

# WHAT IS THE PROBABILITY YOUR VOTE WILL MAKE A DIFFERENCE?

ANDREW GELMAN, NATE SILVER and AARON EDLIN\*

*One of the motivations for voting is that one vote can make a difference. In a presidential election, the probability that your vote is decisive is equal to the probability that your state is necessary for an electoral college win, times the probability the vote in your state is tied in that event. We computed these probabilities a week before the 2008 presidential election, using state-by-state election forecasts based on the latest polls. The states where a single vote was most likely to matter are New Mexico, Virginia, New Hampshire, and Colorado, where your vote had an approximate 1 in 10 million chance of determining the national election outcome. On average, a voter in America had a 1 in 60 million chance of being decisive in the presidential election. (JEL H0)*

## I. INTRODUCTION

What is the probability that one vote will make a difference? This is of interest in the utility theory of voting and also for campaign strategists who must assess the costs and benefits of attempting to persuade or turn out voters in particular states.<sup>1</sup> In a presidential election, the probability that your vote is decisive is equal to the probability that your state is necessary for an electoral college win, times the probability the vote in your state is tied, conditional on your state being necessary.

Broadly speaking, there are three ways of estimating the probability of a decisive vote. The first approach, and the oldest in the literature,<sup>2</sup> is to set up a theoretical model for the joint distribution of the popular vote margin in the separate

jurisdictions (in this case, the 50 states plus the District of Columbia) and then analytically compute the probability of each state being nationally decisive and locally tied. Such models offer important insights but are not directly useful for presidential elections, given that our interest is in the particularities of individual states in 2008 rather than on general properties of an electoral system.

The second approach is purely empirical, an estimate of the form  $y/n$ , where  $n$  is the number of elections in some large historical sample and  $y$  is the number that were tied. Elections are very rarely tied, so one can instead use an estimate such as  $y/(2,000 n)$ , where  $y$  is now the number of elections decided by less than 1,000 votes. Such an empirical analysis can be embedded into a regression of vote differential on election-level predictors. Mulligan and Hunter (2003) use such a model for estimating the average probability of a decisive vote in legislative elections, and Gelman, Katz, and Bafumi (2004) use the distribution of aggregate vote margins to study the dependence of the probability of a tie on the number of voters.<sup>3</sup> Unfortunately, the complex structure and the small number of

\*We thank an anonymous reviewer for helpful comments and the National Science Foundation for financial support.

*Gelman:* Professor, Department of Statistics and Department of Political Science, Columbia University, New York, NY. E-mail [gelman@stat.columbia.edu](mailto:gelman@stat.columbia.edu), <http://www.stat.columbia.edu/~gelman>

*Silver:* Baseball Prospectus, <http://fivethirtyeight.com>. E-mail [538dotcom@gmail.com](mailto:538dotcom@gmail.com)

*Edlin:* Professor, Department of Economics and School of Law, University of California, Berkeley, CA. E-mail [edlin@econ.berkeley.edu](mailto:edlin@econ.berkeley.edu)

1. There is a large literature on utility models for voting; see, for example, Ferejohn and Fiorina (1974) and, for our own view, Edlin, Gelman, and Kaplan (2007). See Stromberg (2008) for a recent analysis of the relation between electoral strategy and the probability of a decisive vote.

2. Examples include Good and Mayer (1975) and Chamberlain and Rothchild (1981); see Gelman, Katz, and Bafumi (2004) for a review of such methods and their relation to computing the empirical probability of decisiveness.

3. An objection sometimes arises about this sort of calculation that one vote never makes a difference, because if the election were decided by one vote, there would be a recount anyway. Gelman, Katz, and Bafumi (2004, p. 674) discuss why this argument is wrong, even for real elections with disputed votes, recounts, and so forth. This can be shown by setting up a more elaborate model that allows for a gray area in vote counting and then demonstrating that the simpler model of decisive votes is a reasonable approximation.

presidential elections rule out a purely empirical approach in this case.

The third method, which we use here, is to set up a forecasting model for the current election. The uncertainties for the 50 statewide elections in such a model should be correlated—or, equivalently, the model should have a national error term, and possibly regional error terms, in addition to any independent state errors. Given a forecast that is a joint probability distribution, one can calculate the probability of decisive vote using simulation or, in a setting such as a national election in which the probability of a tied election is tiny, one can use a mix of simulation and analytic calculations as was done by Gelman, King, and Boscardin (1998).<sup>4</sup>

## II. METHOD

We use a state-by-state election forecast computed on October 24, 2008, based on the latest opinion polls (Silver 2008). Earlier in the campaign, polls are not particularly informative, and it is better to use regression-based forecasts, at least to account for possible national swings (see Lock and Gelman 2008 and, for further discussion, Wlezien and Erikson 2004, and the references therein). By late October, however, polls are close to the actual election outcome and so it is reasonable to construct a forecast distribution using poll aggregation.<sup>5</sup>

4. Another approach, which we do not consider here, is to derive an implicit probability distribution for vote outcomes using the prices on election betting markets. This method might seem appealing at first but it falls apart in practice, at least based on the current state of the betting markets. For example, [www.intrade.com](http://www.intrade.com) assigned McCain a ridiculously high 3% chance of winning the District of Columbia. Prediction markets do have a role as information aggregators, especially for primary elections, where polls are highly unstable and models based on fundamentals do not do particularly well. But for the general election, incorporating information from prediction markets seems hardly worth the effort. Certainly, we can't take the market's implied probabilities at face values.

5. Based on the economy and other factors, political scientists forecasted that the Democrats will receive 51.8% of the two-party vote (Hibbs 2008) or maybe 52.2% (Erikson and Wlezien 2008). Such forecasts have standard errors large enough that the aforementioned researchers gave McCain a roughly 30% chance of winning, given the information available as of June or August. Were we doing our analysis using pre-October polls, we would construct our probability distribution for the election outcome using a weighted average of recent polls and model-based forecasts. (See also Campbell 2008, for further information on recent forecasting models.) After the election, we compared state-by-state forecasts to actual election outcomes and found a very close match, with a difference of less than 1 percentage point in 22 states and a difference of less than 3 percentage points in 39 of the 50 states (Gelman and Sides 2009).

The result of our forecasts is a  $10,000 \times 51$  matrix representing 10,000 computer simulations of the election, with each simulation including the vote margin for Obama (e.g., a margin of  $-0.208$  in Alabama corresponds to Obama and McCain receiving 60.4% and 39.6%, respectively, of the two-party vote in the state). We convert this percentage to a numerical vote margin by multiplying by the voter turnout in the state (which we estimate by taking the two-party turnout from 2004 and adding 4% corresponding to population increase and another 4% to account for an anticipated increase in turnout this year).<sup>6</sup>

As discussed earlier, the simulations for the states are correlated: the joint probability distribution of the 51 election outcomes includes uncertainty about the national swing as well as state-by-state fluctuations.

We use the forecasts to estimate the probability of a decisive vote in two steps. For each state:

1. Let  $E$  be the number of electoral votes in your state. We estimate *the probability that these are necessary for an electoral college win* by computing the proportion of the 10,000 simulations for which the electoral vote margin based on all the other states is less than  $E$ , plus  $\frac{1}{2}$  the proportion of simulations for which the margin based on all other states equals  $E$ . (This last part assumes implicitly that we have no idea who would win in the event of an electoral vote tie).<sup>7</sup>

2. We estimate *the probability that your vote is decisive, if your state's electoral votes are necessary*, by working with the subset of the 10,000 simulations for which the electoral vote margin based on all the other states is less than or equal to  $E$ . We compute the mean  $M$  and standard deviation  $S$  of the vote margin among that subset of simulations and then compute the probability of an exact tie as the density at 0

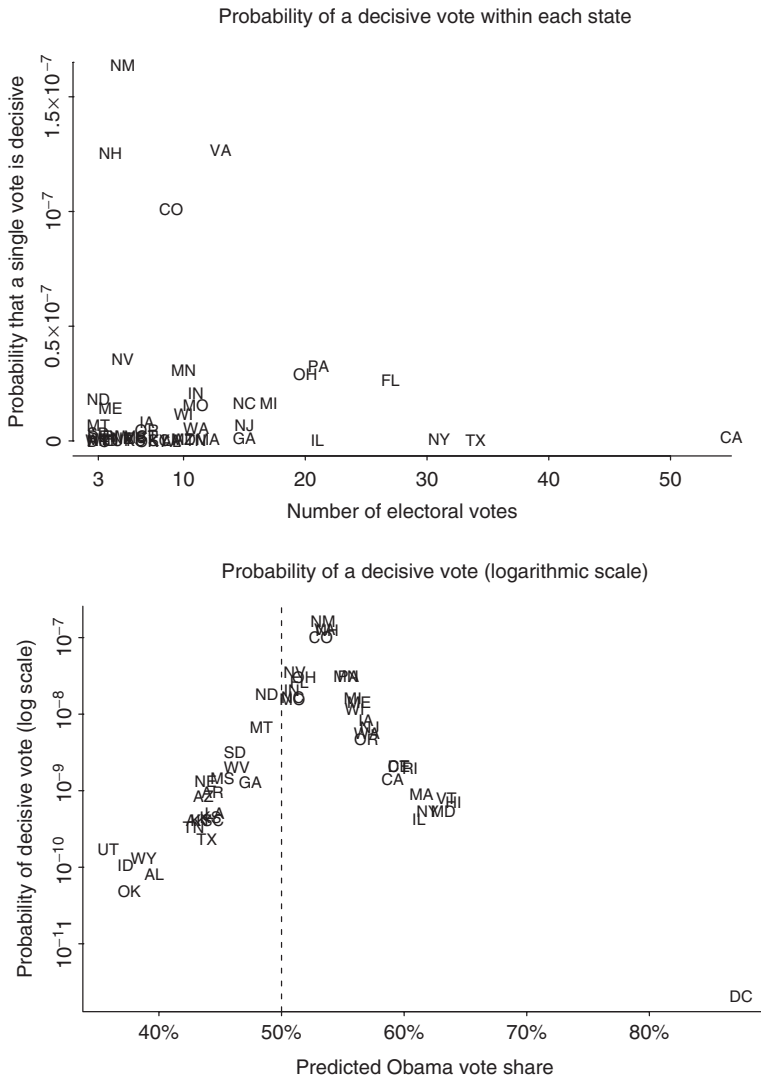
6. More sophisticated turnout models are possible, for example, using recent polls on likely voting. But such adjustments would make only tiny changes to the probability of tie elections. All that is really important here is that we put in a reasonable guess at the total number of votes in each state. In retrospect, there were some differences from 2004 to 2008; for example, Indiana saw an increase in turnout of nearly 5 percentage points, while Ohio's rate of voter turnout decreased slightly (McDonald 2009).

7. We ignored the splitting of Nebraska's and Maine's electoral votes, which retrospectively turned out to be a mistake in 2008, when Obama won an electoral vote from one of Nebraska's districts.



**FIGURE 2**

Probability of a Decisive Vote Plotted versus the Number of Electoral Votes and, on the Logarithmic Scale, versus Obama’s Predicted Vote in Each State



Note: The probability that your vote matters is highest in a set of small and midsize states that are near the national median politically.

in 60 million chance that a randomly selected voter would be decisive.

III. DISCUSSION

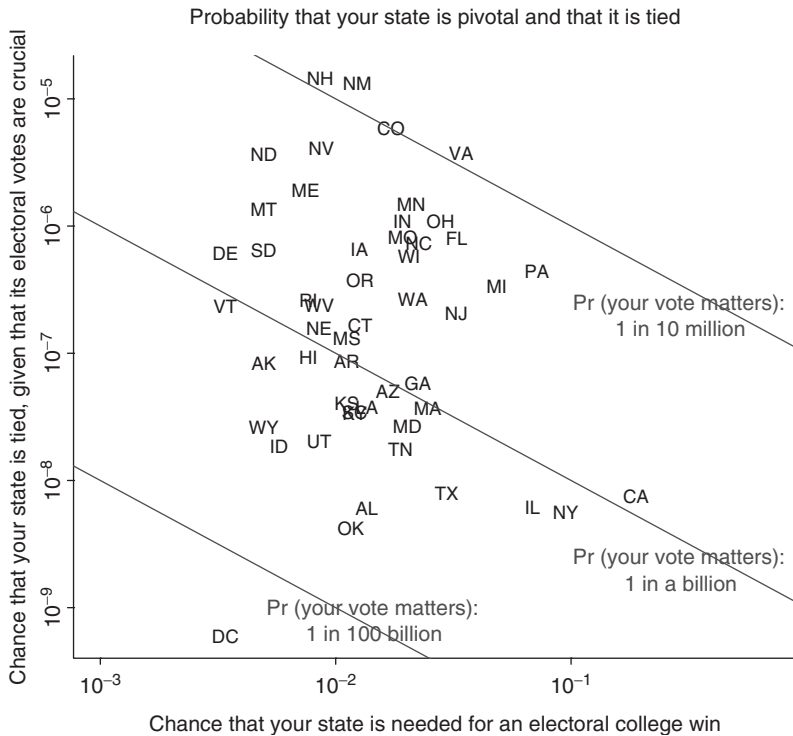
A probability of 1 in 10 million is tiny but, as discussed by Edlin, Gelman, and Kaplan (2007), can provide a rational reason for voting; in this perspective, a vote is like a lottery ticket with

a 1 in 10 million chance of winning, but the payoff is the chance to change national policy and improve (one hopes) the lives of hundreds of millions, compared to the alternative if the other candidate were to win.<sup>9</sup> Different voters

9. Hansen, Palfrey, and Rosenthal (1987) discuss the cost of voting, and Edlin, Gelman, and Kaplan (2008) estimate the dollar-equivalent benefits, considering voting as

**FIGURE 3**

Probability that Your Vote in Any Particular State is Decisive, Factored into Its Two Component Parts: the Probability of Your State’s Electoral Votes Being Necessary, and the Probability of Your State Being Tied, Given That Its Electoral Votes Are Necessary



Notes: The two probabilities multiply to give the chance that a single vote is pivotal. The lines show different levels of this probability, which is at most about 1 in 10 million. However, as discussed by Edlin, Gelman, and Kaplan (2007), it can be rational to vote even when your chance of being decisive is this low.

have different opinions about which candidate is desirable, but many can feel that their preference is not only better for them personally but for millions of others. (On the other hand, for voters in states such as New York, California, and Texas, where the probability of a decisive vote is closer to 1 in a billion, any reasons for voting must go beyond any instrumental rationality.)

We can also consider the incentives for campaigns and voter organizers. We estimate that turning out 1,000 additional voters for your side (or persuading 500 people to change their vote) in New Mexico would have a 1.3% chance of flipping the state in the event (with probability 1.2%) that its electoral votes are

decisive, which combine to a 1 in 6,000 chance of swinging the national election. The same effect could be attained at the national level by persuading 10,000 random supporters to vote or by persuading 5,000 random supporters of the other side to switch.

From a mathematical perspective, it is unsurprising that the states near the national median are those where a vote is most likely to make a difference, and this is consistent with campaigns’ focus on a few key swing states such as New Mexico, Virginia, New Hampshire, and Colorado. In 2008, Ohio, Florida, and Pennsylvania were in the second tier of priority.

**REFERENCES**

equivalent to a charitable contribution. Fowler (2006) gives experimental evidence relating altruism to voter turnout.

Campbell, J. E. “Forecasting the 2008 National Elections.” *PS: Political Science and Politics*, 41, 2008, 679–81.

- Chamberlain, G., and M. Rothchild. "A Note on the Probability of Casting a Decisive Vote." *Journal of Economic Theory*, 25, 1981, 152–62.
- Edlin, A., A. Gelman, and N. Kaplan. "Voting as a Rational Choice: Why and How People Vote to Improve the Well-Being of Others." *Rationality and Society*, 19, 2007, 293–314.
- Erikson, R. S., and C. Wlezien. "Leading Economic Indicators, the Polls, and the Presidential Vote." *PS: Political Science and Politics*, 41, 2008, 703–07.
- Fowler, J. H. "Altruism and Turnout." *Journal of Politics*, 68, 2006, 674–83.
- Gelman, A., G. King, and W. J. Boscardin. "Estimating the Probability of Events that Have Never Occurred: When is Your Vote Decisive?" *Journal of the American Statistical Association*, 93, 1998, 1–9.
- Gelman, A., J. N. Katz, and J. Bafumi. "Standard Voting Power Indexes Don't Work: An Empirical Analysis." *British Journal of Political Science*, 34, 2004, 657–74.
- Gelman, A., and J. Sides. "Election 2008: What Really Happened? And What Does it Mean?" Technical report, Department of Statistics, Columbia University, 2009.
- Good, I. J., and L. S. Mayer. "Estimating the Efficacy of a Vote." *Behavioral Science*, 20, 1975, 25–33.
- Hansen, S., T. R. Palfrey, and H. Rosenthal. "The Downsian Model of Electoral Participation: Formal Theory and Empirical Analysis of the Constituency Size Effect." *Public Choice*, 52, 1987, 15–33.
- Hibbs, D. A. "Implications of the 'Bread and Peace' Model for the 2008 US Presidential Election." *Public Choice*, 137, 2008, 1–10.
- Lange, K. L., R. J. A. Little, and J. M. G. Taylor. "Robust Statistical Modelling Using the t-Distribution." *Journal of the American Statistical Association*, 84, 1989, 881–96.
- Lock, K., and A. Gelman. "Bayesian Combination of State Polls and Election Forecasts." Technical report, Department of Statistics, Columbia University, 2008.
- McDonald, M. "United States Elections Project." 2009. Accessed 20 Feb 2009. <http://elections.gmu.edu/>
- Mulligan, C. B., and C. G. Hunter. "The Empirical Frequency of a Pivotal Vote." *Public Choice*, 116, 2003, 31–54.
- Silver, N. "Frequently Asked Questions." 2008. Accessed 7 Aug 2008. <http://www.fivethirtyeight.com/2008/03/frequently-asked-questions-last-revised.html>
- Stromberg, D. "How the Electoral College Influences Campaigns and Policy: the Probability of Being Florida." *American Economic Review*, 98, 2008, 769–807.
- Wlezien, C., and R. S. Erikson. "The Fundamentals, the Polls, and the Presidential Vote." *PS: Political Science and Politics*, 37, 2004, 747–51.