

# Florida or Ohio? Forecasting Presidential State Outcomes Using Reverse Random Walks\*

Aaron Strauss  
Department of Politics  
Princeton University  
abstraus@princeton.edu

March 2007  
2006-2007 Political Methodology Seminar

## Abstract

The classic Colonel Blotto game offers insights into the problem of allocating resources in battleground states during a presidential election. The key to a effective strategy is determining, *ex ante*, which states will be pivotal on Election Day. Existing Bayesian election models are inappropriate for this game, as they estimate the current standings of the candidates or states rather than the final outcomes. I develop a Bayesian dynamic linear forecasting model that incorporates informative priors from historical regressions, updates based on in-cycle state and national polls, and accounts for the uncertainty of events that take place between the polls' issuance and Election Day. National and state shocks are modeled as a reverse random walk beginning with the final outcome and moving backwards through time. Uncertainty about the final outcome is calculated by combining the random walk's linearly decreasing variance over time, natural poll measurement error, house effects of national polls, and historically stable trends of election results. Using the resulting estimates of the states' standings relative to each other, I simulate electoral vote outcomes and determine the probability of each state being pivotal. I find that early polls can be misleading to such an extent that putting any weight on them produces worse forecasts than solely relying on historical trends.

## 1 Background

As recent history has demonstrated, a presidential campaign's electoral vote strategy can be the difference between winning and losing the White House. The competing campaigns attempt to win the Electoral College battle by gaining asymmetric information and insights about the state vote terrain. One crucial aspect of this challenge is determining, *ex ante*, which state will be pivotal on Election Day.

---

\*The cooperation of Harrison Hickman and Mark Mellman in the sharing of data is greatly appreciated. My thanks also to Ben Lauderdale and Hillary D. Hampton for their help in crafting this paper. Finally, those at the Princeton political methodology seminar were instrumental in providing this paper with proper direction.

I develop a Bayesian dynamic linear forecasting model to assess the likelihood that, given historical trends and current polling, a state will be pivotal. States that are located near the national trend line of two-party vote have the highest probability of making the difference between winning and losing. I model both state and national shocks as reverse random walks from informative priors of the final outcome, and use this specification to weight current polls against stable historical trends. The resulting forecasts are then used as inputs to Electoral College simulations.

## 1.1 Prior Research

A few decades ago, the majority view among political scientists was that the Electoral College privileged large states, since they have a disproportionate effect on the electoral vote. (Mann and Shapley, 1964; Owen, 1975; among others). In an influential work, Brams and Davis (1974) propose that campaigns allocate resources by state using a factor equal to each states' electoral vote, raised to the  $(3/2)$ 's power. For example, consider a presidential campaign determining how to divide its voter contact budget between State A with 16 electoral votes, and State B with 4 electoral voters. Instead of spending four times more money in State A (as would be demanded by a division based solely on electoral votes), Brams and Davis (1974) propose that campaigns allocate resources in proportion to states' electoral vote raised to the  $(3/2)$ 's power. For example, suppose a presidential campaign had to determine how to divide its voter contact budget between state A with 16 electoral votes and state B with 4 electoral voters. Instead of spending four times more money in State A (as would a division based solely on electoral votes), Brams and Davis recommend spending eight times more money in state A than in state B ( $16^{(3/2)}/4^{(3/2)} = 8$ ). The authors admit that this  $3/2$ 's strategy only holds as a pure-strategy equilibrium when (1) states are *ex ante* equally competitive, and (2) campaigns match each other's resource allocation in each state.

Bartels (1985) tests the  $3/2$ 's proposition empirically, analyzing spending by the 1976 Carter campaign. Campaign resources are divided into two groups: instrumental resources (those that are leveraged to win votes, such as advertising) and ornamental resources (such as top-level campaign staff, who are hired primarily for "show" in uncompetitive states). Bartels does find that campaign trips and advertising are disproportionately allocated to larger states, while organizational funding and staffing are spread on a per capita basis.

Yet modern presidential campaigns have not followed such a disproportionate pattern. The elections of 2000 and 2004 saw hardly any advertising spending in the four of the five largest states (California, Texas, New York, and Illinois). The simple reason behind this phenomenon was that these states were not competitive (Johnston, Hagen, and Jamieson, 2004; Merolla, Munger, and Tofias 2005).<sup>1</sup> The level of a state's competitiveness weighs heavily in a candidate's decision to compete actively there. Under some models, larger states have a great chance of having a close outcome (Banzhaf, 1968), but recent empirical evidence indicates that this relationship is small at best (Gelman, Katz And Tuerlinckx, 2002).

Campaigns attempt to assess each state's competitiveness before allocating their resources. Colantoni, Levesque and Ordeshook (1975) put it succinctly: "Identify the states that you and your opponent are certain to carry, and focus your attention on the remaining (competitive) states." However, the two candidates may disagree on which states are competitive and the level of closeness. This uncertainty leads to a manifestation of a Colonel Blotto game (Merolla, Munger, Tofias; 2006).

The zero-sum Blotto game has been well-studied for decades—the most common form of the Blotto game revolves around Colonel Blotto, who must decide how to allocate 100 soldiers across 10 battlefields. His opponent, Colonel Rumsfeld, simultaneously does the same. A battlefield is won with certainty when one player sends more soldiers to that front than her opponent. The winner of the most fronts wins the battle. A straightforward proof (see, for instance, Golman and Page, 2002) shows that no pure strategy equilibrium exists.<sup>2</sup> But in presidential campaigns, not all battlefields are equal; some are sure to be won, and some are sure to be lost—the election hinges on the remainders. The key to victory, then, is determining, before your opponent, which states will be pivotal.

Forecasting which states will be pivotal well ahead of time requires knowledge about the national political environment as well as the relative positions of the states. For instance, if the Democratic candidate were losing by five percentage points in the national horserace, said candidate should focus resources on states where she is also losing by the same five percentage points, and not on states that are currently close. Even if, on Election Day, the Democrat won the competitive state

---

<sup>1</sup>The fourth most-heavily populous state, Florida, saw a great deal of action. The Bush campaign spent moderately in California in 2000; neither campaign advertised in the sunshine state in 2004.

<sup>2</sup>The mixed strategy equilibrium to this game is found in Gross and Wagner (1950).

and the national race stayed as it was, then she would almost certainly still lose the Electoral College.<sup>3</sup> If a national shock changed the race by five points—so her and her Republican opponent were running neck-and-neck—those states that did not appear to be competitive before the shock would become critical.<sup>4</sup>

Campaigns inform their Blotto game choices by tracking both national and state polls and comparing the relative competitiveness between the states and the nation at large. The key to a successful resource allocation is forecasting final state competitiveness given a close national result. Franklin (2001) introduces a model that estimates both the candidates’ standings in the states relative to the national, but does not forecast the final results.

I extend Franklin’s model by introducing an informative prior of the final result along with mid-campaign polling. Gelman and King (1993) report that presidential campaign polls are relatively variable compared to historical trends. I leverage this finding to form informative priors for both the national race and, more importantly, relative state standing. Use of this historical data assumes that even with better information, the Electoral College Blotto game will play out similarly to elections in the past.<sup>5</sup>

Jackman (2005) models the day-to-day shocks of the national Australian election as a random walk. Since, Gelman and King (1993) demonstrate that the final result is known with some certainty, the motion of a campaign cannot be random. Instead, I adjust Jackman’s specification slightly and model shocks as a reverse random walk, starting at Election Day and moving backwards to the start of the campaign. This assumption, which fits the 2004 data well, yields a formula for how to weight current state polls with historical trends. In total, the model presented here allows campaigns to forecast state outcomes by analyzing the polls as they come out of the field.

## 2 The Model

Two models are presented below: the first analyzes national and state polls over a presidential campaign to determine how states vary relative to the national, and the second uses information

---

<sup>3</sup>A similar analysis to Garand and Parent (1991) using campaigns from 1952-2004, shows a slight Republican bias in the Electoral College system. If a Democrat receives 50% of the two-party popular vote, she should expect to receive 49% of the electoral vote.

<sup>4</sup>This assumption that national shocks affect the country uniformly appears to be true in most cases, but with exceptions (Johnston, Hagen, Jamieson, 2004).

<sup>5</sup>For evidence that campaign activity does affect the final vote outcome, see Shaw (1999).

from the first to forecast final outcomes for the states. I develop a Bayesian dynamic linear model in which national and state shocks are considered reverse random walks. State polls are evaluated with reference to the national horserace. When estimating the nationwide candidate preference, I control for the effects of different polling organizations (Jackman, 2005). The forecasting model uses informative historical priors and weights the importance of mid-campaign polls by assuming that variance decreases linearly as Election Day approaches.

## 2.1 Model of National Vote

National vote is estimated using a combination of polls results, survey size, and house effects (bias). Let the total number of national polls be  $n$  and  $i$  index these polls  $1, \dots, n$ . Each poll reports Democratic national two-party vote,  $y_i$ , which varies about some underlying value,  $\mu_i$ , with random error. The variance of each poll,  $s_i$ , is calculated from the assumption that the poll's  $q_i$  respondents are sampled randomly from the population. (With few exceptions, I follow the notation implemented by Jackman, 2005.)

$$\begin{aligned} y_i &\sim \mathcal{N}(\mu_i, s_i^2) \\ s_i^2 &= \frac{y_i(1 - y_i)}{q_i} \end{aligned} \tag{1}$$

While the variance of the poll is based only on the simple “margin of error” formula, pollsters’ results are assumed to be biased in some direction. Question wording, question order, likely voter screeners, and other systematic “house” effects can lead to poll results that are consistently different from the true national two-party Democratic preference. Let the number of days until the Election,  $t = 1, 2, \dots$  index the actual position of the horserace,  $\alpha_t$ . Let the number of survey organizations,  $J$ , be indexed by  $j = 1, \dots, J$  and  $\delta_{ji}$  represent the house effects for the organization  $j$  that fielded poll  $i$ . Polling firms’ bias is simply represented by,

$$\mu_i = \alpha_{t_i} + \delta_{j_i}.$$

As shown, the model is under-identified because all polls could be erroneous by 2 percentage points in the same direction, and the public (and methodologists) would be oblivious. The identifying restriction, then, is that the house effects “balance out”:  $\sum_{j=1}^J \delta_j = 0$ .

Daily events, large and small, effect the course of the presidential campaign. Yet, the movements are not random as scholars are able to predict the result of a presidential election with low uncertainty years ahead of time. Instead of taking the conventional approach of modeling events as a forward random walk (Jackman, 2005), shocks are modeled as random backwards through time,

$$\alpha_t \sim \mathcal{N}(\alpha_{t-1}, \omega^2),$$

with  $\omega^2$  representing the variance of the shocks.<sup>6</sup> (Note again that larger values of  $t$  signify earlier periods of the campaign cycle.)

Several competing forecasts of presidential elections exist (Fair, 1978; Rosenstone, 1983) among others. The choice of models does not matter much for two reasons. First, the final national result is not the quantity of interest. For the Blotto resource allocation game to matter on Election Day, the popular vote must be close. Thus, when simulating which states are pivotal, I assume a tight national horserace. Second, the high poll density, especially at the end of the campaign (15 in the final week) swamps the prior as long as the prior is accurate.

The forecast of Cuzan and Bundrick—calculated in February, 2004—is chosen as the prior for the national vote. The authors predicted that President Bush would be re-elected with 52.2% of the national vote. The standard deviation for their estimate was 2.53%.

$$\alpha_0 \sim \mathcal{N}(0.522, (0.0253)^2),$$

I do not have strong priors on the other two parameters :  $\delta_j$ , or  $\omega$ . The day-to-day shocks of the campaign are assumed to fall within plus or minus 2 percentage points 95% of the time (at a maximum). While specific events (e.g., conventions, debates) might cross this threshold, those are the exceptions, not the norms.<sup>7</sup> The prior on house effects assumes that 95% of bias is less than 20% (in either direction): all reputable polling organizations should easily pass that standard.<sup>8</sup>

---

<sup>6</sup>Alternatively, shocks could be modeled as a Ornstein-Uhlenbeck process (1930), in which shocks both have random and equilibrium-reverting components. Applying both the reverse random walk and Ornstein-Uhlenbeck models the 2004 national data, yields very similar results: point estimates are on average within 0.1% of each other and the standard deviations of Ornstein-Uhlenbeck estimates on average 0.1% larger than those for the reverse random walk model. The point estimates for the Ornstein-Uhlenbeck are 0.005 for the Weiner process multiplier, 0.07 for mean-reversion weight, 49.8% for the mean. As more data become available, I will explore whether the nonlinear variance implied by Ornstein-Uhlenbeck best fits the observations.

<sup>7</sup>Even conventions, which last four days, might not produce effects larger than two percentage points per day.

<sup>8</sup>Because of the identifying restriction, the calculation of the prior is technically off by a factor of  $(J-1)/J$  (see

Formally, the priors are defined as

$$\begin{aligned}\omega &\sim \text{unif}(0, (0.01)^2) \\ \delta_j &\sim \mathcal{N}(0, (0.1)^2)\end{aligned}$$

Each of these parameters can be drawn from a known distribution, conditioned on the others. Let  $D$  represent the data,  $\Theta$  represent all of the parameters, and (for example)  $\Theta_{-\delta_j}$  represent all parameters except  $\delta_j$ . A Gibbs sampler then draws the conditional distributions by (Jackman, 2005),

$$\begin{aligned}f(\alpha_t | \Theta_{-\alpha_t}, D) &\sim \mathcal{N}\left(\left[\frac{y_i - \delta_{j_i}}{s^2} + \frac{\alpha_{t+1} + \alpha_{t-1}}{\omega^2}\right] \left[\frac{1}{s^2} + \frac{2}{\omega^2}\right]^{-1}, \left[\frac{1}{s^2} + \frac{2}{\omega^2}\right]^{-1}\right) \\ f(\delta_k | \Theta_{-\delta_k}, D) &\sim \mathcal{N}\left(\left[\sum_{i \in P_k} \frac{y_i - \alpha_{t_i}}{s_i^2}\right] \left[\frac{1}{\sum_{i \in P_k} s_i^2} + \frac{1}{d^2}\right]^{-1}, \left[\frac{1}{\sum_{i \in P_k} s_i^2} + \frac{1}{d^2}\right]^{-1}\right) \\ \text{let } \tau &\equiv \omega^{-2} \\ f(\tau | \Theta_{-\omega}, D) &\propto g\left(\frac{T-2}{2}, \frac{1}{2} \sum_{t=1}^{T-1} (\alpha_{t+1} - \alpha_t)^2\right) \mathcal{I}(\tau > 10,000),\end{aligned}$$

where  $g$  is the Gamma distribution with alpha and beta parameters, and  $\mathcal{I}$  is an indicator function that constrains the value of  $\omega$  to stay within the boundaries set by the prior: (0,0.01).

## 2.2 Model of State Vote

In estimating state choices, unlike in national vote choice estimates, the goal is *not* to approximate the state of the horserace on a given day  $t$ , but to estimate the final state vote deviation from the national. Polls from the 2004 campaign are analyzed from a historical perspective in order to calibrate the forecasting model. For state  $k = 1, \dots, 50$ , let the final two-party Democratic vote be  $\gamma_{k,0}$  and the state deviation be  $\theta_{k,0}$ .<sup>9</sup> These two values form the arithmetic relation,

$$\theta_{k,0} = \gamma_{k,0} - \alpha_0,$$

where consistent with Section 2.1 the notation above,  $\alpha_0$  is the true two-party national vote on

---

Beck, Jackman, and Rosenthal, 2006). Since I analyze more than 20 organizations, I ignore this complication.

<sup>9</sup>The second subscript is time;  $t = 0$  is Election Day.

Election Day. The challenge here is that the value of  $\theta_{k,0}$  is realized *after* all of the data points (the state polls). Assets to estimating this future value include, sporadic state polls, a steady stream of national polls and strong prior from historical regressions that explain much of the state-to-state variance. Pinpointing the exact value of  $\theta_{k,0}$  within the campaign (i.e., predicting the future precisely) is impossible, but one can estimate the correct uncertainty of  $\theta_{k,0}$  so that campaigns can play the Blotto game optimally. To understand the uncertainty about  $\theta_{k,0}$ , I analyze the 2004 election with  $\theta_{k,0}$  revealed, noting the error in the mid-campaign polls.

Let  $m = 1, \dots, M$  index the state polls conducted during the election season. Given a state poll two-party topline result  $z_m$ , the state's deviation from the national,  $\zeta_m$ , is distributed

$$\zeta_m \sim \mathcal{N}(z_m - \alpha_{t_m}, \psi_m^2), \quad (2)$$

where  $\psi_m^2$  is the convolution of the poll's margin of error and the uncertainty about the estimate of  $\alpha_{t_m}$  from the national model. (Due to a lack of state data and a wide array of state survey organizations, house effects are ignored.) A state's current standing in the polls,  $\zeta_m$ , informs the state's Election Day competitiveness relative to the national vote,  $\theta_{k_m,0}$ . How much information resides in the data point  $\zeta_m$  is an open question. A simplistic model would let  $\zeta$  be an unbiased estimator of  $\theta_{k_m}$  with variance equal to that of equation 2. I label this the “low variance” model,

$$\theta_{k_m}, \zeta_m \sim \mathcal{N}(z_m - \alpha_{t_m}, \psi_m^2)$$

Alternatively, uncertainty about  $\theta_{k_m}$  might also include the fact that polling firms have difficulty determining exactly who will vote in the election. Non-random selection (e.g., older voters answering the phone more often than younger voters) may affect the survey's measurement error in ways not accounted for by the standard “margin of error” calculation. From this perspective, which I term the “constant” model, the variance about  $\theta_{k_m}$  increases by an additional term,  $v^2$ ,

$$\theta_{k_m} \sim \mathcal{N}(\zeta_m, v^2 + \psi^2).$$



These two models, however, neglect to address a key fact of state polls: information later in the campaign is a better forecast than information early in the campaign. In essence, when state-specific shocks occur, they shift the state’s relative competitiveness. These shocks are modeled as a reverse random walk, just as the national shocks are,

$$\theta_{k,t} \sim \mathcal{N}(\theta_{k,(t-1)}, \sigma^2),$$

where  $\sigma^2$  represents the daily variance of the random walk. The variance is assumed to be constant across states; a discussion of that choice follows shortly. One property of random walks is that the variance increases linear over time. In the case of the reverse random walk the variance decreases linearly over time (but increases over the variable  $t$ ). Thus,

$$\theta_{k_m,0} \sim \mathcal{N}(\zeta_m, v^2 + t_m \sigma^2 + \psi_m^2),$$

is the forecasting equation for the “reverse walk” model. (For completeness, the  $v^2$  offset term is included). Under the reverse walk model, in contrast to the previous two models (low variance and offset), forecasting the outcome in the summer of the campaign season relies less on polling than does forecasting in the later month. Since, as Section 4.2 will show, the reverse walk model best fits the data, I focus on that specification.

This reverse walk model is under-identified if both  $\theta_{k_m,0}$  and  $v^2$  are taken as unknown parameters. To solve this problem, I estimate  $v^2$  and  $\sigma^2$  using past election results with the true  $\theta_{k_m,0}$  revealed. Next, the joint distribution of  $v^2$  and  $\sigma^2$  from this “past election” estimate is used to infer the level of mid-election uncertainty about the relative status of each state. Campaigns can use this “in-cycle” estimate to more effectively allocate resource allocation as they gain information from state polls.

**Past election analysis.** In this estimate, the quantities of interest are  $v^2$  and  $\sigma^2$ . I am agnostic about the prior values of both quantities, except that, intuitively, state shocks are less

prevalent than national shocks. Thus, the prior on  $\sigma$  is more restrictive than for  $\omega$ ,

$$\sigma \sim \text{unif}(0, .005)$$

At the maximum value for  $\sigma$ , 0.005, 95% of daily state shocks are less than plus or minus one percentage point. For an example that is possibly more intuitive, consider that the median value, 0.025, 95% of weekly changes in state deviation are within 3.5 percentage points (in either direction). Since in 2004 86% of states stayed within that range of their 2000 state deviation—*four years later*—the prior distribution appears to cover the sensible alternatives.

As for,  $v^2$ , many pollsters might claim that this value is zero, but that may be giving the industry too much credit. I allow for additional standard deviation up to 5%, above and beyond the traditional reported standard deviation of a poll of around 2%.

$$v \sim \text{unif}(0, .05)$$

The model's data is state polls with known mean but unknown variance (because  $\sigma^2$  is unknown) and with revealed final values. For state  $k$ , the joint likelihood function for  $v^2$  and  $\sigma^2$  is

$$\begin{aligned} \text{let } \tau_m^2 &\equiv v^2 + t_m \sigma^2 + \psi_m^2 \\ \Pr(v^2, \sigma^2 | \Theta_{-v, -\sigma}, D) &\propto \prod_{m \in P_k} \frac{1}{\sqrt{\tau_m^2}} \exp \left[ -\frac{(\theta_{k,0} - (z_m - \alpha_{t_m}))^2}{2\tau_m^2} \right] \mathcal{I}(\sigma, v) \end{aligned}$$

Thus, one could estimate  $v^2$  and  $\sigma^2$  for each state, but there is no reason to think that some states have fewer shocks in each election than others. Even if I believed otherwise, this proposition would be impossible to prove with one (or even two) elections. Instead, I simply aggregate across the states,

$$\begin{aligned} \Pr(v^2, \sigma^2 | \Theta_{-v, -\sigma}, D) &\propto \prod_{k=1}^{50} \prod_{m \in P_k} \frac{1}{\sqrt{\tau_m^2}} \exp \left[ -\frac{(\theta_{k,0} - (z_m - \alpha_{t_m}))^2}{2\tau_m^2} \right] \mathcal{I}(\sigma, v) \\ \log(\Pr(v^2, \sigma^2 | \Theta_{-v, -\sigma}, D)) &\propto \sum_{k=1}^{50} \sum_{m \in P_k} \left( \log\left(\frac{1}{\sqrt{\tau_m^2}}\right) + \left[ -\frac{(\theta_{k,0} - (z_m - \alpha_{t_m}))^2}{2\tau_m^2} \right] \right) \mathcal{I}(\sigma, v) \quad (3) \end{aligned}$$

A Gauss-Lagrange integral is used to derive the mean and variance of the two parameters in

this joint distribution; a Metropolis algorithm yields a random draw from the distribution (see Appendix).

**In-cycle estimates.** With knowledge of the parameters of the reverse walk model,  $v^2$  and  $\sigma^2$ , campaigns can estimate the state’s deviation from the national vote on election day,  $\theta_{k,0}$ , using only knowledge of state polls and past electoral performance. Campbell (1992, 2006) demonstrates that even without polling, state deviation from the national is predictable within a few percentage points leveraging just historical data. His regression, with modifications, is replicated in Section 3. This analysis serves as the prior for  $\theta_{k,0}$ ,

$$\theta_{k,0} \sim \mathcal{N}(\hat{\beta}X_k, h^2),$$

where  $\hat{\beta}$  are the regression coefficients estimated with historical data,  $X_k$  are the current campaign conditions for state  $k$ , and  $h^2$  is the variance of the regression. For a poll  $m$  in state  $k$ , the posterior distribution of  $\theta_{k,0}$  is normally distributed with the state poll and prior weighted by their precision.

$$\theta_{k,0} | \Theta_{-\theta_k}, D \sim \mathcal{N} \left( \left[ \frac{\hat{\beta}X_k}{h^2} + \frac{\zeta_m}{\tau_m^2} \right] \left[ \frac{1}{h^2} + \frac{1}{\tau_m^2} \right]^{-1}, \left[ \frac{1}{h^2} + \frac{1}{\tau_m^2} \right]^{-1} \right) \quad (4)$$

The naive, greedy approach would be to update beliefs about equation 4 using all the polls available. However, this method yields inaccurate results for two related reasons. First, prior polls have serial correlation—they are not independent draws from a distribution of variance  $\tau^2$ . Second, since state shocks are modeled as random walks, trends in the past are not indicative of future movements. The only relevant piece of information is the current state standing.

To estimate current competitiveness, one might run a Gibbs sampler, such as the one discussed in 2.1, on the available state data and retrieve the most recent value of  $\theta$ . This process is computationally challenging, as over 1.5 million parameters would have to be estimated; thus, I opt for a simpler approach.<sup>10</sup> Averaging the two most recent polls yields an unbiased estimate of the state’s standing at the time of the recent poll. Using any more polls without a Gibbs Sampler would run afoul of the serial correlation problem above. This current state standing is then ap-

---

<sup>10</sup>Under the comprehensive method, for each day  $t$  of the campaign (246 days), for each state  $k$ , each  $\theta_{k,t}$  would need to be estimated.  $\sum_{t=1}^{246} \sum_{k=1}^{50} (246 - t + 1) = 1,519,050$

plied to equation 4, resulting in an unbiased, and potentially more accurate, forecast of final state competitiveness.

### 2.3 Details of State Forecasting

Equation 4 assumes that  $v^2$  and  $\sigma^2$  (represented by  $\tau^2$ ) are known with certainty. But the past election model returns a joint distribution of these two parameters. Thus, the mean and variance of the normal distribution are integrated over this distribution using a Gauss-Lagrange method.

In-cycle inference, as currently modeled, treats the 50 state draws as independent events. They are not, as a natural constraint exists in addition to the above model. Not all states can deviate from the national vote in one direction; if they could, the national vote would be higher than estimated in the nationwide model. The constraint is specified as,

$$\sum_{k=1}^{50} w_k \theta_{k,0} = 0$$

where  $w_k$  is a weight variable that controls for state turnout. To account for this restriction, I renormalize the state estimates using actual 2004 turnout. I monitor this renormalization factor to ensure the aggregate state estimate is not wildly off from the national.

## 3 Estimating the Prior With a Simple Regression

Campbell (1992, 2006) extends the national presidential election regression equation (Rosenstone, 1983) to the states.<sup>11</sup> He uses state two-party vote percentage in each state as the dependent variable and includes explanatory variables for the economy, home state of the presidential and vice presidential candidates, state legislature makeup, ADA score of members of Congress, and indicator variables for the various regions in elections up to 1988. Campbell adjusts his lagged variables in an iterative process to find the underlying “normal state vote”. His final root mean squared error for the 15 presidential elections since 1948 is about four percentage points.

The analysis in this paper is focused on states’ positions relative to the national vote, so the dependent variable for the replication of Campbell’s work is the state Democratic two-party

---

<sup>11</sup>The District of Columbia is excluded from this analysis and all analyses in this paper, except for Electoral College simulation.

percentage minus the analogous nationwide. Since, I do not adjust my lagged dependent variable, explanatory variables are transformed from static values to the change from the previous time period. Presidential and vice presidential candidate home state advantage is divided by log of the state population instead of using the conventional method of assigning half weight to the large states of California, Texas, New York, and Illinois. The analysis is restricted to data drawn from post-1976 elections (so that most of the indicator variables can be ignored) through 2000 (so that the 2004 outcomes may be forecast).<sup>12</sup> ADA scores and legislature partisan proportions are taken directly from Campbell’s data. The results are displayed in Table 1; the RMSE is similar to Campbell’s over the same time period (3.7%).

<b>Independent Variable</b>	<b>Coefficient</b>	<b>Std. Err.</b>
Lagged Dep. Var.	0.916	0.0332
Economic growth	0.297	0.0975
Pres. cand. state/ $\ln(\text{pop})$	29.5	12.0
VP cand. state/ $\ln(\text{pop})$	30.8	11.5
Ada score	0.0148	.00816
Pct. Dem. Legislature	9.02	0.0120
Constant	-2.44	9.803
$n = 300$ ; $R^2 = 0.73$ ; $RMSE = 3.69$		

Table 1: OLS regression with dependent variable as state two-party Democratic vote minus national vote (on a [-100,100] scale for ease of interpretation). 1980-2000.

Residuals for more recent presidential elections are lower than their decades-old counterparts. To reduce the estimated variance of the prior distribution, I could restrict the analysis to only the past three election. This alteration would be appropriate if the underlying cause of the heteroscedasticity were more stable state elections. On the other hand, if the causes behind aggregate election behavior were changing over time, a restricted regression’s root mean square error (RMSE) estimate might be biased downward. To determine which of the alternatives is most likely, Figure 1 plots each year’s RMSE from the pooled regressions (solid line) and the RMSE of a regression using only data on single campaign (dashed line). If recent election outcomes were not more predictable but instead based on different factors, the dashed line would stay straight, as the within-election variability would be constant. Both lines appear to be decreasing meaning that recent election outcomes are more probably stable. Nevertheless, to be cautious, data from all elections between

<sup>12</sup>The two post-1976 indicator variables in Campbell’s analysis are not included in my regression.

1980 and 2000 are included in the calculation of the prior; restricting the election range is left as an avenue for future work.

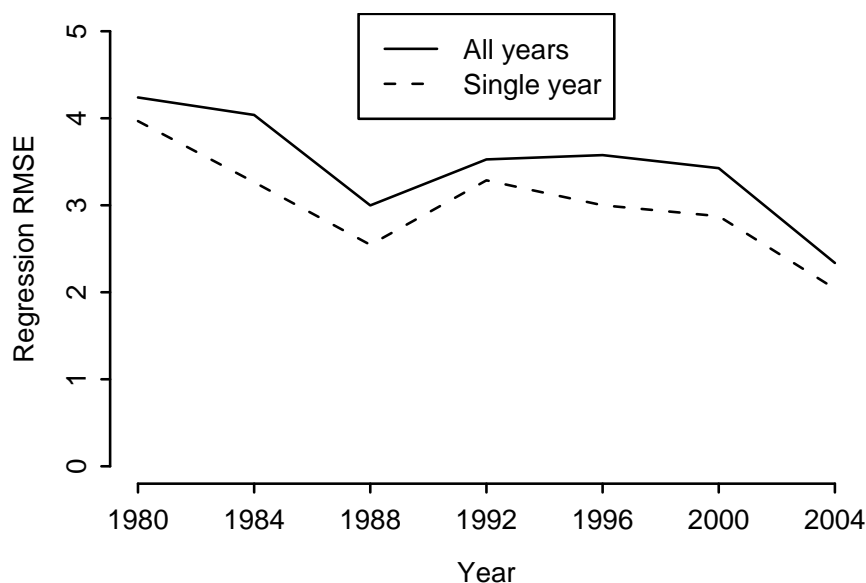


Figure 1: Root mean square error of states in a given year for specified OLS regression. Regression is either based on only that specific year (solid line) all years 1980-2004 (dashed).

## 4 Applying the Polls

Polling during the 2004 presidential campaign is used first to determine that the “reverse walk” model is the most realistic; the data is then used to estimate the model’s parameters. Polls were collected by Kerry campaign staffers primarily via Internet resources, such as National Journal’s Hotline and [realclearpolitics.com](http://realclearpolitics.com). The dataset contains no private Kerry polling, and the time frame begins on March 1, 2004.

### 4.1 National Polls

During the 2004 general election, 22 survey organizations conducted four or more national polls each, for a total of 243 data points. If multiple screeners were used, priority is given to a “likely voters” screen, then “registered voters”, and finally “adults”. If different sets of candidates were included, priority is given to the question that included the most candidates (e.g., Nader).

Figure 2 shows the Gibbs sampler estimate of  $\alpha$  during the campaign for 50,000 iterations

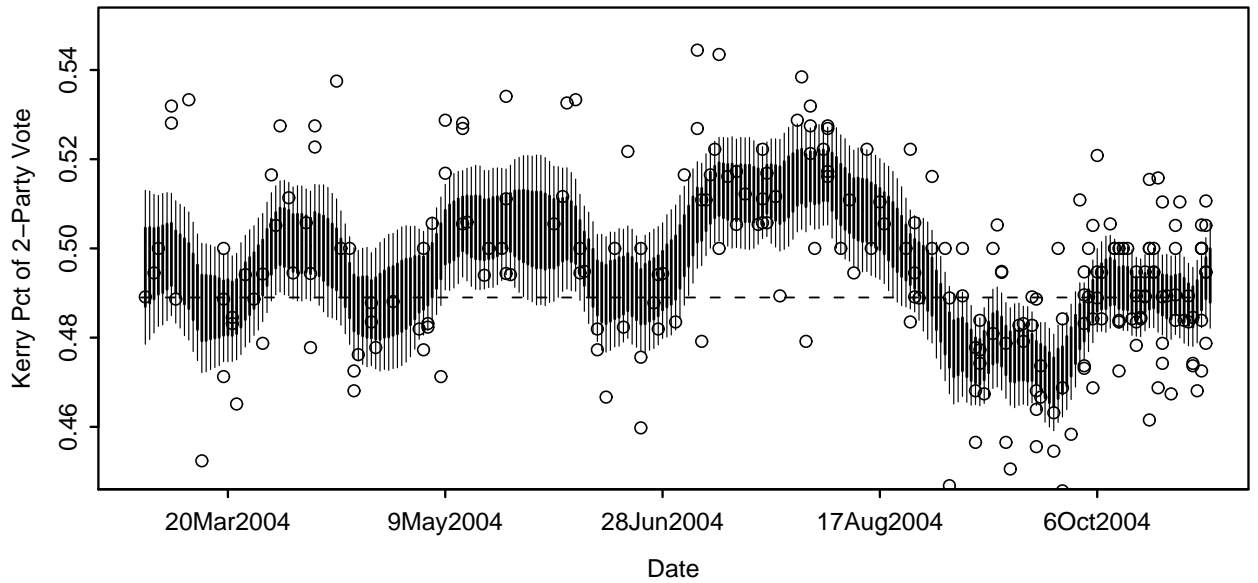


Figure 2: Dynamic linear model estimate of Kerry proportion of two-party vote over the course of the 2004 campaign. Circles represent polls. Horizontal line is final popular vote outcome.

after a burn-in of 5000 and a thinning factor of 100. The bold vertical lines indicate one standard deviation in either direction; the thinner line is the analogous for two standard deviations. The estimate for  $\alpha$  is relatively certain—the average standard deviation is 0.6%.

The polling organization “house effects” are quite small: only two organizations (ARG and Quinnipiac University) have point estimate effect magnitudes greater than a percentage point. The average effect magnitude is 0.4%. The Associated Press has the lowest estimated house effect, followed closely by the Pew Research Center for the People & the Press.

## 4.2 State Polls

Every state was polled at least once during the 2004 presidential campaign, with Florida receiving the most attention and 61 polls. Overall, 690 state survey results were released to the public. If a poll had multiple vote preference question, preference was given to the question that one matched the candidates on the ballot in that state (i.e., Nader included where he met the ballot requirements).

To determine which model type—low variance, offset, or reverse walk—is most appropriate, I analyze the mean square error for the polls over time. The estimate of the national horserace is subtracted from topline Democratic tow-party percentage for each poll; this value is then squared.

This squared error is averaged over a 40-day window and plotted (with 95%  $\chi^2$  confidence intervals) in Figure 3. The dashed line at the bottom of the graph is the baseline variance generated by  $\psi^2$ . Clearly early campaign polls contain more uncertainty than the low variance model would indicate.

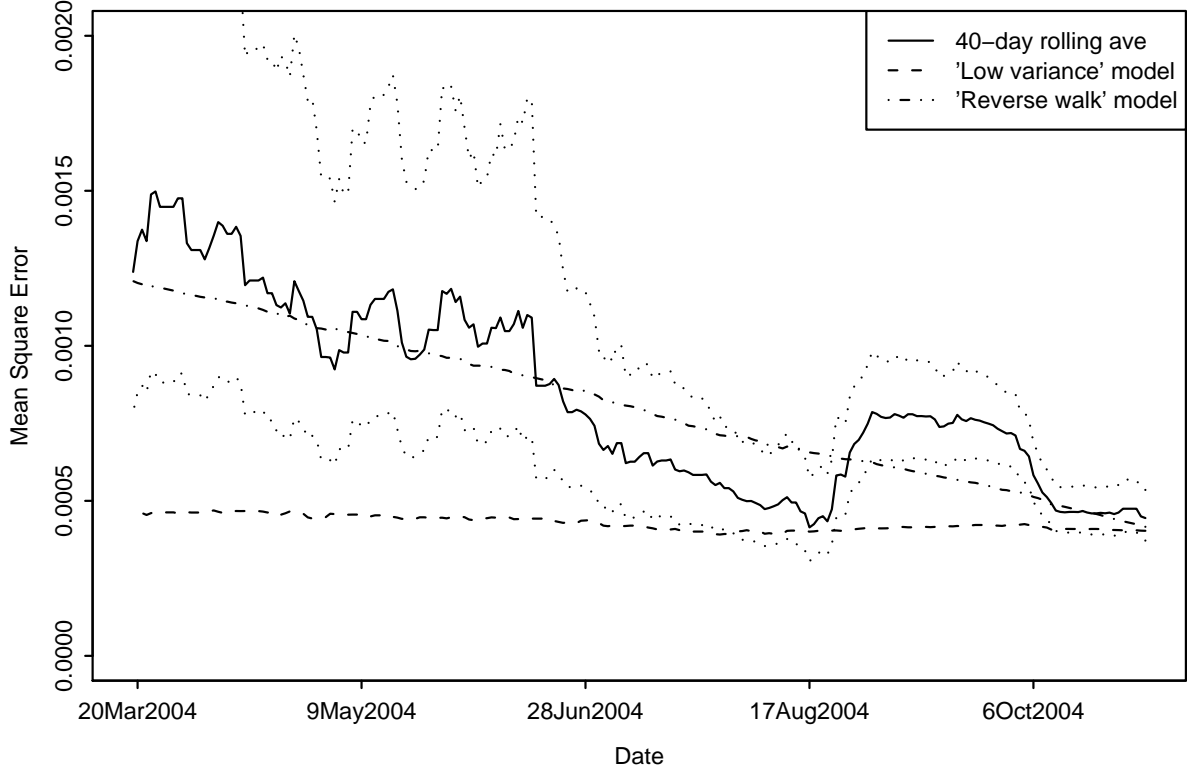


Figure 3: Mean square error of the difference state polls and the final level of state competitiveness relative to the national vote. The reverse walk model, represented by the line moving from the middle-left to lower right, provides the best fit for the data.  $\chi^2$  95% confidence intervals are in gray and dotted.

Given the declining variance in the data, the reverse walk model is much more promising. To estimate  $v$  and  $\sigma$ , Gauss-Lagrange quadrature is employed. The integral's range is from 0 to -20 on the logged- $\sigma^2/v^2$  scale, with  $35^2$  points over the two dimensions. Estimates for  $v$  and  $\sigma$  (in standard deviation “units”) are presented in Table 2 as well as plotted as the dotted line in Figure 3.

A clear interpretation of the point estimate of  $\sigma$  is that a poll 100 days out from the election has an increased margin of error of  $\pm 3.5$  percentage points. The value for  $v$  indicates that polls are only off by 0.3 percentage points more than the reported margin of error. Also, note that the standard error for  $v^2$  is much larger than that of  $\sigma^2$ , providing further evidence that  $\sigma^2$  is more



Parameter	Point Est.	95% Conf. Int.
<b>State Vote Shocks</b>		
$\sigma$	0.00180	[0.00169, 0.00192]
$v$	0.00311	[ $7.28 * 10^{-4}$ , 0.133]
<b>National Vote Shocks</b>		
$\omega$	0.00397	[.00231, .00563]

Table 2: Parameters of the national vote estimate and reverse walk model. Note that “standard deviation” units are reported for clarity.

crucial for the accuracy of the model.

Presidential campaigns often generate news that garners national attentions (e.g., conventions, debates, candidate flubs), and the intuition that national shocks are larger than state shocks is confirmed by the model (Table 2). At the point estimates, national shocks are 2.2 times as large as state shocks; incorporating confidence intervals demonstrates that these two parameters are statistically significant.

## 5 In-cycle Estimation

### 5.1 Forecasting Final State Standings

With posterior distributions for the reverse walk parameters established, I can simulate campaign forecasts at different points in the election season. From March 1st to November 1st, and every 1st and 15th day of the months in between, I forecast the final state deviations. Using these estimates, I simulate potential Electoral College outcomes and predict which states will be pivotal.

This simulation is not perfectly realistic for several reasons, as some information that would not be available to practitioners is allowed into the forecasting. First, the parameter estimates of  $v^2$  and  $\sigma^2$  are based on the full 2004 dataset. A preliminary analysis of the 2000 dataset yields similar parameter estimates, so there is some confirmation of the stability of these distributions.<sup>13</sup> Second, data from the entire campaign is input into the Gibbs sampler to estimate the national horserace. During an actual campaign, the Gibbs sampler would be run every day incorporating any new information. Third, the final turnout results are used as the weights for the state standing renormalizations. A more realistic approach would be to estimate turnout using past turnout and

<sup>13</sup>The 2000 election data is clean and includes several inconsistencies. In future versions of this work, I plan to fix this data and present results from the 2000 polls.

state demographic changes.

In general, the model’s predictions are supported by the final outcomes. The forecast estimates are calculated from the closed form solution and then integrated over the joint posterior distribution of  $v^2$  and  $\sigma^2$  (equation 4). The mean squared errors of the forecasts and the percent of outcomes that fall within the 95% interval are plotted over the course of the campaign in Figure 4. Note that the summer state polls are actually misleading and increase the errors of the forecasts. This results is consistent with Gelman and King’s (1993) finding that the factors ultimately affect election outcomes are predictable, but that these factors are not as important during the summer of presidential campaigns. Thus, they show that early national polls are misleading; there appears to be an analogous phenomenon regarding state voting. Perhaps a more general model than the “reverse walk” approach is needed, where the variance of state standing to the final outcome does not decrease linearly over the course of the campaign. More data and a new underlying theory would be needed to justify this model. Overall, the current model performs well: 96% of all projections fall within the 95% confidence interval, and forecasting error decreases over the course of the campaign.

A modeling change that clearly does adequate forecast the results is aggregating all of the polling data via the “greedy” approach discussed in Section 2.3. Many fewer than 95% of states are forecasted within their 95%-confidence intervals and the mean squared error is larger when one includes polls than when one forecasts drawing only on historical events in all periods except the final period. Attempting to use all the polls is fallacious.

To get a better sense of the forecasts during the campaign, I run Gauss-Lagrange quadrature for each of the 50 Ohio polls. The results, displayed in Figure 5, demonstrate the accuracy of the variance assigned each forecasts. In March, the prior has a large variance and centered below the actual outcome. With each update, the variance decreases, though never so much so that the eventual outcome is not within the 95% confidence interval.<sup>14</sup> By the end of the campaign, the polls are fairly accurate measure of what to expect on Election Day, and the tighter October forecasts reflect this reduced uncertainty.

---

<sup>14</sup>Some of the updates barely meet the 95% confidence interval standard, which is reassuring, since otherwise the interval might be too wide.

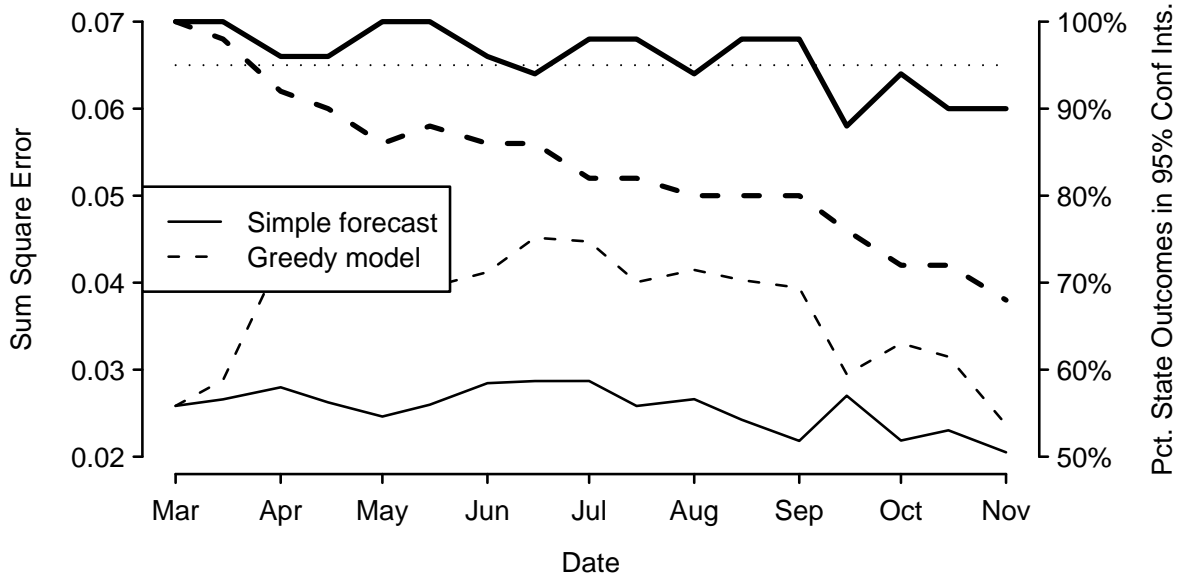


Figure 4: Sum square error (non-bolded lines) and percent of final state outcomes falling within the forecasted 95% confidence interval (bold lines) over the course of the campaign. The simple, two-poll, forecast method (solid lines) outperforms the greedy method (dashed lines) in both metrics.

## 5.2 Simulating the Electoral College

The main goal of this paper is to demonstrate how to forecast state outcomes; yet, these forecasts are little good on their own. Converting state competitiveness to Electoral College importance is crucial to resource allocation. Using the forecasts at the beginning and middle of each month from the previous section, I simulate the Electoral College and find stable pivotal probabilities for each state.

First, I assume a close election: the national vote is a uniform draw between 48% and 52% two-party vote for the Democrat. I apply the state standings relative standing to the national (with uncertainty) vote to this overall horserace number. Finally, I determine the cheapest possible path to Electoral College victory for the “losing” candidate, where cheapest is solely a factor of number of voters. Media costs, overlapping media markets, proportion of undecided are all ignored. If a state is among this path of least resistance, then I consider it a “pivotal” state. Note that multiple states can be pivotal in a single election. Consider the scenario in which a presidential candidate in a neck-and-neck race is six electoral votes away from winning. The most competitive states (that he has not won) are Florida (27 electoral votes), New Hampshire (4 EVs), and New Mexico (5 EVs), which are all a half percentage point below the national vote. Clearly, the cheapest method

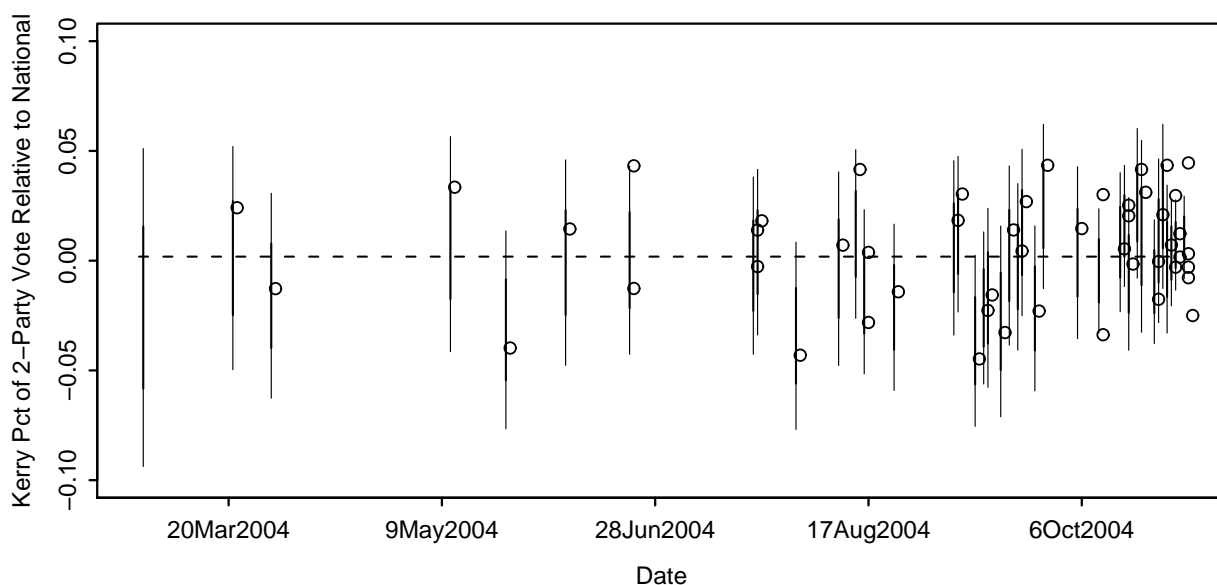


Figure 5: The forecast of final Ohio state standing through the campaign, updated with each additional poll. At every stage, the eventual outcome was within the 95% confidence interval.

to win the Electoral College is to pay for voter advertising, outreach, and persuasion in the smaller two states, because fewer voters must be convinced to defect to the losing candidates side.<sup>15</sup> On the other hand, if the same candidate needs 29 votes to win, the campaign should spend a lot of resources in Florida and New Hampshire. The other restriction placed on the model is that campaigns would not fight in a states with projected margins of losing greater than 4 percentage points. Clearly, this assumption pertains more to the end of campaigns and in elections in which very few of voters are undecided.<sup>16</sup>

After 1,000 Electoral College simulations executed for every two weeks of the campaign, eight states defined the battleground, representing the majority (60%) of pivotal state situations.<sup>17</sup> Figure 6 tracks three of these states—New Mexico, Florida, and Ohio—over time. On Election Day, 2004, Ohio was the crucial state, yet it trailed Florida (a state that ended up solidly Bush) in terms of pivotal likelihood throughout most of the campaign. Note that in the final period, when Ohio converges to the national horserace percentage (Figure 5), its probability of being pivotal jumps.

<sup>15</sup>The possible exception to this rule is candidate visits to media markets, which do not have constant “per voter” cost. I ignore this exception at the present.

<sup>16</sup>There are several more implicit assumptions underlying this method of counting pivotal states. For a fuller discussion of the nature of the Blotto game see Merolla, Munger, and Tofias (2005).

<sup>17</sup>In descending order of percent of the time pivotal: New Mexico, Iowa, West Virginia, Florida, New Hampshire, Ohio, Pennsylvania, Wisconsin.

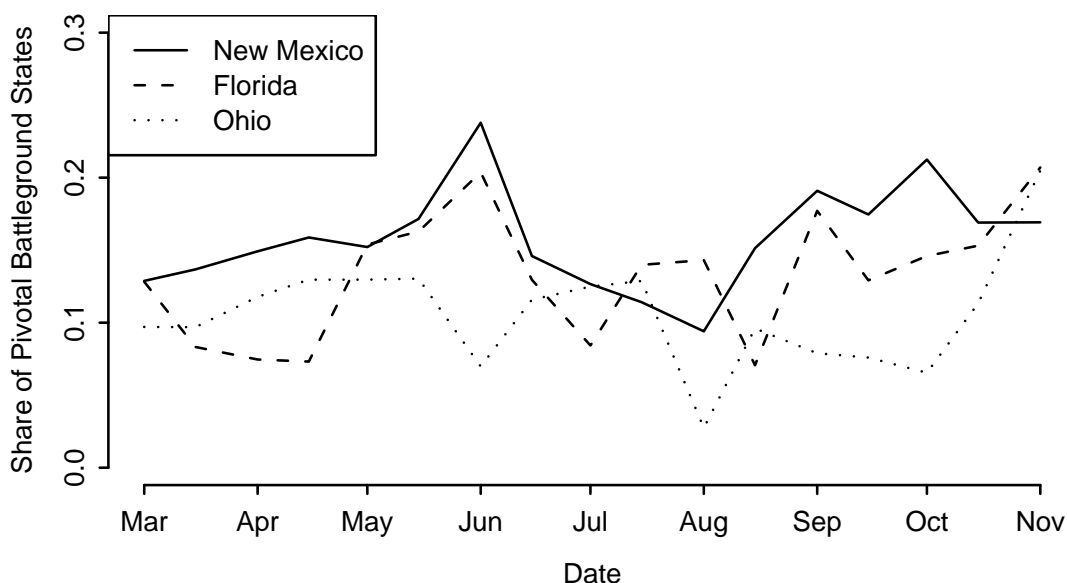


Figure 6: Electoral College simulation results over time. The y-axis value is calculated by dividing the number of times the given state is pivotal over the number of times any of the eight battleground states is pivotal.

In these “probability of being pivotal” calculations, New Mexico and Florida are treated the same if they are both the only pivotal state in a simulation. However, the cost of moving Florida from the loss column to the win column is probably much more expensive than doing the same for New Mexico. Taking cost per voter into account (which the algorithm already calculates) yields a much different resource allocation picture. Figure 7 shows expected costs for the same three states as before; in this calculation, the large states, Florida and Ohio, dominate the landscape.

This procedure confirms the early consensus in the literature that presidential campaigns spend more money on large states rather than small states. What campaigns have become better over the years is determining which states are most competitive and leaving aside the ones that are not. Conditioning on the state’s competitiveness, larger states do attract more campaign investment.

## 6 Discussion and Conclusion

The symmetrical Blotto game has no pure strategy equilibrium. Thus, presidential campaigns hire pollsters in an attempt to gain asymmetric information about the battleground landscape. This knowledge is used to determine where to allocate precious resources, such as advertising dollars.

Relying on current polling data can be hazardous to the health of a campaign. Psychologists

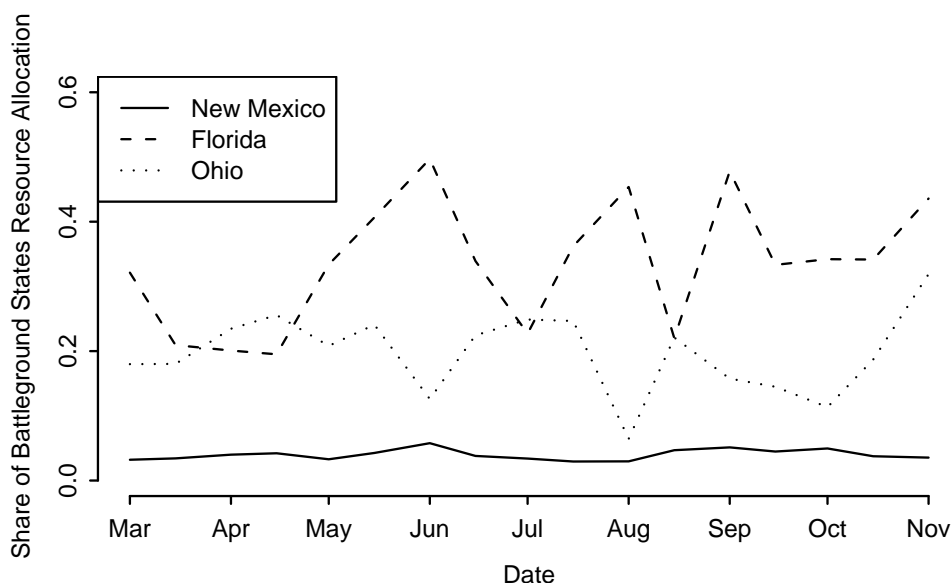


Figure 7: Same Electoral College simulations as Figure 6, but weighted by the cost per voter of winning the state given its current level of competitiveness.

have found that people have a propensity to place too much weight on their current situation when predicting future events—a phenomenon termed “projection bias” (Loewenstein, O’Donoghue, and Rabin, 2003). Bush’s “steal” of West Virginia from Gore in 2000 is an example of one campaign underestimating the potential magnitude of future shocks. Polls conducted during the summer of presidential elections may be misleading to such a degree that practitioners of the Blotto game are better off ignoring them, instead relying on historical regression analysis. When Election Day draws near, however, the polls are do assist with predicting which states will be pivotal.

The forecasting model presented here confirms the work of other scholars that large states do not receive more resources than small states because of a competitiveness advantage (Gelman, Katz And Tuerlinckx, 2002). In fact, small states often offer a cheap path to a few Electoral Votes. Large states, however, do require vast resources to shift the vote a few percentage points relative to other states. Hence, campaigns act sensibly in that they only compete in large states where the vote outcome is near 50%.

Presidential campaigns gain an advantage in the Blotto game of resource allocation *not* by inferring what the current situation is, but by forecasting how the later stages of the campaign may develop. Predicting the future is difficult and should be undertaken with great care. Modeling the current environment as a reverse random walk guides calculations of uncertainty about the future,

thus preventing missteps in the present.

## A Conditional Posterior Distribution of $\sigma^2$

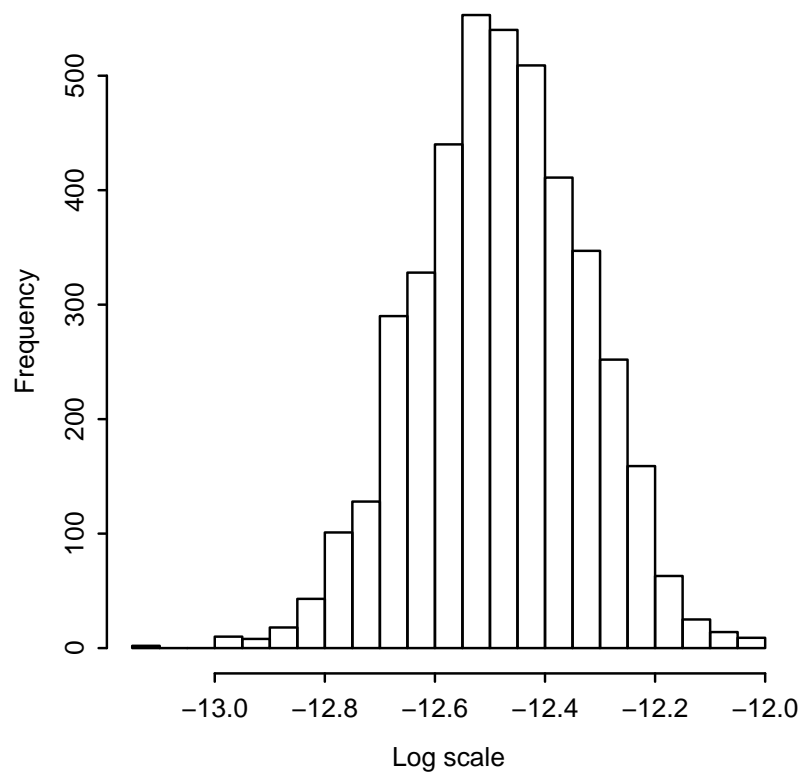


Figure 8: Conditional posterior distribution of  $\sigma^2$  on the log-transformed scale. Results via Metropolis algorithm with 1000 iterations, 5 chains, and 150 burn-in. Jump acceptance rate is 41%.  $v^2$  is set at its point estimate.



## References

- [1] Bartels, Larry M. 1985. "Resource Allocation In a Presidential Campaign". *The Journal of Politics*, Vol. 47, No. 3. (Aug., 1985), pp. 928-936.
- [2] Beck, Nathaniel, Simon Jackman and Howard Rosenthal. 2006. "Presidential Approval: the case of George W. Bush". Presented at the Political Methodology Conference, UC-Davis.
- [3] Brams, Steven J. and Morton D. Davis. 1974 "The 3/2's Rule in Presidential Campaigning". *The American Political Science Review*. Vol. 68, No. 1. (Mar., 1974), pp. 113-134.
- [4] Campell, James E.. 1992. "Forecasting the Presidential Vote in the States" *American Journal of Political Science*, Vol. 36, No. 2. (May, 1992), pp. 386-407.
- [5] Campell, James E., Syed Ali and Farida Jalalzai. 2006. "Forecasting the Presidential Vote in the States, 1948-2004". *Journal of Political Marketing*. Vol 5, Iss 1/2.
- [6] Colantoni, Claude S., Terrence J. Levesque, and Peter C. Ordeshook. 1975 "Campaign Resource Allocations Under the Electoral College" *The American Political Science Review*, Vol. 69, No. 1. (Mar., 1975), pp. 141-154.
- [7] Cuza, Alfred and Charles Bundrick. "Fiscal Effects On Presidential Elections: A Forecast For 2004". Working paper. April 3, 2004. [www.uwf.edu/govt/facultyforums/fiscaleffectsprselect2004.pdf](http://www.uwf.edu/govt/facultyforums/fiscaleffectsprselect2004.pdf). Report of February estimate available: [http://www.forecastingprinciples.com/Political/Cuzans\\_comments.html](http://www.forecastingprinciples.com/Political/Cuzans_comments.html)
- [8] Fair, Ray. 1978. "The Effect of Economic Events on Votes for President." *The Review of Economics and Statistics*. Vol 60, No. 2. pp. 159-173.
- [9] Franklin, Charles H. "Pre-Election Polls in Nation and State: A Dynamic Bayesian Hierarchical Model" Working Paper. Madison May 16, 2001
- [10] Garand, James C. and T. Wayne Parent. 1991. "Representation, Swing, and Bias in U.S. Presidential Elections, 1872-1988" *American Journal of Political Science*, Vol. 35, No. 4. (Nov., 1991), pp. 1011-1031.

- [11] Gelman, Andrew, Jonathan N. Katz and Franci Tuerlinckx. 2002. "The Mathematics and Statistics of Voting Power." *Statistical Science* Vol. 17, No. 4, 420-435.
- [12] Gelman, Andrew and Gary King. 1993. "Why Are American Presidential Election Campaign Polls so Variable When Votes Are So Predictable" *British Journal of Political Science*, Vol. 23, No. 4. (Oct., 1993), pp. 409-451.
- [13] Golman, Russell and Scott E. Page. 2006. "General Blotto: Games of Allocative Strategic Mismatch" Working Paper.
- [14] Gross, O. and R. Wagner. 1950. "A Continuous Colonel Blotto Game." RAND Corporation.
- [15] Jackman, Simon. "Pooling the Polls Over an Election Campaign" *Australian Journal of Political Science*. Vol 4. Iss.40.
- [16] Johnston, Richard, Michael G. Hage, and Kathleen Hall Jamieson. 2004. *The 2000 Presidential Election and the Foundations of Party Politics*. Cambridge University Press
- [17] Merolla, Jennifer, Michael Munger, And Michael Tofias. 2006 "Lotto, Blotto, or Frontrunner U.S. Presidential Elections and The Nature of Mistakes" Working paper. 2006.
- [18] Merolla, Jennifer, Michael Munger, And Michael Tofias. 2005. "In Play: A Commentary on Strategies in the 2004 U.S. Presidential Election" *Public Choice* (2005)
- [19] Owen, Guillermo. 1975. "Evaluation of a Presidential Election Game". The American Political Science Review, Vol. 69, No. 3. (Sep., 1975), pp. 947-953.
- [20] Rosenstone, Steven J. 1983. *Forecasting Presidential Elections*. Yale Univ Press.
- [21] Shaw, Daron. 1999. "The Effect of TV Ads and Candidate Appearances on Statewide Presidential Votes, 1988-96." *American Political Science Review*. Vol 93, No. 2. pp. 345-361.
- [22] Uhlenbeck, G. E. and L. S. Ornstein. 1930. "On the theory of Brownian motion." *Phys. Rev.* Vol. 36, pp. 823-841.