

# *A Unified Method of Evaluating Electoral Systems and Redistricting Plans\**

Andrew Gelman, *Department of Statistics,  
University of California, Berkeley*

Gary King, *Department of Government, Harvard University*

We derive a unified statistical method with which one can produce substantially improved definitions and estimates of almost any feature of two-party electoral systems that can be defined based on district vote shares. Our single method enables one to calculate more efficient estimates, with more trustworthy assessments of their uncertainty, than each of the separate multifarious existing measures of partisan bias, electoral responsiveness, seats-votes curves, expected or predicted vote in each district in a legislature, the probability that a given party will win the seat in each district, the proportion of incumbents or others who will lose their seats, the proportion of women or minority candidates to be elected, the incumbency advantage and other causal effects, the likely effects on the electoral system and district votes of proposed electoral reforms such as term limitations, campaign spending limits, and drawing majority-minority districts, and numerous others. To illustrate, we estimate the partisan bias and electoral responsiveness of the U.S. House of Representatives since 1900 and evaluate the fairness of competing redistricting plans for the 1992 Ohio state legislature.

## **1. Introduction**

We introduce a unified and relatively simple statistical model with which one can evaluate electoral systems and redistricting plans in almost any way for virtually any legislature with two major parties and predominately single-member districts. This model is useful (1) for understanding an election that has already taken place; (2) for predicting a future elec-

\*We have written a computer program to implement the methods derived in this article. The program is called *JudgeIt* and is available from the ICPSR, via "gopher" or "anonymous FTP" from Haavelmo.Harvard.Edu, or by contacting us. *JudgeIt* has been used in redistricting processes in many states and won the 1992 Research Software Award from the American Political Science Association. All data and information necessary to replicate the empirical analyses in this article are available from the ICPSR in a Class V data set listed under our names. Section 6.2 was drawn from Gary King's experience as a consultant to the State of Ohio in 1991-92. We thank Jim Alt, Neal Beck, Mike Lewis-Beck, and Doug Rivers for many helpful comments on this paper and the National Science Foundation for research grant SBR-9223637. Please address correspondence to Gary King, Department of Government, Harvard University, Littauer Center North Yard, Cambridge, Massachusetts 02138; email: gk@isr.Harvard.Edu; phone: (617)495-2027.

tion, possibly subject to a new redistricting plan; and (3) for evaluating a past election under specified counterfactual conditions (e.g., supposing no incumbents had run for reelection). For each of these general situations, our method enables one to make virtually any prediction about or characterization of the pattern of district votes in an electoral system (e.g., the proportion of African American legislators who would lose their seats under a proposed redistricting plan). Some of the electoral system summaries that can be produced by this method include:

- the seats-votes curve (Edgeworth 1898; Butler 1951; Schrodt 1981; Niemi and Fett 1986; Gelman and King 1990b);
- partisan bias and electoral responsiveness (Tufto 1973; Grofman 1983; King and Browning 1987; Brady 1988);
- the expected or predicted vote in each legislative district (Lewis-Beck and Rice 1992; Cain 1985; Born 1985);
- the probability that a given party will win the seat in each district;
- incumbency advantage and other causal effects (Erikson 1971; Alford and Brady 1988; Gelman and King 1990a);
- the expected proportion of incumbents, or others, who will lose their seats (Mayhew 1974; Fiorina 1977; Ferejohn 1977; Jacobson 1987);
- the expected number of women or minority candidates to be elected;
- the likely effects on the electoral system and district votes of proposed electoral reforms, such as term limitations (Benjamin and Malbin 1992; Rothstein and Gilmour 1992), campaign spending limits (Jacobson 1980), and drawing majority-minority districts (Grofman, Handley, and Niemi 1992);
- the contributions of incumbency advantage, or other aspects of the electoral system, to electoral system phenomena such as divided government (Mayhew 1991; Fiorina 1992; Campbell 1992; King and Gelman 1991).

In providing estimates of these theoretical concepts, we apply the most important insight of the field of statistics in this century to the study of legislative elections—the distinction between the data one observes and the theoretical concepts to be estimated with the data. Too often in the legislative elections literature, scholars define a theoretical concept as identical to a measure of it. For example, scholars have defined the vulnerability of incumbents *as* their electoral margin in the last election. Certainly the latter has something to do with the former, but the two are

not the same.<sup>1</sup> Our model enables one to define any theoretical concept related to legislative elections and to make the best use of available data to provide empirical estimates.

Moreover, as soon as one makes the fundamental distinction between theory and data, the indispensable role of quantitative estimates of uncertainty (such as standard errors, confidence intervals, or margins of error) becomes absolutely clear. Perhaps since the distinction between theoretical concepts and data is not always made in legislative elections research, consistent reporting of standard errors is not yet routine. Our model automatically produces these estimates for virtually every quantity calculated and should therefore make reporting easier as well.

We also provide means of evaluating the fit of the model to the data. Since our model cannot be estimated with existing statistical software, we have made available, as an accompaniment to this paper, a general purpose computer program that implements this model. We have evaluated the performance of this model in hundreds of thousands of districts, in dozens of election years in the U.S. Congress and numerous state legislatures. We have also had some limited experience applying the model to data from foreign countries.

This paper had its origins in our attempts to generalize the models and methods developed in a series of articles by King and Browning (1987), King (1989a), Gelman and King (1990a, 1990b), and King and Gelman (1991). The most recent of these models is technically sophisticated and quite computationally intensive; we believed that adding additional features would make it more realistic but would unfortunately produce an even more complicated model. We were right about the former and wrong about the latter: adding additional information produced a more realistic model for which much of the algebra fell out, leaving a surprisingly simple model. We simplified the model further by eliminating features that did not materially affect the substantive conclusions, based on our analyses of congressional and state legislative elections. The resulting model turned out to be useful not only for our original goal but for numerous other applications in the academic literature as well.

Section 2 introduces the basic model; section 3 discusses issues of preliminary estimation. We derive the distribution of votes in actual, predictive, and counterfactual situations derived from this model in section 4. In section 5, we show how to estimate, along with a standard

<sup>1</sup>It is not difficult to imagine an incumbent who wins elections by small margins but, perhaps due to high levels of racial polarization, consistently and predictably wins reelection; similarly, an incumbent who won with a large margin in the last election would be quite vulnerable if he or she were convicted of a felony.

error, any feature or prediction of interest. Examples appear in section 6. Section 7 concludes. Appendixes A–C provide technical details.

## 2. A Model of District-based Electoral Systems

### 2.1. *The Model*

In order to generally distinguish between the data (actual election results) and theoretical quantities of interest, we begin by predicting what could happen, or what would have happened, if the election were held again under specified conditions. More specifically, we define *hypothetical election results* as the set of all possible election outcomes that could have occurred if all political conditions up to the start of the campaign were held constant and the campaign were run again.

To define hypothetical elections formally, we need a probability model to encompass our uncertainty and allow a range of reasonable possibilities for the hypothetical outcomes. This avoids the unreasonable assumption that election outcomes are *exactly* determined and can be forecast without error, given variables measured before the start of the campaign. The model presented here allows us to calculate the (posterior) probability distribution of these hypothetical election results. We think of the observed election result as just one of the possible hypothetical election results that could have occurred. Any specific theoretical quantity of interest and its standard error can then be calculated directly from the distribution of hypothetical election results.

Although our model applies to any two-party system, we use the labels “Democratic” and “Republican” to fix ideas more clearly. We also use a state legislative election as the running example with which to introduce this model. It should be clear, however, that the model applies much more widely. Following the algebraic presentation of the model, we discuss in more detail a substantive interpretation of the model in section 2.2, and the explanatory variables we recommend including in section 2.3. A discussion of alternative arrangements for the treatment of uncontested districts appears in Appendix A.

*Notation.* We assume a legislature comprising  $n$  single-member districts, denoting  $v_i$  as the Democratic proportion of the two-party vote in each district  $i$ , and  $v$  as the set of votes for all districts  $(v_1, v_2, \dots, v_n)$ . The votes  $v$  will be predicted by  $k$  explanatory variables, which can together be written as an  $n \times k$  matrix,  $X$ . The first column of  $X$  should be all ones, corresponding to an intercept term in a regression; the remaining columns should be substantive explanatory variables, which we discuss in section 2.3. The matrix  $X$  is always known, and the votes vector  $v$  is

Table 1. Model Structure

District Number	Actual Election Results	Hypothetical Replications of Each District Election			
		1	2	...	$m$
1	$v_1$	$v_1^{(hyp)1}$	$v_1^{(hyp)2}$	...	$v_1^{(hyp)m}$
2	$v_2$	$v_2^{(hyp)1}$	$v_2^{(hyp)2}$	...	$v_2^{(hyp)m}$
⋮	⋮	⋮	⋮	⋮	⋮
$n$	$v_n$	$v_n^{(hyp)1}$	$v_n^{(hyp)2}$	...	$v_n^{(hyp)m}$
Quantity of interest	$Q$	$Q^{(hyp)1}$	$Q^{(hyp)2}$	...	$Q^{(hyp)m}$

known when evaluating an election that has occurred, but unknown for prediction.

For hypothetical elections, we define a known matrix  $X^{(hyp)}$  of explanatory variables, and an unknown vector  $v^{(hyp)}$  of hypothetical district-level Democratic vote proportions. The vector  $X^{(hyp)}$  is either equal to  $X$  or, to evaluate counterfactual assumptions, is defined as needed.<sup>2</sup> The goal of the analysis is inference about  $v^{(hyp)}$ , given  $X^{(hyp)}$  (and given  $v$  if available). For the next brief subsection, we refer to the individual hypothetical replication  $j$  of the election in district  $i$  as  $v_i^{(hyp)j}$ .

*Conceptual overview.* The goal of our model is to calculate a joint probability distribution for all quantities of interest, such as those listed in section 1. From this, all point estimates and standard errors can be calculated. One way to portray our method of calculating this, as well as the results of the model we are about to describe, is Table 1 (see Rubin 1987). Each row in the table refers to a district, with the district number in the first column and the actual election result in the second. If the problem is one of prediction, instead of evaluation, the actual election result is obviously not known. The remaining columns depict  $m$  hypothetical replications of each district election. Thus, in district 2, the actual Democratic proportion of the two-party vote is  $v_2$  (which might be, say, 0.56) and again is known only if we are evaluating and not predicting. The first hypothetical replication, numbered 1, of the election in district 2 is  $v_2^{(hyp)1}$  (which might be, say, 0.52 or 0.57), an example of what might happen if all conditions up to the start of the general election campaign were the same, but the campaign and balloting were run again. The sec-

<sup>2</sup>For example, one might set the incumbency codes to all zeros to study the likely effect of term limits.

ond hypothetical replication in district 2 is denoted  $v_2^{(\text{hyp})2}$ , which is another draw from the same probability distribution characterizing these hypothetical elections. In this way, we model the uncertainty in electoral results by this variation across hypothetical elections.

The last row in the table is a “quantity of interest,” denoted  $Q$  for the actual election result and  $Q^{(\text{hyp})j}$  for hypothetical election replication  $j$ . This summary statistic is calculated from its corresponding column of real or hypothetical data. This could be virtually anything, but to fix ideas imagine that it is the proportion of incumbents who win reelection. Since for most purposes, incumbency is one of the conditions that, at the start of a general election campaign, we assume to be fixed across hypothetical replications,  $Q^{(\text{hyp})j}$  is calculated by taking the proportion of the district votes in the column greater than 0.5 (for Democratic incumbents) or less than 0.5 (for Republican incumbents) among only the rows that have incumbents.

Once we have a set of hypothetical election results for each district election, and we have decided which summary statistic  $Q$  we wish to calculate and have calculated it, we can use a simple procedure to calculate an overall point estimate and standard error. The point estimate is just the average, and the standard error is the square root of the variance, across the row of  $m$  summary statistics in the table calculated for each hypothetical election result:

$$\bar{Q} = \frac{1}{m} \sum_{j=1}^m Q^{(\text{hyp})j}; \quad (1)$$

$$\text{Var}(Q) = \frac{1}{m-1} \sum_{j=1}^m (Q^{(\text{hyp})j} - \bar{Q})^2. \quad (2)$$

The only task remaining, then, is to estimate the (posterior) probability distribution, with which we can generate these hypothetical election results and, in turn, calculate our point estimates and standard errors. Once we have this distribution, our actual method of calculation and estimation will correspond closely to the procedure illustrated in Table 1.

*The probability model.* We model the district vote outcomes with a random components linear regression of  $v$  on  $X$ ,

$$v = X\beta + \gamma + \varepsilon, \quad (3)$$

where  $\beta$  is a vector of  $k$  parameters that must be estimated from data, and  $\gamma$  and  $\varepsilon$  are two vectors of independent error terms. Strictly speaking, this independence assumption is imposed as a definitional feature of our

model, not assumed as a characteristic of the world; testing whether definitions such as this are “true” makes little sense, since  $\varepsilon$  is *defined* as the part of the error term that is independent for each district vote. The variable  $\varepsilon$  is a traditional random error term;  $\gamma$  is the “random component” error term, which helps correct for the fact that the  $X$  variables do not completely describe the state of the electoral system at the start of the campaign due to the omission of relevant variables and measurement error in the variables included. (We provide a detailed interpretation of  $\gamma$  and  $\varepsilon$  in section 2.2.) For each district  $i$ , the error terms are assigned independent normal distributions,

$$\begin{aligned}\gamma_i &\sim N(0, \sigma_\gamma^2) \\ \varepsilon_i &\sim N(0, \sigma_\varepsilon^2),\end{aligned}\tag{4}$$

with variances  $\sigma_\gamma^2$  and  $\sigma_\varepsilon^2$  that must be estimated. Because  $\gamma_i$  and  $\varepsilon_i$  are independent, equations (3) and (4) are, for some purposes, equivalent to a linear regression of  $y$  on  $X$  with a single error term of variance  $\sigma_\gamma^2 + \sigma_\varepsilon^2$ . For mathematical convenience, we reparameterize by defining a parameter for the total variance and a parameter for the proportion of variance due to  $\gamma$ :

$$\begin{aligned}\sigma^2 &= \sigma_\gamma^2 + \sigma_\varepsilon^2 \\ \lambda &= \frac{\sigma_\gamma^2}{\sigma_\gamma^2 + \sigma_\varepsilon^2}.\end{aligned}$$

The vector  $v^{(\text{hyp})}$  of hypothetical vote proportions is determined by an analogous probability model:

$$v^{(\text{hyp})} = X^{(\text{hyp})} \beta + \delta^{(\text{hyp})} + \gamma + \varepsilon^{(\text{hyp})},\tag{5}$$

where  $\varepsilon^{(\text{hyp})}$  is a new vector of  $n$  independent error terms with variance  $\sigma_\varepsilon^2$ , and  $\delta^{(\text{hyp})}$  is a known constant used to model statewide partisan swing. The hypothetical outcome,  $v^{(\text{hyp})}$ , which generates the numerical results for the columns in Table 1, differs from the actual  $v$  in three ways:

1. The explanatory variable matrix  $X$  is replaced by  $X^{(\text{hyp})}$ , to recognize that we may wish to specify different conditions under which the hypothetical election may be run (such as no incumbents running).
2. A constant,  $\delta^{(\text{hyp})}$ , is added to allow a statewide partisan swing to be specified. One can specify either  $\delta^{(\text{hyp})}$  or a corresponding value for the expected average district vote,  $E(\bar{v}^{(\text{hyp})})$ .<sup>3</sup>

<sup>3</sup>This is true, since  $\delta^{(\text{hyp})} = E(\bar{v}^{(\text{hyp})}) - (1/n) \sum_{i=1}^n (X^{(\text{hyp})} \beta)_i$ .

3. The new error term,  $\varepsilon^{(\text{hyp})}$ , models the fact that, even if the variables in  $X$  were unchanged, we would not expect  $v^{(\text{hyp})}$  to be identical to  $v$ . Across many hypothetical elections,  $\gamma$  remains unchanged, while  $\varepsilon$  varies.<sup>4</sup>

The parameters of this model to be estimated— $\sigma^2$ ,  $\lambda$ , and  $\beta$ —are not usually of primary interest in evaluating electoral systems and redistricting plans (although  $\beta$  is in some cases of interest in evaluating causal effects). Instead, we define all the quantities of interest, including the seats-votes curve, district vote predictions, and the like, in terms of the distribution of hypothetical election outcomes  $v^{(\text{hyp})}$ , given the average district vote  $\bar{v}$  or the actual election outcomes  $v$  when available (which, in turn, depend on the parameters). The specific method for calculating this conditional distribution from this model is described in section 4.

## 2.2. Interpretation of the Model

*Aggregate partisan swing.* The parameter  $\delta^{(\text{hyp})}$ , or the corresponding value of  $E(\bar{v}^{(\text{hyp})})$ , in this model is a notational convenience that allows us to vary the average district vote in a hypothetical (or predicted) election, without affecting the relative positions of the districts. This partitioning reflects the common result that it is often quite easy to predict which districts will vote more Republican than others, but it is harder to forecast exactly what the average vote will be across districts; put differently, given the average vote across districts, it is easy to predict the vote proportions within each district.

According to the model, expected votes differ across legislative districts at any one time (as governed by  $X\beta + \gamma$ ). Over time, the districts swing along with the statewide mean (due to the scalar parameter  $\delta^{(\text{hyp})}$ ) but only on average (due to the random error term  $\varepsilon$ ). Another way to put this is that districts move on average with the statewide mean, but, for any given statewide issue producing a swing toward a party, any district may move with or against the statewide trend (as indicated by a scatter plot of district votes in two successive years; see, e.g., Figures 1 and 2 in King and Gelman 1991). This allows us to account for local differences in the appeal of the candidates and other factors simultaneously with statewide swings.

The stochastic model is interpreted slightly differently for prediction and evaluation: for prediction, we ask how many seats *will* the Democrats

<sup>4</sup>Another way to think of this model is that, for each district  $i$ ,  $v_i$  and  $v_i^{(\text{hyp})}$  are not independent, due to the random variable  $\gamma_i$  they share. Given the explanatory variables  $X$ ,  $X^{(\text{hyp})}$ , and  $\delta^{(\text{hyp})}$ , and the parameters  $\beta$ ,  $\sigma^2$ , and  $\lambda$ , we can combine equations (3) and (5) to find that the theoretical correlation between  $v_i$  and  $v_i^{(\text{hyp})}$  is  $\lambda$ .



win with an average of 45% of the votes, say, and in evaluation we ask, how many seats *would* they have won if essentially the same election campaign had been run again.<sup>5</sup> The only difference between evaluation and prediction is that we observe one of the possible hypothetical election outcomes for the former (the actual vote, the second column in Table 1) and do not observe any for the latter. Even after taking into account the information in our model expressed through the explanatory variables  $X$ , observing the election outcome will in general help us to some degree in characterizing the conditional distribution of hypothetical election outcomes; thus, under our model, inference about observed election systems will generally be sharper than predictions.

*Error terms.* To understand the division between  $\gamma$  and  $\varepsilon$  (or, equivalently, the relative values of  $\sigma_\gamma^2$  and  $\sigma_\varepsilon^2$ ), consider the two extreme cases. If  $\sigma_\gamma^2 = 0$ , then  $\lambda = 0$ , and there is no *systematic* difference among districts except that described by  $X\beta$ . Consider two districts,  $i$  and  $j$ , with identical values of  $X$  but different outcomes in the election:  $v_i = 0.6$  and  $v_j = 0.7$ . Without  $\gamma$ , all the difference between  $v_i$  and  $v_j$  must be attributed to the random error  $\varepsilon$ , and as a result, the distributions of hypothetical election results in districts  $i$  and  $j$  will be the same. For real electoral systems, the explanatory variables in  $X$  cannot be perfect, and it is usually wrong to ignore the additional information that the Democrats did better in district  $j$  in the observed election.

At the other extreme, if  $\sigma_\varepsilon^2 = 0$ , then  $\lambda = 1$ , and every hypothetical election outcome  $v^{(\text{hyp})}$  must be *identical* to the observed  $v$ , except for the constant shift,  $\delta^{(\text{hyp})}$ . This is just the uniform partisan swing assumption (Butler 1951), which requires that individual districts move in lock-step with statewide swings. As distinct from the *assumption* of uniform partisan swing, there is also a uniform partisan swing *method*, which is only strictly appropriate when the corresponding assumption is correct. Unfortunately, this assumption does not apply to any known electoral system (again, see Figures 1 and 2 in King and Gelman 1991); in characterizing an electoral system, only the variation that persists in future election results deserves to be labeled “systematic” in the electoral system. In principle, the uniform partisan swing method can work in some cases even though the assumption is false, but it never produces honest estimates of uncertainty and is always less statistically efficient than the method we describe here.

<sup>5</sup>Of course, the model can address numerous other substantive questions; this one highlights the difference between what will happen in prediction and what would have happened in evaluation.

A final way of understanding the separation of  $\gamma$  and  $\epsilon$  is by imagining that an extremely large number of explanatory variables were available in  $X$ —not just previous election results but also campaign spending, the effectiveness of every election campaign, lists of all campaign events, the results of all the campaign polls, the weather on election day, and so on. In this case,  $v$  could be predicted with essentially zero error. In this extreme situation, we would *not* want to assume that  $X$  is unchanged in hypothetical elections. It would make more sense to split  $X$  into two sets of variables: variables like past election results, incumbency, and party support in each district, which would not change if the election campaign were to be held again; and variables like campaign poll results and the weather on election day that would change. Between these extremes, some dividing line must be drawn that then *defines* the set of hypothetical elections and thus the electoral system.

In the ideal world in which  $X$  contains all possible variables, we would lump the first set of variables with  $\gamma$  as the systematic component and the second set with  $\epsilon$  as the stochastic component. When, as in practice, the variables in  $X$  are not complete, we can attribute the residuals in the regression to the unobserved variables, and the residual variance may be partitioned *statistically*, as we do, into that due to  $\gamma$  and  $\epsilon$ .<sup>6</sup>

### 2.3. Explanatory Variables

Whether we seek to predict the future or evaluate the past, our model requires explanatory variables that help predict votes  $v$  from explanatory variables  $X$ . (Section 4.2 continues the discussion of explanatory variables in the context of defining counterfactual scenarios.) Our model uses these explanatory variables to help statistically partition election results into systematic and stochastic components. Since the purpose is not estimating causal effects, the rules for choosing explanatory variables are different than usual. The main immediate goal is to choose variables that would help in forecasting future votes.<sup>7</sup>

When used for forecasting, our model will perform only as well as the variables chosen, and we take the main insight about the choice of

<sup>6</sup>The distinction between  $\gamma$  and  $\epsilon$  can also be understood from the perspective of various philosophies about the role of random error in statistical analyses; see King (1989b, 1991b).

<sup>7</sup>By comparison, Gelman and King (1990a) did not control for any consequences of incumbency in their estimation of incumbency advantage. In the current application, we are just trying to predict, and we have no objection to controlling for variables such as campaign spending that are determined after, and in part as a consequence of, the incumbent's decision of whether to run for reelection. We should be careful, of course, not to interpret the coefficients as causal estimates.

variables from the large existing forecasting literature (see Lewis-Beck and Rice 1992 for a review). For a given set of explanatory variables, our model will not produce better forecasts on average, for example, but it will give more accurate estimates of the uncertainty of those forecasts.

On the basis of prior research, it is clear that the variables in  $X$  should certainly include *past legislative election results*, when available.<sup>8</sup> When including a previous legislative election, it also makes sense to include variables for *incumbency*,  $inc(t)_i$  (defined as 1 if a Democratic incumbent is running; 0 if no incumbent is running; and  $-1$  if a Republican incumbent is running), and *uncontestedness*,  $unc(t)_i$  (defined as 1 if a Democrat is running uncontested; 0 if the district election is contested; and  $-1$  if a Republican is running uncontested), for district  $i$  in election  $t$ . For election  $t$ , the variable  $inc(t)$  is the set of all district values,  $inc(t)_1, \dots, inc(t)_n$ , and similarly for  $unc(t)$ .

Uncontestedness is important because uncontested elections do not fit any linear model unless explicitly controlled for (or adjusted as in Appendix A). An incumbency status variable is important if there is a large incumbency advantage because without such an indicator, election results from a mixture of incumbent-controlled and open districts would not fit the assumptions of our linear model with independent error terms. In addition, including the incumbency variable usually improves the predictive power of the model. In years following redistricting, incumbency status is often unclear, but it is still better to estimate the variable roughly, for each district, than to ignore it. If one is concerned about assuming that incumbency status will remain the same after redistricting, perhaps because of nonpartisan population base changes, one could include interaction effects so that the effect of incumbency would vary by population base or other variables.<sup>9</sup>

<sup>8</sup>When predicting an election just following a redistricting, the past election results will have to be reaggregated into the new districts, if possible. However, it usually makes little sense to do this without taking into account differences between the districts. By far the most important problem in creating voting data for new districts is combining districts with different incumbency status. To deal with this problem, we suggest the following procedure: (1) estimate the incumbency advantage from historical voting data using the method described in Gelman and King (1990b); (2) obtain legislative voting data at the precinct (or election district) level; (3) subtract the incumbency advantage from the incumbent candidate in precincts represented by incumbents and give those votes to the opposition party candidate (thus creating an estimate of what the electoral data at the precinct level would have been in an open-seat election); (4) aggregate the precincts up to the new districts; and finally, (5) add back the incumbency advantage based on the incumbency status in the new districts. One could do better still by also correcting for differences in candidate quality or other factors, if it really seems worth the effort.

<sup>9</sup>The explanatory variable here could be the proportion of the people who *move* into or out of the district, for current districts, or proportion of people who *are moved* into or

As discussed in Gelman and King (1990a), a variable for *party control* in the district is also useful:  $party(t)_i$  (defined as 1 if the sitting incumbent is a Democrat; -1 if the sitting incumbent is a Republican; and 0 otherwise—seat controlled by a third party). This variable controls for possible bimodality in the partisan strengths of districts.<sup>10</sup> It is especially important to include the party variable when the incumbency variable is inaccessible; without correcting for at least one of the two, most recent U.S. electoral systems would almost certainly be bimodal (Mayhew 1974). When using a statewide variable, such as the race for state comptroller, for prediction, it is of course unnecessary to include incumbency, as it affects all districts equally, but it may be desirable to adjust the district-level votes for comptroller by correcting for the “friends and neighbors” effect due to the local popularity of the statewide candidates.

Uncontestedness, incumbency, and party control should also, of course, be used when predicting future elections. Party control should always be available, but uncontestedness and the decision of the incumbent about whether to run will be more difficult to gather. We have found, however, that redistricters almost always have this information, so it is in principle not difficult if one can gain access.<sup>11</sup>

Statewide or nationwide election results, broken down by legislative districts, are also useful (when available). Other plausibly useful variables

---

out of the district by redrawing the district lines, for forecasts following redistricting. However, in our experience, nonpartisan base changes have little effect and are not especially useful for forecasting.

<sup>10</sup>A regression model fit to a distribution with two clusters of partisans will wrongly fit a single linear model to both clusters, and as a result, will overestimate the “regression effect.” Or to put it another way, a distribution with two clusters, corresponding to the two parties, does not fit a linear model without including different intercepts and a common slope. See King (1991a) for a graph of this and an explanation, and Gelman and King (1990b) for a description of the procedures we have developed to accommodate this problem even with bimodality.

<sup>11</sup>In most states, these data are produced by state computers and officially available to the public, but in practice, obtaining relevant information on the computer can require some inside contact. If data gathering becomes difficult, it makes the most sense in the framework of our model to make a guess about the incumbency status and uncontestedness of each district in the predicted election, and then draw conclusions conditional on the guess. Uncertainty in future electoral conditions may be expressed by considering several possibilities for incumbency and uncontestedness and evaluating the predictions for each. A standard assumption to start with is that all incumbents run, and incumbency status in new districts is determined by their residence (the actual figure is that about 85% of state legislators run for reelection, and 90% among U.S. House members; the decision to run again is independent of the legislator’s share of the vote in the previous election), all contested seats stay contested, and all uncontested seats likewise remain so. One should also remember that following redistricting, more seats tend to be contested, and fewer incumbents run.

are campaign contributions received before the start of the campaign (perhaps transformed to the log scale to better fit the linear model), demographic characteristics of the voting age population in the district, party registration figures (or relative proportions of votes cast in Democratic and Republican presidential primaries), and measures of candidate quality.

### 3. Preliminary Estimation

The purpose of the model (equations 3–5) is to estimate a quantity of interest  $Q$  from  $v^{(\text{hyp})}$ , for an existing, future, or counterfactual election, given a matrix of explanatory variables,  $X^{(\text{hyp})}$ , a shift parameter,  $\delta^{(\text{hyp})}$  (or the corresponding value of the expected average district vote,  $E(\bar{v}^{(\text{hyp})})$ ), and the actual election outcomes, if available. We can do so using equation (5), as long as we know the parameters,  $\beta$ ,  $\sigma^2$ , and  $\lambda$ . Our approach is to estimate the model parameters using the regression equation (3), using current and past elections as data. Unlike in many applications of regression modeling in the social sciences, the parameters  $\beta$ ,  $\sigma^2$ , and  $\lambda$  are not themselves the goal of our analysis but rather intermediate quantities, used for estimating the distribution of hypothetical election quantities.

When the current election,  $v$ , has been observed, we estimate  $\beta$  by simply regressing  $v$  on  $X$ , yielding  $\hat{\beta}$  and the usual least squares variance matrix,  $\hat{\Sigma}_\beta$ . For prediction, when  $v$  has not yet occurred, we run the same regression in the most recent year for which we have electoral results.<sup>12</sup> The parameter  $\sigma^2$  is estimated in the above regression by the usual least squares estimate of the “standard error of the regression,”  $\hat{\sigma}^2 = e'e / (n - k)$ , where  $e$  is the vector of residuals from the regression of  $v$  on  $X$ ;  $n$  is the number of observations; and  $k$  is the number of columns in  $X$ .

The remaining parameter to be estimated,  $\lambda$ , is needed when using the model for evaluation and counterfactual evaluation. To estimate  $\lambda$ , we use the Democratic proportion of the two-party vote in the election following the one used to define  $v$  and regress it on  $v$  and the original explanatory variables,  $X$ . Since we treat our analysis as conditional on uncontestedness and incumbency status, we also include the values of these two variables for the next election period. The regression coefficient on  $v$  is our estimate of  $\lambda$ , which may be thought of as an intuitive estimate of the proportion of variation due to  $\gamma$ , by directly estimating

<sup>12</sup>For our later calculations, we summarize our knowledge of  $\beta$  as a Bayesian multivariate normal posterior distribution (assuming a “noninformative” improper uniform prior distribution; see Box and Tiao 1973), with mean vector  $\hat{\beta} = (X'X)^{-1}X'y$  and covariance matrix  $\hat{\Sigma}_\beta = \hat{\sigma}^2(X'X)^{-1}$ .

how much the actual vote predicts the next election over and above the predictive power of previous values of the explanatory variables.<sup>13</sup> Rare individual estimates of  $\lambda$  above one or below zero are truncated to within this range. If the next election result is not available,  $\lambda$  can be estimated from a recent pair of election years.<sup>14</sup>

*Pooling estimates across election years.* In practice, more accurate estimates of  $\sigma^2$  and  $\lambda$  may be formed by averaging the estimates obtained in separate regressions from several election years from the same legislature. Since pooling in this way reduces the uncertainty in the estimates to essentially zero, we use the pooled point estimate instead of its distribution. Thus, we act as if we know  $\sigma^2 = \hat{\sigma}^2$  and  $\lambda = \hat{\lambda}$  for all further calculations.<sup>15</sup>

Since they are conditional on the explanatory variables,  $X$ , the estimates of  $\sigma^2$  and  $\lambda$  should be pooled only for elections for which roughly the same information is available about  $X$ . Typically, election years will be divided into two classes: those immediately following redistricting periods (and the first election in the data set), and all the others. In the former, the explanatory variables will not usually include the vote in the previous election. It makes sense to calculate two estimates of each  $\sigma^2$  and  $\lambda$  by pooling within each type of election.<sup>16</sup>

We do not advocate pooling the estimates of  $\beta$ , since these parameters are usually more volatile over time. In general, we want to estimate an electoral system using the data closest at hand, pooling only for the hyperparameters that seem to vary slowly over the years.

Empirically, our estimates of  $\sigma$  for U.S. legislative elections have been around 0.06, and almost always between about 0.02 and 0.12,

<sup>13</sup>The regression estimate is formally justified by the fact that  $\lambda$  is the coefficient of  $v$  in the expected value of the distribution for  $v^{(\text{hyp})}$  in equation (7).

<sup>14</sup>Another method of estimating  $\lambda$  is based on the fact that the theoretical correlation between  $v$  and  $v^{(\text{hyp})}$  is exactly  $\lambda$ . Although we do not observe  $v^{(\text{hyp})}$ , we can use the empirical correlation (truncated to be positive) between the residuals of the regressions predicting  $v$  from  $X$  in two successive years. We have found that the two estimators yield roughly the same estimates of  $\lambda$  in almost all empirical examples.

<sup>15</sup>To use a distribution, instead of assuming the pooled point estimate has no uncertainty, one would need the posterior distribution of  $\sigma^2$ . This is equivalent to the distribution of  $\hat{\sigma}^2 \times N(n-k)/\chi_{N(n-k)}^2$ , where  $\chi_{N(n-k)}^2$  is a chi-square random variable with  $N(n-k)$  degrees of freedom, and  $N$  is the number of election years used in the pooling. Since  $N$  is usually quite large (number of districts multiplied by the number of election years), this distribution has a very small variance, making our assumption in the text empirically reasonable. Our extensive empirical analyses, not presented here, confirm this point in practice.

<sup>16</sup>If precinct data are available and it is possible to follow the procedure in note 8, then all election years can be pooled for estimating  $\sigma^2$  and  $\lambda$ .

indicating that our explanatory variables account for the district-level vote to within plus or minus about six percentage points. Estimates of  $\lambda$  have generally ranged from 0.2 and 0.9 (most commonly near 0.6), indicating that the proportion of the variation, not explained by  $X$ , that recurs in the next election varies greatly with the electoral system and the explanatory variables used. We also find that results are usually very robust to even moderate changes in estimates of  $\lambda$  and  $\sigma$ .

#### 4. The Distribution of Hypothetical Votes

In this section, we give the distribution of  $v^{(\text{hyp})}$ , so that we can generate multiple hypothetical election results, as in Table 1, and eventually point estimates and standard errors of our quantities of interest. For prediction,  $v$  is unknown, so the goal is the unconditional distribution,  $P(v^{(\text{hyp})})$ . For evaluating an election that has already occurred, we use the information in  $v$ , which is available, and derive the distribution of  $v^{(\text{hyp})}$  given  $v$ ,  $P(v^{(\text{hyp})}|v)$ . In either case, the distribution for  $v^{(\text{hyp})}$  is implicitly conditional on  $X^{(\text{hyp})}$  and  $\delta^{(\text{hyp})}$  as well as  $\sigma^2$  and  $\lambda$  (since we take  $\sigma^2$  and  $\lambda$  as fixed after their estimation in section 3). We describe how to calculate quantities of interest and their uncertainty from these distributions in section 5.

##### 4.1. The Predictive Distribution of Future Elections

We give the predictive distribution first because it is a simpler special case. Predictive uncertainty has two components: first, the fundamental variability represented by the parameter  $\sigma^2$ , the variance of the district election results in  $v$ , conditional on the explanatory variables  $X$ , and second, the uncertainty due to our estimation of  $\beta$ , as modeled with the distribution in equation (10). If we had an infinite number of electoral districts, estimation uncertainty would drop to zero, but  $\sigma^2$  would not change. We include the variance due to estimating  $\beta$  in our analysis, which in practice is a relatively smaller addition to the  $\sigma^2$  from the model, although it does vary across districts, unlike  $\sigma^2$ .<sup>17</sup>

The technical derivation appears in Appendix B, ultimately producing the following predictive distribution of hypothetical election results:

$$P(v^{(\text{hyp})}) = N(v^{(\text{hyp})} | X^{(\text{hyp})} \hat{\beta} + \delta, X^{(\text{hyp})} \Sigma_{\beta} X^{(\text{hyp})'} + \sigma^2 I). \quad (6)$$

This is the unconditional Normal distribution of  $v^{(\text{hyp})}$ , from which we can calculate the hypothetical election results, as in Table 1, and then the

<sup>17</sup>Another possible source of variance is uncertainty in  $X$  and  $\delta$ , which we do not model mathematically. Instead, if these explanatory variables are unknown, we would just compare several analyses, with the uncertain variables set at different reasonable values.

distributions of various summaries of the electoral system. The parameters  $\lambda$  and  $\gamma$  do not appear in equation (6) and are therefore unnecessary for prediction.<sup>18</sup> This is a familiar result in econometrics (e.g., Goldberger 1991, 175–76) and indicates that “predicted values” can be calculated as usual with the possible addition of a statewide swing parameter:  $X^{(\text{hyp})} \hat{\beta} + \delta$ . The uncertainty of these predictions is given by the distribution and variance in equation (6).

#### 4.2. *Evaluating an Existing Electoral System under Actual or Counterfactual Conditions*

We now present the distribution of hypothetical votes for historical elections. Since the distribution of hypothetical votes for evaluating counterfactual conditions produces actual evaluations as a special case, we save space by including only the general result here.

When designing counterfactual scenarios, one should be careful in specifying  $X^{(\text{hyp})}$  and deciding exactly what conditions to be held constant. For example, what if a term limitation initiative had swept the nation before the last election, and all the incumbent members of Congress were forced to retire? How would this have affected the electoral system?<sup>19</sup> The only practical difficulty is precisely defining the conditions under which the election would be held—in our model, the explanatory variables,  $X^{(\text{hyp})}$  and statewide swing  $\delta^{(\text{hyp})}$  (or  $E(\bar{v}^{(\text{hyp})})$ ). One possible definition of this particular counterfactual scenario is as follows. Obviously, no incumbents run for reelection, so the column of  $X^{(\text{hyp})}$  indicating incumbency status should have all zeros. With all seats open, it is also reasonable to assume that all the districts are contested. Now, instead of setting  $E(\bar{v}^{(\text{hyp})})$ , we can just set the aggregate partisan swing,  $\delta^{(\text{hyp})}$ , to zero, making the assumption that nothing else systematic changes but incumbency and contestedness. With all these variables set, the hypothetical elections can be modeled as in this section.

In general, when designing counterfactual scenarios, it is important to control only for variables that happened *before* the intervention of interest (see King 1991b). For example, when supposing that no incumbents will run for reelection, measures of campaign contributions and

<sup>18</sup>To account for the uncertainty from estimating  $\sigma^2$  explicitly, the normal distribution would be changed to a Student's  $t$  distribution with  $n - k$  degrees of freedom. However, pooling the estimate of  $\sigma^2$  across elections, as we recommend, makes the degrees of freedom large, and the normal distribution in this situation approximates the  $t$  quite well.

<sup>19</sup>In the absence of the term limit threat, King and Gelman (1991) asked what the electoral system would have been like had the incumbency advantage suddenly disappeared, in order to understand the effect of the incumbency advantage on postwar U.S. House elections.



candidate quality, for both incumbents and challengers, should be discarded (or appropriately modified). Campaign contributions and the decision of opposing candidates of quality to run for election are both largely decided after the incumbent's reelection intentions are known and in part the consequence of incumbency advantage.

The technical details of our derivation appear in Appendix C. The result, which appears complicated at first, is the following Normal probability distribution of hypothetical district election results given observed votes:

$$P(v^{(\text{hyp})} | v) = N(v^{(\text{hyp})} | \lambda v + (X^{(\text{hyp})} - \lambda X) \hat{\beta} + \delta^{(\text{hyp})}, (1 - \lambda^2) \sigma^2 I + (X^{(\text{hyp})} - \lambda X) \Sigma_{\beta} (X^{(\text{hyp})} - \lambda X)'). \quad (7)$$

A good way to understand this result is to focus on the special case where we are analyzing the actual election, and therefore  $X^{(\text{hyp})} = X$  and  $\delta^{(\text{hyp})} = 0$ . In this case, the expected Democratic proportion of the two-party vote in district  $i$  from equation (7) simplifies to

$$E(v_i^{(\text{hyp})} | v) = \lambda v + (1 - \lambda) X \beta, \quad (8)$$

a weighted average of the observed vote,  $v$ , and the vote predicted from past elections,  $X\beta$ .

Considering only  $v$  (the correct estimate if  $\lambda = 1$ ) produces the discredited uniform partisan swing assumption, while the other extreme of using only  $X\beta$  (if  $\lambda = 0$ ) wrongly ignores the information from the current election, which is obviously relevant to estimating the current electoral system even after taking into account one's predictions. To put it another way,  $\lambda$  is a parameter that determines which estimator of the form  $\lambda v + (1 - \lambda) X \beta$  is most effective at estimating the district-level results of another election,  $v^{(\text{hyp})}$ , from the same electoral system (i.e., if everything were the same up to the start of the campaign and the campaign and election were run again). Thus, in characterizing the expected value of the distribution  $P(v^{(\text{hyp})} | v)$ , we clearly want to use our prediction  $X\beta$ , and we also want to use the systematic and persistent aspects of  $v$  not predicted by  $X\beta$ . The question is how much to weight  $v$  and  $X\beta$ ;  $\lambda$  is the weight and our method of estimating  $\lambda$  provides an answer.

We also use the entire distribution, including the variance of the predictive equation, to generate hypothetical election results like those in Table 1 and ultimately to obtain expressions of the uncertainty associated with various summaries of the electoral system.

## 5. Calculating Summaries of Electoral Systems

Once the distribution of hypothetical election results has been obtained, using either equation (6) for prediction or equation (7) for actual

or counterfactual evaluation, we calculate various summaries of interest of the electoral system and their standard errors. These may include, for example, district-level vote and seat forecasts, partisan bias, electoral responsiveness, the number of incumbents who are reelected, and others listed in section 1.

The general methodology we introduce to accomplish these tasks is called “Bayesian simulation,” although it could as easily have been called “approximating conditional probability distributions by drawing random numbers.” It is also intuitively explained in Table 1. To explain the exact mathematical procedure intuitively, section 5.1 shows how to calculate one feature of the electoral system—forecasts of district vote proportions—that can be calculated analytically. Section 5.2 demonstrates how to calculate these and other summaries with Bayesian simulation. We then turn to other features of an electoral system for which Bayesian simulation works, but no analytic solution is available. No new assumptions are introduced in this section; we show only how to re-express the simulations of district-level vote proportions in more useful formats. Bayesian simulation should be widely applicable to many political science problems, beyond those we describe here.

### *5.1. District-level Vote and Seat Predictions: Analytic Solution*

We can predict district vote proportions using our estimate of the expected vote in each district,  $E(v_i^{(\text{hyp})})$ , the average over the range of possible hypothetical election results,  $v_i^{(\text{hyp})}$ . We calculate this from the expected value of the multivariate normal distribution from either equation (6) for prediction or equation (7) for actual or counterfactual evaluation. For example, the vector of district-level predictions for the Democratic proportion of the two-party vote is  $X^{(\text{hyp})}\hat{\beta} + \delta^{(\text{hyp})}$ . The variables in  $X^{(\text{hyp})}$  are chosen at the start, and  $\hat{\beta}$  is estimated as part of the procedure. Only the constant,  $\delta^{(\text{hyp})}$ , needs to be specified. For most purposes,  $\delta^{(\text{hyp})}$  will be set to zero, but to simulate statewide swings it will also be useful to set it to other values. For example, we could set  $\delta^{(\text{hyp})}$  so that the average district vote takes on several specified values, such as those of the last several elections, to see how the predictions depend on the statewide partisan swing.

To calculate standard errors for the predictions, we can use the standard deviation of the same multivariate normal distribution. However, in this case, and most others, one can calculate two types of standard errors, based on either the predictive uncertainty or only the uncertainty in the expected value. For district-level vote predictions, total predictive uncertainty is probably most relevant. However, if one wishes a summary of only the uncertainty in where the expected vote is, then only the estimation uncertainty should be used.

The standard errors based on the total predictive uncertainty are merely the square roots of the diagonal elements of the variance matrix in equation (6),  $X^{(\text{hyp})} \Sigma_{\beta} X^{(\text{hyp})'} + \sigma^2 I$  (or equation 7 for actual or counterfactual evaluation). For standard errors based only on the estimation uncertainty in the expected district vote, we would take the square root of the diagonal elements of only the first term in the variance matrix, such as  $X^{(\text{hyp})} \Sigma_{\beta} X^{(\text{hyp})'}$  for prediction.

Another useful way to summarize the uncertainty for district-level predictions is to focus not on votes but on the probability that the election in district  $i$  goes for the Democrat (or Republican),

$$\begin{aligned} P(\text{Democrat wins}) &= P(v_i^{(\text{hyp})} > 0.5) \\ &= \Phi \left[ \frac{E(v_i^{(\text{hyp})}) - 0.5}{\text{var}(v_i^{(\text{hyp})})^{1/2}} \right], \end{aligned} \tag{9}$$

where  $\Phi$  is the standard normal cumulative distribution function. When aggregated over all districts, these probabilities provide a better measure of Democratic strength than counting the number of districts with  $E(v_i^{(\text{hyp})}) > 0.5$ . For example, if all the districts in a state have  $E(v_i^{(\text{hyp})}) = 0.49$  and  $\text{var}(v_i^{(\text{hyp})}) = 0.07^2$ , then, for each district,  $P(\text{Democrat wins}) = \Phi(-0.01/0.07) = 0.44$ . However, we certainly would not expect every Republican candidate in the state to win; with enough districts, even small probabilities will produce at least some Democratic wins.

## 5.2. An Introduction to Bayesian Simulation

Bayesian simulation allows us to summarize the electoral system using observable quantities of immediate interest. In contrast, just reporting regression coefficients or some other model estimate may require convoluted reasoning to translate into something with substantive political meaning.

For many summaries, one can calculate exact analytic solutions directly from the expected values in either equation (6) or equation (7), as was done in section 5.1. In general, though, simulation is the easiest approach to computing estimates and standard errors of summaries of the distributions we have derived. It is frequently easier, computationally faster, and simpler to understand than the analytic solution.

For example, suppose one has a distribution  $f(y)$  and wishes to calculate its expected value. The formal ‘‘analytic’’ method would be to compute the integral,  $E(Y) = \int_{-\infty}^{\infty} yf(y) dy$ , which is difficult or impossible in many cases, especially if  $y$  represents a large vector of dependent random variables. Fortunately, if random draws can be obtained from this distribution (a task that is often easy, especially with modern computers), it

is possible to use sampling theory to approximate the expected value to an arbitrary degree of precision. That is, one merely takes a large number of random draws from this distribution and averages them. The desired degree of precision can be attained by merely increasing the number of draws. We apply this same logic to our distributions and theoretical features with any degree of complexity.

Simulating hypothetical elections from our model requires three steps:

1. Create a large number,  $m$ , of simulated hypothetical election results,  $v^{(\text{hyp})}$ , by taking random draws from the multivariate normal distribution (in equation 6 or 7). Each election result is a vector of length  $n$ ,  $v^{(\text{hyp})} = (v_1^{(\text{hyp})}, \dots, v_n^{(\text{hyp})})$  and is portrayed as a column in Table 1.
2. For each simulated vector of hypothetical election results  $v^{(\text{hyp})}$ , calculate the desired descriptive summary (e.g., vote in district 4,  $v_4^{(\text{hyp})}$ ; expected seats given average district vote set at 0.5,  $E(\bar{s} | \bar{v} = 0.5)$ ; or the proportion of districts with between 0.25 and 0.5 probability of a Democratic win). This is  $Q^{(\text{hyp})j}$  in Table 1.
3. Calculate the mean and standard deviation of the  $m$  summaries (as in equations 1 and 2); these constitute the point estimate and its standard error, respectively.

If it were possible to use an infinite number of simulations ( $m \rightarrow \infty$ ), this procedure would yield exact standard errors and point estimates. For most purposes,  $m = 100$  suffices to produce sufficiently precise approximations.

Perhaps the only unfamiliar technical issue involved here is how to draw values randomly from one of the two multivariate normal distributions. Fortunately, the solution to this problem is quite simple. To create a single simulated hypothetical election outcome (one of the  $m$  required in the description, above):<sup>20</sup>

1. Draw one random vector,  $\beta^\circ$ , from the posterior distribution for  $\beta$  (see equation 10).
2. If the current election,  $v$ , has been observed, insert  $\beta^\circ$  into the distribution for  $\gamma | \beta$  (see equation 15), and draw a vector,  $\gamma^\circ$ , from

<sup>20</sup>One might think that draws from the posterior distribution of  $v^{(\text{hyp})}$  might easily be obtained directly from the appropriate normal distribution (equation 6 for prediction or equation 7 for observed elections), using some method like the Cholesky decomposition applied to the covariance matrix. However, that approach requires inverting an  $n$ -dimensional matrix, a task that is not desirable or even possible here because the covariance matrixes are singular.

this distribution. In prediction, for which  $v$  has not yet occurred, skip this step.

3. Finally, for prediction, insert  $\beta^\circ$  into the distribution  $P(v^{(\text{hyp})}|\beta)$  (see equation 11), or for evaluation or counterfactual evaluation insert  $\beta^\circ$  and  $\gamma^\circ$  into the distribution  $P(v^{(\text{hyp})}|\gamma, \beta, v)$  (see equation 14), set  $\delta^{(\text{hyp})}$  as you choose, and draw a single random vector,  $v^{(\text{hyp})}$ , from the appropriate distribution.<sup>21</sup>

We now apply Bayesian simulation to calculate several more complicated quantities of interest, along with their standard errors.

### 5.3. *The Distribution of Seats Given Votes*

Following the notation of Gelman and King (1990b) and King and Gelman (1991), we label  $\bar{v}$  and  $\bar{s}$  as the average district vote and the average proportion of seats obtained by the Democrats in the current election, and  $\bar{v}^{(\text{hyp})}$  and  $\bar{s}^{(\text{hyp})}$  for the corresponding hypothetical quantities. To obtain the seats-votes curve, we estimate the distribution of  $\bar{s}^{(\text{hyp})}$ , given  $\bar{v}^{(\text{hyp})}$ , for a range of values of  $\bar{v}^{(\text{hyp})}$ . Of course, our predictions and estimates are only as good as our model, and it is only reasonable to expect the model to hold over a plausible range of aggregate outcomes. For example, if  $\bar{v}^{(\text{hyp})}$  is currently 0.5, and has been near that for many years, we would not use our model to predict what would happen in the case of  $\bar{v} = 0.2$ ; such a change would imply an electoral transition beyond the scope of the historical data on which the model is based. We typically calculate the seats-votes curve for a range of votes centered at the current value, such as every percentage point in the range  $[\bar{v}^{(\text{hyp})} \pm 0.15]$ , or perhaps at a conventional range such as  $[0.4, 0.6]$ .

Whether for a predicted or hypothetical election, we can easily obtain the distribution of  $\bar{s}^{(\text{hyp})}$ , given  $\bar{v}^{(\text{hyp})}$ , for each of a range of values,  $\bar{v}^{(1)}, \dots, \bar{v}^{(l)}$ , as follows:

1. Using the appropriate distribution (prediction of a future election or actual or counterfactual evaluation of a past election), simulate a set of  $m$  election results,  $v^{(\text{hyp})}$ , given some arbitrary value,  $\bar{v}^{(0)}$  (or the corresponding value of  $\delta^{(\text{hyp})}$ ), for the average district vote.
2. For  $j = 1, 2, \dots, l$ :
  - a. Add the constant,  $\bar{v}^{(j)} - \bar{v}^{(0)}$ , to each component of each simulated vector,  $v$ , produced in step 1. No new simulations are required.

<sup>21</sup>The distributions in equations (11) and (14) have diagonal covariance matrices, so this last step requires draws from independent normal distributions.

- b. Calculate  $\bar{s}^{(\text{hyp})}$  for each simulated election; the set of these  $m$  values estimates the distribution of  $\bar{s}^{(\text{hyp})}$ , given  $\bar{v}^{(\text{hyp})} = \bar{v}^{(j)}$ . In particular, the sample mean and standard deviation of the  $m$  values provide estimates of  $E(\bar{s}^{(\text{hyp})} | \bar{v}^{(\text{hyp})})$  and  $\sqrt{\text{var}(\bar{s}^{(\text{hyp})} | \bar{v}^{(\text{hyp})})}$ , respectively.

One can then plot the seats-votes curve with dotted lines at two standard deviations above and below the seats-votes curve, to indicate the region where, according to the model, 95% of predicted elections will fall.

In addition to summarizing our uncertainty in any given prediction, the standard errors make our predictions testable: we can run the model on a past election, make a prediction, and then see whether about 95% of the election outcomes fall in the 95% region. Repeating this over many past elections, we can see whether our assessments of uncertainty are accurate.

#### 5.4. Bias and Responsiveness

We can calculate bias and responsiveness (along with standard errors) with the same simulations used for the seats-votes curve. We define four quantities of interest:

- Responsiveness as  $\bar{v}^{(\text{hyp})}$  ranges from 0.45 to 0.55: the average difference,  $[E(\bar{s}^{(\text{hyp})} | \bar{v}^{(\text{hyp})} = 0.55) - E(\bar{s}^{(\text{hyp})} | \bar{v}^{(\text{hyp})} = 0.45)]$ , divided by the vote swing,  $0.55 - 0.45$ . If the electoral system one is studying regularly produces average district votes at or near this interval, then this is a reasonable estimate to evaluate the responsiveness of the electoral system. For example, a responsiveness of 2.3 means that a 1% increase in the average district vote for a party across districts will translate into an expected 2.3% increase in seats for the same party in the legislature (plus or minus the estimated standard error). The responsiveness is of interest to people concerned that redistricting plans will be drawn to protect incumbents from vote swings. In addition, scholars at least since Mayhew (1974) and Erikson (1971) have been interested in the responsiveness of the U.S. House and other legislatures.
- Responsiveness at a chosen value  $v_0$  (such as the actual average district vote): the average difference,  $[E(\bar{s}^{(\text{hyp})} | \bar{v}^{(\text{hyp})} = v_0 + 0.01) - E(\bar{s}^{(\text{hyp})} | \bar{v}^{(\text{hyp})} = v_0 - 0.01)]$ , divided by the vote swing, 0.02. One must be careful with this measure, however, since it tracks with  $v_0$ ; even if the electoral system is unchanging, a change in  $\bar{v}$  moves the system to a new point on the seats-votes curve, and a new local responsiveness. On the other hand, setting  $v_0$  to the observed  $\bar{v}$  is an unquestionably realistic point.

- Average partisan bias between  $\bar{v}^{(\text{hyp})} = 0.45$  and  $\bar{v}^{(\text{hyp})} = 0.55$ . We define partisan bias as the deviation from partisan symmetry. For example, if one party is able to translate 55% of the average district vote into 75% of the seats in the legislature, then it would be symmetric for the other party, if it were to receive 55% of the average district vote, to also receive 75% of the seats. We follow King and Browning (1987), King (1989a), and Gelman and King (1990b) and define partisan bias as the proportion of the seats in the legislature the Democrats receive over and above what is fair. For example, if partisan bias is  $-0.05$ , then the Democrats receive 5% fewer seats in the legislature than they should under the symmetry standard (and the Republicans receive 5% more seats than they should).
- Partisan bias at  $\bar{v}^{(\text{hyp})} = 0.5$ : the average value of  $\bar{s}^{(\text{hyp})}$ , given  $\bar{v}^{(\text{hyp})} = 0.5$ , minus 0.5; that is,  $E(\bar{s}^{(\text{hyp})} | \bar{v}^{(\text{hyp})} = 0.5) - 0.5$ . This is interpreted in the same manner as above, but is most appropriate for very competitive electoral systems. This partisan bias statistic is the expected proportion of the seats over 0.5 that the Democrats receive when they receive exactly half the average district vote.

In all these cases, we calculate the summary for each simulated hypothetical election result,  $v^{(\text{hyp})}$ , and then use the average of the  $m$  summaries as the point estimate and their standard deviation as the estimate of uncertainty (the standard error).

The two definitions of bias presented above will generally produce very similar results because it is rare to see an electoral system whose districts are all clustered around 0.5. In most situations, we prefer bias averaged from 0.45 to 0.55 because it allows for a realistic range of average district vote,  $\bar{v}$ . For the same reason, we prefer responsiveness measured over a 10% vote swing; again, except at the extremes, swing has little effect on the estimates.

We allow responsiveness to be defined at any value of the average district vote; 0.5 is often a natural reference point, but it is frequently more reasonable to use a value nearer to typical election outcomes. For example, the average Democratic share of the vote in U.S. House districts has fluctuated around 0.55 for decades, and many southern state legislatures have never seen the average Democratic vote fall below 0.6.

In contrast, we define partisan bias as a departure of the seats-votes curve from partisan symmetry, which only makes sense when centered at  $\bar{v} = 0.5$ . In a state in which the Democrats have never achieved more than  $\bar{v} = 0.2$ , for example, we would not trust any model to make assumptions about what might happen if they suddenly received half the vote.

This is especially important when considering minority representation (see King, Bruce, and Gelman 1993).

### 5.5. Other Summaries

The flexibility of our Bayesian simulation-based method allows us to estimate any summary of the predicted or hypothetical elections, not just the specific quantities discussed above.

For example, suppose we are evaluating a redistricting plan and are interested in the likely reelection rate: the proportion of incumbents, of both parties, who will win their bids for reelection. For purposes of the prediction, we shall assume that we know which incumbents will run for reelection, and that the statewide partisan swing,  $\delta^{(hyp)}$ , is zero. The following procedure, again using the  $m$  simulated predicted elections  $v^{(hyp)}$ , creates an estimate and standard error for the reelection rate.

1. Simulate  $m$  vectors,  $v^{(hyp)}$ , of hypothetical elections, given the specified values of  $X^{(hyp)}$  and  $\delta^{(hyp)}$ .
2. For each set of hypothetical election results, compute the proportion of incumbents who win reelection. That is, count the number of hypothetical districts with incumbents whose vote,  $v_i^{(hyp)}$ , exceeds 0.5 (for Democrats) or is less than 0.5 (for Republicans) and divide by the number of incumbents running.
3. The resulting set of  $m$  proportions approximates the distribution of the reelection rate. If we wish, we can summarize the distribution with estimates of its expected value and corresponding standard error by calculating the mean and standard deviation of the  $m$  values.

To include a more realistic assessment of uncertainty, one should repeat the above analysis with a range of plausible values of  $\delta^{(hyp)}$  (or  $E(\bar{v}^{(hyp)}) = \delta^{(hyp)} + \sum_{i=1}^n (X^{(hyp)}\beta)_i$ ), rather than merely setting  $\delta^{(hyp)} = 0$ . Reasonable values of average district vote may be obtained by examining a graph of  $\bar{v}$  in recent elections, or more formally using a forecasting procedure. Repeating the above analysis with  $k$  different values of  $\delta^{(hyp)}$  yields  $km$  simulated vectors  $v^{(hyp)}$ , from which means and standard deviations of all quantities can be estimated, as before.

For another, more complicated, example, suppose we are interested in the number of "marginal seats" in an electoral system, which we shall define (for convenience) as seats with at least one chance in four of changing parties if the election were to be repeated with the same pattern of uncontested seats and incumbents. We shall assume, to be specific, that in the hypothetical repeated election, a statewide partisan swing of up to two percentage points from the current value,  $\bar{v}$ , could occur in



either direction. We can estimate the number of marginal seats, so defined, as follows.

1. Perform the preliminary estimation as described in section 4.
2. Simulate  $\beta^\circ$  and, if  $v$  has been observed,  $\gamma^\circ$ , using steps 1 and 2 of section 5.2.
3. Perform the following steps 100 times, to obtain 100 simulated vectors,  $v^{(\text{hyp})}$ .
  - a. Simulate one hypothetical election result,  $v^{(\text{hyp})}$ , with average district vote equal to the current  $\bar{v}$ , following step 3 of section 5.2.
  - b. Add the shift,  $\delta^{(\text{hyp})}$ , to all districts in the vector  $v^{(\text{hyp})}$ , where  $\delta^{(\text{hyp})}$  is drawn at random from a uniform distribution with range  $[-0.02, +0.02]$ .
4. For each district, count the number of times the seat changed parties (i.e., that  $v_i^{(\text{hyp})} > 0.5$  if the observed  $v_i < 0.5$  and vice versa). If a district changed parties in at least 25 of 100 simulations, it is "marginal." Compute the proportion of marginal districts out of the  $n$  districts.
5. Repeat steps 2–4  $m$  times, to create  $m$  simulations of the electoral system and  $m$  corresponding simulated proportions of marginal seats in each hypothetical election. Our point estimate of the proportion of marginal seats is the average of the  $m$  proportions, and their standard error is the corresponding standard deviation over the  $m$  simulations.

## 6. Examples

### 6.1. The U.S. House of Representatives since 1900

In this section, we estimate partisan bias and electoral responsiveness for the House of Representatives in nonsouthern states during this century. To begin, we define  $X$  as including a constant term along with  $inc(t)$ ,  $unc(t)$ ,  $party(t)$ ,  $v(t-1)$ ,  $inc(t-1)$ , and  $unc(t-1)$ . In years following redistricting (i.e., those ending in "2"), we do not include the variables corresponding to time  $(t-1)$ , since they are not available without a great deal of work matching and disaggregating districts between redistrictings. We ignore the fact that some states performed additional redistrictings in other years.<sup>22</sup>

<sup>22</sup>We imputed uncontested district election results at 0.25 and 0.75 and discarded the few cases of multimember districts and districts with third-party victories. Had there been many anomalous district elections, we would have to use more complicated methods to handle the missing data, but it did not seem worth the effort here.

We estimate  $\sigma^2$  and  $\lambda$  for as many of the years as possible. (The parameter  $\lambda$  cannot be estimated for the final election before a redistricting.) We then pooled estimates for years following redistricting and also for all other years. For years following redistricting, our estimates are  $\hat{\sigma} = 0.091$  and  $\hat{\lambda} = 0.663$ , and for all other years,  $\hat{\sigma} = 0.065$  and  $\hat{\lambda} = 0.560$ . Recall that  $\sigma$  is a proportion, so 0.065 means that we can forecast future district-level results to within about 6.5% of the vote (plus the error in estimating  $\beta$ ). The relative values of these parameters are reasonable: in years following redistricting, compared to other years, we have less information (fewer variables) in  $X$ ; as a result, predictive uncertainty, as measured by  $\sigma$ , is higher. In addition, when estimating hypothetical election results, the variables in  $X$ , representing past elections, are less useful immediately following redistricting, and so  $\lambda$  is higher, as it should be. (The parameter  $\lambda$  is highest when  $v^{(\text{hyp})}$  is estimated solely from  $v$  and is lowest when only  $X$  is used.)

Figure 1 plots partisan bias (with  $E(\bar{v}^{(\text{hyp})})$  from 0.45 to 0.55). A standard error bar appears at the top of the graph. Most small year-to-year changes in bias are within a standard error or two, indicating stochastic variation, rather than a systematic change in the underlying electoral system. However, the long-term systemic trends are quite substantial. The evidence demonstrates unambiguous bias in favor of the Republican party from the 1930s until the mid-1960s. After that, the trend toward Democratic bias is consistent and sustained. The pattern during the first two decades of this century is less clear, and the standard errors are much larger. These substantive results are close to those produced by the more complicated, and computationally intensive, model of King and Gelman (1991) for the period they examined, 1946–86. The explanation for the trend since the mid-1960s is the growth in incumbency advantage from about 2% to nearly 12% by the late 1980s.

Figure 2 displays the estimates of electoral responsiveness (with  $E(\bar{v}^{(\text{hyp})})$  set to the actual average district vote,  $\bar{v}$ ) of the U.S. House during the same period, evaluated at the observed average district vote. For comparison, the figure also presents this same summary calculated with the assumption of uniform partisan swing. On average over many years, the two methods produce roughly the same results, but the uniform partisan swing estimate is much more variable, an indication that it is fitting transient year-to-year changes instead of the long-term systematic patterns of interest. This relative “statistical inefficiency” reflects the fact that uniform partisan swing uses only the information in  $v$ , whereas our model uses information in  $v$  as well as in the predicted values  $X\beta$  (see equation 8). As a consequence, uniform partisan swing might yield reasonable estimates on average over a century, but it can be far from

Figure 1. Partisan Bias in the U.S. House, Non-South

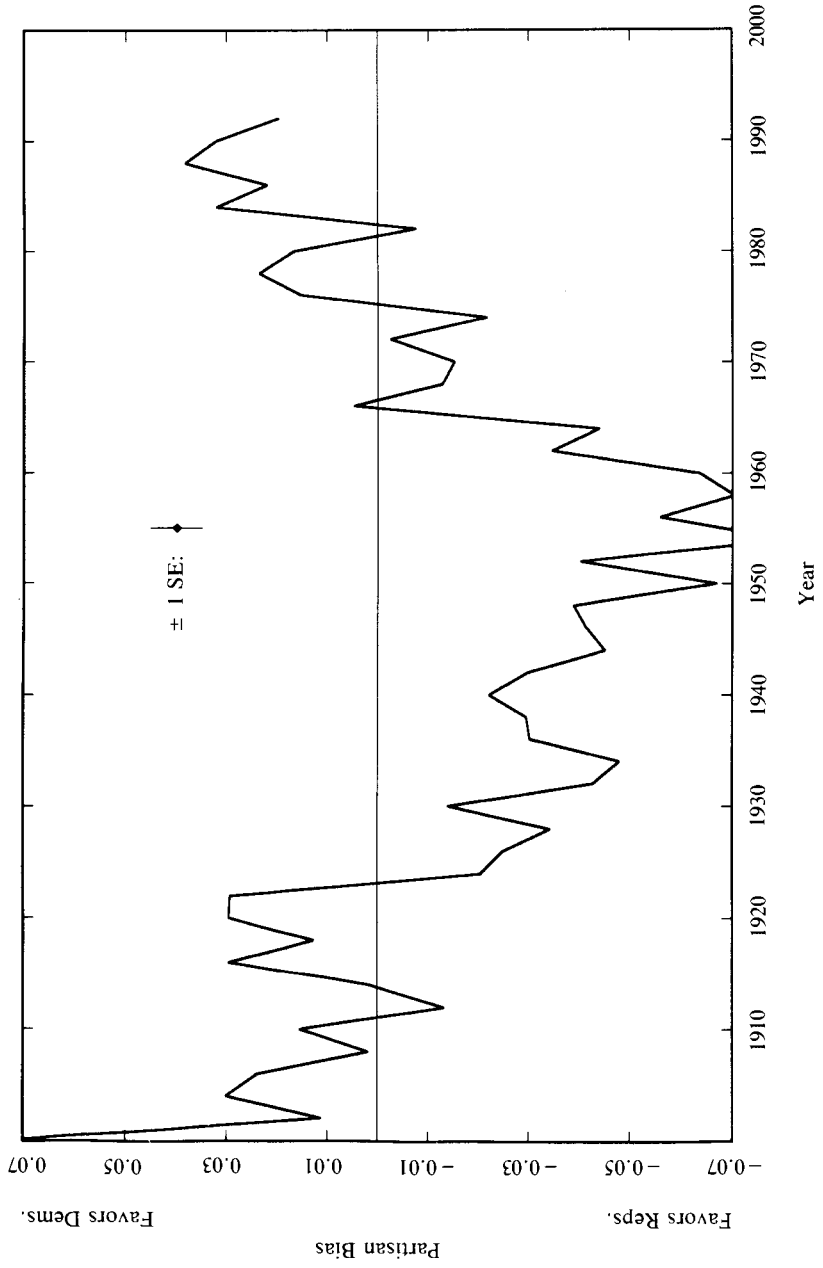
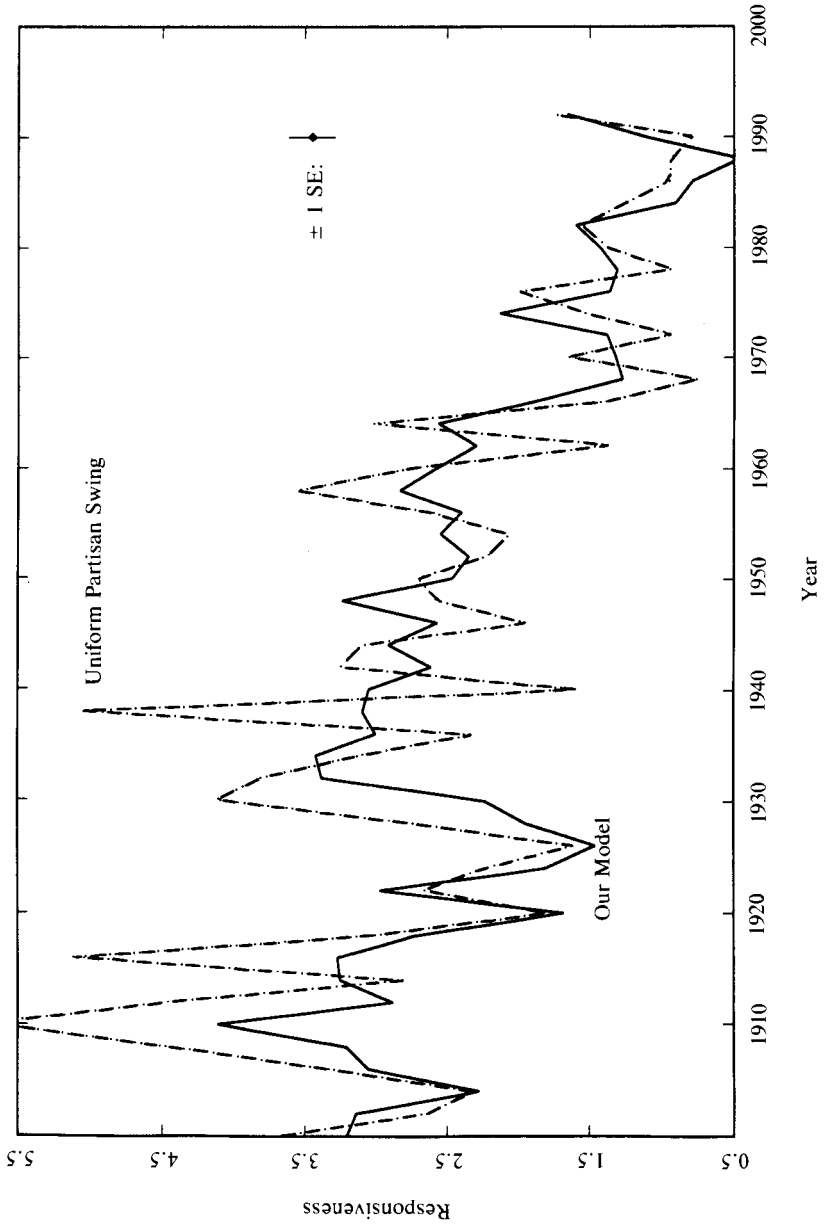


Figure 2. Responsiveness in the U.S. House, Non-South



accurate for any one year, or even as an average over a small number of years. In addition, the implied standard errors of uniform partisan swing are all zero, which corresponds to the incredible assumption that, given  $v$  and  $\bar{v}^{(\text{hyp})}$ , the hypothetical election outcomes,  $v^{(\text{hyp})}$ , are exactly known and implies that all summaries of the electoral system are known perfectly. Our method is estimated one year at a time. Thus, the smoothness of the resulting estimates is not an artifact but is due to a lower level of estimation error resulting from the much larger quantity of information extracted from the same data in each election used to make the desired inference.

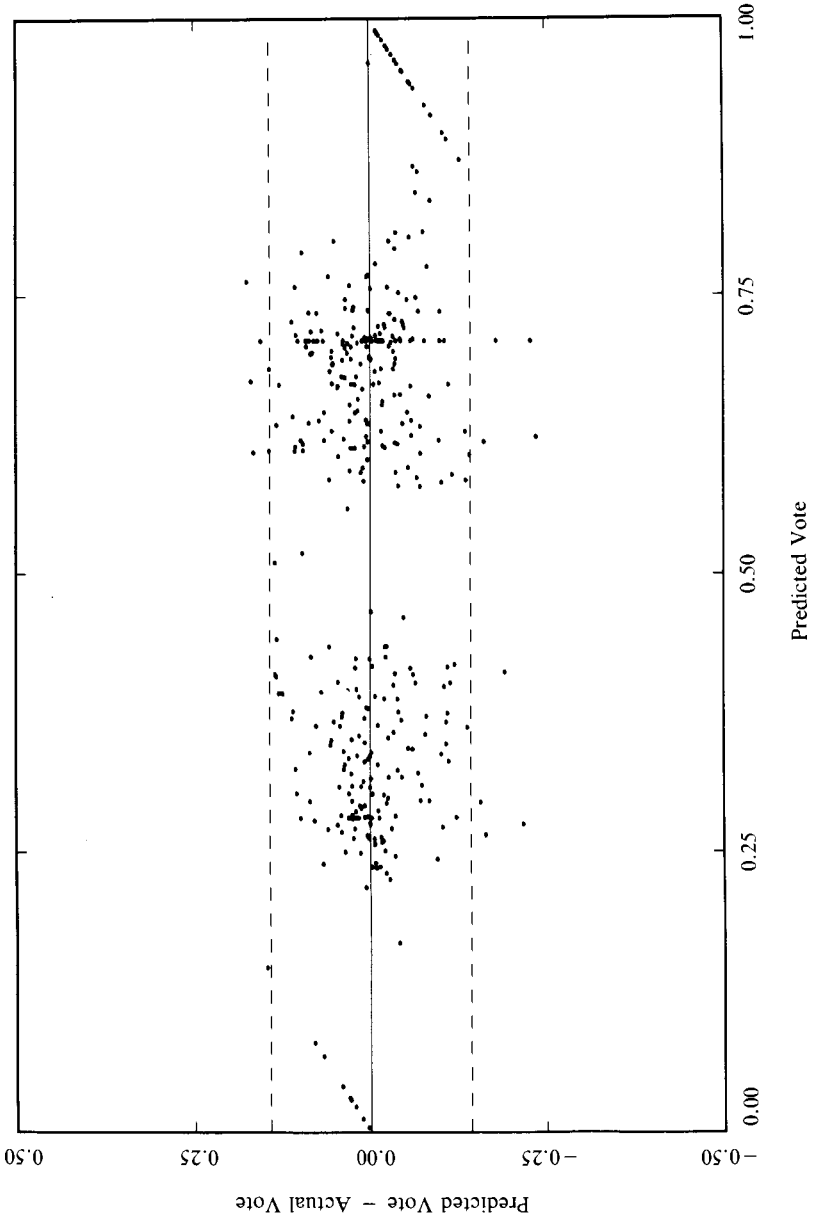
*Evaluating predictive uncertainty estimates.* To illustrate the accuracy of forecasts made with our model, and most important, the accuracy of the uncertainty estimates implied, we predict the U.S. House of Representatives in 1988. As an example, we use as explanatory variables district election results of 1986 and incumbency and uncontestedness of each district for 1986 and 1988, variables that are typically available months before the election. In state legislative elections, our forecasts will frequently be better because better explanatory variables, such as votes for statewide offices broken down into legislative districts, are usually available. We could also improve forecasts of these U.S. House data with other variables, such as campaign spending, scandals, candidate quality, and the like, but our primary purpose here is to evaluate the model's uncertainty estimates, not to produce the best possible predictions. As usual, we only predict the *relative* district election results; for convenience, we shall assume that the nationwide partisan swing,  $\delta^{(\text{hyp})}$ , is zero.

To apply our model, we first estimate the parameter  $\beta$  by predicting 1986 from 1984, which is, in fact, the same estimation done above for evaluating the electoral system. We then draw from the distribution of predicted 1988 elections, using the estimated  $\beta$  and the pooled estimate of  $\sigma^2$  obtained above.

For each district,  $i$ , we label the means of the simulated predictions as  $\hat{v}_i$  and the actual election result as  $v_i$ . We would like to evaluate our model by comparing the prediction errors,  $v_i - \hat{v}_i$ , to the standard errors of the district election predictions.<sup>23</sup> Figure 3 plots the prediction errors versus the observed election outcomes. A solid line appears at the nationwide partisan swing, and dotted lines are displayed at plus and minus

<sup>23</sup>Before making this comparison, we correct for the nationwide partisan swing, which we do not try to predict with our model. From 1986 to 1988, the average Democratic district vote swung from 57.2% to 56.6%.

Figure 3. Predicting the 1988 U.S. House Election



twice the average predictive standard error. If the uncertainty estimates are exactly correct, 95% of the prediction errors would fall within the dotted lines on average; the actual result in this case was 96.4%.<sup>24</sup>

We also evaluated the accuracy of our uncertainty estimates by comparing the average of the squared actual prediction errors from the prediction model with the estimated prediction variance from the model (the average of the diagonal elements of the variance in equation 7). The average of the observed squared prediction errors was 0.065<sup>2</sup>, which is slightly smaller than the average predicted variance, 0.071<sup>2</sup>.

The fact that the predictive standard errors in Figure 3 and the predictive error variance just calculated are so close to their theoretical values is very strong confirmation of our model and of the accuracy of the uncertainty estimates it produces. Moreover, our extensive analyses of thousands of other elections and with different sets of explanatory variables give the same consistent support for our model and its uncertainty estimates.

Some may view the uncertainty estimates reported here as large. In part, this is because almost all existing election forecasters calculate uncertainty estimates incorrectly, and so those presented here are among the first correct estimates published (see Beck 1992; Greene 1993). They can be reduced by including explanatory variables based on more detailed knowledge of the district election results, such as campaign spending, scandals, candidate quality, relative campaign effectiveness, and the like. In our experience, the state of knowledge about elections is unlikely to reduce our uncertainty by more than half the illustrative values we report here. Most important for present purposes is that the figures reported here demonstrate that our model produces accurate uncertainty estimates. With this model and computer program, scholars can process correctly whatever explanatory variables they are innovative enough to collect.<sup>25</sup>

<sup>24</sup>If the absolute values of the prediction errors were consistently too low, our model would be overstating our predictive uncertainty, yielding overly cautious predictions. High absolute prediction errors would imply that the predictions are less accurate than expected, yielding overconfident predictions. The former wastes information; the latter risks incorrect inferences.

<sup>25</sup>In an earlier version of this paper, Figure 3 had four points that appeared to be true outliers because they stood more than four standard deviations away from their predicted values. In fact, these turned out to be coding rather than prediction errors, and our model does fit quite well. In practice, the existence of a few real outliers, were they to occur, would not greatly affect most of the electoral summaries we calculate, such as bias, responsiveness, number of incumbents to lose reelection, etc.

## 6.2. *Redistricting the Ohio Legislature*

Legislatures and courts are often called upon to evaluate competing redistricting plans. We show here how our method can be applied to this problem by estimating partisan bias for elections held in the lower house of the Ohio state legislature from 1968 to 1990 and by predicting partisan bias for each of three redistricting plans proposed for the legislature in 1992.

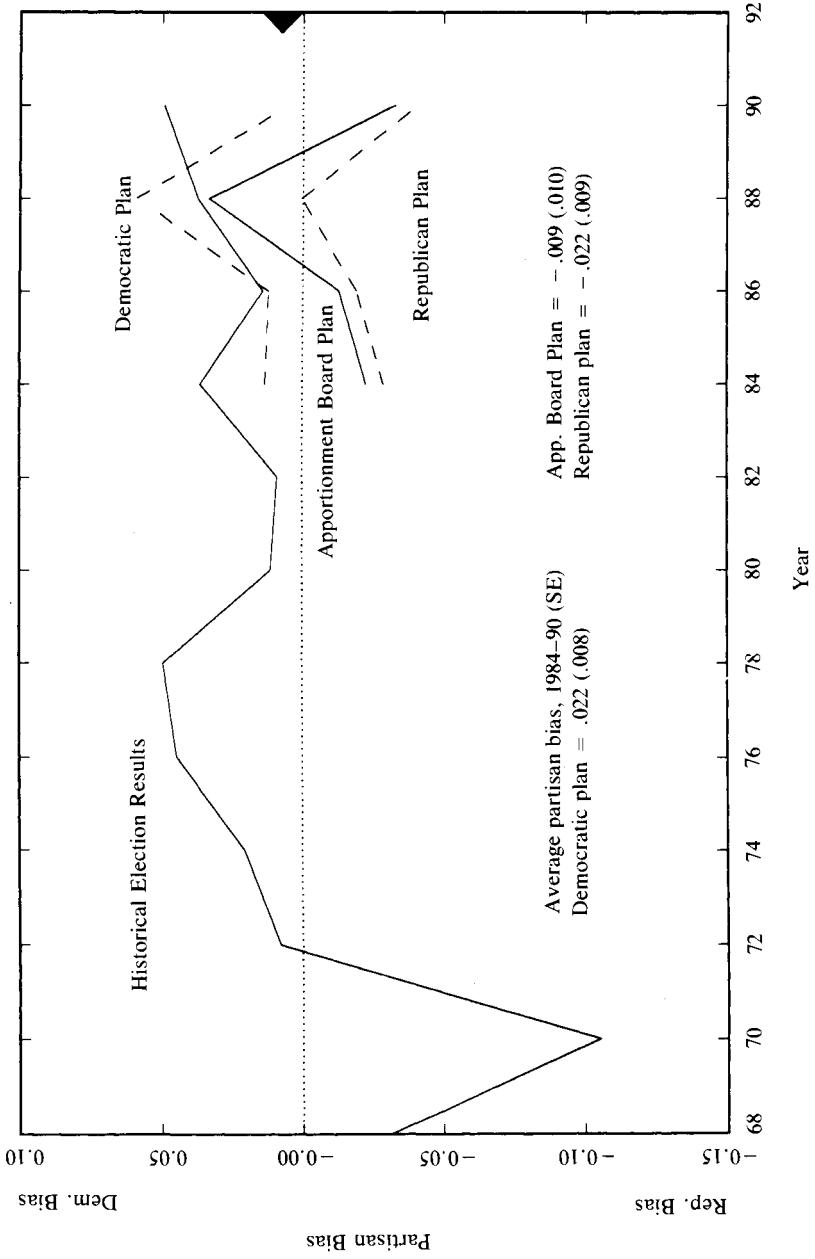
For the current districts and the proposed redistricting plans, we use the following explanatory variables: incumbency, uncontestedness, previous vote for the state House, all available statewide races broken down into current or proposed legislative districts, and the proportion of African Americans of the voting age population in the district. We estimate the parameters,  $\beta$ ,  $\sigma^2$ , and  $\lambda$  with these explanatory variables for current districts, the latter two pooled over the years.

During the period 1968–90 to which our data apply, elections have been held under three redistricting plans. The long solid line in Figure 4 demonstrates the consequences of these redistricting plans by tracing the degree and direction of partisan bias (estimated from 0.45 to 0.55 average district vote) in the Ohio State House of Representatives. In 1971 the redistricting process was controlled by the Democratic party, which managed to implement a substantial gerrymander, one that has been found to have caused the largest effect on partisan bias ever noted in the academic literature (Gelman and King 1990b). The Democrats controlled the redistricting process in 1981 and made few changes from the 1971 plan; partisan bias did not change significantly. As a result, the electoral system in the Ohio State House favored the Democratic party consistently over the last 20 years, giving them 3%–4% more seats (and the Republicans 3%–4% fewer seats) than they would have been expected to receive in an electoral system without bias.

In the 1992 apportionment process, the Republicans controlled three of the five seats on the Apportionment Board. They delegated the mapping process almost entirely to James Tilling, secretary to the board and staff director of the Senate Republicans. According to his account, Tilling knew the current districts heavily favored the Democrats, and he set out to rectify this situation. However, he reports having decided to draw fair districts, rather than turning a Democratic gerrymander into a Republican gerrymander. His reasoning was that this would produce a fair result for the voters of Ohio, but it would also benefit his party, in comparison to the existing districts. Of course, it would not benefit the Republicans as much as it could have, so the Republican party produced a competing



Figure 4. Partisan Bias in the Ohio House



plan and tried to get it adopted. The Democrats tried the same, but the Apportionment Board stuck with Tilling's plan.<sup>26</sup>

Figure 4 also displays the predicted 1992 results for the three proposed redistricting plans using corresponding explanatory variables based on reaggregated electoral results from 1984 to 1990. We use the variability in our explanatory variables from 1984 to 1990 in this way to portray our uncertainty in their actual values, and the actual electoral conditions, that could obtain in 1992. Thus, our predictions for partisan bias in 1992 are plotted on the right side of Figure 4 at each of these four election years.

As can be clearly seen, the partisans' plans favor their own parties, and the Apportionment Board plan lay in between. A summary of the results, with standard errors of these summaries, also appears on the figure.<sup>27</sup> On average, both the Democratic and Republican plans would bias the electoral system by about 2.2% in their favor. The Tilling plan, adopted by the Apportionment Board, is forecast to have approximately zero partisan bias. The standard errors indicate that our estimates are fairly precise; both partisan plans are clearly distinguishable from zero bias, whereas the Tilling plan cannot be distinguished from zero. Interestingly, even the Democratic plan was slightly less biased toward the Democrats than the actual results from the late 1980s. Perhaps the Democrats wanted to draw a plan that would be reasonable enough to be accepted by the courts; both parties also appeared more concerned about other issues, such as protecting the particular incumbents currently in office.

We finished the analyses in this paper well before the 1992 elections and widely distributed the results of an earlier draft; they were also presented in court and are on the public record. After the election, we estimated partisan bias from the actual election results. This result, a partisan

<sup>26</sup>In response to a suit brought by the Apportionment Board, the Ohio state court declared the districts constitutional according to Ohio law. The Democratic party then brought suit before a three-judge federal panel charging the board with creating a political gerrymander under the guise of the Voting Rights Act. However, the plaintiffs offered no evidence about partisan bias and, after the information in Figure 4 was presented at the trial, effectively withdrew their partisan gerrymandering claim. The Court invalidated the Apportionment Board plan, essentially on grounds of reverse discrimination. This decision was stayed by the U.S. Supreme Court, meaning that the 1992 elections for the Ohio House were held under the Ohio Apportionment Board's redistricting plan. The Supreme Court later reversed the lower court, leaving the districts in place for the rest of the decade (*Voinovich v. Quilter* 1993).

<sup>27</sup>The standard error of each *point* is of course larger than the standard error of the *average* reported in the figure.

bias of 0.007 (with a standard error of 0.010) differs from the predicted bias by a trivial amount, not statistically distinguishable from our prediction of zero. This result appears as a solid triangle at the right of Figure 4. This result therefore provides further evidence in support of the model and its uncertainty estimates. Moreover, an analysis of the individual district forecasts produced correct seat predictions in every district.

## 7. Discussion

Instead of summarizing the contributions of this paper, we focus in this concluding section on how it differs from the previous method proposed in the literature, published in King and Gelman (1991). The most important improvement in our model is the introduction of explanatory variables, which add considerable additional information to the analysis and give us substantial leverage and statistical efficiency for most problems. Including explanatory variables also enables us to integrate the study of counterfactual evaluations directly and to make predictions based on stable and verifiable relationships in existing legislative elections data. Any information lost by introducing various simplifications is more than compensated for by the information in these explanatory variables. If explanatory variables are not available, then the King and Gelman (1991) method should be used instead of the one we propose here (e.g., Jackman, forthcoming).

Because our previous model had no explanatory variables, it had to model the well-known bimodal pattern of district vote outcomes, and it did so by fitting it to a mixture of normal distributions. By requiring the inclusion of either partisan control or incumbency status as explanatory variables, we avoid this complication: conditional on these explanatory variables, voting data are unimodal and reasonably modeled by a single normal distribution and the random components regression model we introduce.

We model district vote proportions directly, without the logit transformation used in King and Gelman (1991). The loss is very slight because contested district vote proportions above 0.8 or below 0.2 are rare. This greatly simplifies the analysis, eliminating the third-order Taylor series approximations previously used to compute expected values.

Finally, we singly impute votes in uncontested districts (Appendix A), avoiding the complication of estimating a distribution for what could have happened had the elections been contested, as in King and Gelman (1991).

Since our approach can be evaluated in new legislative elections data, as we did in our examples here, it will always outperform methods of forecasting and evaluation based on plausible but unchecked assump-

tions. One example is the use of "baseline" votes, the votes for a low-visibility office elected statewide and broken down into legislative districts, for forecasting or, in combination with uniform partisan swing, for evaluations (as suggested by Backstrom, Robins, and Eller 1978).<sup>28</sup> Since the baseline vote can be used as an additional predictor (another column in  $X$  in our model) and since our method enables the relationship between this variable and the legislative elections of interest to be estimated rather than assumed, our approach is guaranteed to do better. In addition, our approach has the advantage of using the same model to make district vote predictions and statewide evaluations, guaranteeing internally consistent results.

While one can use this model (and the associated computer program) for many different types of analyses of legislative elections, there are a number of topics for future research that remain worthy of study. The model does not deal with multimember districts, primaries, nonpartisan elections, or multiparty elections. It is also not yet equipped to handle other types of electoral systems, such as the various types of proportional representation, more common in other countries.

*Manuscript submitted 5 October 1992*

*Final manuscript received 10 August 1993*

#### APPENDIX A

##### Vote in Uncontested Districts

Depending on the goal of the analysis, one may consider adjusting the vote,  $v_i$ , in uncontested districts. This is especially important in evaluating a redistricting plan, since one wants to evaluate the effect on actual or potential voters; decisions taken away from the voters, such as when candidate decisions make voting for one party impossible, require special treatment. Adjustment is particularly important when studying electoral systems with a small number of districts, such as a congressional delegation in one state, or with a large number of uncontested seats.

Adjustment is used to make the votes reflect what the election would have been like had the district election been contested. For prediction of election outcomes, any reasonable adjustment preserves the same result: a win for the single contesting candidate. In contrast, when we summarize an electoral system by the seats-votes curve, bias, and responsiveness, or the expected proportion of marginal seats, the assignment of votes to uncontested districts has an effect because it alters  $\bar{v}$  and, thus, the average district vote, without changing the number of seats won by either party. Imputing values other than zero or one tends to reduce  $\bar{v}$  for the party that wins more uncontested seats, thus increasing the apparent

<sup>28</sup>Statewide measures of partisan bias and electoral responsiveness, and some other issues, are discussed by King and Browning (1987), Cain (1985), Grofman (1983), King (1989a), Niemi (1985), and Niemi and Fett (1986).

partisan bias in their favor. That is, imputation reduces the number of apparently “wasted” votes in uncontested districts.

For many applications, we believe some imputation for uncontested districts is essential. One can examine the votes in districts before and after being uncontested, or look at district totals in statewide races and observe partisan strength that is generally closer to 25% and 75% than zero and one. In defining the properties of the electoral system or evaluating the effects of redistricting, it seems reasonable to be interested in how many seats a party receives as a function of its aggregate statewide support. The fact that a seat is uncontested is an indication of strong support in a district, but certainly not 100% support.

There is, of course, room for debate on what values to impute. Three relatively simple options include keeping the vote proportions at zero and one; imputing common values such as 0.25 and 0.75 for Republican and Democratic uncontested districts, respectively; or imputing values from elections in the same districts but different years.<sup>29</sup>

After considerable experimentation, we suggest the following method. First, we temporarily impute values 0.25 and 0.75 for Republican and Democratic uncontested seats. Second, we regress the vote  $v$  on the same explanatory variables  $X$  as in our model (equation 3). If the legislative district vote for one or more years is included among the explanatory variables, we temporarily impute 0.25 and 0.75 for these too, even though the focus of this regression is creating imputations for the dependent variable. Finally, we create an imputation for each uncontested district  $i$  by calculating  $x_i\beta + \epsilon_i$ , where  $x_i$  is row  $i$  of the explanatory variable matrix  $X$ , and  $\epsilon_i$  is a single random draw from a normal distribution with mean zero and variance estimated as part of this regression. We then use these imputations for uncontested districts, along with the actual vote in the contested districts, in all subsequent analyses.

This procedure produces more accurate and efficient imputations and is logically consistent with the rest of our model, while incorporating the uncertainty in making these imputations in any final estimate or summary. The uncertainty is introduced in our method by the addition of a random  $\epsilon_i$  for each district instead of using only the fitted values  $(x_i\beta)$ , a version of Rubin’s (1987) “multiple imputation” method for nonresponse in surveys.<sup>30</sup>

## APPENDIX B

### Deriving the Predictive Distribution of Future Elections

The predictive distribution of  $v^{(hyp)}$  discussed in section 4.1 can be derived by classical methods in econometrics. However, we present a different, Bayesian derivation of this result here in order to introduce this method in a relatively familiar case. As a result, explaining the distributions we derive for actual and counterfactual evaluation in Appendix C, which require Bayesian analyses, will be considerably easier.<sup>31</sup>

<sup>29</sup>King and Gelman (1991) arrive at the approximate value 0.75 in congressional elections to signify the strength of a party running uncontested, by taking the average of all election results in districts, in the elections preceding uncontestedness (but with the same incumbency status). That article also demonstrates how to use an entire distribution for uncontested elections to incorporate the uncertainty in this process; however, we have found that single imputation, as we recommend in this article, is simpler and does not substantially alter substantive inferences.

<sup>30</sup>A single imputation is sufficient in this context, since we are usually interested in summaries that average over all the districts. Independent single imputation across uncontested districts is roughly equivalent to a multiple imputation for the entire system.

<sup>31</sup>Bayesian analysis is based on the following principles: (1) all unknown quantities,

To simplify our derivations, we introduce the following statistical notation for the multivariate normal distribution. If a random vector,  $z$ , is normally distributed with mean vector  $\mu$ , and covariance matrix  $\Sigma$ , we write its probability density function as

$$P(z) = N(z | \mu, \Sigma).$$

For example, the posterior distribution for  $\beta$  from the preliminary estimation in section 3 may be written as

$$P(\beta) \sim N(\beta | \hat{\beta}, \hat{\Sigma}_\beta). \tag{10}$$

The derivation of  $P(v^{(hyp)})$  requires three essential steps, which will be repeated for the derivations in Appendix C. In step 1, we identify the distribution of  $v^{(hyp)}$  given  $\gamma$  and  $\beta$ , which is given by equations (4) and (5):

$$P(v^{(hyp)} | \gamma, \beta) = N(v^{(hyp)} | X^{(hyp)}\beta + \gamma + \delta^{(hyp)}, (1 - \lambda)\sigma^2 I).$$

We cannot use this distribution directly because it conditions on two unknowns,  $\gamma$  and  $\beta$ . (Recall that  $\lambda$  and  $\sigma^2$  are assumed known after estimation.) The remaining two steps in this derivation average over the distributions of these two parameters.

In step 2, we average over the distribution of  $\gamma$ , which is simply  $P(\gamma | \beta) = N(\gamma | 0, \lambda\sigma^2 I)$ , from equation (4).

$$\begin{aligned} P(v^{(hyp)} | \beta) &= \int_{-\infty}^{\infty} P(v^{(hyp)} | \gamma, \beta) P(\gamma | \beta) d\gamma \\ &= N(v^{(hyp)} | X^{(hyp)}\beta + \delta^{(hyp)}, \sigma^2 I). \end{aligned} \tag{11}$$

Finally, in step 3, we average over the uncertainty in  $\beta$  (equation 10), using the same rule from probability theory:

$$\begin{aligned} P(v^{(hyp)}) &= \int_{-\infty}^{\infty} P(v^{(hyp)} | \beta) P(\beta) d\beta \\ &= \int_{-\infty}^{\infty} N(v^{(hyp)} | X^{(hyp)}\beta + \delta^{(hyp)}, \sigma^2 I) N(\beta | \hat{\beta}, \hat{\Sigma}_\beta) d\beta \\ &= N(v^{(hyp)} | X^{(hyp)}\hat{\beta} + \delta, X^{(hyp)}\hat{\Sigma}_\beta X^{(hyp)'} + \sigma^2 I). \end{aligned} \tag{13}$$

The last line here is the final unconditional distribution of  $v^{(hyp)}$ , which we use to calculate the hypothetical election results, as in Table 1, and then the distributions of various summaries of the electoral system.

### APPENDIX C

#### Deriving the Distribution under Actual or Counterfactual Electoral Conditions

We now derive the distribution of  $P(v^{(hyp)} | v)$  in a manner directly analogous to that in Appendix B. In step 1, we identify the distribution of  $v^{(hyp)}$  given  $\gamma$  and  $\beta$  and, unlike prediction, observed votes  $v$ . Actually, once  $\gamma$  and  $\beta$  are known,  $v$  gives no additional

---

including  $v^{(hyp)}$ ,  $\beta$ , and so on, are treated as random variables, and they have a joint probability distribution; (2) if we have two variables (or sets of variables),  $A$  and  $B$ , and we have observed  $B$ , then we are interested in the conditional distribution of  $A$ , given  $B$ :  $P(A|B) = P(A, B)/P(B)$ ; (3) if two variables (or sets of variables),  $A$  and  $B$ , are both unknown, then the distribution of  $A$  alone is  $P(A) = \int P(A, B)dB = \int P(A|B)P(B)dB$ .

information about  $v^{(\text{hyp})}$  because  $\epsilon$  and  $\epsilon^{(\text{hyp})}$  are independent, and so the distribution of  $v^{(\text{hyp})}$  is still given by equation (5):

$$P(v^{(\text{hyp})}|\gamma, \beta, v) = N(v^{(\text{hyp})}|\lambda X^{(\text{hyp})}\beta + \gamma + \delta^{(\text{hyp})}, (1 - \lambda)\sigma^2 I). \quad (14)$$

Step 2 requires averaging over  $\gamma$ , but in this case the distribution can be improved by conditioning on  $v$ —the key difference between this derivation and that for prediction in Appendix B. We condition on  $v$  with Bayes's theorem:

$$\begin{aligned} P(\gamma|\beta, v) &= \frac{P(v|\gamma, \beta)P(\gamma|\beta)}{P(v|\beta)} \\ &\propto N(v|\lambda X^{(\text{hyp})}\beta + \gamma, (1 - \lambda)\sigma^2 I)N(\gamma|0, \lambda\sigma^2 I) \\ &= N(\gamma|\lambda(v - X^{(\text{hyp})}\beta), \lambda(1 - \lambda)\sigma^2 I). \end{aligned} \quad (15)$$

To complete step 2, we use this result to integrate out  $\gamma$ :

$$\begin{aligned} P(v^{(\text{hyp})}|\beta, v) &= \int_{-\infty}^{\infty} P(v^{(\text{hyp})}|\gamma, \beta, v)P(\gamma|\beta, v)d\gamma \\ &= N(v^{(\text{hyp})}|\lambda v + (X^{(\text{hyp})} - \lambda X)\beta + \delta^{(\text{hyp})}, (1 - \lambda^2)\sigma^2 I). \end{aligned} \quad (16)$$

Finally, we integrate out  $\beta$  for step 3:

$$\begin{aligned} P(v^{(\text{hyp})}|v) &= \int_{-\infty}^{\infty} P(v^{(\text{hyp})}|\beta, v)P(\beta|v)d\beta \\ &= N(v^{(\text{hyp})}|\lambda v + (X^{(\text{hyp})} - \lambda X)\hat{\beta} + \delta^{(\text{hyp})}, \\ &\quad (1 - \lambda^2)\sigma^2 I + (X^{(\text{hyp})} - \lambda X)\Sigma_{\beta}(X^{(\text{hyp})} - \lambda X)'). \end{aligned}$$

## REFERENCES

- Alford, John, and David W. Brady. 1988. "Partisan and Incumbent Advantage in U.S. House Elections, 1846–1986." Working paper 11. Rice University.
- Backstrom, Charles, Leonard Robins, and Scott Eller. 1978. "Issues in Gerrymandering: An Exploratory Measure of Partisan Gerrymandering Applied to Minnesota." *Minnesota Law Review* 62:1121–59.
- Beck, Nathaniel. 1992. "Forecasting the 1992 Presidential Election: The Message Is in the Confidence Interval." *Public Perspective* 3:32–34.
- Benjamin, Gerald, and Michael J. Malbin. 1992. *Limiting Legislative Terms*. Washington, DC: Congressional Quarterly Press.
- Born, Richard. 1985. "Partisan Intentions and Election Day Realities in the Congressional Redistricting Process." *American Political Science Review* 79:305–19.
- Box, George E. P., and George C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesley.
- Brady, David W. 1988. *Critical Elections and Congressional Policy Making*. Stanford: Stanford University Press.
- Butler, David E. 1951. Appendix. In *The British General Election of 1950*, ed. H. G. Nicholas. London: Macmillan.
- Cain, Bruce. 1985. "Assessing the Partisan Effects of Redistricting." *American Political Science Review* 79:320–33.
- Campbell, James E. 1992. "Divided Government, Partisan Bias, and Turnout in Congress-

- sional Elections: Do Democrats Sit in the 'Cheap Seats'?" Photocopy, Louisiana State University.
- Edgeworth, Francis Y. 1898. "Miscellaneous Applications of the Calculus of Probabilities." *Journal of the Royal Statistical Society* 61:534-44.
- Erikson, Robert S. 1971. "The Advantage of Incumbency in Congressional Elections." *Polity* 3:395-405.
- Ferejohn, John A. 1977. "On the Decline of Competition in Congressional Elections." *American Political Science Review* 71:166-76.
- Fiorina, Morris. 1977. *Congress: Keystone of the Washington Establishment*. New Haven: Yale University Press.
- . 1992. *Divided Government*. New York: Macmillan.
- Gelman, Andrew, and Gary King. 1990a. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34:1142-64.
- . 1990b. "Estimating the Electoral Consequences of Legislative Redistricting." *Journal of the American Statistical Association* 85:274-82.
- Goldberger, Arthur S. 1991. *A Course in Econometrics*. Cambridge: Harvard University Press.
- Greene, Jay P. 1993. "Forewarned before Forecast: Presidential Election Forecasting Models and the 1992 Election." *PS* 26:17-21.
- Grofman, Bernard. 1983. "Measures of Bias and Proportionality in Seats-Votes Relationships." *Political Methodology* 9:295-327.
- Grofman, Bernard, Lisa Handley, and Richard Niemi. 1992. *Minority Representation and the Quest for Voting Equality*. New York: Cambridge University Press.
- Hanushek, Eric, and John Jackson. 1977. *Statistical Methods for Social Scientists*. New York: Academic Press.
- Jackman, Simon. N.d. "Measuring Electoral Bias: Australia, 1949-1993." *British Journal of Political Science*. Forthcoming.
- Jacobson, Gary C. 1980. *Money in Congressional Elections*. New Haven: Yale University Press.
- . 1987. "The Marginals Never Vanished: Incumbency and Competition in Elections to the U.S. House of Representatives." *American Journal of Political Science* 31:126-41.
- . 1990. *The Electoral Origins of Divided Government: Competition in U.S. House Elections, 1946-1988*. Boulder, CO: Westview Press.
- King, Gary. 1989a. "Representation through Legislative Redistricting: A Stochastic Model." *American Journal of Political Science* 33:787-824.
- . 1989b. *Unifying Political Methodology*. New York: Cambridge University Press.
- . 1991a. "Stochastic Variation: A Comment on Lewis-Beck and Skalaban's 'The R-Square.'" *Political Analysis* 2:185-200.
- . 1991b. "'Truth' Is Stranger than Prediction, More Questionable than Causal Inference." *American Journal of Political Science* 35:1047-53.
- King, Gary, and Robert X. Browning. 1987. "Democratic Representation and Partisan Bias in Congressional Elections." *American Political Science Review* 81:1251-76.
- King, Gary, John Bruce, and Andrew Gelman. 1993. "Standards of Racial Fairness in Legislative Redistricting." In *Classifying by Race*, ed. Paul E. Peterson.
- King, Gary, and Andrew Gelman. 1991. "Systemic Consequences of Incumbency Advantage in U.S. House Elections." *American Journal of Political Science* 35:110-38.
- Lewis-Beck, Michael S., and Tom W. Rice. 1992. *Forecasting Elections*. Washington, DC: Congressional Quarterly Press.



- Mayhew, David R. 1974. "Congressional Elections: The Case of the Vanishing Marginals." *Polity* 6:295-317.
- . 1991. *Divided We Govern: Party Control, Lawmaking, and Investigations, 1946-1990*. New Haven: Yale University Press.
- Niemi, Richard G., and Patrick Fett. 1986. "The Swing Ratio: An Explanation and an Assessment." *Legislative Studies Quarterly* 11:75-90.
- Rothstein, Paul, and John B. Gilmour. 1992. "Term Limitation in a Dynamic Model of Partisan Balance." Photocopy, Washington University, St. Louis.
- Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley.
- Schrodt, Philip A. 1981. "A Statistical Study of the Cube Law in Five Electoral Systems." *Political Methodology* 7:31-53.
- Tufte, Edward. 1973. "The Relationship between Seats and Votes in Two-Party Systems." *American Political Science Review* 67:540-54.