International Journal of Forecasting 31 (2015) 965–979

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Under-performing, over-performing, or just performing? The limitations of fundamentals-based presidential election forecasting

Benjamin E. Lauderdale[a,*], Drew Linzer[b]

[a] Department of Methodology, London School of Economics, United Kingdom
[b] Department of Political Science, Emory University, United States

ARTICLE INFO

Keywords:
Electoral forecasting
U.S. presidential elections
Bayesian statistics

ABSTRACT

U.S. presidential election forecasts are of widespread interest to political commentators, campaign strategists, research scientists, and the public. We argue that most fundamentals-based political science forecasts overstate what historical political and economic factors can tell us about the probable outcome of a forthcoming presidential election. Existing approaches generally overlook the uncertainty in coefficient estimates, decisions about model specifications, and the translation from popular vote shares to Electoral College outcomes. We introduce a Bayesian forecasting model for state-level presidential elections that accounts for each of these sources of error, and allows for the inclusion of structural predictors at both the national and state levels. Applying the model to presidential election data from 1952 to 2012, we demonstrate that, for covariates with typical levels of predictive power, the 95% prediction intervals for presidential vote shares should span approximately ±10% at the state level and ±7% at the national level.

> *It's a paradox. The economy is in the doldrums. Yet the incumbent is ahead in the polls. According to a huge body of research by political scientists, this is not supposed to happen.*
>
> [Niall Ferguson, *Newsweek*, September 10, 2012.]

> *There's a meme out there that Barack Obama's narrow-but-persistent lead is somehow happening despite the state of the economy. I'd say it's largely the reverse—the economy's doing pretty meh and Obama's doing about what you'd expect based on that.*
>
> [Matthew Yglesias, *Slate*, September 11, 2012.]

1. Introduction

One of the primary aims of U.S. presidential election forecasting is to generate expectations about the election outcome prior to the campaign. Candidates and party organizations use these expectations to formulate campaign strategies, while pundits and commentators use them to assess whether the candidates are over- or under-performing in the polls, relative to the current economic and political climate. In recent years, political scientists and economists have developed a variety of regression-based statistical models for predicting future vote outcomes on the basis of historical relationships between "fundamental" conditions and past election results. However, because there is little theoretical consensus as to which fundamental variables are best for prediction, and because many of these models are fitted to as few as 15 pre-

* Corresponding author.
E-mail addresses: b.e.lauderdale@lse.ac.uk (B.E. Lauderdale), drew@votamatic.org (D. Linzer).

http://dx.doi.org/10.1016/j.ijforecast.2015.03.002
0169-2070/

vious elections, the model predictions often diverge substantially. In 2012, the forecasts published in the *PS: Political Science and Politics* symposium on the presidential election indicated Democratic two-party vote shares ranging from 45.5% to 53.8%: anywhere from a decisive loss to a decisive win for President Obama (Campbell, 2012). The reported probabilities of an Obama victory corresponding to these vote predictions ranged from as low as 0.10 to as high as 0.88. Given this variation, it is not clear what anybody should have believed initially about the election outcome.

In this paper, we argue that, while there is value in building and testing fundamentals-based presidential election forecasting models, it is a mistake to take the predictions from any individual model too seriously. Given the limited amount of historical election data currently available, many different models will be justifiable empirically, and each specification will produce a distinct forecast. Moreover, as we show, most of the presidential election forecasts published fail to account for the full range of estimation and specification uncertainty in their underlying models, leading forecasters to overstate the degree of confidence in their expected election outcome. Regardless of a model's point prediction, the predicted probability of a Democratic or Republican victory for most models should be much closer to 0.5 than what is typically reported. This is particularly true in close elections (such as 2000, 2004, or 2012), where the historical data do not decisively indicate either a win or a loss for the incumbent party.

We catalog and describe three major sources of uncertainty that are commonly overlooked when forecasting presidential elections. First, many forecasts misrepresent the total, combined uncertainty in their coefficient estimates and model residuals. We find that these factors alone translate into posterior 95% prediction intervals for the national major-party vote that should span at least ten percentage points. Second, most forecasts neglect the uncertainty associated with the process by which researchers arrive at their model specification. Specification searches are well known to lead to pseudo-parsimonious, overfitted models that overstate the confidence about which variables are most predictive. Bayesian Model Averaging has been proposed as a solution (Bartels & Zaller, 2001; Montgomery & Nyhan, 2010), but it addresses these problems only partially.

Finally, most forecasting models ignore a key institutional feature of U.S. presidential elections, which is that they consist of 51 separate but correlated state-level elections, with outcomes that are aggregated through the Electoral College. This introduces a small, but non-negligible, additional uncertainty into election forecasts—at least, if the goal is to predict the national election winner, rather than the popular vote. At the same time, researchers often underestimate the influence of national-level vote swings on state-level election outcomes. Prediction errors at the state level will be correlated not only over time within each state, but also across states by election. Ignoring this correlation structure can lead to dramatic errors in reporting the uncertainty about future elections.[1]

[1] A potential additional source of uncertainty, though one that we do not consider further, is measurement error in the predictors themselves, especially if the measurement quality has varied over time.

Despite these areas of concern, we find evidence that conventional fundamentals-based presidential forecasting models, which extrapolate from a multitude of economic and political factors, are theoretically robust. There is both practical and scientific merit in discovering which variables, or types of variables, are correlated with election outcomes. When models disagree, analysts can gain insights into the ways in which different modeling assumptions lead to different expectations about the probable election outcome. It can also be instructive to aggregate or combine forecasts from different model specifications.

To produce fundamentals-based forecasts that are more realistic and are accompanied by appropriate statements of uncertainty, we introduce a novel Bayesian presidential forecasting model that can incorporate predictors at the national level as well as the state level. The model, which is based on state-level vote outcomes, includes both state- and election-specific random effects to account for structural features of presidential elections. We impose priors on the coefficients for predictor effects that enable us to consider many more independent variables than would be possible in a classical approach. Once the model has been estimated, the posterior distributions of the forecasted state-level and national-level popular and electoral vote shares reflect the coefficient uncertainty, specification uncertainty, and Electoral College uncertainty. The posterior from the model can also constitute a historical prior from which one can begin to incorporate polling data about a forthcoming election (e.g., Linzer, 2013).

We apply the model to state- and national-level presidential election data from 1952 to 2012. Because we do not know which political and economic performance measures are the most predictive, we investigate a series of specifications that employ different combinations of independent variables, as well as random placebo predictors. Our results demonstrate that there is not sufficient historical evidence to warrant strong, early-campaign assessments about the probable outcome of a presidential election. They also highlight the fundamental difficulty of estimating the effects of national-level variables from only 16 elections. However, past elections provide a fair amount of evidence about relative state-level vote outcomes, and our model enables a reliable estimation of the effects of state-level predictors, such as presidential and vice-presidential candidate home states and party convention locations.

2. Problems with existing forecasting methods

In the standard approach to forecasting U.S. presidential elections, researchers specify a linear or non-linear multiple regression model with past elections' vote outcomes (expressed as the incumbent party candidate's share of the major-party vote) as the dependent variable, y, and a small set of political or economic "fundamentals" as the independent variables, X. Fitting the model to data from elections $1 \ldots T-1$ produces coefficient estimates $\hat{\beta}$ that indicate the effects of each predictor. Researchers then insert the observed values of the independent variables for the current election, x_T, into the fitted model equation, to calculate a predicted vote share, $\hat{y}_T$. This prediction is used to estimate the ultimate quantity of interest: the probability that either candidate will win the presidency. Of the 13 presidential election forecasting models published in the 2012 *PS* symposium, 12 followed this procedure (Campbell, 2012).

We show that, in practice, most researchers overstate the confidence in their fundamentals-based forecasts, exaggerating the probability that either candidate would be expected to win. The reasons for this are both statistical and substantive. Statistical errors relate to inaccurate reporting of the uncertainty in a fitted model, which may be only one of many models considered by the researcher. Substantively, many models oversimplify institutional features of the U.S. presidential election system that are relevant to prediction.

2.1. Statistical errors

2.1.1. Coefficient uncertainty

Most forecasters are not explicit as to the way in which they derive candidates' win probabilities from their fitted models. For models that predict the national vote outcome using linear regression, it appears that researchers generally assume a normal distribution of errors around a mean $\hat{y}_T$, with a standard deviation equal to $\hat{\sigma}$, the estimated conditional standard deviation of y given X. The probability that the incumbent party candidate will win is taken to be $1 - \Phi((0.5 - \hat{y}_T)/\hat{\sigma})$.

However, this calculation neglects models' estimation uncertainty. When predicting y_T at a new observation x_T, the standard error of prediction σ_p and the $1-\alpha$ prediction interval (PI) around $\hat{y}_T$ are:

$$\sigma_p(x_T) = \hat{\sigma}\sqrt{1 + x_T'(X'X)^{-1}x_T} \tag{1}$$

$$\mathrm{PI}_{1-\alpha} = \hat{y}_T \pm t_{\alpha/2,n-k-1} \cdot \sigma_p(x_T). \tag{2}$$

Note that $\sigma_p(x_T) > \hat{\sigma}$; the uncertainty in an out-of-sample prediction is always greater than the point estimate of the standard deviation of the residuals. However, of all of the forecasts of the 2012 election summarized by Campbell (2012), only Klarner (2012) addresses estimation uncertainty explicitly. The others do not appear to use the proper prediction interval.

To illustrate the consequences of this distinction, consider the well-known Time-for-change forecasting model of Abramowitz (2008). Fitting the model to presidential election results from 1948 to 2008, and inserting observed values of the independent variables from 2012, we calculate Obama's predicted share of the national two-party vote to be $\hat{y} = 52.25\%$, with $\hat{\sigma} = 1.98$. If this point prediction is treated as being estimated without error, it implies that Obama would win the popular vote with a probability of 0.87. However, once the estimation uncertainty in β is taken into account, the probability of an Obama victory falls to 0.68, with a 95% prediction interval of (47.5, 57.0), covering outcomes all the way from a decisive defeat to a near-landslide victory. This is a strikingly weaker conclusion than when we ignored the coefficient uncertainty: two to one odds in favor of Obama, rather than seven to one.

Forecasting models other than linear regressions may not have a ready formula for the prediction interval.

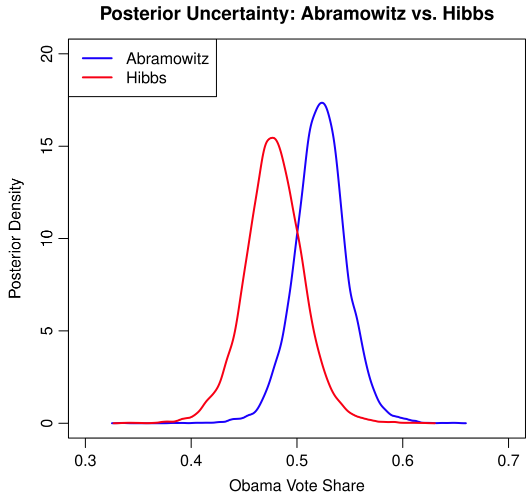

Fig. 1. Posterior predictive distributions for the 2012 Obama vote share under two three-parameter specifications.

In these cases, it is possible to generate equivalent calculations for any likelihood model by either analytic approximation (King, 1991) or simulation (King, Tomz, & Wittenberg, 2000). An analysis by Bayesian posterior simulation takes this uncertainty into account automatically, and is our preferred method, given that the multivariate normal approximation to the likelihood may not hold with small numbers of observations. The Hibbs (2012) forecast, for example, is based on a nonlinear regression model that includes a parameter for weighting the relative importance of more recent versus less recent changes in real disposable income. We simulate the Hibbs model's Bayesian posterior prediction interval for the 2012 election, placing uniform priors on all parameters. Whereas Hibbs (2012) reports a 10% probability of Obama victory, with a point forecast that Obama would receive 47.5% of the major-party vote, we calculate a 21% posterior probability of an Obama win. Thus, the odds against Obama decline from nine to one to four to one.

Once the estimation uncertainty has been taken into account, the posterior predictive distributions for the Abramowitz and Hibbs models overlap considerably, despite the five-point difference in their point forecasts (Fig. 1). Perhaps as important as this methodological point is the resulting substantive point about how little information is contained in the national-level election data. With three estimated parameters, and using data starting in 1948 or 1952, the models' 95% posterior predictive intervals for the national vote shares span at least 10 percentage points in the two-party vote, or 20 percentage points of margin.

2.1.2. Specification uncertainty

There is no consensus as to which combination of variables is most appropriate for predicting U.S. presidential elections. In the past, researchers have employed a diverse range of economic indicators, including the real per capita GDP, real disposable personal income, and unemployment rates, among others; while these, in turn, have been measured in the election year, in each year of the

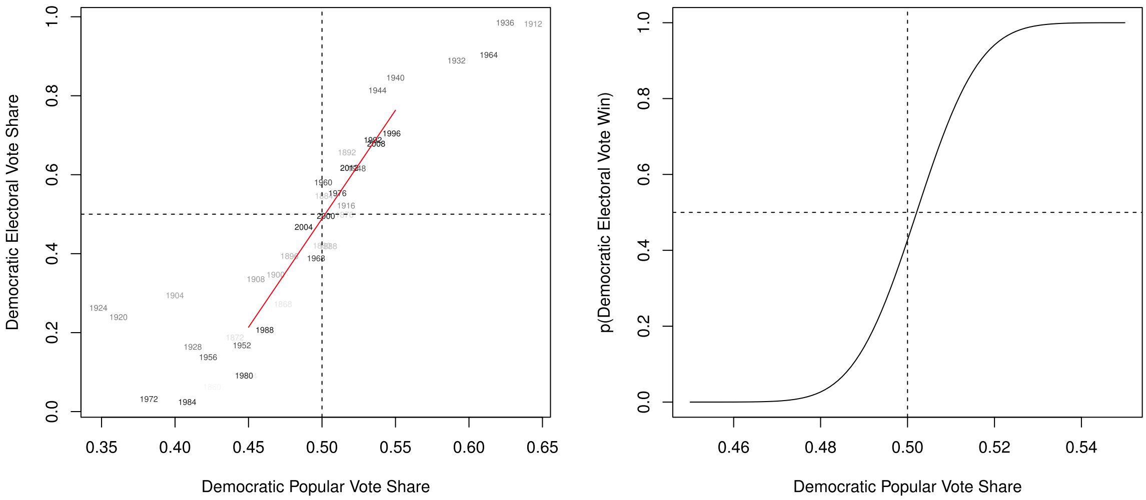

Fig. 2. Left: Electoral vote against the popular vote for U.S. presidential elections since 1856, with more recent elections shown in darker text. Right: Approximate probability of electoral vote victory as a function of the popular vote share, for a Democratic candidate, based on the linear regression in the left panel.

president's term, or as a weighted average over the entire term. Likewise, there are many ways in which the incumbency status of a candidate or party might matter: as the number of terms in office for the incumbent party, whether the incumbent party has held office for at least two terms, or whether the incumbent president is running for reelection (Abramowitz, 2012; Achen & Bartels, 2004; Berry & Bickers, 2012; Holbrook, 2012). All of these variables are intercorrelated, and there are plausible theoretical justifications for each; yet we have very little evidence upon which to base a selection. The only honest assessment that we can make is to say that we cannot determine which variables are best, except to the extent that the data can adjudicate between them. Unfortunately, because of the small size of most election datasets, different model specifications will often fit the data similarly well.

One approach to resolving this specification uncertainty is the use of Bayesian Model Averaging (BMA) (Bartels, 1997; Montgomery & Nyhan, 2010). BMA takes a set of forecasting models, each with its own combination of independent variables, and estimates a hierarchical model in which the data are assumed to come from exactly one of those specifications, unknown to the researcher ex ante. BMA then averages over the uncertainty about which one of these models generated the data.[2] Montgomery, Hollenbach, and Ward (2012) used this approach to construct an ensemble prediction from the other forecasting models published in the October 2012 *PS* symposium. While BMA provides a more accurate way of accounting for the specification uncertainty, it is still based on the assumption that one of the constituent models is correct. This is unlikely to be the case for U.S. presidential elections, where the outcome depends upon a large number of interrelated factors.

2.2. Substantive errors

2.2.1. Popular vs. electoral votes

U.S. presidential elections consist of 51 distinct state-level elections. Although it is more commonplace (and simpler) to forecast national-level than state-level vote shares, this ignores the intermediate role played by the Electoral College. The standard argument for forecasting only national vote outcomes is that reversals between the national vote and the electoral vote are unlikely. However, this has happened in three presidential elections out of 57; 5% is not all that rare![3]

Aggregating votes through the Electoral College adds an uncertainty to presidential election outcomes that purely national-level forecasting models will miss. The easiest way to see this is to examine a plot of electoral vote shares versus popular vote shares in past elections (Fig. 2). Despite the limited availability of historical data, it is evident that there is a non-trivial variation in electoral vote shares in the region around a tied popular vote election. If we run a regression on the close elections (vote shares from 0.45 to 0.55) and assume normal residuals, we can calculate an approximate probability of an Electoral College reversal. It appears that an Electoral College reversal of the popular vote is not highly improbable when the margin of victory in the popular vote is fewer than three points, corresponding to Democratic candidate vote shares of between 0.485 and 0.515. Therefore, the uncertainty in a forecast of the election winner based on national vote shares should factor in not only uncertainty about the point prediction of the vote share, but also the possibility that predicted vote shares of above 0.5 may nevertheless result in Electoral College losses (and vice-versa).

2 Imai and Tingley (2012) propose using a hierarchical mixture model in which different observations could arise from different data generating processes. Their approach is statistically similar to the hierarchical categorical model assumed by BMA, but is offered as a tool for model comparison rather than for capturing model uncertainty.

3 One could argue with both the numerator and the denominator. All three of the reversals were in disputed elections, which is hardly surprising given that a close election makes a reversal much more likely; and the popular vote is unsystematic for many of the early elections, so the denominator should perhaps be reduced.

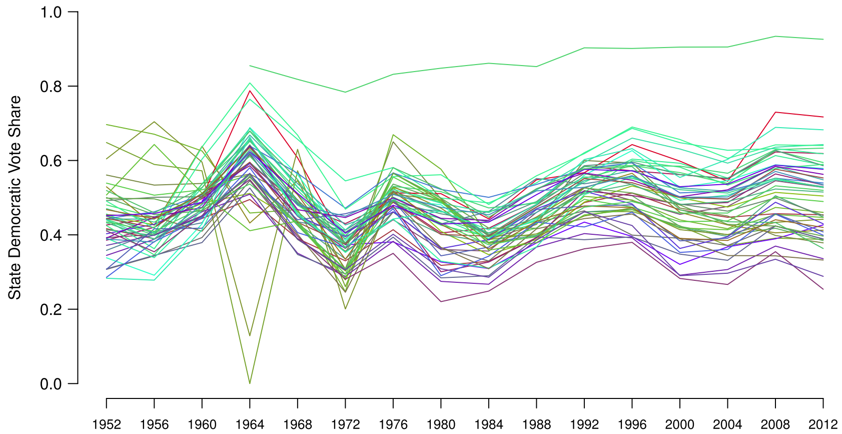

Fig. 3. State-level Democratic two-party vote, 1952–2012. State trendlines are colored by geographic location, with proximate states being more similar. (For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2.2. National-level swing

It is straightforward to incorporate the Electoral College into an election forecast if a model predicts state-level vote shares rather than national-level vote shares. However, the 51 state-level elections are not independent, either within or across election years. From election to election, national-level swings induce correlated errors across states, as vote shares at the state level tend to rise and fall in tandem (Fig. 3). The exceptions are primarily Southern states between 1964 and 1972. Some of the near-parallel movement is predictable from national-level variables, but some is not. The unpredictable portion can be conceptualized as either a year-specific random effect or an error correlation across states by election. The consequence of this is that a pooled linear regression of state-year vote shares as a function of national- or state-level covariates will overstate the confidence in the model prediction. Of the three state-level analyses in the 2012 *PS* symposium, two ignored national-level swings (Berry & Bickers, 2012; Jerôme & Jerôme-Speziari, 2012), while one included year-level random effects (Klarner, 2012).

Individual states also have persistent tendencies to support either Democratic or Republican presidential candidates. These partisan leanings cannot be addressed by state fixed effects, as many have changed gradually but substantially over the last half-century. Fig. 4 shows state-level presidential vote shares since 1956 as a function of vote shares in the previous election. The most important predictor of the relative vote across states is the relative vote across states in the previous election. As such, Berry and Bickers (2012) and Klarner (2012) specify models using a lagged dependent variable. However, the degree of stability in the ordering of states' presidential vote shares has increased substantially since 1976, implying that the coefficient on the lagged state-level vote is itself variable. Thus, it is not surprising that Klarner (2012) estimates a much weaker lag coefficient (0.85) on his analysis of 1952–2008 than do Berry and Bickers (2012) on their analysis beginning in 1980, when more state-level economic data are available (0.99). While more years of data are typically preferable, the inclusion of the much more fluid political geography of the period before 1976 may have led Klarner (2012) to underestimate the coefficient on the lagged dependent variable for prediction in the current era of near-uniform swing, with consequences that are not clear.

2.3. Summary

Statistical forecasts are superior to ad hoc forecasts because they use historical data systematically to make out-of-sample predictions; but forecasts are only as reliable as the assumptions upon which they are based. While no model can capture all of the relevant features of a problem, we have demonstrated that the limitations of the forecasting models published for the 2012 U.S. presidential election have non-trivial consequences for the quality of their predictions.[4] The most severe problems arise with the reporting of uncertainty, but many of the existing models can be improved further by refocusing their analysis on the state level, and employing more flexible specifications that are truer to the known features of presidential elections.

3. A state-level Bayesian forecasting model

We describe a highly customizable Bayesian regression approach for forecasting U.S. presidential elections that addresses each of the statistical and substantive issues that we have raised. However, although the model explicitly integrates a range of features of the U.S. election system, variable selection remains a critical concern. When forecasting elections, there are many more plausible predictors than observed elections. In a likelihood framework, it is not possible to estimate the effects of all of these variables simultaneously. Instead, we place a Bayesian prior on the regression coefficients that functions similarly to ridge or lasso regressions for linear models (Tibshirani, 1996;

4 We would be remiss if we did not recognize Klarner (2012) as being subject to the fewest of these criticisms.

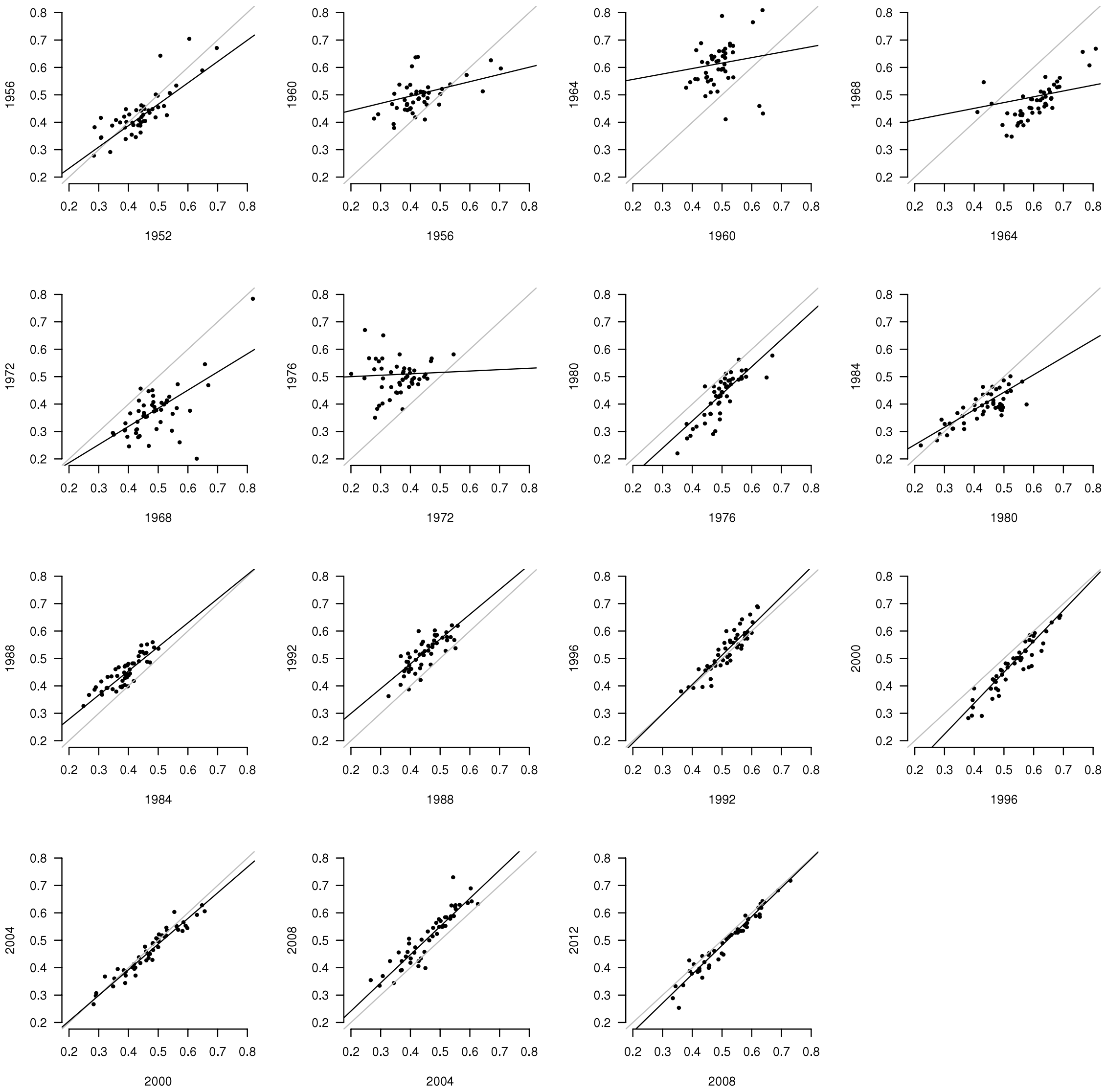

Fig. 4. State-level Democratic two-party vote shares in consecutive presidential elections, 1952–2012. Grey lines correspond to no change; black lines show a total-vote-weighted regression line.

Tikhonov, 1943).[5] These regularizations of the standard regression model impose different forms of the assumption that none of the coefficients are especially large.

Our estimation strategy allows the data to reveal which of a potentially large set of candidate variables are the most predictive, thus allowing forecasters to remain agnostic about the economic and political factors that "really" matter for predicting elections. For example, it is possible to include several economic performance measures at the state and national levels, even if that leads to a model with more predictors than election years. We do not expect to be able to make strong statements about the exact way in which each variable contributes to the forecast, but this is not the primary objective. Rather, our goal is to limit the model dependence across alternative specifications, and properly convey the uncertainty in any given election forecast. To the extent that we are nonetheless interested in which variables predict the outcome best, we will learn something—but without the false confidence that arises from a specification search.

3.1. Specification

Let y_{st} denote the Democratic share of the major-party vote in state s and election t. Extrapolating from elections $t = 1 \dots T - 1$, we wish to forecast the outcome of an upcoming election T, represented by unknown vote shares y_{sT}. We model y_{st} as a function of the national- and state-level covariates X and Z, and the state- and election-level

5 Ridge regressions and lasso regressions are equivalent to Bayesian regressions estimated by posterior maximum with independent Gaussian (ridge) and Laplacian (lasso) priors on each of the regressors (Park & Casella, 2008).

random effects α and δ:

$$y_{st} = \alpha_{st} + \sum_k \beta_k x_{kt} + \sum_l \gamma_l z_{lst} + \delta_t + \epsilon_{st}. \quad (3)$$

The coefficients β_k are the effects of national-level variables on vote shares in all states, while the coefficients γ_l are the effects of state-level variables on the relative vote shares between states.

The random effect α_{st} captures states' persistent tendency to vote for one party or the other. States' party affinities drift over time, so we allow the magnitude of α_{st} to vary from one election to the next as a Bayesian dynamic linear model (West & Harrison, 1997):

$$\alpha_{st} \sim \mathcal{N}(\alpha_{s,t-1}, \sigma_\alpha) \quad (4)$$

$$\sigma_\alpha \sim \mathcal{N}_{1/2}(\sigma). \quad (5)$$

α_{st} can be conceptualized as a state's "normal vote", yet one that evolves over time. Values of α_{st} that are closer to one indicate states that have a Democratic lean, while values of α_{st} that are closer to zero indicate states that have a Republican lean. The rate of change in these normal votes is determined by an estimated σ_α, over which we place a half-normal prior (Gelman, 2006). In addition, we impose a normalization constraint on the α_{st} values, requiring them to have a population-weighted mean equal to 0.5, so that they can be interpreted directly as the expected Democratic vote share for each state in a tied election, before any election-year modeled or un-modeled shocks.

A second random effect, δ_t, captures common, national-level shifts in states' vote shares in election t, beyond what is predicted by the national-level covariates X. We let

$$\delta_t \sim \mathcal{N}(0, \sigma_\delta) \quad (6)$$

$$\sigma_\delta \sim \mathcal{N}_{1/2}(\sigma). \quad (7)$$

A failure to include this election-specific effect would lead to over-confident forecasts, as it would rule out the possibility of idiosyncratic shifts of all states towards the same party. Positive values of δ_t indicate years in which state-level Democratic vote shares exceeded expectations relative to the variables in X. Negative values of δ_t indicate the reverse.

The model allows researchers to include both national-level variables X and state-level variables Z as predictors of the election outcome. However, to preserve the proper interpretation of β_k and γ_l, the predictor variables must be transformed in two ways. First, our specification requires all variables to be standardized so that the scales of the variables are comparable. Second, because the model is expressed in terms of the Democratic Party vote share (y_{st}), it is necessary to multiply variables that predict the incumbent vote share relative to the challenger vote share by an incumbent party indicator P_t, which equals 1 if the incumbent president in election t is Democratic and -1 if the incumbent president is Republican. This ensures that "performance" variables affect Republican and Democratic candidates symmetrically, as in a typical incumbent party vote regression. Positive coefficients on performance variables indicate predictors that have a positive association with incumbent party vote shares.

Variables can be included in the model directly (that is, without multiplying by P_t) if they are expected to predict the Democratic vote share relative to the Republican vote share, independently of incumbency status. We term these "non-performance" or "partisan" variables. Thus, for example, if one thought that high unemployment always hurt Republicans relative to Democrats, one would include it without multiplying by P_t; but if one thought that high unemployment always hurt incumbents relative to challengers, one would multiply it by P_t. For state-level non-performance variables such as candidate home states or party convention locations, the coding for a state is 1 for the Democrat and -1 for the Republican, or 0 otherwise. We assume that non-performance variables have equivalent effects on Democratic and Republican candidates. Finally, we include a constant term $x_{0t} = P_t$ to capture any advantage to the candidate from the incumbent presidential party when all other x values are zero.[6]

To guard against over-estimating the effects of the predictors and over-fitting the model, we assume priors for the national-level and state-level regression coefficients that are normally distributed with mean zero, and place a half-normal prior on the standard deviation of that distribution:[7]

$$\beta_k \sim \mathcal{N}(0, \sigma_\beta) \quad (8)$$

$$\sigma_\beta \sim \mathcal{N}_{1/2}(\sigma) \quad (9)$$

$$\gamma_l \sim \mathcal{N}(0, \sigma_\gamma) \quad (10)$$

$$\sigma_\gamma \sim \mathcal{N}_{1/2}(\sigma). \quad (11)$$

This approach allows us to use many more predictive variables than would be possible in a standard analysis, as we are estimating the distribution of coefficient magnitudes and constraining the coefficients accordingly. Our expectation that some (but not all) of these economic and political variables matter is reflected in their common prior.

The final error term, ϵ_{st}, is a conventional state-election-specific random effect (an uncorrelated error term) capturing any remaining variation in y_{st}:

$$\epsilon_{st} \sim \mathcal{N}(0, \sigma_{\epsilon,t}) \quad (12)$$

$$\sigma_{\epsilon,t} \sim \mathcal{N}_{1/2}(\sigma). \quad (13)$$

We allow the magnitude of the state-level errors to vary across elections, as there is substantial historical variation in the degree to which states follow their previous behavior. The most striking example is the 1964 election, in which the pattern of the South voting more Democratic than the rest of the country reversed itself for the first time since the end of Reconstruction. The state-level 1964 vote

6 Since P_t is the same in every state, no incumbent party intercept is included in Z, as it would be redundant because of the one in X.

7 This is a more parsimonious prior than that imposed by a Laplace/double-exponential distribution. The double-exponential distribution is a scale-mixture of normals with an exponential distribution on the variance (Park & Casella, 2008), which is equivalent to a scaled χ^2_2 distribution. The half-normal distribution on σ_β corresponds to a χ^2_1 distribution on σ^2_β, which places more density near zero and at large values of β.

shares have almost no correlation with those from 1960, but many of these deviations failed to survive to 1968. To the extent that these shifts were part of the beginning of the southern realignment, this is captured by the evolving state-level normal votes α_{st}, but there was also a great deal of fluctuation in the state-level votes that was not persistent.

The standard deviations σ_α, σ_β, σ_γ, σ_δ, and $\sigma_{\epsilon,t}$ of the priors for the various additive terms in the model all depend on a common hyperparameter σ, which describes the aggregate variation in state vote shares over time across all sources of variation. We place a standard uniform prior on σ.

3.2. Estimation and forecasting

To generate an election forecast, we observe the predictors x_{kt} and z_{lst} through to the current election, T, and historical election results y_{st} through to election $T-1$. We then assume that the degree of state-level noise $\sigma_{\epsilon,T}$ in the upcoming election will be the same as in the previous election. This is an important assumption, as the fluctuation in the relative Democratic two-party vote shares across states has been quite low since 1976, and our forecasts depend upon this remaining the case. In comparison, states' election-to-election fluctuations from 1956 to 1972 were very large, which meant that predicting the relative Democratic vote share would have been tremendously difficult. In principle, it would be possible to model the historical trend in this variable, but we do not pursue this additional complexity. If a shock were to upset the more stable pattern in recent years, it would undermine our ability to make confident projections.

We treat the "missing" state-level vote shares y_{sT} and the other quantities of interest in election T as a set of unknowns to be estimated. We simulate the posterior distribution of these and the other parameters of the model using a Markov chain Monte Carlo algorithm implemented in JAGS (Plummer, 2008; R Core Team, 2013). Because there is a substantial posterior correlation across parameters, a large number of iterations is required in order to achieve a reliable sample from the posterior distribution.

The fitted model produces many quantities of interest. The state-level Democratic popular vote share forecasts are y_{sT}. States won by the Democratic candidate are a binary variable $W_{sT}=1 \iff y_{sT}>0.5$. To construct the national two-party vote forecast from the state-level y_{sT}, we need to make an assumption about the relative numbers of two-party votes in each state. Relative state vote totals change very slowly over time, so we assume that the new election will have the same relative turnout as the previous election. Letting the fraction of the national two-party turnout be Q_{sT} for state s in election T, the national Democratic popular vote is $\sum_s Q_{sT}y_{sT}$. Letting E_{sT} represent the number of electoral votes for state s in election T, the national Electoral College result is $E=\sum_s E_{sT}W_{sT}$. The proportion of posterior draws of E that are 270 or greater – a majority of the 538 electoral votes – is taken as the Democratic candidate's probability of victory.

3.3. Features, limitations and extensions

The regularization of national-level variable coefficients allows researchers to include all plausible predictors, and thus capture the real uncertainty about which matter and how. However, while our model makes it easier to be honest about the forecasting uncertainty, it does not prevent researchers from fishing for variables that will yield strong forecasts, if desired. It is impossible to have a modeling framework that is responsive to the data but will not make strong predictions for at least some predictor. One can still choose which variables to include, and this can still influence the resulting forecast. With a sufficiently clever specification search, one might come up with extreme predictions. However, these specifications will be harder to find and more difficult to defend with our approach. No one using our framework can reject the inclusion of an additional variable with their favored three or four by appealing to the limitations of classical regression with small numbers of correlated predictors.

There are many potential generalizations to the model that could be used to capture additional features of U.S. presidential elections. First, it is possible to introduce correlations across states as a function of geography. For example, the large deviations from historical trends in the 1964 state-level elections were strongly geographically concentrated. We explored the idea of implementing this extension using a variogram approach, but the added model complexity had a minimal effect on the forecasts. Including this variation might be more useful if one was supplementing historical data with polling data, as it would enable state-level polls not only to indicate the state-level vote relative to the national-level vote, but also to indicate something about the probable vote in neighboring states.

Second, our specification assumes that state-level predictors have non-persistent effects: the persistent component of the state-level vote α is not modeled using data. This is in contrast to lagged dependent variable models, which (when the lag coefficient is large) assume that nearly all of the effect of state-level variables persists into future elections. We suspect that the truth is somewhere in between the two: the effects of presidential performance or candidate home states fade, but persist in part through the accumulation of party reputations, for example. Modeling α with performance variables could address these possibilities. Further, the inclusion of demographic information in a model for α could increase the predictive performance of the model in cases where demographic trends are shifting a state predictably relative to others, as has occurred recently with the increasing Latino population in the Southwest, and the increasing population in Virginia and North Carolina with higher education.

Finally, we assume an additive structure in our model, and implicitly impose a uniform swing on the state vote shares by including national-level variables and random effects in the state-level vote specification. Since 1976, this has been a very good approximation to reality; before 1976, it was less so. The almost complete inversion of the relative state-level votes between 1956 and 1964 is not captured in the model in any direct way, but manifests itself as large standard deviations $\sigma_{\epsilon,t}$ on the error term for some of the elections in that period. While this has limited consequences for forecasting in the current period of stable relative state votes, it reveals that our model might perform quite poorly at predicting relative state-level votes in the face of a substantial realignment. Of course, any purely historical analysis would perform poorly in the event of a sudden realignment, which is why updating historical models with polling data is vital for the development of successful forecasts close to an election (e.g., Linzer, 2013).

4. Application: the 2012 presidential election

We use the model to forecast the 2012 U.S. presidential election based only on data that would have been available in the summer of the election year. Our dataset contains data on every presidential election since 1952. The model results include predictions of the state and national vote outcomes, the Electoral College vote, and the uncertainty in each of those estimates, which then produces an estimate of the probability that Democrat Barack Obama would be re-elected over Republican challenger Mitt Romney. We then compare our predictions to the known outcome. In fitting the model, we also estimate the substantive effects of a range of theoretically motivated national- and state-level structural variables on the vote outcome. To test the sensitivity of our forecasts to the choice of predictors, we perform a pair of robustness checks, first fitting the model with no predictors, and then adding placebo predictors that are generated randomly so as to have no systematic association with vote outcomes.

4.1. Election predictors

At the national level, we use a series of economic performance measures drawn from previous election studies: growth in personal disposable income in the final two years of the president's term; the national unemployment rate in the third and fourth years; the national inflation rate in the third and fourth years; and the growth of GDP from the first to the second quarter of the election year, as per Abramowitz (2012). To capture the political context, we include three variables: whether the sitting president is running; the number of previous terms in office for the incumbent party; and whether the incumbent party has already been in office for two (or more) consecutive terms. Because of the high degree of correlation between these political variables, we do not expect the model to reveal which matter "most". We also include the Hibbs (2012) military fatalities variable.[8]

Most state-level economic variables are not available for presidential elections prior to 1980. One exception is the state-level average personal income growth over the entire term, which we include.[9] We also examine the effects of the home states of the presidential and vice presidential candidates, the location of the presidential nominating conventions, and the party of a state's governor.

4.2. Predictor estimated effects

Estimates of the effects of the national-level predictors are subject to a large amount of posterior uncertainty; none of the coefficients have central 95% posterior intervals that exclude zero (Fig. 5). Nonetheless, these variables collectively provide predictive power, and the posteriors of the individual coefficients are consistent with the current theory and evidence on retrospective voting. The strongest predictor of the vote among the variables that we have included in the model is disposable income growth in the election year (year 4), followed by income growth in the third year of the president's term, and the growth in GDP from the first to the second quarter of the election year. Candidates from the incumbent party tend to perform better in their first attempt at reelection, but this advantage diminishes the longer the party has been in office. Military fatalities are associated negatively with voting for the incumbent party presidential candidate. There is little evidence that unemployment or inflation rates have much predictive value, though higher unemployment in year 4 and higher inflation in years 3 and 4 are associated negatively with incumbent vote shares.

The effects of the state-level variables are estimated with a much higher level of precision. Transforming the normalized coefficient estimates back into their original units, we estimate that, in their home states, presidential candidates win an additional 3.6% of the vote, with a 95% posterior credible interval of (2.2%, 5.0%). Candidates also receive an extra 1.8% of the vote in the home state of their vice presidential candidate, with a 95% posterior credible interval of (0.6%, 3.1%). There is little evidence that the location of the national convention is predictive, and, if anything, candidates do slightly better in states where the opposing party controls the governorship.

While we find evidence that state-level deviations from national economic trends predict state-level variations in vote shares, the substantive magnitude is not especially large, given the observed level of variation in economic performance across states. An additional 1% of growth in a given state over the four-year period is associated with a 0.1% increase in the incumbent party candidate's vote share. As there is typically no more than a 30 percentage-point difference between the highest- and lowest-growth states during a single term, the maximal effect of the relative income growth is no more than 3% of the vote, and will normally be much less. It is possible that other state-level economic variables could reveal more substantial effects.

Estimates of the Democratic normal vote, α_{st}, illustrate how states' underlying voting tendencies have evolved over time, distinct from political or economic forces that are specific to any single election (Fig. 6). Trends in α_{st} are smoothed relative to the observed y_{st} (Fig. 3), and exclude national shifts and other covariate effects. While some states have maintained a consistent partisan tilt, such as Ohio, others have moved dramatically towards or away

[8] Unlike Hibbs (2012, p. 636), we would have no objection to including candidate-specific effects, if we had any way of assessing the political appeal, charisma, or overall "quality" of candidates or their campaigns in a consistent manner. Unfortunately, no such a measure is available at present.

[9] There are various ways of accounting for missing data in state-level variables such as per-capita GDP growth or unemployment rates, but this would require a far more complicated model.

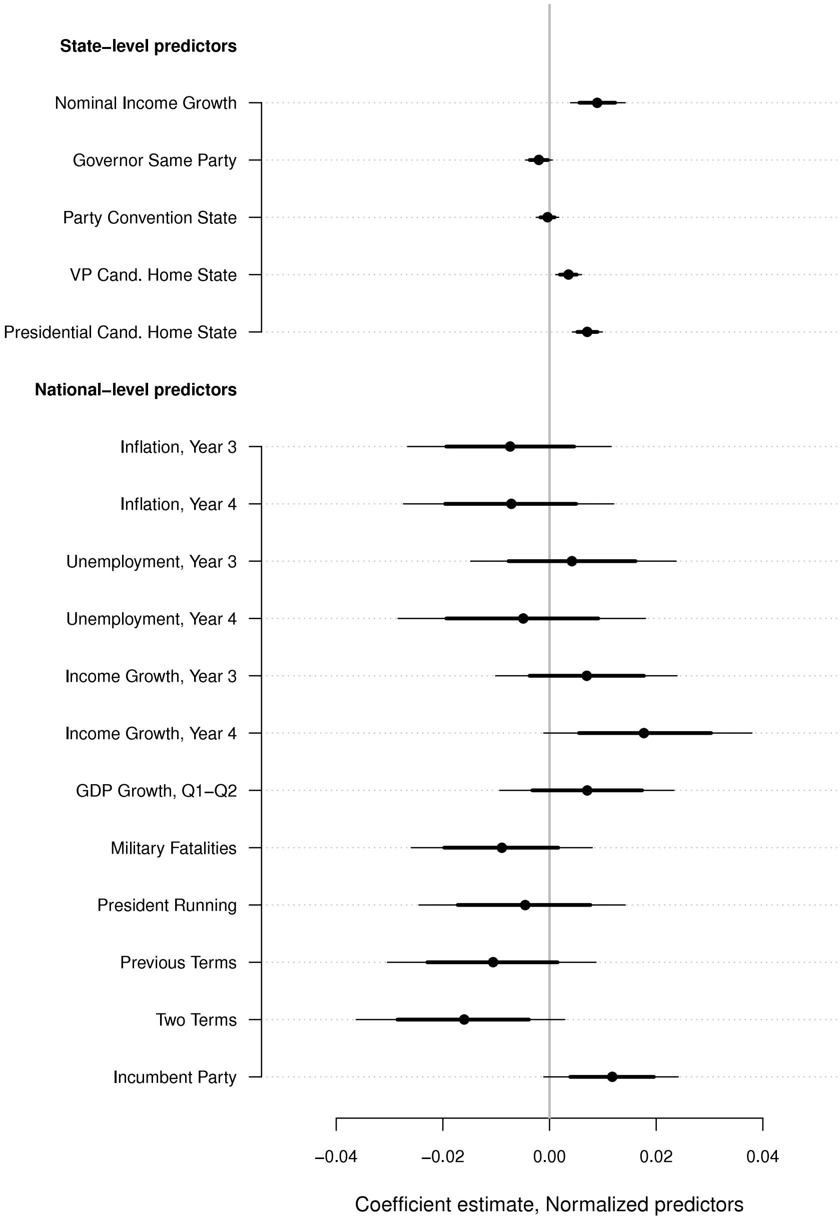

Fig. 5. Coefficient plot of normalized state- and national-level predictors for forecasting 2012 state-level Democratic presidential vote shares. The bold segments indicate posterior 80% credible intervals, and the thin segments indicate 95% credible intervals.

from voting for Democratic candidates for president (e.g., Maryland and Utah). Alabama transitioned from being one of the most Democratic states in 1952 to being one of the most Republican in 2008. Due to the dynamic model, states' forecasted values of α_{sT} for 2012 are expected to be their respective estimates from 2008, but with increased levels of posterior uncertainty.

4.3. Vote forecasts

The mean posterior forecast from the model is that President Barack Obama would be reelected in 2012, with 52.1% of the national major-party popular vote and 313 electoral votes. The actual result was 52.0% of the major-party vote and 332 electoral votes. Considering the level of predictive uncertainty in the popular vote and electoral vote forecasts, the accuracy of the predictions for 2012 is better than we can expect to occur in an average year. The 95% posterior interval for the popular vote share spans nearly 15 percentage points, $(0.446, 0.592)$, while the 95% posterior interval for the electoral vote is $(151, 455)$ (Fig. 7). The election outcome falls almost exactly on the contour of the 50% ellipse in the joint posterior distribution of the popular and electoral votes, in line with a well-calibrated prediction.

The posterior probability of an Obama victory under the specified model is 0.72. This is not especially strong

Fig. 6. Estimates of states' Democratic normal votes, α_{st}, 1952–2008, with forecasts for 2012. The highlighted states illustrate stability, as well as variation over time.

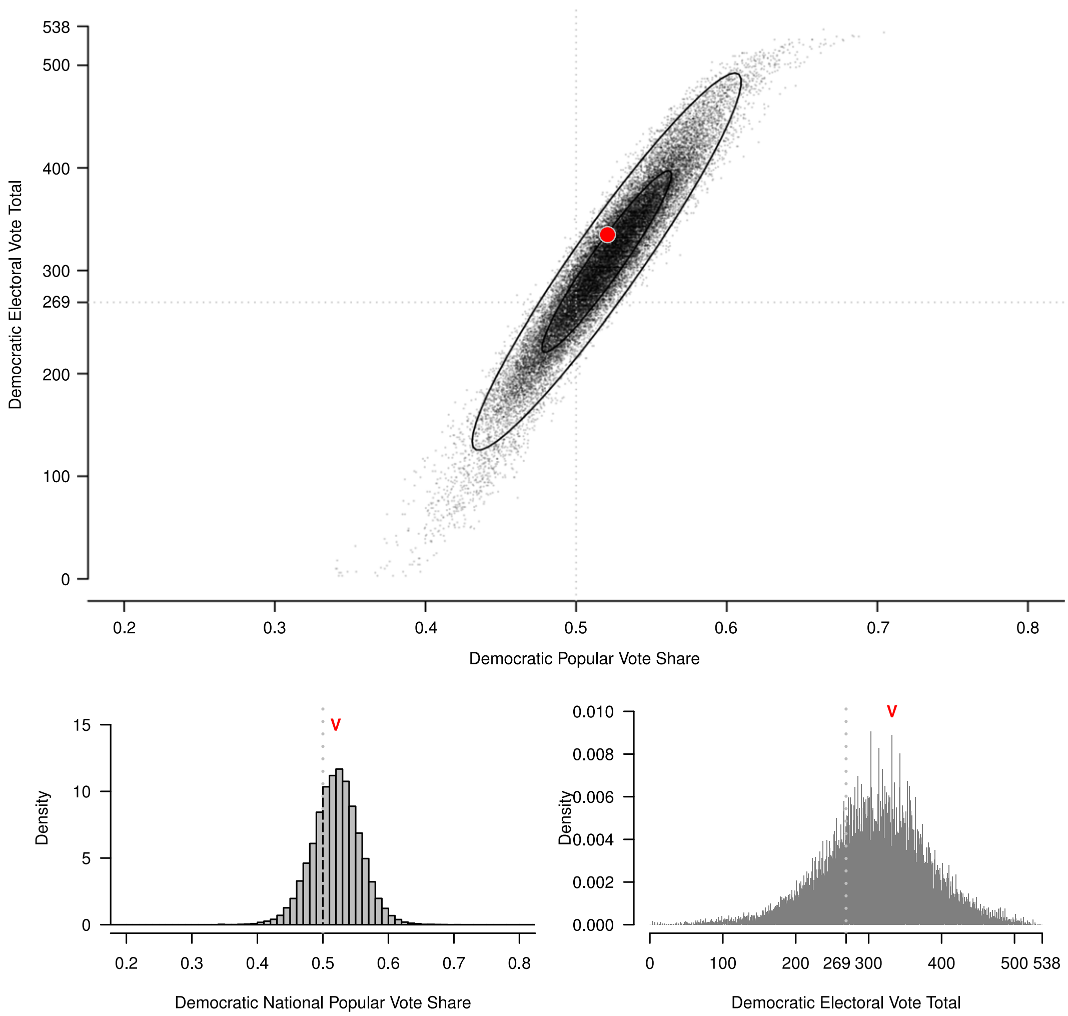

Fig. 7. Top: The joint posterior distribution of popular and electoral votes, overlaid with approximate 50% and 95% ellipses. The red point shows the actual election result. Bottom left: The marginal posterior distribution of the national Democratic major-party popular vote share. Bottom right: The marginal posterior distribution of the electoral vote for Obama. 2012 election results are marked V. (For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

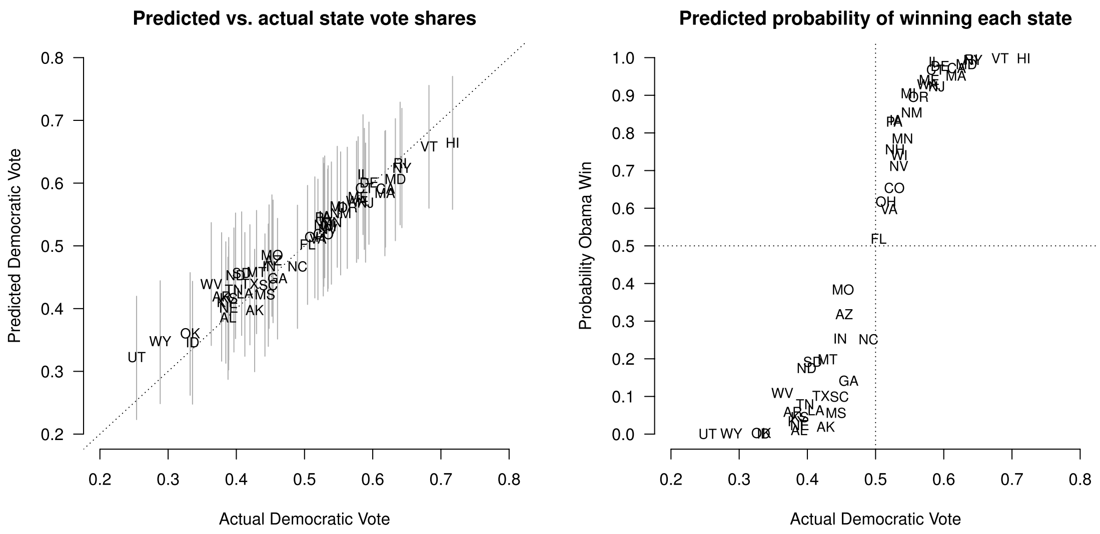

Fig. 8. Left: State two-party Democratic vote share forecasts, with 95% posterior intervals, compared to actual results. Right: Posterior probabilities that Obama will win each state. Washington, D.C. is not shown.

evidence for an Obama victory. As was noted by Mayhew (2008), the reelection rate for incumbent presidents over the course of U.S. history is approximately two in three (preceding the 2012 contest, $21/31 = 0.68$). Thus, if one knew nothing else beyond the fact that Obama was the current incumbent, and one treated all of U.S. history as being equally informative about his chance of reelection, one would arrive at very nearly the same prediction. The probability of an Electoral College reversal is 8%, and is close to symmetric with respect to parties. The predicted probability of an Electoral College tie is only 0.5%.

The state vote predictions are aligned closely with the actual results (Fig. 8). States where the Democratic vote shares were very high tend to be slightly under-predicted, and states where the Democratic vote shares were very low tend to be over-predicted. The average error across all 50 states and Washington, D.C. is 2.1 percentage points. This low level of error can be attributed largely to the national vote being so close to the mean posterior prediction. It is also a result of the 2012 election being a very easy election to predict at the state level, as the changes in states' vote shares compared to 2008 were as close to a uniform swing as had been observed since 1944 (Jackman, 2014).

As with the national forecasts, there is a substantial amount of uncertainty in the state forecasts, despite their accuracy for 2012. The span of the 95% posterior credible intervals for the predicted state vote shares is approximately ± 10 percentage points. However, this large amount of uncertainty corresponds to reasonable probabilities of an Obama or Romney victory in each state (Fig. 8). Florida is correctly rated a toss-up, at 0.52 chance of an Obama victory. States in which Obama is favored slightly, such as Virginia, Ohio, and Wisconsin, have probabilities of winning that range from 0.60 to 0.75. Only states in which there was little doubt of an Obama victory – e.g., Massachusetts, California, and Illinois – have win probabilities above 0.95. The same is true for states in which Romney was favored to win.

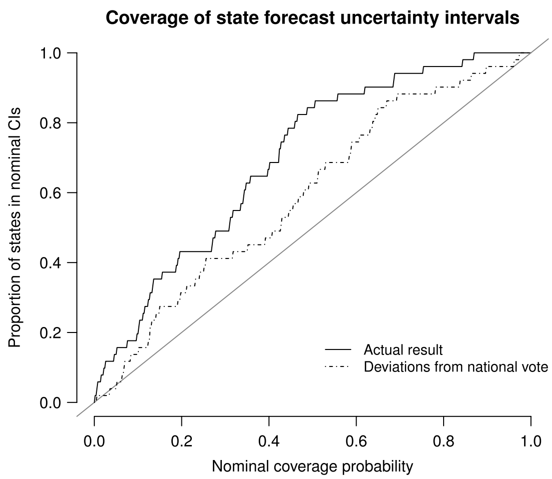

Fig. 9. Actual and nominal coverage rates of posterior credible intervals for state vote forecasts. The gray line indicates perfectly calibrated uncertainty intervals; e.g., the 80% posterior credible intervals include 80% of states' observed election outcomes. For 2012, the raw vote forecasts were under-confident. The calibration of the state uncertainty intervals is more accurate for states' deviations from the national-level vote.

Our uncertainty estimates for the 2012 state forecasts appear under-confident when presented as a collection of 51 intervals (Fig. 8). Obama won every state in which the model predicted at least a 50% chance of victory, and lost every other state. This is unlikely to happen if the state-level forecasts are independent, but the historical pattern of a near-uniform swing implies that much of the posterior variation for individual states is not independent. To assess the accuracy of the state forecasts' uncertainty estimates, we calculate the proportion of states whose observed vote outcomes fall within the credible intervals implied by a range of nominal posterior probabilities (Fig. 9). For the raw state-level votes, the posterior intervals are under-confident by up to 35%; that is, the posterior intervals were "too wide", and contained more of the state vote outcomes than expected. However, if we assume a uniform swing, we find that the states' posterior credible intervals would not have been over-confident unless the national vote forecast had been wrong by more than 3% in either direction. This is a reasonable "cushion" for avoiding over-confidence. If we condition on the national-level election outcome by calculating the coverage rates of the state-level deviations from the national vote, we find that the uncertainty in the state-level estimates is much closer to the theoretical ideal (Fig. 9).

4.4. Sensitivity to variable selection

The forecasts produced by our model are based upon the historical relationships between election outcomes and the predictor variables in X and Z. However, as we have noted, although there is a theoretical basis for including general measures of political context and economic well-being, there is no consensus as to which exact variables should be included in the model. (We certainly have no reason to believe that the predictors that we have chosen are the best.) It is therefore important that the forecasts not be overly sensitive to any particular set of predictors. Using our framework, combinations of variables in X and Z that are equally defensible on theoretical grounds should not lead to highly divergent forecasts. Here, we report several sensitivity analyses with respect to the inclusion and exclusion of national-level and state-level variables, and the inclusion of randomly generated placebo predictors.

4.4.1. Model without predictors

We first estimate a baseline, "null" model that includes no national-level or state-level predictors. This allows us to assess how much information was added to the forecast by our chosen covariates. Drawing only upon historical voting patterns, the null model predicts that Obama would receive 50.0% of the national two-party popular vote, and 272 electoral votes—a tie. More importantly, the 95% credible interval around the vote forecast ranges from 38% to 62%; a span of 24% of the two-party vote. This mirrors the range of Democratic popular vote outcomes observed from 1952 to 2008: a low of 38% in 1972 and a high of 61% in 1964. For Obama's electoral vote forecast, the 95% predictive interval under the null model is (31, 485); a span of 454 electoral votes. Again, without any structural factors to explain the observed variation in election results, the model returns a predictive distribution that matches the historical distribution of outcomes (Democratic candidates received 17 electoral votes in 1972, and 486 electoral votes in 1964). In contrast, the width of the 95% intervals using the explanatory variables above was 14% of two-party vote and 304 electoral votes. We therefore gain some predictive power from our covariates, but much uncertainty remains.

4.4.2. Model with placebo predictors

Our framework is designed to avoid the false discovery of relationships between covariates and election outcomes where none exist. In such cases, chance associations in the data could lead to election forecasts that diverge misleadingly from the null model. To assess the variability in our model's forecasts given arbitrary predictors, we conduct a placebo test that replaces the explanatory variables used in the analyses above (e.g., Fig. 5) with simulated variables drawn randomly from a standard normal distribution. Although the placebo predictors have no systematic relationship with observed election outcomes, the number of past elections is small enough that different sets of random values will generate different forecasts. The primary quantity of interest is the dispersion of these forecasts. We fit the model for 400 sets of placebo predictors, holding the numbers of variables at the state and national levels constant.[10] For each simulation, we record the predicted Obama share of the national popular vote and the electoral vote.

The model performs remarkably well when presented with uninformative explanatory variables, despite the large number of national-level variables relative to the number of elections. The median forecasts of the vote share and the electoral vote are the same as the null model: a tie. Of the 400 vote share forecasts, 90% are between 47% and 53%, and half are between 49% and 51%. For the electoral vote, 90% of forecasts are between 209 and 323, and half are between 249 and 282. Just as importantly, the posterior credible intervals around the point predictions are as wide as under the null model with no predictors. For the forecasted vote shares, the average range of the 95% credible intervals is 24%, while the average range for the electoral vote forecasts is 445. Because of this large amount of posterior uncertainty, 90% of the simulated predicted probabilities of an Obama victory are between 0.30 and 0.70, and half are between 0.44 and 0.56. It is therefore very unlikely that our model will issue strong forecasts as a result of over-fitting to covariates that contain no predictive information.

4.4.3. Model with real and placebo predictors

We now simulate a series of forecasts based upon a combination of placebo predictors and structural variables from our primary analysis. This tests the sensitivity of the model forecasts to the choice of predictors, as many indicators are equally justifiable theoretically. For each simulation, we randomly replace six of the "real" national-level predictors and three of the state-level predictors with placebo variables drawn from a standard normal distribution. It is good if the model predictions are consistent regardless of exactly which of the economic and political indicators happen to be included. To maintain the model's interpretation with respect to parties, we keep the incumbent party constant in all simulations.

The simulation results suggest a compromise between the forecasts from our main model (Fig. 7) and forecasts based entirely on placebo variables. The median forecast for Obama is 51.4% of the national two-party popular vote, 297 electoral votes, and a 0.64 probability of victory. Relative to the models with only placebo predictors, the inclusion of real predictors reduces the posterior uncertainty

[10] In an extended set of simulations, we find that changing the number of placebo predictors does not affect the results that we present.

in the forecasts, but also increases the range of the point predictions across the simulations.[11] The average 95% posterior credible interval around the national election forecast is 18% of the popular vote and 361 electoral votes. Across repeated simulations, half of the vote forecasts are between 50% and 53%, and half of the electoral vote forecasts are between 263 and 325. The implied probabilities of an Obama reelection range from 0.49 to 0.77 in half of the simulations. As expected, the choice of predictors affects the forecasts, but the majority of forecasts remain in a reasonable range.

5. Conclusion

The ability to predict U.S. presidential election outcomes accurately – and to state the uncertainty in those forecasts properly – is of considerable practical value and popular interest. For political scientists, it is key to the testing of theories of voter behavior, the consequences of incumbency and policy performance, and the efficacy of election campaigns. Although many scholars have proposed variables that might be associated with election outcomes, the forecasting of presidential elections has received remarkably little methodological attention.

While the standard regression-based approach to election forecasting has had some success, its track record is highly inconsistent. We have demonstrated that much of this inconsistency is a result of the ways in which existing models neglect key institutional features of the U.S. election system, focus on national- rather than state-level election outcomes, and understate the uncertainty associated with model specification, estimation, and out-of-sample inference. Few models utilize random effects or other techniques that are well-suited to prediction based on limited data. Considering the innovations in Bayesian statistics and regularization methods like the lasso over the last 30 years, the methodological approaches in this relatively public-facing area of political science are due for an update.

The Bayesian approach to forecasting is attractive for several reasons. First, Bayesian simulation accounts for uncertainty in the model parameters automatically, by integrating over their posterior distributions when constructing quantities of interest. Second, through the use of shrinkage priors over the model parameters, we can accommodate the fact that there are more national-level predictors than presidential elections in the data. Third, the Bayesian approach facilitates the inclusion of both national-level effects for each election, and state-level effects that change gradually over time. Finally, forecasting – especially from limited data – is fundamentally subjective. We have tried to state our priors and modeling assumptions clearly, and to provide justifications for them. Other researchers may make different choices, so our approach offers a framework within which to declare these assumptions in a rigorous manner.

[11] We obtain similar results from a simulation that removes a random subset of predictors in each trial, rather than replacing them with placebo variables.

Our key point is that there is simply not a lot of evidence in the historical record that we can use for making predictions about U.S. presidential elections on the basis of economic and political conditions. It is important to be honest about the limited strength of the evidence. This paper has attempted to account thoroughly for the major uncertainties associated with predicting state-level (and therefore Electoral College) election results. As we show, it is not possible to make very precise predictions with respect to the national-level vote. Thus, when we see an election result, we cannot make highly confident statements about why candidates performed as well as they did, or that they over- or under-performed relative to expectations. Until more elections have been observed, our expectations about forthcoming U.S. presidential elections cannot be very strong.

Acknowledgments

For comments and feedback on earlier versions of this article, the authors thank Matthew Dickinson, Mark Huberty, Michael Lewis-Beck, John Sides, Andreas Graefe, and conference participants at the Center for Advanced Studies, LMU München and the 2013 Midwest Political Science Association Annual Conference.

References

Abramowitz, A. (2008). Forecasting the 2008 presidential election with the time-for-change model. *PS: Political Science and Politics*, *41*(4), 691–695.

Abramowitz, A. (2012). Forecasting in a polarized era: the time for change model and the 2012 presidential election. *PS: Political Science and Politics*, *45*(4), 618–619.

Achen, C.H., & Bartels, L.M. (2004). Musical chairs: pocketbook voting and the limits of democratic accountability. In *Annual meeting of the American Political Science Association*.

Bartels, L. M. (1997). Specification uncertainty and model averaging. *American Journal of Political Science*, *41*, 641–674.

Bartels, L. M., & Zaller, J. (2001). Presidential vote models: a recount. *PS: Political Science and Politics*, *34*(1), 9–20.

Berry, M. J., & Bickers, K. N. (2012). Forecasting the 2012 presidential election with state-level economic indicators. *PS: Political Science and Politics*, *45*, 669–674.

Campbell, J. E. (2012). Forecasting the 2012 American national elections. *PS: Political Science and Politics*, *45*, 610–613.

Ferguson, N. (2012). Why Obama Is Winning. *Newsweek*. September 10. http://www.newsweek.com/niall-ferguson-why-obama-winning-64627.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, *1*(3), 515–533.

Hibbs, D. A. (2012). Obama's reelection prospects under bread and peace voting in the 2012 US presidential election. *PS: Political Science and Politics*, *45*, 635–639.

Holbrook, T. M. (2012). Incumbency, national conditions, and the 2012 presidential election. *PS: Political Science and Politics*, *45*, 640–643.

Imai, K., & Tingley, D. (2012). A statistical method for empirical testing of competing theories. *American Journal of Political Science*, *56*(1), 218–236.

Jackman, S. (2014). The predictive power of uniform swing. *PS: Political Science and Politics*, *47*, 317–321.

Jerôme, B., & Jerôme-Speziari, V. (2012). Forecasting the 2012 US presidential election: lessons from a state-by-state political economy model. *PS: Political Science and Politics*, *45*, 663–668.

King, G. (1991). Calculating standard errors of predicted values based on nonlinear functional forms. *The Political Methodologist*, *4*(2), 2–4.

King, G., Tomz, M., & Wittenberg, J. (2000). Making the most of statistical analyses: improving interpretation and presentation. *American Journal of Political Science*, *44*, 341–355.

Klarner, C. E. (2012). State-level forecasts of the 2012 federal and gubernatorial elections. *PS: Political Science and Politics*, *45*, 655–662.

Linzer, D. A. (2013). Dynamic Bayesian forecasting of presidential elections in the states. *Journal of the American Statistical Association*, *108*(501), 124–134.

Mayhew, D. R. (2008). Incumbency advantage in US presidential elections: the historical record. *Political Science Quarterly*, *123*(2), 201–228.

Montgomery, J. M., Hollenbach, F. M., & Ward, M. D. (2012). Ensemble predictions of the 2012 US presidential election. *PS: Political Science and Politics*, *45*, 651–654.

Montgomery, J. M., & Nyhan, B. (2010). Bayesian model averaging: theoretical developments and practical applications. *Political Analysis*, *18*(2), 245–270.

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, *103*(482), 681–686.

Plummer, M. (2008). Just another Gibbs sampler. http://www-ice.iarc.fr/~martyn/software/jags/.

R Core Team (2013). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, http://www.R-project.org/.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B*, *58*(1), 267–288.

Tikhonov, A. N. (1943). On the stability of inverse problems. *Doklady Akademii Nauk SSSR*, *39*(5), 195–198.

West, M., & Harrison, J. (1997). *Bayesian forecasting and dynamic models*. Springer.

Yglesias, M. (2012). Obama's winning because of the economy, not despite it. *Slate*. September 11. http://www.slate.com/blogs/moneybox/2012/09/11/obama_s_election_lead_it_s_because_the_economy_s_doing_well_not_despite_economic_problems_.html.