over past election outcomes. In the period 1900–1992, there were 20,597 U.S. House elections, out of which 6 were decided by fewer than 10 votes, 49 by fewer than 100 votes, 293 by fewer than 500 votes, and 585 by fewer than 1,000 votes. This suggests a frequency probability of about .5/20,597 that a single vote will be decisive in a randomly chosen U.S. House election. This number is of course much less than our upper bound of 1/10,000, because most of the elections were not close.

For U.S. presidential elections, a similar rough calculation reveals that 18% of the state election results v_i in our dataset lay between .48 and .52. This suggests for a state with n_i voters an estimated probability of $.18/(.04n_i)$ for the event that v_i is exactly .5 if n_i is even or exactly $.5 - .5/n_i$ if n_i is odd. We can perform a similar calculation for the probability that a state is decisive in the electoral college; of the 11 presidential elections in 1948–1988, 2 were close enough that switching 50 electoral votes would decide the election. This suggests for a state with e_i electoral votes an estimated probability of about $1/2(2/11)(e_i/50)$ that a vote that a decisive in a state will swing the national election. (The factor of $1/2$ applies because we are considering the effect of casting a vote, not the effect of switching a preference from one party to the other.) Multiplying the two factors yields a combined probability of $.008e_i/n_i$ that an individual vote will be decisive. For example, a voter in a medium-sized state with 10 electoral votes and a turnout of 2 million would have an estimated probability of 1 in 25 million of casting a decisive vote. This number is consistent with our estimates based on the forecasting model averaging over all election years.

For the presidential elections, we present the foregoing approximate frequency calculations as a numerical check. For the substantive political analysis, we prefer the forecast-based estimates, because they condition on relevant information about the closeness of the election, the voting pattern in each state, and so forth, as discussed in Section 2.1.

6. CONCLUDING REMARKS

6.1 Implication for the Study of the Electoral College and Voting in General

Like all other researchers, we estimate the (prospective) probability that a single vote will affect the outcome of the U.S. presidential election to be very low, typically of order of magnitude 1 in 10 million, rising to as much as about 1 in 1.5 million for some small states in some close elections (e.g., Nevada in 1960 and Alaska in 1976) and less than 1 in 100 million for all states in landslide elections such as 1972.

Contrary to Banzhaf (1968) and Brams and Davis (1973, 1974), we do not find a "bias" in favor of large states. The largest biases are in favor of most of the small states (because all states receive a minimum of three electoral votes no matter how small their population) and against voters in states such as Utah, and in the District of Columbia, who have virtually no change of deciding the presidential elec-

tion, because of their atypical voting behavior, not the size of their states.

Our results and general methodology are of obvious interest to candidates deciding how to allocate their campaign resources and states concerned about attracting the attention of prospective presidents. In general, the probability of influencing the election outcome by mobilizing N supporters to vote in a single state is roughly N times the probability that a single vote in that state will be decisive, and so state-by-state campaign efforts can be chosen to maximize that probability, with the optimal decision varying as the campaign progresses and the election forecasts change. This point has been discussed by Brams and Davis (1973, 1974). Similarly, the probability of swinging the election by changing the preferences of N voters in a single state is roughly $2N$ times the probability that a single vote in that state will be decisive.

In addition, our results are of interest to rational choice theorists interested in the rationality of the decision of the individual citizen whether to vote; of course, one must also account for the possibility that the voter may influence other, nonpresidential contests at the ballot box.

6.2 Mathematical Discussion of our Results and Comparison to Methods Not Based on Forecasts

The probability of a tie in a state is on the order of $1/n_i$, and the probability that a state will be decisive given that a tie occurs is (crudely) proportional to e_i, which is roughly proportional to n_i (except in the smallest states). Therefore, we expect the product of these two factors to be approximately constant, with a slight advantage to the smallest states. To illustrate, Figure 3 plots for 1992 the log-probability that a state will be decisive given that it is tied versus the log-probability that it will be tied. Most of the points lie close to the dotted line indicating a probability of 10^{-7}.
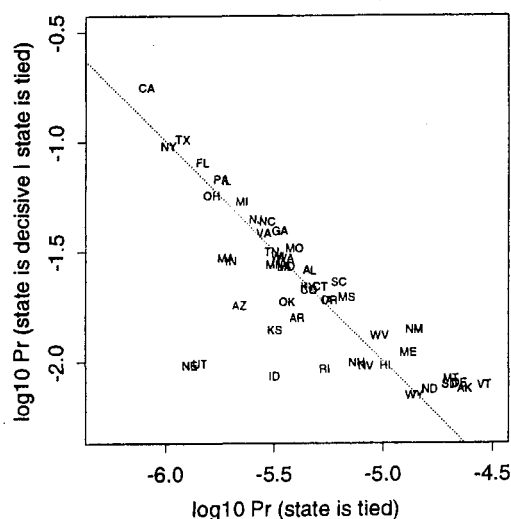

Figure 3. Probability That a State Is Decisive Given Tied Versus the Probability That the State Is Tied for 1992 Plotted on a Log Scale. ⋯, product of 10^{-7}.
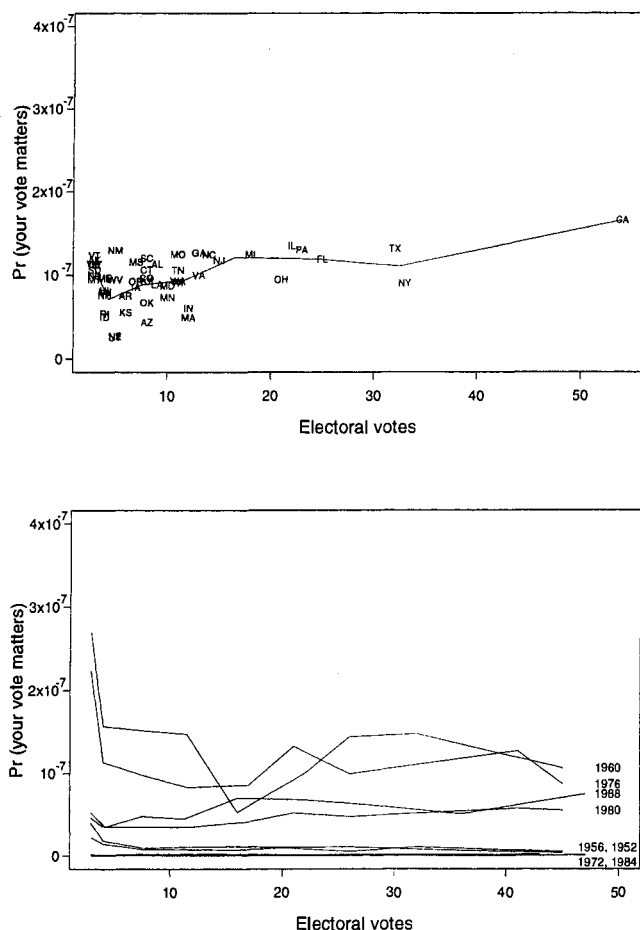

Figure 4. 'Same Plot as in Figure 1 for the Model With θ Set to 1 (i.e., State-Level Variance Inversely Proportional to Turnout).

Many of the theoretical models in the literature (see Sec. 1) assume that the standard deviation of v_i in a state is proportional to $1/\sqrt{n_i}$. Our model can replicate this assumption by fixing the value of θ to be exactly 1; see expression (5). We performed this computation to investigate whether our findings would change measurably with such an assumption. Figure 4 shows the results for 1992 and previous years: the probability that a single vote will be decisive increases slightly for the very largest states, but only slightly and not to the extent anticipated by the binomial-based models. This is because the forecasting model has several variance components, and the regional and national errors do not, of course, vary by state size. Our results are not as sensitive to the parameter θ as one might fear. Future analysts thus may wish to opt for the simpler homoscedastic regression-based forecasts of Gelman and King (1993).

Another possible modeling choice is the compound binomial: modeling an expected vote outcome u_i using a linear model as is done in this article and then modeling votes by a binomial distribution, $n_i v_i \sim \text{Bin}(n_i, u_i)$. Although this class of models seems reasonable, we do not adopt it because in practice, the turnout in U.S. elections is so large that the binomial variability is negligible compared to the forecast uncertainty in the model. For example, in 1992, turnout in all states was greater than 160,000, and $\sqrt{(.5)(.5)/160,000} = .00125$, as compared to statewide

error terms of about .03. Boscardin and Gelman (1996) also considered a generalization of the compound binomial model, fitting an error variance of the form $(\sigma_1^2 + \sigma_2^2/n_i)$. Results were very similar to those obtained from the power-law variance model shown here.

To return to more substantive concerns, we consider how the results would change as better information is added so as to increase the accuracy of the forecasts. In most states this will have the effect of reducing the chance of an exact tie; that is, adding information will bring the probability that one vote will be decisive even closer to 0. However, for a state that is close to evenly divided, the resulting probability will continue to increase as more information is added. In reality, one cannot achieve arbitrary precision in the forecasts. Even for the most knowledgeable observers on the morning of election day, there is quite a bit of uncertainty in the day's outcome.

6.3 Empirical Forecasting Versus Mathematical Modeling—Implications for Decision Theory and Public Choice

The probability of an unlikely event, such as an individual's vote being decisive in a nationwide election, can be estimated in a straightforward fashion as a byproduct of any forecasting system that includes forecasting uncertainty. The results are model dependent, but the use of forecasting models is a strength, because the models can be checked for accuracy and improved if they do not forecast well. For the case of presidential elections, we use extensions of standard forecasting methods to determine the probability of a vote being decisive for each state and find results that make good political sense, but contradict many published findings in this field that are based on mathematical models not fit to actual elections.

An alternative approach would be to attempt to assess subjective probabilities directly. For example, one could poll individual voters to determine their perceived probabilities that the election will be a tie. However, people are notoriously poor at assessing probabilities that are close to 0 (see Kahneman, Slovic, and Tversky 1982). If interested in the effect on campaign decisions, one could interview campaign organizations to determine their internal forecasts or use the prognostications of informed commentators, although political science forecasting models outperform even the most eloquent media pundits.

Decision theorists have long noted the need for estimating subjective probabilities for expected utility calculations (see, e.g., Savage 1954). This is difficult when the events in question are so rare that they have never been observed to occur, and especially difficult in nonexperimental research where collecting more data is either infeasible or impossible. Our application demonstrates the utility of bringing related information to bear on improving the estimation of the probability of rare events. This is a useful addition to the tendency, at least in political science, to obtain probabilities through formal models with only minimal empirical input. The example of the probability of decisive voting illustrates

the conceptual and substantive gains that can be made by returning to a forecasting basis for modeling uncertainties in decision making.

REFERENCES

Aldrich, J. H. (1993), "Rational Choice and Turnout," *American Journal of Political Science*, 37, 246–278.

Abbott, D. W., and Levine, J. P. (1991), *Wrong Winner: The Coming Debacle in the Electoral College*, New York: Praeger.

Banzhaf, J. P. (1968), "One Man, 3.312 Votes: A Mathematical Analysis of the Electoral College," *Villanova Law Review*, 13, 304–332.

Barzel, Y., and Silberberg, E. (1973), "Is the Act of Voting Rational?," *Public Choice*, 16, 51–58.

Beck, N. (1975), "A Note on the Probability of a Tied Election," *Public Choice*, 23, 75–79.

Belin, T. R., Gjertson, D. W., and Hu, M. Y. (1995), "Summarizing DNA Evidence When Relatives Are Possible Suspects," Technical Report, University of California Los Angeles, Dept. of Biostatistics.

Belin, T. R., and Rubin, D. B. (1995), "A Method for Calibrating False-Match Rates in Record Linkage," *Journal of the American Statistical Association*, 90, 694–707.

Bernardo, J. M. (1984), "Monitoring the 1982 Spanish Socialist Victory: A Bayesian Analysis," *Journal of the American Statistical Association*, 79, 510–515.

Bernardo, J. M., and Giron, F. J. (1992), "Robust Sequential Prediction From Non-Random Samples: The Election-Night Forecasting Case" (with discussion), in *Bayesian Statistics 4*, eds. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, New York: Oxford University Press, pp. 61–77.

Bier, V. M. (1993), "Statistical Methods for the Use of Accident Precursor Data in Estimating the Frequency of Rare Events," *Reliability Engineering and System Safety*, 41, 267–280.

Boscardin, W. J., and Gelman, A. (1996), "Bayesian Regression With Parametric Models for Heteroscedasticity," *Advances in Econometrics*, 11A, 87–109.

Brams, S. J., and Davis, M. D. (1973), "Models of Resource Allocation in Presidential Campaigning: Implications for Democratic Representation," *Annals of the New York Academy of Sciences (Democratic Representation and Apportionment: Quantitative Methods, Measures, and Criteria)*, 219, 105–123.

—— (1974), "The 3/2's Rule in Presidential Campaigning," *American Political Science Review*, 68, 113–134.

Campbell, J. E. (1992), "Forecasting the Presidential Vote in the States," *American Journal of Political Science*, 36, 386–407.

Chamberlain, G., and Rothchild, M. (1981), "A Note on the Probability of Casting a Decisive Vote," *Journal of Economic Theory*, 25, 152–162.

Feddersen, T. (1992), "A Voting Model Implying Duverger's Law and Positive Turnout," *American Journal of Political Science*, 36, 938–962.

Ferejohn, J., and Fiorina, M. (1974), "The Paradox of Not Voting: A Decision Theoretic Analysis," *American Political Science Review*, 68, 525.

Gelman, A., and King, G. (1993), "Why are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?" *British Journal of Political Science*, 23, 409–451.

Jackman, R. W., (1987), "Political Institutions and Voter Turnout in the Industrial Democracies," *American Political Science Review*, 81, 405–423.

Kahneman, D., Slovic, P., and Tversky, A. (1982), *Judgment Under Uncertainty: Heuristics and Biases*, New York: Cambridge University Press.

Margolis, H. (1977), "Probability of a Tie Election," *Public Choice*, 31, 134–137.

Martz, H. F., and Zimmer, W. J. (1992), "The Risk of Catastrophic Failure of the Solid Rocket Boosters on the Space Shuttle," *The American Statistician*, 46, 42–47.

Merrill, S. (1978), "Citizen Voting Power Under the Electoral College:

A Stochastic Model Based on State Voting Patterns," *SIAM Journal of Applied Mathematics*, 34, 376–390.

Riker, W. H., and Ordeshook, P. C. (1968), "A Theory of the Calculus of Voting," *American Political Science Review*, 62, 25–42.

Rosenstone, S. J. (1983), *Forecasting Presidential Elections*, New Haven, CT: Yale University Press.

Sanchez, S. M., and Higle, J. L. (1992), "Observational Studies of Rare Events: A Subset Selection Approach," *Journal of the American Statistical Association*, 87, 878–883.

Savage, L. J. (1954), *The Foundations of Statistics*, New York: Wiley.

Sudbury, A. W., Marinopoulos, J., and Gunn, P. (1993), "Assessing the Evidential Value of DNA Profiles Matching Without Using the Assumption of Independent Loci," *Journal of the Forensic Science Society*, 33, 73–82.