2-2010

Electoral Voting and Population Distribution in the United States

Paul Kvam
University of Richmond, pkvam@richmond.edu

Recommended Citation

Kvam, Paul H. "Electoral Voting and Population Distribution in the United States." *Chance: A Magazine for People Interested in the Analysis of Data* 23, no. 1 (February 2010): 41-47. doi:10.1007/s00144-010-0009-y.

Electoral Voting and Population Distribution in the United States

Paul Kvam

In the United States, the electoral system for determining the president is controversial and sometimes confusing to voters keeping track of election outcomes. Instead of directly counting votes to decide the winner of a presidential election, individual states send a representative number of electors to the Electoral College, and they are trusted to cast their collective vote for the candidate who won the popular vote in their state.

Forty-eight states and Washington, DC, employ the winner-takes-all method, each awarding its electors as a single bloc. Maine and Nebraska select one elector within each congressional district by popular vote and the remaining two electors by the aggregate, statewide popular vote. Due to this all-or-nothing outcome, the winner of the popular vote will not always be the candidate who wins the most electoral votes—such as in the 2000 election.

There are numerous critics of the current system who point out how its weaknesses are exacerbated by the naturally uneven spread of the country's population across the 50 states and that every state—no matter how sparsely populated—is guaranteed to have at least three electoral votes out of the total 538. Two of the votes correspond to two congressional senators that represent each state, independent of the state population.

Under the current rules, the value of a vote differs from state to state. A large state such as California has an immense effect on the national election, but, compared to a sparsely populated state such as Alaska, is grossly under-represented in the U.S. senate, where all senators have an equal vote. Arnold Barnett and Edward Kaplan, in their 2007 *CHANCE* article, "A Cure for the Electoral College?" called the Electoral College "the fun-house mirror of American politics" and suggested a weighted voting system that would mitigate the problem caused by the present winner-take-all rule.

An Examination of the Impact of California

Figure 1 illustrates the allotment of electoral votes from 2000–2008. California has 55 out of the total 538. That's already more than the 11 Vice President Walter Mondale achieved in the 1980 election against President Ronald Reagan, more than the 49 President Jimmy Carter achieved against Reagan in 1980, more than the 17 Sen. George McGovern achieved against President Richard Nixon in 1972, and more than the 52 Sen. Barry Goldwater achieved against President Lyndon Johnson in 1964. Given that 270 votes are needed to secure the majority, the candidate who wins the most votes in California, even if by a slim majority, garners more than 20% of the electoral votes needed to win the national election.

The 538 electoral votes in the 2008 presidential election were distributed among 50 states and the District of Columbia (DC). To simplify the language, we will treat DC as a state, because it has a minimum allotment of three electoral votes. So, we will refer to 51 states in this scenario. Consider a hypothetical situation within states in which candidates have an equal chance of winning the popular vote (and thus all of a state's electoral votes) and state results are decidedly independent of one another. As mentioned previously, we are examining the impact of disparities, but not modeling the actual U.S. political landscape.

Suppose candidate A has won the popular vote in California over candidate B. If candidate A wins m out of 50 of the remaining states, then the electoral votes of candidate A will be the sum of the electoral votes of those m states plus the votes in California. Excluding California, the average number of votes per state is 9.66, and the standard deviation is 7.25 votes. The mean number of votes out of the remaining for candidate A can be calculated as the mean of the conditional mean. That is, if candidate A wins M states, the average number of votes is $9.66M$. If M—the number of states—has a binomial

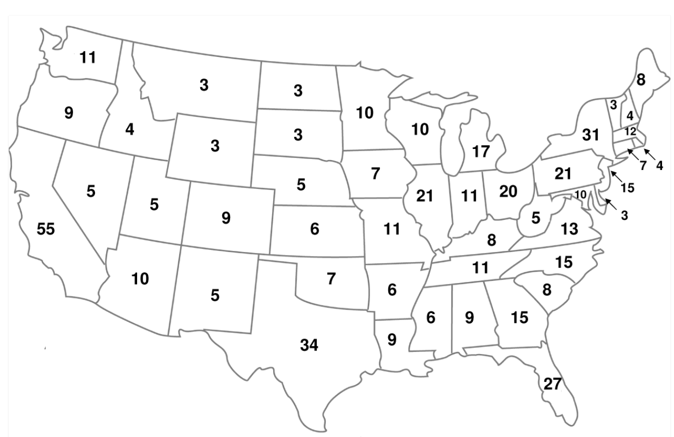

Figure 1. Electoral votes in 2008 for 50 states and the District of Columbia (not in figure: Hawaii (4 votes) and Alaska (3 votes))

distribution with $n = 50$ independent trials and probability of going for candidate A of $p = 1/2$, then the expected value of M is $np = 50(1/2) = 25$. Consequently, the expected number of additional votes for candidate A, under the simple coinflip model, is $9.66(25) = 241.5$.

The calculation for the standard deviation of the number of votes is slightly more complicated, but uses a familiar formula from introductory probability theory. In words, the variance of the number of votes is the mean of the conditional variance plus the variance of the conditional mean. The standard deviation is the square root of that number. In symbols,

$$\sqrt{E(Var(Votes \mid M)) + Var\,(E(Votes \mid M))}. \qquad (1)$$

The variance of the votes given the number of states going for candidate A, assuming the states decide independently, is $Var(Votes \mid M) = 7.25^2M$. The expectation of this is $7.25^2(25) = 1314.06$. As before, $E(Votes \mid M) = 9.66M$. The variance of M from the binomial distribution is $np(1-p) = 50(1/2)(1\text{-}1/2) = 12.5$. A multiplier such as 9.66 in a variance is squared, so $Var(E(Votes \mid M)) = 9.66^2(12.5) = 1166.45$. Then, the standard deviation is the square root of the sum of the two parts:

$$\sqrt{1314.06 + 1166.45} = 49.80.$$

If the distribution of the number of additional electoral votes for candidate A is roughly normal, then the probability that candidate A gets the needed 215 additional votes to win is approximately

$$\begin{aligned} P(Votes \geq 215) &\approx 1 - \Phi\left(\frac{215 - 241.5}{49.80}\right) \\ &= 1 - \Phi\,(\text{-}0.532) = 1 - 0.297 = 0.703. \end{aligned} \qquad (2)$$

In this scenario, the candidate who wins California will win the general election about 70% of the time. The importance of California as a big state is clear, even if the simple assumptions—such as a 50/50 coinflip for the candidate who wins California to win each other state—do not match political reality. No other state would have this large of a conditional probability.

Turning to another issue, can a probability model describe the uneven distribution of electoral votes across the states? If so, what can that model tell us about how the Electoral College disparity has changed since the electoral voting system was introduced 220 years ago?

A Multinomial Model for the Distribution of Electoral Votes

We will consider modeling the distribution of electoral votes to the 51 states. A multinomial distribution is a probability distribution for the number of items from a fixed total distributed to categories or classes. There is a probability for each category, which in this application are the 51 states. Under a multinomial model, the votes are distributed to the states independently of one another and with fixed probabilities of going to each state. To model the distribution of votes, it will make more mathematical sense to consider only the votes corresponding to the congressional seats that are not already guaranteed with minimum state population. Specifically, each state has two elected members of the Senate (the state population does not matter) and at least one member in the House of Representatives. Additional representatives allotted to the state are based on population of the state and the constraint that the total number in Congress is 538.

Let $X_1,\ldots,X_n$ be the electoral votes allocated to the n states from the remaining $k = 385$ possible votes. Then, $k = \sum X_i = 385$. Now the maximum from the $n = 51$ states in the 2008 election is California, with 52, and eight sparsely populated states (again, including DC) have none. Can the multinomial model reasonably explain this distribution of votes?

A simple model is one for which votes are distributed to states independently with equal probability. This is a multinomial model with constant chances for the 51 states, which will not suffice. Under this model, states would get 7.55 votes on average, with a standard deviation of 2.72 votes. The equal probability model does not characterize the extreme variability that allows eight states to end up with no votes while a single state ends up with 52.

Another approach is to treat the probabilities as unknown parameters and estimate them based on the actual observed electoral vote distributions. The maximum likelihood estimates of the probabilities are simply the actual vote counts divided by the number of total votes. In California, for example, the empirical probability is 52/385, or 13.5%. This model fits exactly, but it does not help us learn anything about the distribution beyond what we already know. In 2008, using this framework, we have exactly one sample—the actual number of electoral votes in the nation.

A Multivariate Polya Model

The multivariate Polya distribution (MPD) is closely related to the multinomial distribution when each bin (state) has a probability of receiving each of the $k = 385$ votes. The 51 probabilities add to one. The probabilities, themselves, are modeled as arising from a mixing distribution. With more than two categories, it is called a Dirichlet distribution and is a model for probabilities adding to one. That is, the probabilities are thought of as arising independently from a single distribution. The resulting mixture distribution for the electoral vote counts is MDP. In the special case when $n = 2$, the model for the counts is a binomial probability model and the binomial probability parameter is assigned a beta distribution. The resulting mixed distribution is called a beta-binomial distribution. The multivariate Polya distribution has probability mass function

$$p\mathbf{x}\,(x_1,\ldots,x_n) = \frac{\Gamma(n\beta)}{\Gamma(n\beta + k)} \; \frac{k!}{\prod_{j=1}^{n} x_j!} \; \prod_{j=1}^{n} \frac{\Gamma(x_j + \beta)}{\Gamma(\beta)}\,, \qquad (3)$$

where $\sum x_j = k$ and β is called the contagion parameter. The contagion parameter β regulates how evenly k votes will be distributed among n equally likely states.

To see how this works, consider a ball and urn problem in which n urns each contain β balls. Another ball is randomly placed in one urn with probability determined by the proportion of the balls in the urn. To start, each urn has an equally likely chance of getting the first ball. The second ball, however, is more likely to go to the same urn that received the first ball. If β is large, the chance does not increase much, but β also can

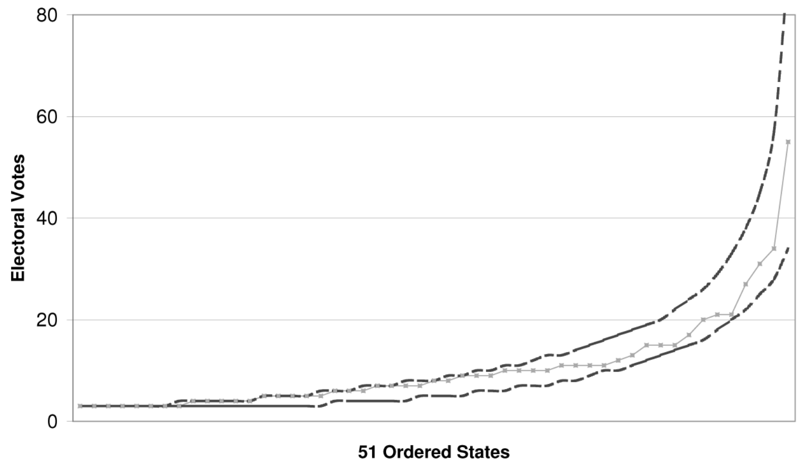

Figure 2. Goodness-of-fit test result for U.S. 2008 electoral vote data and the multivariate Polya distribution

be a fraction, and the first ball randomly placed in an urn can have a greater effect.

For example, as β goes to infinity ($\beta \uparrow \infty$), the contagion disappears and the distribution of X converges to the multinomial distribution with equal probabilities $(n; k; n^{-1}\mathbf{1})$, where $\mathbf{1}$ is an n-vector of ones. Conversely, as β decreases to zero ($\beta \downarrow 0$), the contagion causes all the votes to be distributed to just one state, creating a maximum amount of unevenness. These opposite extremes for β also represent the extreme difference in multivariate Polya distributions in terms of entropy, including Kullback-Leibler divergence.

Using the likelihood in (3), the maximum likelihood estimator for b corresponding to the 2008 U.S. electoral map is $\hat{\beta}$ = 0.507, with an approximate 95% confidence interval of (0.307, 0.836) based on the likelihood ratio. To ensure the MPD adequately fits the election data, we will apply a heuristic goodness-of-fit test.

Because the expected values for each state are identical, a standard chi-square test is generally ineffective. Alternatively, a test based on the expected values for the n order statistics (the values of $X_{1,\ldots,}X_n$ sorted in ascending order) of X can effectively distinguish MDP data from more general categorical distributions. A test can be constructed using simulation to construct 95% confidence intervals for each order statistic based on MPD simulations. Simulating the MPD is relatively easy, as it can be constructed as a multinomial distribution with mixing parameter $(p_{1,\ldots,}p_n)$ having a Dirichlet distribution. A Dirichlet vector can be composed from a vector of independent gamma random variables (with identical scale parameter) divided by their sum.

In Figure 2, the goodness-of-fit results are plotted for the 51 states, which are ordered according to the number of electoral votes. (California is represented by the right-most plotted point.) The 95% confidence intervals are based on 105 simulations. The apparent suitable fit reinforces the assertion that the contagion model can characterize the way electoral votes are distributed differently across the states. This simplifies our modeling effort because the contagion model requires only one parameter to be estimated, instead of 51 separate proportions.

Previous U.S. Elections

Since the first U.S. presidential election in 1789, when President George Washington essentially ran unopposed, the number of states has increased from its original 10 (listed in Table 1) to the current 51. The population (both absolute and relative) of those existing states in 1789, as well as most future states, has changed dramatically in the course of the last 200 years. We are interested in finding out how the distribution of electoral votes across the states has evolved, in terms of the contagion parameter, over this same course of time. Figure 3 shows how the estimate of β has changed, along with the number of states participating in the federal election, since 1789.

As the estimated contagion β has decreased over time, state populations have polarized toward small groups of heavily populated states and large groups of sparsely populated ones. The MPD model fit also changes from year to year, sometimes due to a large number of states that were constituted at the time of the census. To go along with Figure 3, Figure 4 displays the 95% confidence intervals for the contagion parameter for each electoral vote configuration in the United States since 1789. In earlier elections, in which there is less certainty about the finiteness of the contagion parameter, the upper confidence

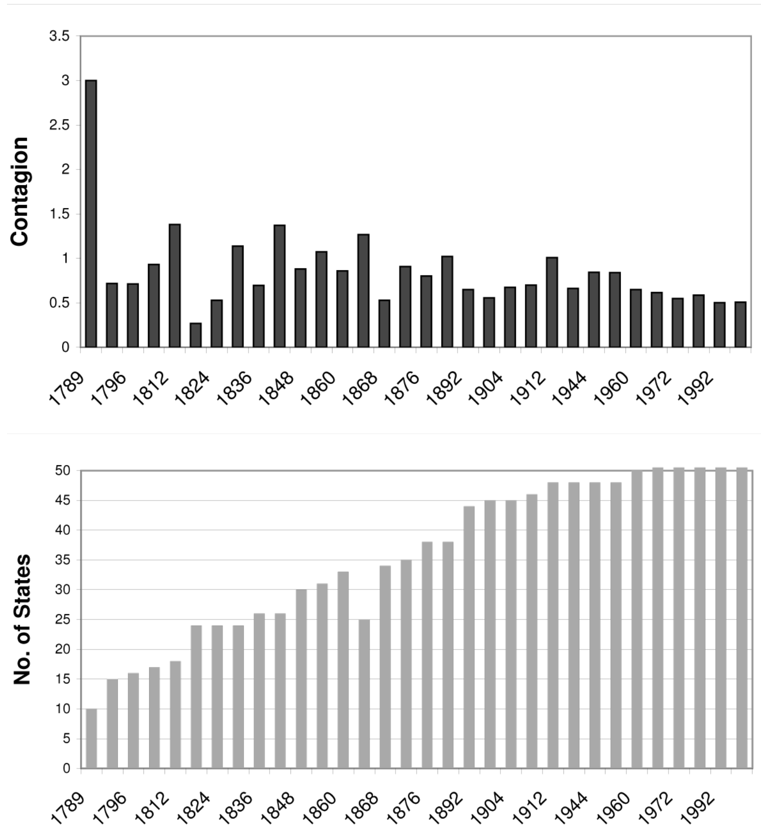

Figure 3. Estimated MPD contagion parameter and number of states in past elections (1789–2008)

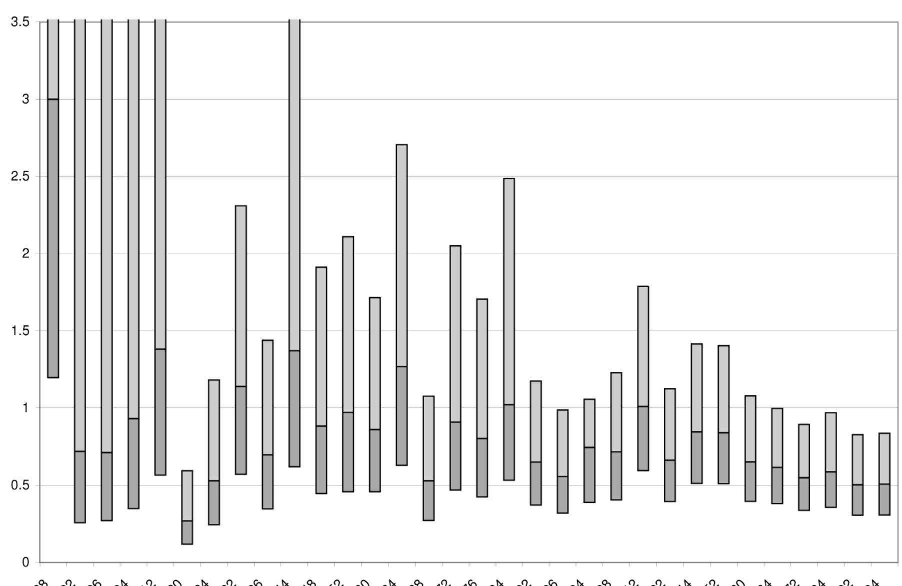

Figure 4. Estimated contagion parameter (with 95% confidence interval) for MPD model in past elections (1789–2008)

Table 1—Electoral Vote Distribution for the 1789 and 1792 U.S. Federal Election

	Electoral Votes	
State	**1789**	**1792**
Virginia	10	21
Massachusetts	10	16
Pennsylvania	10	15
New York		12
North Carolina		12
Connecticut	7	9
Maryland	6	8
South Carolina	7	8
New Jersey	6	7
New Hampshire	5	6
Georgia	5	4
Kentucky		4
Rhode Island		4
Delaware	3	3
Vermont		3

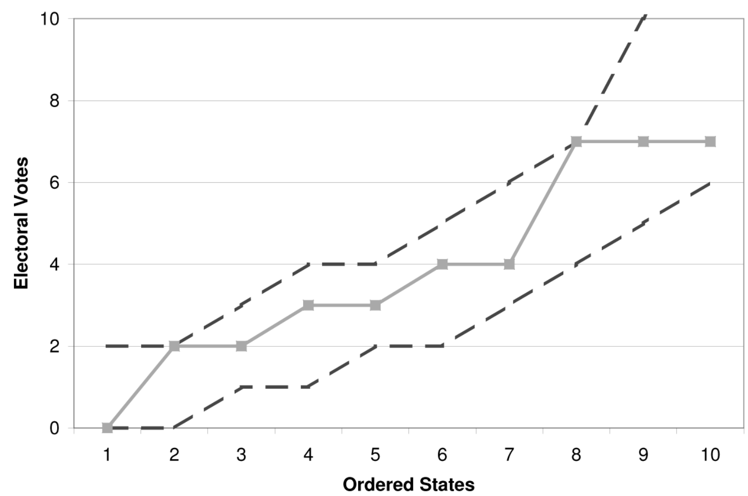

Figure 5. Goodness-of-fit test result for U.S. 1789 electoral vote data and the multivariate Polya distribution

limit can become arbitrarily large. For this reason, the upper bounds in Figure 4 are cut off at $\beta \geq 3.5$.

The wider intervals corresponding to some earlier elections (including all elections before1820) suggest the MPD fit is better in more recent years. The MPD goodness of fit for the1789 data is illustrated in Figure 5. The plot does not belie a fit to the MDP. A goodness of fit for the multinomial distribution (with equal probabilities), however, has a chi-square test statistic of 13.6, with significance $P(\mathbf{X}_9^2 \geq 13.6) = 0.14$. This is the best fit among all the election cycles that follow; all other significance values of the multinomial goodness-of-fit test are less than 0.01, indicating a clear lack of fit.

Given the somewhat uniform spread of the population across the 10 participating states in the 1789 election (see Table 1), it would be hard to believe the founding fathers or anyone else considered future elections based on five times as many states and such a dramatically less even population distribution. The contagion parameter (β), which represents a metric of population evenness or stability, is estimated to be largest ($\hat{\beta} = 3$) for the first election, but in every election after 1789, the estimate was much smaller—almost always less than one.

Table 1 shows that by the election of 1792—when Washington was re-elected—five more states were added to union and population disparity between the states increased greatly. Likewise, the contagion parameter decreased from 3.0 to 0.72. While Virginia accumulated a large population with a dominant number of electoral votes (21) in 1792, several smaller states had only the minimum number of three votes.

There are two noticeable drops in the contagion parameter in Figures 3 and 4. One was the 1824 election. Shortly before the 1820 election, five sparsely populated western territories gained statehood (i.e., Alabama, Illinois, Indiana, Mississippi, and Missouri). In this election, President James Monroe defeated President John Quincy Adams and was re-elected as president. The other drop is in 1868. In the 1864 election, several southern states did not participate in President Abraham Lincoln's re-election.

The Effect of the Contagion Parameter

If the contagion parameter is sufficiently small, it becomes more likely that one or two states will dominate in terms of electoral votes. With the current rules that add two electoral votes to the minimum one each state gets based on relative population, the electoral college will never satisfy the 80/20 rule. That is, if 20% of the states have almost all the country's population, they can never garner even 80% of the electoral votes with 51 states. Currently, our most populated 10 states have less than half the total number of electoral votes. Even if we consider only the 385 electoral votes assigned after each state already has the three minimum electoral votes, simulation shows that the 80/20 rule is satisfied only with a contagion parameter of $\beta \leq 0.22$.

To illustrate the influence the contagion parameter wields on a presidential election outcome, we again look at how the population disparity creates big states that dominate the others in terms of electoral votes. We consider a population in which the two candidates have an equal chance of winning any individual state and examine the probability that the candidate who wins the largest state wins the election. This is how the "California effect" was computed earlier. We found

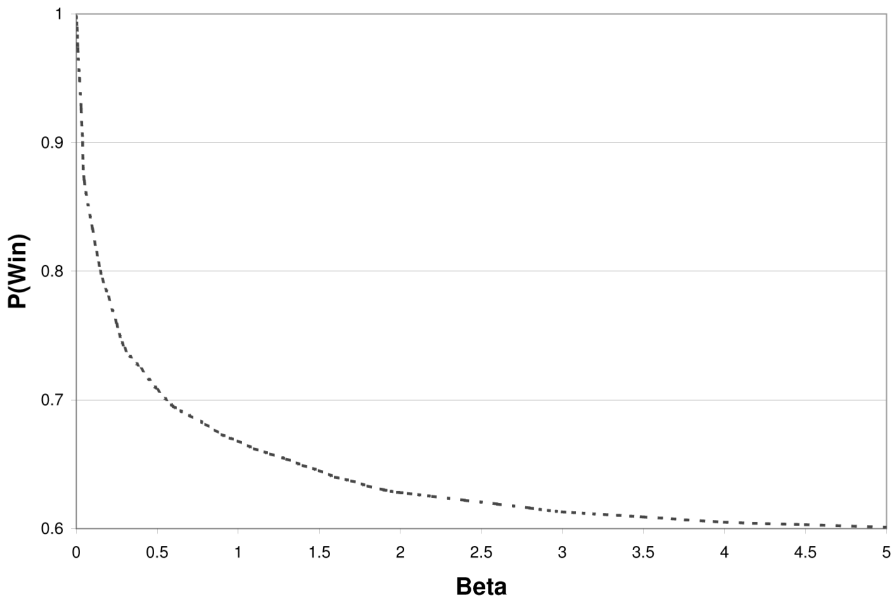

Figure 6. The conditional probability (based on 10^6 simulations) a candidate will win the Election (vertical axis) given he or she has already won the largest state. Along the horizontal axis, the contagion parameter, β, ranges from 0 to 5.

that if a candidate wins California and has a 50/50 chance of winning each of the remaining states, that candidate has a 69.4% chance of winning the election.

We again let $X = (X_1,\ldots,X_n)$ have a multivariate Polya distribution where $k = 358$ votes are allotted to $n = 51$ states. The probability of winning the election, given the largest state is already won, increases as β decreases. Figure 6 shows how the probability of winning the general election goes to one as β decreases to zero. Furthermore, as β increases to infinity, the probability will correspond to the analogous model based on the multinomial distribution with equal bin probabilities. In that case, the largest state is not much larger than most of the other states, and the probability of winning the election is around 0.58.

Advertising Dollars and Presidential Elections

In 2008, the University of Wisconsin Advertising Project estimated that more than 50% of the cost of television advertisements was spent in just 10 states. In 2000, FairVote's Presidential Elections Reform Program reported that the advertising money spent in Florida alone exceeded that of 45 states and the District of Columbia combined. The unequal advertising spending reflects how the electoral vote share is unequal across states and the outcome in some states is decided by small margins.

From Figure 6, the probability of winning with $\beta = 0.5$ is about 0.71, which is in agreement with the approximation based on California being the biggest state in the 2000–2008 electoral vote distribution. If we go back to the first election in 1789, where the population was more evenly spread across 10 states and $\hat{\beta}=3.0$, the probability decreases to 0.62.

Achieving Evenness in the Electoral College

During the U.S. Constitutional Convention in 1787, the framework for the Electoral College was created, delegating votes to states in proportion to its representatives in Congress. The United States was still an agrarian society at the time of the first federal election, and the new nation's population was spread out somewhat uniformly among most of the 10 voting states.

It can be argued that this balance was part of the original design that emphasized constitutional rights as a mixture of state-based and population-based government. In "Federalist No. 39," President James Madison proposed to have a Congress with two houses, one based on population (the House of Representatives) and one based on states (the Senate), with the election of the president also relying on this mixture of two government modes.

The balance achieved in the first election can be characterized in the statistical model based on the multivariate Polya distribution, where the contagion parameter was estimated as $\hat{\beta} = 3.00$. To achieve this kind of voting balance with our current population distribution, a lot of state lines would have to be redrawn. For example, if the United States collapsed North and South Dakota into one state, removing a star from the flag, this would move the contagion parameter from its current 0.5071 to 0.6461. To gain any substantial increase for β, all of the states with just three or four electoral votes would have to be conjoined with larger states. The 23rd Amendment, which guarantees at least three electoral votes to DC, would

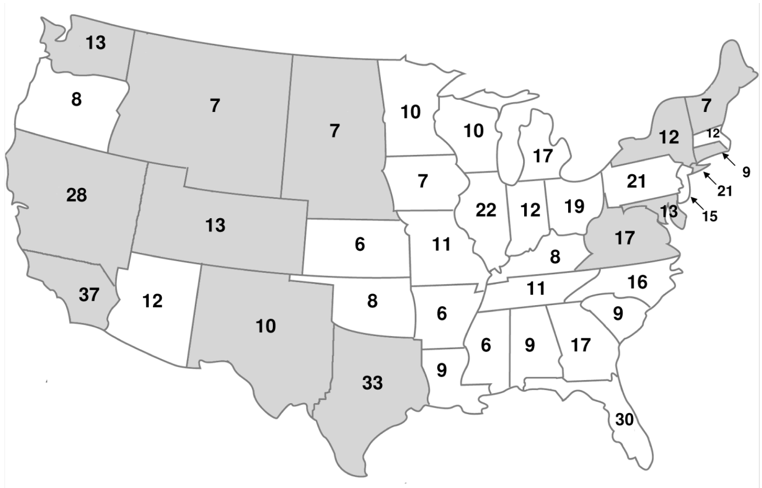

Figure 7. The reconfigured United States, with 39 states and contagion parameter $\hat{\beta}$ = 3.0. Modified states are shaded gray. Note that Alaska and Hawaii, not pictured, also are modified.

need rectification, and, more likely, DC would be swallowed by Maryland or Virginia.

To show how difficult this balance would be to achieve, following is an efficient (albeit facetious) 12-step master plan listed below that will increase the contagion parameter to its original $\hat{\beta}$ = 3.0:

Split California into two states (i.e., Northern California and Southern California) based on the county lines starting between San Bernardino and Tulare counties

Split off the New York City metropolitan area (including all of Long Island) from the state of New York

Split off west Texas, according to the vertical borders implied by its congressional districts, joining it with New Mexico

Join Montana, Wyoming, and Idaho

Add Alaska to Washington

Add Hawaii and Nevada to Northern California

Join Utah and Colorado

Join North Dakota, South Dakota, and Nebraska

Join Maine, Vermont, and New Hampshire

Join Virginia and West Virginia

Add Delaware and DC to Maryland

Join Connecticut and Rhode Island

The results of this splitting and mixing are displayed in Figure 7, based on population estimates for 2007. The New York City metropolitan area, as a new state, would garner 21 electoral votes, while the remains of upstate New York would have 12. Among the 39 states in the reconfiguration, the lowest number of electoral votes for any state is six. No additional pairing significantly increased $\hat{\beta}$ past 3.0.

Discussion

The reconfigured union of states displayed in Figure 7 serves to illustrate how much the population of the United States has changed since the Constitutional Convention. Barnett and Kaplan suggested that any constitutional amendment to change the electoral system directly has little or no chance of being approved because it would require a large number of small states to vote against their own interests. Their aim is to go through a side door by tweaking the system via a weighted vote share. The authors show that the results would be significantly closer to matching a national popular vote.

An additional metric was provided here that reflects how the unevenness in population distribution translates into inequality of voting power between states. The key is to view the current distribution of population and electoral votes as one of many possible realities, depending on how the population distributes itself into states, or how states are defined with respect to population density.

Further Reading

Barnett, A., and E. H. Kaplan. 2007. A cure for the electoral college? *CHANCE* 20(2):6–9.

David, F. N., and D. E. Barton. 1962. *Combinatorial chance*. Port Jervis: Lubrecht & Cramer, Ltd.

Gainous, J., and J. Gill. 2002. Why does voting get so complicated? A review of theories for analyzing democratic participation. *Statistical Science* 17(4):383–404.

Kvam, P. H., and D. Day. 2001. The multivariate Polya distribution in combat modeling. *Naval Research Logistics* 48(1):1–17.

Pearson, C. 2006. *Who picks the president?* Takoma Park: FairVote.

Peterson, D. W. 2008. Putting chance to work: Reducing the politics in political redistricting. *CHANCE* 21(1):22–26.

University of Wisconsin Advertising Project, *http://wiscadproject.wisc.edu*.