The Optimal Allocation of Campaign Funds in House Elections

Devin Incerti

September 13, 2015

Abstract

Do the Democratic and Republican parties optimally allocate resources in House elections? This paper answers this question by estimating Stromberg's (2008) probabilistic voting model and comparing actual spending patterns to the amount that should have been spent under the model. The model depends crucially on forecasts of the vote in each district that account for both district and national uncertainty. I employ two types of forecasting models—a Bayesian hierarchical model and a state-space model that incorporates all available polling data and uses the hierarchical model as a prior. The correlation between actual spending and the amount that should have been spent is over 0.5 in each non-redistricting year from 2000 to 2010 and has generally increased over time. Surprisingly, these correlations are consistent across different types of campaign donors including the Democratic Congressional Campaign Committee and the National Republican Congressional Committee; various political action committees; and individuals. There is also evidence that spending patterns are based on maximizing total seats rather than the probability of winning a majority of seats.

1 Introduction

An important component of campaign strategies in legislative elections is the optimal allocation of resources across legislative districts. In this paper, I tackle this problem with a model that generates precise estimates of the marginal value of additional campaign resources in each district. I apply this model to U.S House elections and compare the amount of campaign funds that should have been spent on each candidate with actual spending.

The allocation of campaign resources is viewed as a competition for seats between Democrats and Republicans. This competition is analyzed using the probabilistic voting model of Strömberg (2008), which can be explicitly solved and directly estimated. Like Snyder (1989), I consider two different assumptions about goals: first, parties maximize the expected number of seats won, and second, parties maximize the probability of winning a majority of seats. In equilibrium, the parties spend the most on the closest elections under the first assumption, whereas under the second assumption, an additional factor—the probability that a seat is pivotal—impacts the marginal value of a dollar in a district.

The marginal value of additional spending under Stromberg's model depends on predicted vote shares and accounts for uncertainty at the district and national levels. It is therefore easily integrated with forecasting methods that account for national swings. In this paper, I use two types of forecasting models. The first model is a Bayesian hierarchical model that uses information about districts and candidates using yearly data to provide a forecast as of September 1st of each election year. The second model is a state-space model that uses the hierarchical model as a prior and incorporates all available district and national polls. Unlike the first model, the latter is capable of providing real time forecasts at any date during a campaign. The hierarchical model is used to forecast each non-redistricting year election from 2000 to 2010 and the state-space model is used to analyze the 2010 election at various stages of the campaign.

The empirical analysis uses these forecasts to estimate Stromberg's model, which is then

used to identify the districts that should have received the most contributions. The correlation between the amount that should have been spent according to the model and actual spending is over 0.5 in each non-redistricting election year from 2000 to 2010 and has generally increased over time. Surprisingly, these correlations are consistent across different types of campaign donors: the spending patterns of party committees, such as the Democratic Congressional Campaign Committee (DCCC) and the National Republican Congressional Committee (NRCC), whose strategies should be the most likely to be coordinated and concerned with electoral success, are *no* more consistent with the model's predictions than the spending patterns of political action committees (PACs) or individuals. There is also no evidence that spending for Republican candidates differs from spending for Democratic candidates. Finally, as one might expect, these correlations become even stronger when incorporating polling data using the state-space model: the correlation reaches a peak of over 0.8 when comparing spending in the final month of the 2010 campaign with equilibrium spending based on a forecast made using all information up until that date.

When elections are reasonably close, the parties' goals have a relatively minor effect on optimal spending strategies; when elections are lopsided, the opposite is true. Between 2000 and 2010, only one election—the 2008 election—was lopsided enough to yield spending strategies that were significantly sensitive to party goals. During that election, actual spending is highly correlated ($r \approx 0.664$) with a seat maximization based spending strategy but not ($r \approx 0.353$) one based on maximizing the probability of winning a majority of seats.[1] This provides evidence that parties are more concerned with whether an election is close than whether it is pivotal.

Overall, campaign spending seems to be consistent with a world in which the two major parties in the United States try to maximize expected seat share. However, this does not mean that other factors do not help explain observed differences in spending across districts as well. For example, incumbents and candidates running in open seats both raise more

[1] r refers to Pearson's correlation coefficient.

funds than challengers running against incumbents. Similarly, party leaders raise more funds than their counterparts and members of the House Committee on Financial Services raise considerably more funds from financial firms than other incumbents not on the committee. That being said, the explanatory power of these other factors pales in comparison to the explanatory power of Stromberg's model and it is only in cases in which actors have the largest incentives to consider other factors—like financial service firms buying access to candidates—that they are on close to equal footing.[2]

2 Related Literature

This paper builds on a number of other studies focusing on the strategic allocation of resources in campaigns. These studies date back to the work of Brams and Davis (1973; 1974), who look at how presidential campaigns should allocate campaign resources to maximize their expected electoral vote. Using a model in which each candidate has an equal probability of winning the popular vote in each state, they argue that presidential campaigns should allocate resources roughly in proportion to the 3/2 power of the number of electoral votes in that state, which means that larger states should receive a disproportionate share of campaign resources. This result was then challenged by Colantoni, Levesque and Ordeshook (1975), who conclude that a modified proportional rule that accounts for the closeness of an election fits the data better.

One drawback of these earlier studies is that they assume that parties maximize the expected electoral vote rather than the probability of winning the election. Brams and Davis posited that the implications of this distinction would be relatively minor. Aranson, Hinich and Ordeshook (1974) support this notion by showing that the two goals are equivalent if the game is symmetric—which implies that the expected number of seats won by each party must be that same. However, as discussed by Snyder (1989), this symmetry assumption is strong and not realistic in real world settings. He shows that when this is not the case, the

[2]Members of the House Committee on Financial Services receive a higher proportion of funding from the financial services industry than any other committee receives from a single industry.

two goals yield different equilibria.[3]

More recently, Strömberg (2008) has developed a more general model that can reasonably be applied to actual campaigns. He uses the more reasonable assumption that presidential candidates maximize the probability of winning a majority of electoral college votes (i.e. winning the election) and then calculates the number of times that presidential candidates should have visited each state. He finds a very strong correlation (≈ 0.9) between these visits and observed visits.

The main difference between my paper and these other studies is that I focus on campaign spending in U.S. House elections rather than presidential campaigns. One consequence of this is that the parties' goals are more ambiguous because it is not clear whether they should maximize expected seat share or the probability of winning a majority of seats. In this sense, my substantive focus is closest to Snyder's 1989, although unlike his paper, I test my results empirically. My finding that parties maximize the expected number of seats and not the probability of winning a majority of seats is consistent with Jacobson and Kernell's (1985*b*) assertion that every congressional seat is valuable so parties should aim to maximize seats. This finding is also related to an interesting hypothesis set forth by Snyder (1989) that the leading party might want to be as certain as possible to try to win a majority while the trailing party might simply try to win as many seats as possible to improve future chances at controlling the legislature. Although interesting, I find no evidence that the parties play different strategies in U.S. House elections.

Methodologically, my paper is closest to Stromberg's since I use his model. The only major methodological differences between my paper and his are related to the forecasting methods. The two most significant of these differences are that first, I forecast House elections which are less predictable than presidential elections, and second, I use a technique that can update forecasts in real time as new polls become available. A third but relatively minor difference is that I utilize Bayesian techniques which account for uncertainty in the estimation

[3]These equilibria are consistent with the one's found in this paper.

of the district and national shocks and allow for easier estimation of non-linear functions of the parameters.

The foundations of Stromberg's model date back to the probabilistic voting models of Lindbeck and Weibull (1987) and Dixit and Londregan (1996) used to analyze electoral competition. In these models, two competing political candidates must determine which interest groups should receive favors in order to maximize their probability of winning an election. When the political candidates do not differ in the efficiency in which they make transfers to various group, these models yield a "swing voter" equilibrium—similar to "close election equilibrium" in this paper—where interest groups that are the most politically central receive the most favors.[4]

The results here also shed some light on current debates regarding the motivations of campaign donors. These motivations are important because, as Stratmann (2005) notes, the predicted determinants of campaign contributions depend on the assumptions regarding contributor goals. Four objectives commonly cited in the literature are that contributions are a consumption good, an investment in policy, a means to buy access to a politician, or a way to influence an election (e.g. Ansolabehere, Snyder and de Figueiredo 2003; Stratmann 2005).

This paper's finding that parties should (and do) contribute the most to districts that have the largest probability of being close is consistent with one of the more robust findings in the literature—that contributors spend more on close elections (Kau, Keenan and Rubin 1982; Jacobson 1985*a*; Poole and Romer 1985; Stratmann 1991).[5] In addition, the result that the financial industry donates more to members of the House Committee on Financial Services backs up research showing that candidates serving on congressional committees raise more

[4]See Johansson (2003) for an empirical test of the Lindbeck and Weibull and Dixit and Londregan models.

[5]These studies have two main problems that this paper avoids. First, the closeness of an election is typically measured with an ex-post measure of the electoral margin or the lagged vote from the previous election. This does not mimic the decision of contributors who must make choices prior to election day and have considerably more information available to them then the vote in the previous election. Second, since they are not driven by theory, they do not provide any guidance the functional form of the relationship between the closeness of an election and spending, which should depend on the uncertainty (and probability distribution) of the predicted vote.

money (Grier and Munger 1991; Romer and Snyder Jr 1994; Kroszner and Stratmann 1998). However, the explanatory strength of Stromberg's model relative to committee membership and other influence variables suggests that while both election-motivated and influence-motivated giving are important, election-motivated giving is likely to be more common.

The results in this paper depend on the assumption that additional contributions increase the probability that a candidate will win an election. While it seems difficult to believe that this would not be true, there is a large literature in political science and economics examining this question. Research findings are inconsistent, although more recent studies that have addressed biases have found that campaign spending does impact the vote.

The origins of this literature stem from Jacobson's (1978; 1980; 1985*a*) findings that campaign spending by challengers in congressional elections is very important but that incumbent spending has almost no effect on election outcomes. Subsequent work has refined these early studies by accounting for the endogeneity of candidate spending, which Jacobson's model does not address.[6] This is typically done by using instrumental variables (Gerber 1998; Green and Krasno 1988), including better control variables (Green and Krasno 1988) or addressing simultaneity biases (Erikson and Palfrey 2000). For instance, Green and Krasno (1988) instrument for incumbent spending in the previous election and control for candidate quality by creating an eight point scale based on various traits.[7] In contrast to Jacobson's earlier findings, they find that incumbent spending has a positive and statistically significant effect on the vote in House elections. Erikson and Palfrey (2000) reach similar conclusions by using a game theoretic model to show that simultaneity biases can be eliminated by only analyzing close races. After limiting analyses to these cases, they find the effect of both challenger and incumbent spending to be substantial.[8]

[6] Another explanation proposed by John Ferejohn and Morris Fiorina and mentioned in Jacobson (1985*a*) for Jacobson's finding that incumbent spending makes no difference is that there are almost no cases where incumbents do not respond to lavish spending by challengers with generous spending of their own.

[7] A related article is Gerber (1998), which uses challenger wealth as an instrument for challenger spending in Senate elections.

[8] One study that attempts to reduce biases but finds no effect of either challenger or incumbent expenditures on the vote is Levitt (1994). He controls for unobserved candidate characteristics by limiting his study to elections in which the same two candidates face each other more than once and taking first differences.

3 Model

This section describes a version of Stromberg's model suitable for analyzing U.S. House elections. The model is essentially the same as the one used in Strömberg (2008) to analyze presidential elections except for two minor changes that account for differences in the electoral settings. First, all districts are worth one seat while states in presidential elections are weighted according to their electoral votes. Second, parties maximize both the probability of winning a majority of seats (which is the equivalent of winning a majority of electoral college votes in presidential elections) and the expected number of seats. The second maximization problem is not considered in Strömberg (2008) because presidential candidates are concerned with winning the election, not maximizing their electoral votes.

3.1 Set Up

The model considers electoral competition between two parties, labeled Republican R and Democrat D.[9] During the campaign, each party must decide how to optimally allocate funds across the 435 Congressional districts. More formally, party $J = D, R$ must choose expenditures in district i, e_i^J, subject to the resource constraint,

$$\sum_{i=1}^{435} e_i^J \leq E^J, \tag{1}$$

where E^J is the amount of money party J has to spend on candidates.

The share of votes received by party D in district i is assumed to depend on four primary factors: spending by the national parties, predetermined characteristics of the district, the national political climate, and unknown shocks. The effect of the choice variable, e_i^J, is assumed to be an increasing concave function, $u(e_i^J)$, so that the effect of spending decreases with the amount of spending. The predetermined district characteristics and the national

The finding may lack external validity though since challenger quality is likely correlated with the probability of running in multiple elections. In addition, spending between candidates is unlikely to vary much from one election to the next which might lead to imprecise estimates.

[9] Third party candidates are ignored.

climate are known before the spending decision is made and can be used to make a prediction, V_i, of party D's vote share. Finally, there are two sources of uncertainty, a national error, δ and a district specific error, ϵ_i. The national errors represent uncertain national swings that affect all districts equally and the district errors are unpredictable swings unique to each district. Both error terms are independently drawn from normal distributions,

$$h(\delta) = N(\delta|0, \sigma_\delta^2) \tag{2}$$

and

$$g_i(\epsilon_i) = N(\epsilon_i|0, \sigma^2). \tag{3}$$

Letting $u(e_i^D) - u(e_i^R) = \Delta u_i$, party R will consequently win a district if,

$$\Delta u_i + V_i + \delta + \epsilon_i \leq 1/2. \tag{4}$$

The probability of a victory by party R conditional on expenditures, e_i^D and e_i^R, and the national swing, δ, is therefore,

$$G_i(1/2 - \Delta u_i - V_i - \delta),$$

where $G_i(\cdot)$ is the cumulative distribution function (CDF) of ϵ_i. It follows that if s_i is an indicator variable equal to 1 if party R wins a district and 0 if party D wins, then $s_i = 1$ with probability $G_i(\cdot)$ and $s_i = 0$ with probability $1 - G_i(\cdot)$. The total number of Republican seats is $S = \sum_{i=1}^{435} s_i$. Furthermore, since the $G_i(\cdot)$ are independently (but not identically) distributed conditional on δ, S follows a Poisson binomial distribution with mean

$$\mu_S = \mu_S(\Delta u, \delta) = \sum_{i=1}^{435} G_i(\cdot), \tag{5}$$

and variance,

$$\sigma_S = \sigma_S(\Delta u, \delta) = \sum_{i=1}^{435} G_i(\cdot)(1 - G_i(\cdot)), \tag{6}$$

where $\Delta\mu = (\Delta\mu_1, \Delta\mu_2, \ldots, \Delta\mu_{435})$ represents the utility differences resulting from any allocation of campaign resources across districts by the two parties.

3.2 Party Goals

Optimal strategies depend on the objective of the national parties. Unlike in presidential campaigns where the goal is clearly to win the election, the goals of the parties in House campaigns are less straightforward. As a result, I consider two plausible objective functions.

The first objective function assumes that parties simply maximize the expected number of House seats. For party R, this is just the expectation of μ_S over the national shocks,

$$\mathbb{E}\left[S(\Delta\mu)\right] = \int \sum_{i=1}^{435} G_i(1/2 - \Delta u_i - V_i - \delta)h(\delta)d\delta. \tag{7}$$

A second possibility is that parties maximize the probability of winning a majority of seats. For party R this is,

$$P^R(\Delta\mu) = \int \Pr\left(\sum_{\mathrm{i}=1}^{435} \mathrm{s_i} > 218\right) \mathrm{h}(\delta)\mathrm{d}\delta. \tag{8}$$

This function is more difficult to maximize than equation 7 because it is not additively separable across districts; that is, party R's optimal strategy in one district depends on its strategy in all other districts. In order to solve the problem analytically, Strömberg (2008) suggests calculating the approximate probability of winning instead. Since the national and district shocks are independent, this can be done by using the Lyapunov Central Limit Theorem, which does not require random variables to be identically distributed. Using this approximation, the number of seats won by party R, $S = \sum_{i=1}^{435} s_i$, is asymptotically normally

distributed with mean μ_S and variance σ_S. The approximate probability of party R winning the election is then,

$$P^R(\Delta\mu) = \int 1 - \Phi\left(\frac{218 - \mu_S(\Delta u, \delta)}{\sigma_S(\Delta u, \delta)}\right) h(\delta) d\delta, \tag{9}$$

where $\Phi(\cdot)$ is the standard normal cumulative density function.

3.3 Equilibrium

Let $e^j = (e^j_1, e^j_2, \ldots, e^j_{435})$ be the allocation of campaign spending across districts by party j and $f^R(e^D, e^R)$ be the objective function used by party R. A Nash equilibrium in pure strategies (e^{D*}, e^{R*}) is then characterized by,

$$f^R(e^D, e^{R*}) \geq f^R(e^{D*}, e^{R*}) \geq f^R(e^{D*}, e^R), \tag{10}$$

where e^D and e^R satisfy the budget constraint in equation 1. The game has a unique interior NE which satisfies,

$$\frac{\partial f^J}{\partial e^J_i} = Q_i u'(e^{J*}) = \lambda^J, \tag{11}$$

where $Q_i = \partial f^J / \partial \Delta u_i$ and λ^J is the Lagrange multiplier for party j.[10] Since $u'(e^J)$ is decreasing in e^J, the parties allocate more resources to districts with higher values of Q_i. The value of Q_i will of course depend on the choice of the objective function.

When parties maximize the expected number of seats, Q_i is easy to calculate and is just,

$$Q_i = Q_i^{seats} = \int g_i(1/2 - \Delta u_i - V_i - \delta) h(\delta) d\delta. \tag{12}$$

Not unexpectedly, districts with a predicted vote share close to 1/2 should receive the most expenditures. Spending should also be more concentrated when the error terms, σ and δ,

[10] For a proof of uniqueness see Strömberg (2008).

are smaller. If on the other hand, parties maximize the probability of winning a majority of seats, Q_i is more intricate. For party R, it is,

$$Q_i = Q_i^{maj} = -\int \left(\frac{\partial \Phi(\cdot)}{\partial \mu_S} \frac{\partial \mu_S}{\partial \Delta u_i} + \frac{\partial \Phi(\cdot)}{\partial \sigma_S} \frac{\partial \sigma_S}{\partial \Delta u_i} \right) h(\delta) d\delta \tag{13}$$

$$= \int \frac{1}{\sigma_S} \phi(x(\delta)) g_i(\cdot) h(\delta) d\delta + \int \frac{1}{\sigma_S} \phi(x(\delta)) x(k, \delta) \left(1 - 2G_i(\cdot)\right) g_i(\cdot) h(\delta) d\delta, \tag{14}$$

where $x(\delta) = (218 - \mu_S) / \sigma_S$. The first term is the effect of an increase in spending on the mean number of Republican seats while the second term is its effect on the variance. Parties have an incentive to influence the variance because they want to increase the probability of a desirable outcome. The trailing party will want to increase the variance to increase the probability of a major change in the election outcome while the leading party will want to do just the opposite. The trailing party can increase the variance by spending more on districts in which its candidate it losing and the leading party can decrease the variance by spending more on districts in which it is winning. Intuitively, the leading party does not need to worry about districts in which it is losing because it only needs to make sure that it holds on to the one's it is leading in.

As shown by Strömberg (2008), an alternative interpretation of equation 13 is that it is the probability that a district is 1) decisive in whether or not a party wins a majority of seats and 2) the district is a swing district. Following Stromberg, I call such a district a "decisive swing district". The probability of being a swing district is the probability that an electoral race is tied (or at least very close), while a district is decisive if winning (or losing) that district would make the difference between winning (or losing) a majority of seats. The idea that parties should spend more money on swing districts is consistent with the first order condition in equation 12. The idea that parties should spend more on decisive (also known as pivotal) districts differentiates the two maximization problems.

3.4 Functional Form

In order to solve for equilibrium spending, e_i^{J*} and calculate Q_i, it is necessary to make an assumption about the functional form of $u(e_i^J)$. One functional form particularly amenable to empirical analysis is the logarithmic form, $u(e_i^J) = \theta \log(e_i^J)$, which results in the first order condition for district k and party J,

$$e_k^{J*} = \frac{Q_k}{\sum Q_i} E^J. \tag{15}$$

Each party spends the same fraction of the budget on district k, but e_i^{R*} only equals e_i^{D*} if both parties have identical budgets so that $E^R = E^D$.[11] Equation 15 implies that Q_i is evaluated at $\Delta u_i = \theta \log(E^D/E^R)$, which reduces to $\Delta u_i = 0$ when the budgets are equal. If θ, E^D or E^R are unknown (and time-varying), then this term will be incorporated into the national shock, δ, during estimation.

4 Estimation

To calculate Q_i for each district in each election, it is necessary to estimate the variances of the district and national shocks, σ^2 and σ_δ^2, as well as the two-party vote in each district, V_i. In the section I describe a Bayesian methodology that can estimate these parameters using historical political and economic information and when available, polling data. The historical information provides a forecast of the election as of September 1st in each election year. Polling data from September 1st up until election day is then used to update the forecasts as the campaign progresses. This allows for an examination of the relationship between post-August spending and Q_i at the election year level and whether campaign strategies responds to new polls (which change the values of Q_i).

[11] For a formal proof that $e_i^{D*} = e_i^{R*}$ in equilibrium when $E^R = E^D$ see Strömberg (2008)

4.1 A Bayesian Hierarchical Model

Previous research has shown that national elections are highly predictable from one year to the next using standard regression techniques (e.g. Campbell 1992; Gelman and King 1993; Kastellec, Gelman and Chandler 2008). These "historical" models are often said to use "fundamentals"— such as the political ideology of citizens in a political area or the state of the economy—to forecast the vote for a given election. In Stromberg's model, these fundamentals are predetermined characteristics and national trends that can be used to make a prediction, V_{iy}, of the Democratic share of the two-party vote in district i and election year y. To move from theory to empirics, I assume that this prediction is a linear function of a matrix of explanatory variables, X_{iy}. Using equation 4, the Democratic share of the two party vote, say v_{iy} is then,

$$v_{iy} = X_{iy}\beta + \delta_y + \epsilon_i. \tag{16}$$

Since δ_y and ϵ_i are assumed to be normally distributed, this can be estimated using the Bayesian hierarchical model,

$$v_{iy} \sim N\left(X_{iy}\beta + \delta_y, \sigma^2\right) \tag{17}$$

$$\delta_y \sim N(0, \sigma_\delta^2), \tag{18}$$

where β is a vector of coefficients. δ_y is modeled as a random effect and is centered at 0 because X_{iy} includes an intercept. The Bayesian estimation is completed with uniform hyperprior distributions on β, σ and σ_δ.

4.2 Incorporating Polls with a Bayesian DLM

In recent elections, polling firms have begun polling certain races for the House of Representatives. Since these polls provide (hopefully) useful information about the likely vote in

a district, it would be wise to include them in the forecasting model. The most straightforward way to do this would be to simply include them as additional columns in the data matrix X_{iy}. However, this is problematic because (1) polling data at the district level are a relatively recent phenomenon, (2) firms do not poll all districts, (3) polls are measured with error and may differ from the true state of public opinion at any given time and (4) districts that are polled are typically polled multiple times during the election campaign.

One strategy that can help alleviate these issues is to treat polls as additional data points, rather than as independent variables in a historical regression. This technique is commonly used by researchers and media outlets forecasting presidential and Senate elections.[12] The primary difficulty is that the polls are highly correlated and should not be treated as independent data. As a result, forecasters attempt to average the polls so that the most informative ones receive the most weight. This weighted average of polls can then, in turn, be combined with information from a regression analysis, with weights that should be based on their respective variances.

As noted by the founder of http://fivethirtyeight.com/, Nate Silver, the variability of a poll's forecast can be thought of as a function of three major components: sampling error, temporal error and pollster-induced error. The first two terms are relatively straightforward: sampling error occurs because each poll is based on a sample from the electorate and temporal error is due to uncertainty about opinion shifts between the date a poll is taken and election day. The final term, pollster-induced error, can be thought of as the error left over after accounting for sampling and temporal error. A major part of this residual term can be attributed to house effects, or time-invariant biases specific to certain pollsters. However, this error can also occur due to other polling difficulties such as undecided voters, respondents who will not vote in the actual election or respondents that do not express their true voting intentions.

[12] See, for instance, the Huffington Post's forecasts at http://elections.huffingtonpost.com/, which are based on Simon Jackman's poll-tracking model; Drew Linzer's forecasts at http://votamatic.org/; neuroscientist Sam Wang's website http://election.princeton.edu/; and Nate Silver's http://fivethirtyeight.com/.

In this paper, I utilize a state-space framework that can sequentially adjust forecasts as new polls become available. This model-based poll averaging approach is very similar to the the state-space poll-tracking model employed by Simon Jackman (Jackman 2005, 2009) and the forecasting model based on reverse-random walks used by Drew Linzer (Linzer 2013)—which built on the idea used in (Strauss 2007).

One important feature of House elections that must be accounted for is that district polling is sporadic both over time and across districts. The incomplete nature of the data is important for calculating district and national errors because national errors based solely on district polling use less than half of all districts at any given date and the subset of districts available to calculate national errors changes over time (since different districts are polled at different times). One way around this is to to use the generic congressional vote as a measure of the national vote, since, unlike district polls, polling firms begin conducting polls of the generic congressional vote for the next election at a consistent rate almost as soon as election results are in.

To incorporate the national polls, I separate the vote in each district into the national vote and the district vote relative to the national vote as in Lock and Gelman (2010) and Strauss (2007). This separation allows me to use all available polling data to decompose national and local variation as required by the theoretical model.[13]

The model of the national vote follows the model employed in Jackman (2005) and Jackman (2009), which provides a framework for "pooling" the polls over the course of a campaign. To set notation, let $t = 1, \ldots, T$ represent days of the campaign where $t = 1$ corresponds to the first day of the campaign season and $t = T$ is election day. Furthermore, let k index a poll with sample size N_k so that the number of respondents, n_k, from poll k

[13]Another strategy is to model the correlation between the national polls and the district polls in a multivariate time-series model (see Jackman (2012) for a brief explanation of a model that does this). This approach is not taken here because it is not consistent with the theoretical model used in this paper.

who support the Democratic candidate follows a binomial distribution,

$$n_k \sim \text{Bin}(\text{N}_\text{k}, \pi_\text{k}), \tag{19}$$

where π_k is the proportion of voters from poll k who intend (or report to intend) to vote for a (generic) Democratic candidate. Since the sample size of each poll is relatively large, the observed proportion of respondents who report that they intend to vote democratic, $y_k = n_k/N_k$ is approximated well by a normal distribution,

$$y_k \approx N(\pi_k, \sigma_k^2), \tag{20}$$

where $\sigma_k^2 = y_k(1 - y_k)/N_k$. However, the parameter of underlying interest is not π_k, but the actual state of national opinion at time t. The π_k are consequently modeled as a function of two components: the actual state of opinion, μ_t, and a house effect, λ_j, specific to polling firm $j = 1, \ldots, J$,

$$\pi_k = \mu_{t[k]} + \lambda_{j[k]}. \tag{21}$$

Equation 21 is not identified because one could shift $\mu_{t[k]}$ up/down and $\lambda_{j[k]}$ down/up by the same constant without changing the value of π_k. As a result, I use the identifying restriction that the house effects sum to zero, $\sum_j \lambda_j = 0$.

As currently specified, the model only provides a snapshot of national opinion on any given day. To forecast the election, it is necessary to estimate μ_T, which is an estimate for national opinion on the day of the election. Since forecasts are made on days $t' < T$, it is therefore necessary to make assumptions about the movement of μ_t from one day of the campaign to the next. Since there is no reason to expect there to be any trends in polling, it is reasonable to expect μ_t to follow a random walk, so that the full model can be written

as,

$$y_k = \mu_t + \lambda_j + v_k,\ v_k \sim N(0, \sigma_k^2) \tag{22}$$

$$\mu_t = \mu_{t-1} + w_\mu,\ w_\mu \sim N(0, \sigma_\mu^2), \tag{23}$$

where σ_μ^2 is an estimate of the daily change in μ_t. As shown in Section B.1, equation 22 and equation 23 form a state-space model, or more specifically, (since the model is linear and errors terms are Gaussian) a dynamic linear model (DLM). Equation 22 is known as the observation equation while equation 23 is known as the state equation.

A model for the district vote relative to the national vote proceeds in the same manner as the model for the national vote, but with a few additional differences. The first difference is that it is impractical to correct for house effects because there are only a few polls published per polling firm.[14] The second difference is that national opinion at time t is not actually observed, so the relative vote cannot be observed either. In practice, this is not a large problem because national opinion, μ_t, is estimated very precisely using equation 22 and equation 23 due to the abundance of large national polls.[15]

For the model of the relative district vote, let l index district polls and continue to let i index a district. Define the deviation of a district poll from the national vote as $d_l = y_l - \mu_{t[k]}$. The state-space model for the relative district vote is then,

$$d_l = \xi_{it} + v_l,\ v_l \sim N(0, \sigma_l^2) \tag{24}$$

$$\xi_{it} = \xi_{i,t-1} + w_\xi,\ w_\xi \sim N(0, \sigma_\xi^2), \tag{25}$$

where $\sigma_l^2 = y_l(1 - y_l)/N_l$, N_l is the sample size of the lth poll, $\xi_{i[l]t[l]}$ is an estimate of the deviation of opinion in state i from national opinion at time t, and σ_ξ^2 captures the variance

[14]Pollsters tend to focus on specific districts; there are consequently hundred of pollsters, but only a few polls from each one. This stands in contract to polls of the national vote, for which there are far fewer pollsters and many more polls published per pollster.

[15]The standard deviation of μ_t is typically around 0.004.

of day to day movements in ξ. Equation 24 and equation 25 are just a simple multivariate extension of the model of the national vote that ignores house effects.

The overall forecast of the two-party vote for Democrats from each district is $\mu_T + \xi_{iT}$. Separate forecasts of μ_T and ξ_{iT} can basically be estimated using a Kalman filter[16], which the caveat that Bayesian MCMC techniques are needed to estimate the variance of the state equations and the house effects in the model for the national vote. The Kalman filter is instructive because it quantifies the relative weight attached to previous versions of the states and new polls. For example, using the Kalman filter, the mean and variance of the latent states in district i on day t given polls up to day t, $Y_{1:t}$ are,

$$m_{it} = E(\xi_{it}|Y_{1:t}) = \left[\frac{m_{i,t-1}}{C_{i,t-1} + \sigma_\xi^2} + \sum_{l \in \mathcal{P}_{it}} \frac{d_l}{\sigma_l^2}\right] \cdot C_{it}, \tag{26}$$

$$C_{it} = \text{Var}(\xi_{it}|Y_{1:t}) = \left[\frac{1}{C_{i,t-1} + \sigma_\xi^2} + \sum_{l \in \mathcal{P}_{it}} \frac{1}{\sigma_l^2}\right]^{-1}, \tag{27}$$

where $\mathcal{P}_{it}$ refers to the set of all polls published for district i on day t.[17] The mean of ξ_{it} is thus a weighted average of its mean on the previous day, $m_{i,t-1}$ and the deviation of all new polls from national opinion, $\sum_{l \in \mathcal{P}_{it}} d_l$, with weights proportional to their respective precisions. The precision of each poll is equal to the inverse of its sampling error and the precision of m_{t-1} is the inverse of the sum of its variance and error in the movement of the states from one period to the next. The variance of ξ_{it} is just the inverse of the poll precision plus the prior (m_{t-1}) precision. The interpretation for μ_t is identical except that new estimates are a weighted average of prior states and new polls less the democratic bias of the polling firm (i.e. $y_k - \delta_{j[k]}$).[18]

To incorporate information from the historical regression, I treat forecasts from the hier-

[16]The Kalman filter is commonly used in engineering to track the movement of objects such as satellites or aircraft that are measured with noisy data. It is also frequently used in Macroeconomics and in political science to track public opinion.

[17]There is almost never more than one poll for a given district and time period of a reasonably short duration (such as two weeks). For national elections, there are almost always multiple polling firms surveying on a given day or time period.

[18]See Appendix B for additional details.

archical model described above as election day polls for both the national and district models. The two-party vote given to the national pseudo poll is the mean average district vote from the hierarchical model and the Democratic vote shares given to each district's pseudo poll is the mean of the district forecast from the regression less the mean of average district vote. The corresponding variances are calculating using the posterior predictive distributions of these quantities.[19] The regression forecasts in the national (district) models consequently receive weights proportional to the regression forecast errors of the average district vote and the relative district vote.

When estimating μ_T and ξ_T prior to the election, there will be gaps between the last published poll and the pseudo poll from the regression. The Kalman filter helps bridge this gap by pushing the latent states forward toward election day. The relative weights received by the regression analysis and the polls depends largely on the number of days until the election. For instance, due to the random walk assumption, the precision of m_{T-1} for a forecast of the national vote made on day t' is equal to $1/\left[C_{i,T-t'} + (T-t') \cdot \psi_k^2\right]$, which is linearly decreasing in time and in the day to day movements of the states. It follows that the regression analysis receives more weight when the election day is far away and when there is more movement in the polls.

The Kalman filter described in this section assumes that the variances of the states are known. Since, in practice, this is clearly not the case, I estimate all of the model parameters jointly using Bayesian methods. To do so, I assign the unknown variance parameters, σ_μ^2 and σ_ξ^2, inverse gamma priors. The house effects, λ_j, are given a normal prior centered at 0. The posterior density is simulated with a Gibbs sampler, which is described in Appendix B.

5 Data

The analyses in this paper utilize four main sets of data: data on nationwide variables, data on representatives from the U.S House, campaign contributions data, and polling data. The

[19] See Section 6.2 for more details.

main nationwide variables collected are the president's net approval rating and the Gallup generic Congressional ballot. Both the presidential approval rating and the Gallup generic ballot were obtained from the Roper Center Public Opinion Archives. The net approval ratings is the percentage of survey respondents who approve of the president minus the percentage who disapprove. The ratings are based on polls conducted by various polling organizations multiple times each month. The generic ballot is based on a survey question that asks voters whether they intend to vote for a generic (does not include candidate names) Republican or Democrat in the House election. Survey results for this question are more difficult to obtain than presidential approval ratings since each poll must be searched for individually so I consequently restrict the generic ballot questions to those asked by Gallup in August (the month the forecast is being made).

The second dataset consists of data from three sources: data on House elections from 1946 - 2012 obtained from Gary Jacobson; the committee assignments of members of Congress from each district (Stewart III and Woon 2015); and DW-Nominate scores from http://voteview.com/. The Gary Jacobson data provides information on a number of important characteristics of House elections at the district level. These include the Democrat's share of the two party vote in both House and presidential elections, whether an incumbent is running for reelection (and the party of the incumbent), and whether a challenger has previously held office. The DW-nominate scores are the first dimension scores originally developed by Keith Poole and Howard Rosenthal, which can be interpreted as the liberal-conservative divide in modern politics (Poole and Rosenthal 1997, 2011).

The third dataset contains campaign finance data provided by the Center for Responsive Politics (CRP) at https://www.opensecrets.org/. CRP obtains the data from the Federal Election Commission and adds value to it by cleaning and categorizing the data. The data covers campaign contributions from individuals (above $200) and Political Action Committees (PACs).[20] The data can be separated by date, individual employer, and PAC

[20]Contributions from party committees are included in the PAC table.

type, which allows me to track contributions during different times of the campaign and to focus on spending by different types of contributors.

The final dataset covers district and national polls for the 2010 House election. The district polls are all of those used by the New York times to forecast the 2010 House election and the national polls are the polls listed for the generic congressional vote on the website http://www.realclearpolitics.com/. Each poll contains the date that the poll was taken, its sample size and the proportion of respondents favoring the Democrat and Republican candidate.

The rest of this section describes the specific uses of these datasets in more detail. For more detailed descriptions of the variables and information on sources see the data appendix (Appendix D).

5.1 Forecasting Variables

The model used in this paper includes variables (and error terms) at the national and district level.[21] These variables were obtained from the nationwide data and the data on representatives from the U.S. House. Forecasts are made as of August of each election year using post 1980 data so all variables are measured before September 1st.

Recall that the primary dependent variable is the Democratic share of the two party vote. This is not the only possible choice for the response variable but it is frequently used in the literature. Other choice that are essentially identical include the incumbent party's share of the vote or the margin of victory for the incumbent candidate.

House elections are highly persistent from one election to the next so they can be predicted quite accurately using only the lag of the district vote. The lag is of course unavailable without substantial reaggregation in redistricting years so, as is common in the literature, the analysis excludes years ending in 2.[22] The lagged vote is tied in many respects to

[21] Regional variables were also considered but they did not improve the fit of the model and only increased model complexity.

[22] Examples of other studies that deal with redistricting years in this manner include Gelman and King (1990), Kastellec, Gelman and Chandler (2008), and Gelman and Huang (2008).

individual candidates and is not surprisingly a much stronger predictor of the vote when incumbents are running for re-election than in open-seats (Gelman and Huang 2008). My model consequently includes the lag of the presidential vote in each district in addition to the lag of the House vote. The presidential vote is subtracted from the nationwide presidential vote to control for national trends and can be interpreted as a measure of the degree to which a district leans toward one party or another.[23][24]

One of the largest and most consistent findings in the political science literature on American elections is that, all else equal, incumbent candidates receive more votes than challengers (Gelman and King 1990; Lee 2001; Ansolabehere and Snyder Jr 2002; Gelman and Huang 2008). To model this incumbency effect, I use the variable, *Incumbent*, which is equal to 1 if the incumbent is a Democrat, -1 if the incumbent is a Republican and 0 if there is no incumbent running. $+1, 0, -1$ variables are used repeatedly in the analysis and are always equal to +1 for Democrats and -1 for Republicans; they should be thought of as dummy variables that are constrained to have the same impact on the vote for both Democratic and Republican candidates (recall that the dependent variable is the Democratic candidate's share of the two-party vote).

Since more experienced candidates tend to do better at the polls, I include two $+1, 0, -1$ experience variables: *Freshman incumbent* and *Previous office holder*. *Freshman incumbent* is equal to +1(-1) if a Democratic (Republican) incumbent was elected for the first time in the previous election and 0 otherwise. *Previous office holder* is equal to +1 or -1 if a Democrat or Republican challenger had previously held office and was running against either an incumbent or a challenger without previous experience; it is equal to 0 in all other situations including open seats in which both candidates had previous political experience.

The model also accounts for the ideology of candidates relative to the ideology of vot-

[23]This variable is very similar to the Cook Partisan Voting Index (Cook PVI), which compares the two-party vote in the past two presidential elections to the nation's average share of the same presidential vote. The main difference is that I only include the previous presidential election in order to deal with redistricting.

[24]Models that interacted both the lagged vote and the centered presidential vote with whether incumbents were running were also considered, but did not lead to improvements in the fit of the model.

ers in their districts. Candidate ideology is measured using the first dimension DW-nominate score—which measures the political ideology of candidates and ranges on a liberal-conservative scale from -1 to 1. A measure of district ideology on the same scale as the DW-nominate score was created by taking a weighted average of the DW-nominate scores of the most recent presidential candidates for each party, with weights equal to each party's share of the district two-party vote in the most recent presidential election.[25] The variable used in the model, *Relative 1st dimension DW-nominate score*, is the DW-nominate score minus the district ideology score.[26] In incumbent districts, this variable will have a positive sign if candidates do better at the polls when they have more moderate voting records (perhaps because they appeal more to independent voters), conditional on winning the primary election. In open seats, this variable will be positive if voters tend to elect a Democratic candidate when the previous candidate(s) had a conservative voting record and a Republican candidate when the previous candidate(s) had a liberal voting record.[27][28]

The three nationwide variables are the president's average August approval rating, the August generic ballot and an indicator variable for whether the election is being held in a midterm year.[29] The *August generic ballot* variable is an average of generic ballot polls in the month of August in each election year. Both the midterm variable and the presidential approval variable are multiplied by +1 if the president is a Democrat and -1 if Republican. The impetus for the *midterm election* variable is previous research showing that the party of the president usually loses seats during midterm election (e.g. Erikson 1988).

[25]Michael Dukakis does not have a DW-nominate score so his ideology was imputed using the mean score among Democratic presidential candidates between 1980 and 2010.

[26]This variable outperformed another variable which interacted the absolute value of the DW-nominate with the party controlling the seat, suggesting that voting moderation is more strongly associated with higher vote shares than a candidate's ideological distance from the ideology of his or her voters.

[27]In open districts, the DW-nominate score is equal to the average DW-nominate score among all representatives in the previous Congress. Most districts only have one representative per congressional session but certain events such as deaths create scenarios in which this is not the case.

[28]In OLS regressions, the coefficient on the interaction of the relative DW-nominate score variable with whether a seat is open is essentially zero, implying that the effect of the variable does not differ by whether a seat is open or not.

[29]A variable including second quarter GDP was also considered but it was not statistically or economically significant.

Summary statistics for these variables are presented in Table 1. Variables rarely exceed one in absolute value and are thus on a similar scale to the dependent variable.

Table 1: Summary Statistics for Forecasting Variables

Variable	Min	Median	Max
Dem. share of district vote in last election	0.077	0.536	0.971
Relative Dem. share of presidential vote in last election	−0.320	−0.019	0.538
Incumbent	−1.000	1.000	1.000
Relative 1st dimension DW-nominate score	−0.897	−0.149	1.100
Freshman incumbent	−1.000	0.000	1.000
Previous office holder	−1.000	0.000	1.000
August presidential net approval rating	−0.567	−0.030	0.369
August generic ballot	0.467	0.513	0.615
Midterm election	−1.000	0.000	1.000

Notes: *Relative Dem. share of presidential vote in last election* is the deviation of the Democratic share of the presidential vote in each district from the national vote in the most recent presidential election. *Incumbent* is equal to +1 or -1 depending on whether a Democrat or Republican is running for re-election and 0 otherwise. *Relative 1st dimension DW-nominate score* is the 1st dimension DW-nominate score minus a measure of district ideology. *Freshman incumbent* is defined in the same was as *Incumbent* but is only equal to +1 or -1 if the incumbent was elected for the first time in the previous election. *Previous office-holder* is equal to +1 (-1) if a non-incumbent Democrat (Republican) had previously held office and 0 otherwise (or if the seat was open and both candidates had previously held office). *August presidential net approval* is the president's average net approval rating in August multiplied by +1 if the current president is a Democrat and -1 if Republican. *August generic ballot* is an average of generic ballot polls in the month of August. *Midterm election* is equal to 0 in non-midterm years and +1 (-1) if the president is a Democrat (Republican) and it is a midterm election year. Other variables such as second quarter GDP growth and the party currently controlling a district were also considered, but were not included in the final model because they were not statistically significant and they did not improve the fit of the model.

5.2 Campaign Spending

House candidates receive funds from three primary sources: party committees, PACs, and individuals.[30] Campaign donors can, in turn, spend on the behalf of candidates by either making a direct contribution, a coordinated expenditure or an indirect expenditure for or against a candidate.[31] Independent expenditures are not limited by law but cannot be coordinated with campaigns. By contrast, direct contributions and coordinated expenditures can be coordinated with a campaign but are limited by federal law.[32]

[30]Organizations—such as firms, unions or trade associations—that wish to spend money to influence federal elections must create a separate source of funds known as a PAC.

[31]A common example of an independent expenditure is a TV advertisement praising a candidate or criticizing an opponent. Before 2010, corporations and unions could only fund independent expenditures from their PACs; however in 2010, the U.S. Supreme Court rule in *Citizens United v. Federal Election Commission* that corporations and unions could use their own treasuries to raise money for independent expenditures.

[32]Coordinate expenditures, are, as the name suggests, expenditures made on the behalf of campaigns that can be discussed with the campaign.

Although campaign fund-raising is a year-round endeavor for political candidates, it follows a cyclical pattern and naturally heats up as election day nears. This is shown in Figure 1, which plots average daily spending by PACs in election years by month across the 2000 - 2010 elections. Expenditures are fairly constant from January until August and then jump considerably in September before peaking at around $4 million dollars per day in October and then decreasing slightly in November.[33]

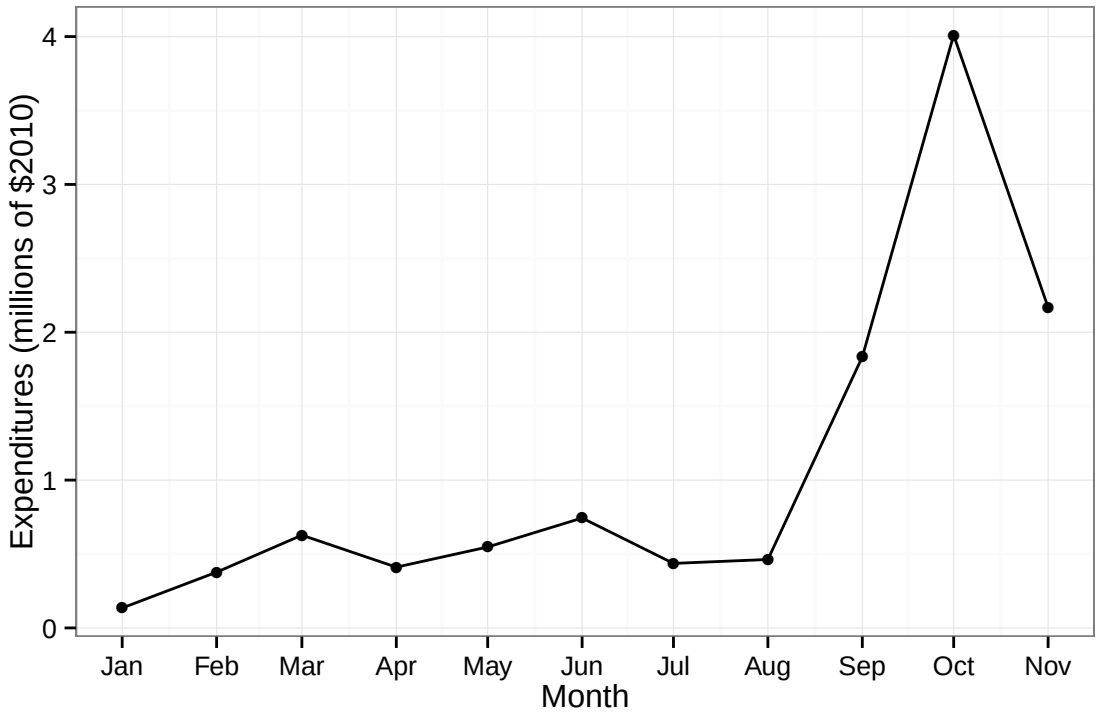

Figure 1: Average Daily Expenditures by PACs on Candidates by Month, 2000 - 2010

Notes: Expenditures are the average (across elections from 2000 to 2010) of the sum of expenditures on candidates by all PACS by month. Expenditures in non-election years are not included. Spending is in millions of 2010 dollars. November spending only include expenditures prior to election day.

Political parties are the focus of this paper, so it is worth examining the spending patterns of party committees in detail. The official campaign arms of the Democratic and Republican party in the House are the Democratic Congressional Committee (DCCC) and the National

[33]November expenditures only include spending that occurred prior to election day. Between 2000 - 2010, the latest election occurred on November 7th.

Republican Congressional Committee (NRCC). The principal goal of these organizations is to assist House candidates and help elect them to Congress. Direct contributions and coordinated expenditure are severely limited by contributions limits, so almost all spending is made independently of candidates. The Democratic National Committee (DNC) and Republican National Committee (RNC) sometimes spend money on House candidates as well, but the amount is trivial compared to the amount spent by the DCCC and NRCC.

While the national party committees have clear incentives to maximize the electoral success of their respective parties, it is unclear whether other organization share this same goal. Herrnson (2009) has argued that the reach of political parties extends considerably further than with the national party committees. He views parties as multilayered coalitions that include actors not usually considered as components of the party. I adopt this classification scheme and separate parties into three separate categories: national party committees, party-connected committees, and allied PACs.

The national party committees are the DCCC, NRCC, DNC and RNC. Party-connected committees consist of state and local party committees, candidate committees, and House candidate's leadership PACs. These committees have incentives to adopt spending strategies based on private or local goals, but Herrnson (2009) argues that they attempt to advance their party's collective goals as well. Spending by state and local party committees is also tied to the desires of the national parties because a significant portion of their receipts come from transfers from the DNC and RNC.[34] Finally, following Herrnson (2009), I consider allied PACs to be PACs that spend over 90% of their funds on one party's candidates.

Post-August spending for these groups and other non-aligned PACs are shown in Table 2. The table reports spending per district averaged over all non-redistricting year elections from 2000 - 2010. The national party committees spend considerably more than all other groups as national party committee spending is between 45 and 50 percent of all PAC and party spending for both the Democratic and Republican parties. Allied PACs and other

[34]The state and local party committees receive most of their funds from individual contributions and transfers from affiliated party committees.

non-aligned PACs are the next largest spenders. PACs are considerably more likely to be allied with Democrats than Republicans, although spending by non-aligned PACs is similar for both Republican and Democratic candidates. Spending by party-connected committees makes up less than 10% of total PAC and party committee spending, making them the least prolific spending group. In total, slightly more is spent on Democrats than Republicans.

Table 2: Average Post-August Spending by Political Parties per District in Non-Redistricting Years, 2000 - 2010

	Democratic party		Republican party	
	$	Fraction of total	$	Fraction of total
National party committees	114,773.40	0.457	107,813.99	0.501
Party-connected committees	21,567.06	0.086	20,115.03	0.093
Allied PACs	58,572.38	0.233	29,828.05	0.139
Other PACs	56,311.25	0.224	57,516.67	0.267

Notes: 2002 is omitted because it is a redistricting year. Spending consists of direct contributions, coordinated expenditures and independent expenditures for a canditate or against the opposing candidate in 2010 dollars. Spending figures are per district means averaged across the 2000 to 2010 non-redistricting year elections. Expenditures prior to September 1st are excluded. National party committees include the DNC, DCCC, RNC and RCCC. Party-connected committee are candidate committees or leadership PACs. Allied PAC consist of PACs that have spent 90% or more of their funds on one political party. Other PACs are all other PACs that do not fit into the other three categories.

Candidates also receive a large percentage of their funds from individuals. Contributions from individuals donating over $200 are itemized but contributions from smaller donors are not. During the 2009 - 2010 election cycle, only 0.26% of the U.S. population gave more than $200, but total contributions for these individuals summed to $1.9 billion while contributions from all other individuals were only $1.1 billion.[35] Individuals who donated over $200 spent $99,189.29 and $101,575.8 per district on Democratic and Republican candidates after August 31st in non-redistricting years from 2000 - 2010 respectively. Since contributions for individuals who donated less than $200 are not itemized, I cannot make a comparable calculation for donations from these individuals. However, if one assumes that spending on House candidates mirrors total spending, then contributions from small donors should have totaled around $58,000.[36] Total PAC and party committee spending on House candidates is

[35] See https://www.opensecrets.org/bigpicture/.

[36] Individuals donating less than $200 spent $1.1 billion/$1.9 billion ≈ 0.58 as much as larger donors in 2010. These larger donors, in turn, spend around $100,000 per district so smaller donors likely spent close to $0.58 \cdot \$100,000 \approx \$58,000$.

therefore larger than spending from individuals and the national party committees are the largest single source of funds for House candidates.

5.3 Polling Data

A steady stream of polls asking about the generic congressional vote during the 2010 election were released starting as early as December of 2008. All told, the website Real Clear Politics gathered 270 polls published by 30 different polling organizations over 670 days with the last poll conducted 4 days prior to the November 2nd election.[37] Most pollsters only published a few polls (median = 3.5), but the two most prolific pollsters, Rasmussen and Gallup, conducted 49% of all polls. The median sample size across surveys was 1600 and the average percentage of undecided respondents was 15%.

District polls were considerably less prolific. Polls were published by 113 different pollsters for 174 out of the 435 Congressional districts and the median number of polls per district in districts with at least 1 poll was 3. The median sample size for these polls was 400.

Figure 2 provides summaries of the frequencies at which district polls were fielded. As shown in panel (a) of the figure, polls were considerably more prevalent in close races (those in which the final vote share was close to 0.5), although few to no polls were published in some close races. This is important for forecasting the election because it means that the districts that are the most difficult to predict are the ones that have the most polling data. This matters for optimal spending as well since it reduces uncertainty about the closeness of a race and provides parties a better sense of which districts to target when attempting to influence the variance of the election.[38]

Panel (b) of the figure shows the cumulative number of polls during each day of the campaign. Nearly 2/3 of the polls were fielded after the start of September, when the rate of polling began accelerating quickly. This temporal pattern is consistent with the monthly

[37] Polls are typically conducted over multiple days. I consider the poll date to be the midpoint between the start and end date of polling (when the poll is conducted over an even number of days I use the median day closest to the start day).

[38] Recall that the trailing (leading) party in total seats has an incentive to spend more on districts in which it is losing (winning) if parties are trying to win a majority of seats.

pattern in contributions, suggesting that the final two months of campaigns capture the most crucial moments.

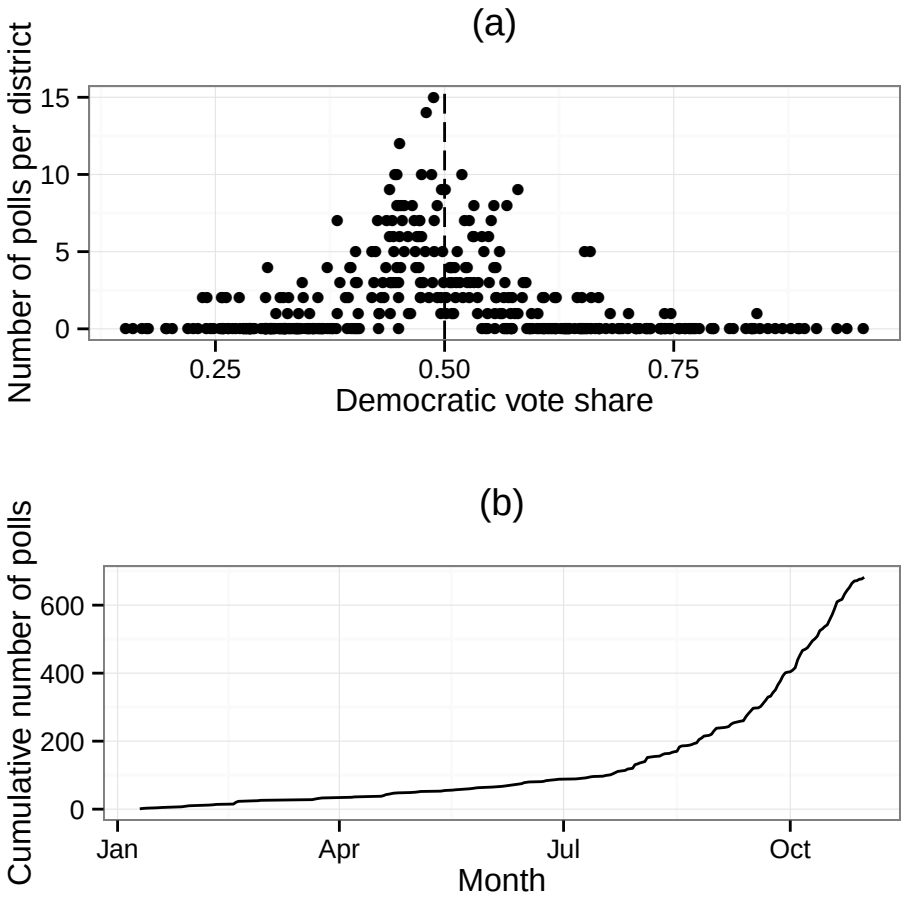

Figure 2: Summary Plots for District Polls

6 Forecasting the Two-Party Vote and Calculating Q_i

6.1 Results from Bayesian Hierarchical Model

The hierarchical model was implemented using Stan, which uses Hamiltonian Monte Carlo sampling (see Appendix A for Stan code).[39] The model was fit 5 separate times using

[39]Information about Stan is available at http://mc-stan.org/.

post 1980 data to forecast (out-of-sample) the 2000, 2004, 2006, 2008 and 2010 elections.[40] Uncontested seats were dropped from the model because parties would have known which districts were uncontested by September 1st.[41][42] The posterior distribution for each of the 5 models was simulated using the last half of 5 chains of 2,000 iterations each. The Gelman and Rubin potential scale reduction factor $\hat{R}$ was approximately 1 for all model parameters in all 5 fits, which suggests that the Markov Chain successfully converged each time.

Posterior means and quantiles for the parameters using data from 1980 - 2008 (which was used to forecast the 2010 election) are shown in Table 3. All of the variables (except the August generic ballot) are precisely estimated and have 95% credibility intervals that do not include zero. The posterior densities for all parameters are all approximately normally distributed, which is consistent with the central limit theorem in a Bayesian context (see Figure C.1 in Appendix C).[43]

The signs of the variables are generally as expected. For example, the president's party does worse in midterm years, challengers with previous political experience receive more votes and voters tend to prefer representatives with less extreme DW-nominate scores. The sign on the freshman incumbent variable may be somewhat of a surprise though, as less experienced incumbents are actually predicted to do better than their more experienced counterparts. Finally, the estimated incumbency effect is consistent with previous research, although it is a little closer to typical estimates of the incumbency effect in recent years (of around 10%) than the more involved method used by Gelman and Huang (2008) (which found the incumbency effect to be closer to 8%).[44]

[40] For example, the 2000 election was fit using elections from 1980 to 1998; the 2004 election was fit using elections from 1980 to 2000 (recall that the model excludes years ending in 2); and so on.

[41] Signature-filing deadlines in districts are all before September 1st.

[42] I also dropped the few races with third party candidates. These do not impact the calculates of Q_i^{maj} because they all occurred before 2000. Independents Bernie Sanders of Vermont and Virgil Goode of Virginia are classified as a Democrat and a Republican respectively.

[43] The distribution on the national error is somewhat of an exception because it is somewhat right-skewed. This likely occurs because the number of national elections is small and the the error terms are truncated at zero.

[44] Gelman and Huang (2008) estimate candidate level incumbency effects using a multilevel model. They also include a model for the party of the incumbent which depends on the district vote lagged two periods.

Table 3: Posterior Inferences for Parameters in Forecasting Model, House Elections 1980 - 2008

		Posterior quantiles		
Variable	Mean	2.5%	Median	97.5%
Coefficients				
Intercept	0.343	0.196	0.344	0.489
Dem. share of district vote in last election	0.418	0.394	0.418	0.442
Relative Dem. share of presidential vote in last election	0.337	0.313	0.337	0.360
Incumbent	0.094	0.089	0.094	0.099
Relative 1st dimension DW-nominate score	0.060	0.049	0.060	0.070
Freshman incumbent	0.012	0.006	0.012	0.018
Previous office holder	0.032	0.028	0.032	0.037
August presidential net approval rating	0.061	0.026	0.061	0.095
August generic ballot	−0.079	−0.355	−0.080	0.199
Midterm election	−0.044	−0.059	−0.044	−0.029
Variance terms				
District error (σ)	0.061	0.060	0.061	0.062
National swing (σ_δ)	0.013	0.007	0.012	0.026

Although parameters have only been reported for the model using data from the 1980 to 2008 elections, the model was actually fit 5 separate times in order to imitate 5 separate real time forecasts using data that would have been available to a political analyst in August *prior* to each election. Each model has its own unique posterior distribution for the parameters. In practice, the district parameters do not vary much from one election to the next but that national parameters (which are estimated using much less data) become more precise over time.

In the Bayesian context, the forecast for an election in year $\tilde{y}$ is a probability distribution over $v_{i\tilde{y}}$, called a posterior predictive distribution, which uses data from election years $y < \tilde{y}$ to form a distribution for the unobserved (to the political analyst at the time) election in year $\tilde{y}$. Given the persistent nature of the data, it is not surprising that the model is able to forecast elections quite accurately. Indeed, if the forecast, f_{iy}, is the mean of the posterior predictive distribution of v_{iy}, then the root mean square forecast error (RMSFE), $\sqrt{E[(f_{iy} - v_{iy})^2]}$, is 0.056 and the winner is only picked incorrectly 7% of the time.

There is also evidence that the decomposition of the error into a district component and

a national component captures uncertainty in elections quite well. To see this, consider Figure 3, which is a histogram of the posterior predictive distributions of the fraction of seats won by Democrats in each of the five non-redistricting year elections from 2000 to 2010.[45] The posterior predictive distribution for the number of seats is formed by simply counting the number of districts that had vote shares over 0.5 for each of the $5,000$ draws from the posterior predictive distributions of $v_{i\tilde{y}}$ for each forecast year $\tilde{y}$. The actual fraction of Democratic seats is usually within the 95% credible interval, but the intervals are not unreasonably wide either.[46] The 2010 election is a bit of an outlier since nearly all of the close elections broke toward the Republicans and large number of Democratic incumbents were defeated by Republican challengers.[47] The next part of this paper will examine whether polling data can help improve this forecast.

6.2 Results from the DLM

Although the DLM allows the states to vary on a daily basis, days are divided into two week periods in order to reduce the computational burden.[48] When a district (in the district model) or polling firm (in the national model) has multiple polls during a period I take an average of the polls weighted by sample size. This average poll is then given a sample size equal to the sum of the sample sizes of the individual polls. Aggregating the polls in this manner does not have a substantial impact on the forecasts for a number of reasons. Firstly, academic research typically shows that frequent polling provides very little information (e.g. Gelman and King 1993; Lock and Gelman 2010). This is likely to be especially true in a window as short as two weeks. Secondly, there are very few cases in which a particular

[45] Uncontested seats are not included in these percentages.

[46] If there was no national shock these 95% credibility intervals would be far to narrow because all of the errors would be independent.

[47] The majority party typically holds an electoral advantage because they tend to win a greater number of seats than votes (Tufte 1973). This Democrat's did not enjoy this partisan bias in 2010.

[48] For the district model using daily data, a separate parameter must be estimated for each district each day. For instance, if one started using polling data in July, then (since there are 174 districts with polling data), approximately $174 \times 90 = 15,660$ parameters (in addition to the variance parameter) would need to be estimated. This estimation would then need to be done each time a forecast was made so that if one wanted daily forecasts, $90 + 90 \times 15,660 = 1,409,490$ parameters would need to be estimated!

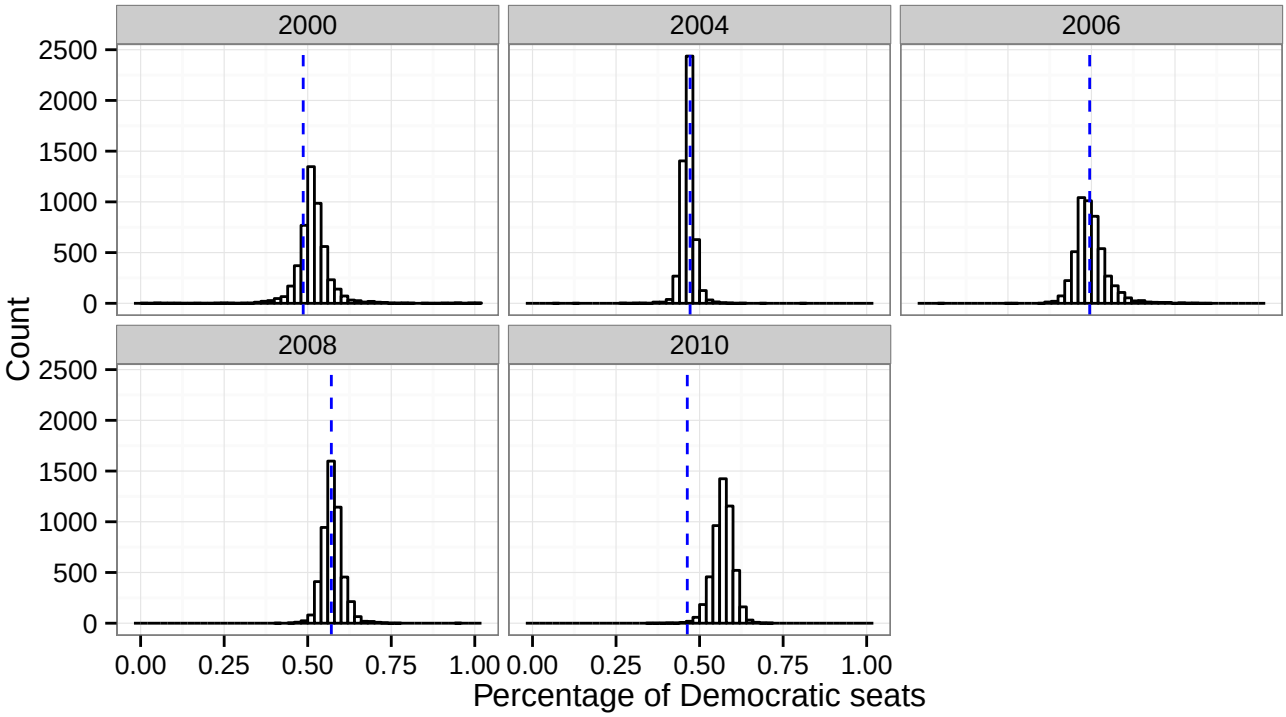

Figure 3: Forecasts of the Number of the Percentage of Democratic Seats

Notes: The histograms are from the posterior predictive distribution of the hierarchical model for each election year.

district was polled more than once in any two week span.

To remain relatively agnostic about biases, I set the prior variance for the House effects σ_λ^2, equal to 0.05^2. This allows for very biased polling firms because plus or minus 2 standard deviations from 0 is plus or minus 10 percentage points. In addition, the inverse gamma priors for the unknown variance parameters, σ_μ^2 and σ_ξ^2, were assigned shape parameters equal to 9 and scale parameters equal to 0.02, corresponding to a mean of 0.05^2. The prior implies that period to period movements should be around 5 percentage points, but its effect is in practice dwarfed by the data. Election day polls were added to both the national and district models. The final poll in the national model is given a mean and standard deviation of 0.514 and 0.019 respectively, which are the mean and standard deviation of the posterior

predictive distribution of the average district vote from the hierarchical model.[49] Similarly, the mean and standard deviation of the election day poll for the district model are the mean and standard deviation of the posterior predictive distribution of the difference between the vote in each district and the average district vote. For reference, the mean of this standard deviation is 0.061, with is nearly identical to the mean of the district level error, σ.

The posterior density of both the national and district DLM's were simulated using $6,000$ iterations of the Gibbs Sampler with a burn-in of $1,000$ iterations. The model is (retrospectively) estimated every two weeks during the month's of September and October. Polling data after a desired forecasting date is removed to ensure that only data that would have been available at the time is used. This yields four separate forecasts of the 2010 election in addition to the prior from the hierarchical model: a mid September (1.5 months prior to election day), late-September (1 month prior to election day), mid-October (2 weeks prior to election day) and late-October (just before the election) forecast. Trace plots of the parameters suggest that the Markov Chain converged each time and are available upon request.

Estimates of the square root of the variance parameters, σ_μ and σ_ξ are shown in Table 4. The table reports estimates listed by the date the model was estimated. The posterior quantiles for σ_μ are consistent across dates and yield 95% credible intervals ranging from around 0.015 to 0.023 with a median of 0.018. Posterior estimates of σ_ξ are more variable but are generally pretty stable; the estimates do however increase slightly as election day nears.

Summaries of model's forecasts are are presented in figures 4 and 5. Figure 4 focuses on the forecast of the national vote. The dark dotted line is the DLM forecast (mean of the posterior distribution of μ_T) and the error bars are 95% credible intervals; from top to

[49] Uncontested seats are assumed to be known with certainty. Uncontested Democratic (Republican) seats are given a vote share of 0.75 (0.25). These values are based on those reported in King and Gelman (1991) and Gelman and King (1994) and derived from vote shares received by districts in the last year before they became uncontested and the first election after they became uncontested. For another paper that uses this strategy see Kastellec, Gelman and Chandler (2008).

Table 4: Posterior Quantiles for Square Root of DLM Variance Paramters

	National model (σ_μ)			District model (σ_ξ)		
Months to election	2.5%	Median	97.5%	2.5%	Median	97.5%
0.0	0.015	0.018	0.023	0.020	0.022	0.024
0.5	0.015	0.018	0.023	0.019	0.022	0.024
1.0	0.015	0.018	0.022	0.018	0.021	0.023
1.5	0.015	0.018	0.023	0.017	0.019	0.022

Notes: Months to election refers to the date that the model was estimated.

bottom, the light dotted lines are the regression based forecast from the hierarchical model, the actual average district vote and an average of polls from the latest available two-week period.[50] Forecasts are reported for the standard model that include the regression estimate as a final pseudo poll (the "prior" model) and an alternative specification that does not (the "no prior" model). Standard errors are considerably smaller for the "prior" forecast than the "no prior" forecast although the difference shrinks as election day gets closer because there is less and less temporal error in the polls. Due to this temporal error, the forecast of the "prior" model does not deviate from the regression forecast until mid-October, but as temporal error decreases more weight is placed on the polls and the forecast is eventually nearly identical to the actual average district vote. On the other hand, the "no prior", or polls only model, hugs the polling average closely and predicts that the national vote will be considerably more Republican. The "no prior" model lies slightly above the polling average though because the model accounts for House effects (see Figure C.4 in the appendix) and the most prolific pollsters (Rasmussen Reports and Gallup) leaned Republican during the campaign.

Figure 5 displays forecasts of the fraction of seats won by the Democrats after excluding uncontested seats. The upper light dotted line labeled "regression forecast" is identical to the mean of the histogram for the 2010 election in Figure 3 and the reported fraction of actual seats won excludes uncontested seats as well. The forecast (and 95% credible intervals) are

[50]The actual average district vote includes uncontested seats. As noted in the footnote above, these calculations are made by imputing the vote share in uncontested Democratic seats to 0.75 and uncontested Republican seats to 0.25.

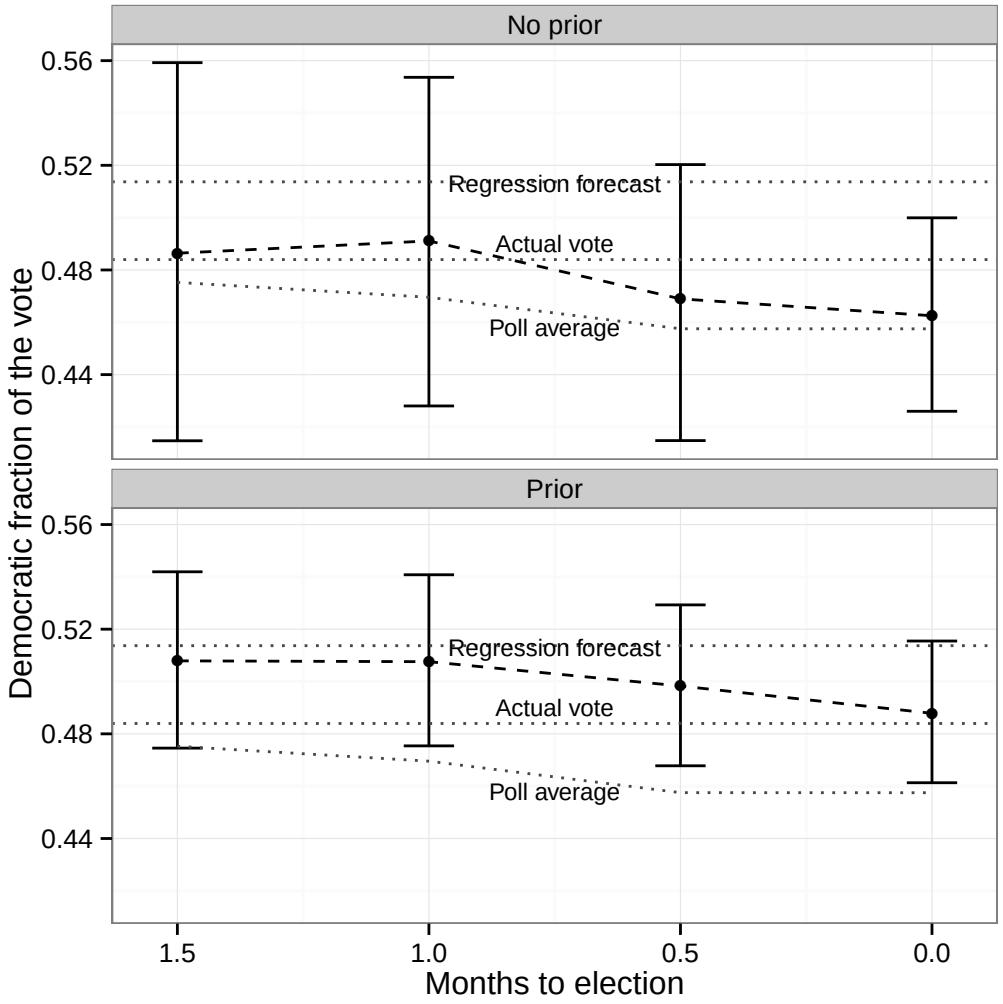

Figure 4: Forecasts of the 2010 National Vote by Date

Notes: The plot labeled "No prior" plots results from the DLM that does not include a final national pseudo poll based on the regression analysis; the plot labeled "Prior" includes this final pseudo poll. Vertical bars are 95% credible intervals for the forecast of the national vote from the DLM (points are median of the simulated vote). The upper dotted line is the point estimate of the regression based forecast (the mean of posterior distribution of the average district vote from the hierarchical model). The middle dotted line is the actual average district vote for Democrats. The lower dotted line is an average of all national polls during each time period (i.e. the poll average at the 2 month mark is all polls from days between 2 and 1.5 months before the election).

calculated by taking the fraction of seats with a predicted vote share greater than 0.5 for each simulated draw of $\mu_T + \xi_{iT}$.[51] Like the forecast of the national vote, the "prior" model does

[51]Recall that the states, ξ_{it} are estimated from district polls relative to national opinion at time t and that national opinion is, in turn, estimated as the mean of the posterior distribution of μ_t. The posterior mean of μ_t from the model estimated for late-October is shown in Figure C.3 in the appendix.

not begin to move away from the regression forecast until mid-October and the "no prior" forecast changes considerably more from period to period. The "no prior" model provides a forecast of the Democratic seat share that is very close to the actual Democratic seat share while the "prior" model overestimates the electoral success of the Democrats.

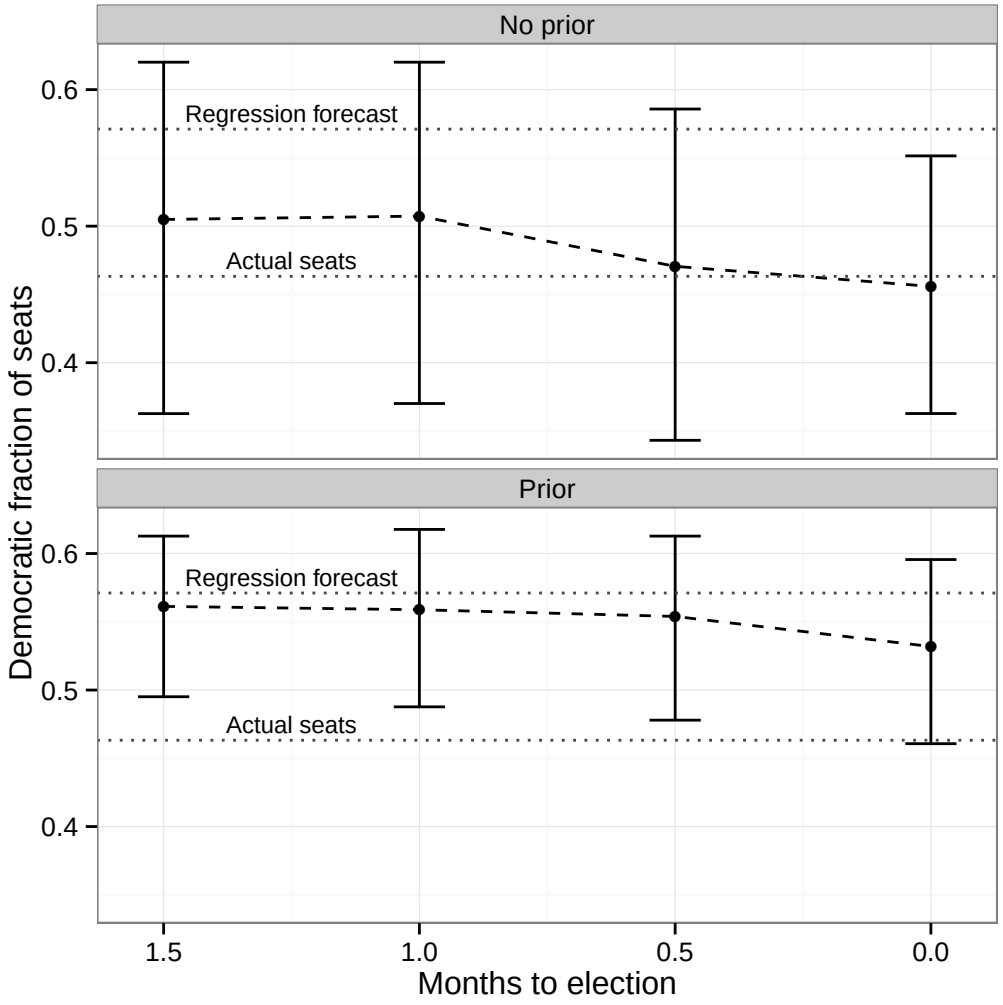

Figure 5: Forecasts of the Democratic Fraction of Seats in the 2010 Election

Notes: The plot labeled "No prior" plots results from the DLM that does not include a final pseudo polls based on the regression analysis; the plot labeled "Prior" includes these final pseudo polls. Vertical bars are 95% credible intervals for the forecast of the Democratic fraction of seats from the DLM (points are median of the simulated vote). The upper dotted line labeled "Regression forecast" is the mean of the posterior distribution of the Democratic seat share from the hierarchical model. The lower dotted line is the fraction of seats won by Democrats. All calculations omit uncontested seats.

It is tempting to use this evidence to discount the regression based prior but it is worth

remembering that i) 2010 is an outlier election, ii) nearly all of the close elections broke toward the Republicans and iii) comparing the average district vote to the actual district vote is only one of many possible ways to evaluate forecasts. Two alternative evaluation measures are shown in Figure 6.

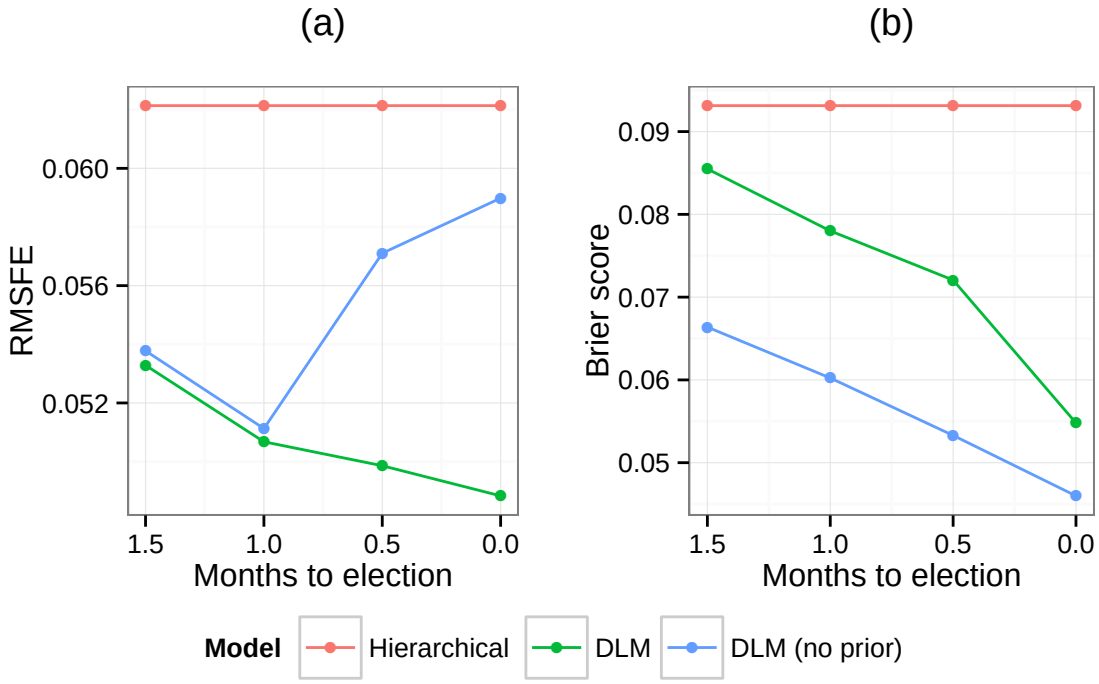

Figure 6: Evaluations of 2010 Election Forecasts

Notes: Figures (a) and (b) plot the root mean square forecast errors and brier scores respectively for the different forecast models by forecast date.

The first summary measure is the standard root mean square forecast error (RMSFE) described earlier in the paper. Here the standard DLM (with a prior) does considerably better than DLM with no prior. Both however have considerably lower RMSFE's than the regression based forecast from the hierarchical model. Strangely, the RMSFE's do not decrease during the campaign.

The second summary measure is the Brier score, which is often used to evaluate the accuracy of binary probabilistic forecasts. For this paper, the Brier score at time t can be written as $BS_t = \frac{1}{408}\sum_{i=1}^{408}(f_{it} - a_{iT})^2$, where 408 is the number of contested House seats; f_{it} is the forecasted probability that the Democratic candidate would win district i at time t,

and a_T is a binary variable equal to 1 if the Democratic candidate won the actual seat and 0 if the Republican candidate won. Higher Brier scores indicate worse forecasting performance. As shown in the figure, the Brier scores for both DLM's decline in a linear fashion over time and are considerably smaller than the Brier score of the hierarchical model. In contrast to the RMSFE, the Brier sore suggests that the DLM without a prior is a better predictor than the DLM with a prior.

The results suggest that forecasts with the DLM's are more accurate than forecast with the hierarchical model alone, but it is somewhat disconcerting that the RMSFEs increase as the election day gets closer. Figure C.5 in the appendix shows that this is being caused by districts with no polling data: the RMSFEs decrease (increase) over time in districts with (without) at least one poll. The poor performance of the RMSFEs in districts without polls is an artifact of the downward trend in the forecast of the national Democratic vote (see Figure 4). This is illustrated in appendix Figure C.6, which plots the mean prediction error (MPE)—defined as $E[f_{it} - v_{iT}]$ where v_{iT} is the observed vote on election day—by forecast date. The plot shows that forecasts become increasingly biased toward Republicans and that the bias is only severe in districts without polls (since the poll based forecasts in districts without polls are driven entirely by the biased national polls). The bias increases the overall RMSFE but not the overall Brier score because the Brier score is more sensitive to changes in forecasts of close elections (which are the most likely to have polling data) than lopsided ones while the RMSFE is equally sensitive to changes in both.

6.3 Calculating Q_i

Armed with parameter estimates from the hierarchical model and the DLM, one can use the estimates of V_i, σ and σ_δ to calculate the Q_i's from equations 12 and 13. In the theoretical model, these parameters are assumed to be known with certainty. But, when using a Bayesian approach this is not the case since the parameters have their own probability distributions.

Accounting for this additional uncertainty with the hierarchical model is straightforward:

I simply calculate Q_i by averaging over all of the unknown parameters. In other words, I add extra uncertainty to the parties maximization problems by integrating the objective functions over $p(\beta, \delta_t, \sigma, \sigma_\delta)$ instead of just integrating over $h(\delta_t)$.

Q_i is calculated in a similar manner when the model is estimated using the DLM. Because I separate the national vote and the district vote relative to the national vote, there is a probability distribution for each component. Due to the normality assumptions, both the forecasts of the national vote, μ_T, and the district vote, ξ_{iT} are normally distributed. σ can consequently be estimated with the standard deviation of the posterior distribution of ξ_T. A posterior distribution for $V_i + \delta$ is simulated by summing the mean of the posterior distribution of ξ_T and each simulated draw from the posterior distribution of μ_T. In total, this provides a range of forecasts depending on the national error as well as an estimate of the district level error as required by the theory.

As discussed in Section 3.3, party goals can have an important impact on Q_i. The implication of these goals are illustrated in Figure 7, which shows how Q_i depends on forecasted Democratic vote shares. When parties want to maximize the expected number of seats, they should spend the most resources on the closest elections. This is clear in the leftmost plot in which Q_i follows a bell shaped pattern peaking when when the forecasted vote share is 0.5.

The strategies are more intricate when the parties maximize the probability of winning a majority of seats. In this case, the parties should spend the most resources on decisive swing districts, or districts that are most likely to be close when winning that district will cause one party to win one more seat than the other party. This, in turn, implies that the trailing party should spend more resources in districts that they are losing in an effort to make the election more unpredictable. These incentives are best illustrated in the 2008 election when the Democratic party was expected to win a large majority of seats.[52] As shown in the plot, the Republican's optimal strategy (assuming they were only concerned with winning a majority of seats), was to to spend the most resources in districts with a

[52]Forecasts from the hierarchical model predict that the Republicans would have won, on average, 177 of the necessary 218 seats for a majority; in reality, they won 178 seats.

predicted Democratic vote share over 0.5. Since the number of seats won by both parties would have only have been close if the national swing shifted drastically in the Republican's favor, Q_i is maximized for districts with a vote share close to 0.6. Interestingly, the cost of spending on districts that Republicans were expected to barely win has such an adverse effect on the variance of the election that Q_i is actually negative in these districts.

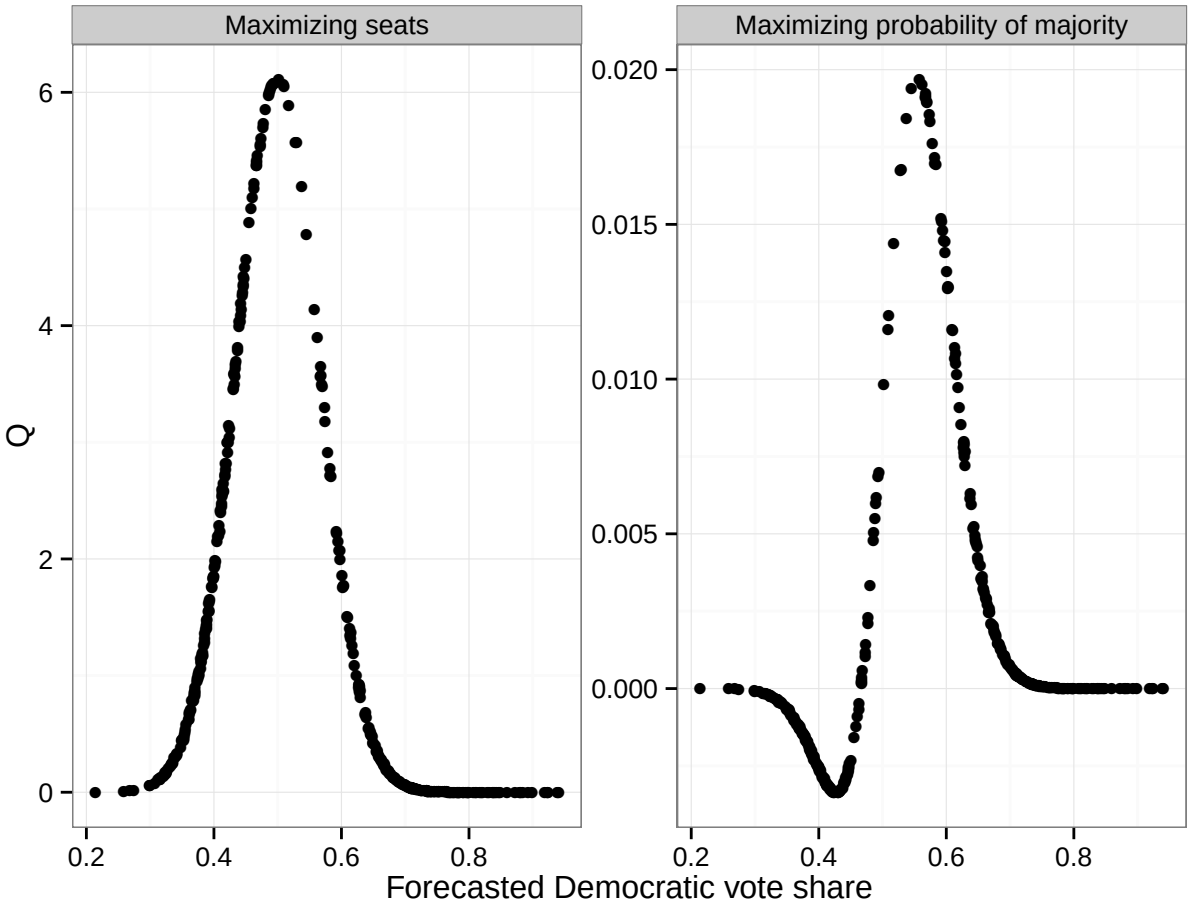

Figure 7: Q_i versus the mean forecast of the Democratic vote in the 2008 election

Notes: The forecasted Democratic vote share in each district is the mean of the posterior predictive distribution from the hierarchical model.

7 Relationship Between Q_i and District Spending

This section examines the empirical relationship between the estimates of Q_i and actual spending patterns by political parties, PACs and interest groups. The exact relationship between Q_i and spending depends on the functional form of $u(e_i^J)$. It would be preferable to estimate a utility function and its parameters using empirical estimates of the effect of spending on votes, but results from the empirical literature are very imprecise (see the discussion in Section 2). Instead, I use the logarithmic utility function analyzed in Section 3.4, which yields an equilibrium in which $= e_i^{J*}/\sum e_i^{J*} = Q_i/\sum Q_i$. That is, I estimate the regression equation,

$$\frac{e_{iy}^J}{\sum e_{iy}^J} = \gamma \frac{Q_{iy}}{\sum Q_{iy}} + X\beta + \eta, \tag{28}$$

where e_{iy}^J is observed spending in district i and year y and X contains other covariates that affect district spending.

7.1 Correlations

Before moving to regression analyses, I examine simple correlations between the Q_i's and actual spending. Figure 8 plots the correlation between Q_i and spending by contributor type and the goal of the parties. The contributor types are broken down according to the party groups discussed in Section 5.2. The top and bottom panels show the correlation with spending between Q_i^{maj} and Q_i^{seats} respectively.

The plot highlights two primary aspects of the data that are worth discussing. First, there are few differences in spending patterns between contributor types. Surprisingly, the correlations for party affiliated groups are not significantly higher than the correlations for other groups. In addition, spending by individuals is actually the most correlated with the Q_i's among all groups.

Second, the correlations are not very sensitive to party goals except in 2008. In Section 6.3

I showed that assumptions about party goals impacted the values of Q_i greatly in 2008 because the Democrats were expected to win the House by almost 80 seats. The 2008 election therefore provides a natural experiment that can be used to identify the goals that the parties actually have. The huge dip in the correlation between spending and Q_i^{maj} but not between spending and Q_i^{seats} is telling. I view this as strong evidence that parties maximize the expected number of seats rather than the probability of winning a majority of seats. Of course, another possibility is that parties attempt to maximize seats when the election is lopsided but try to win a majority of seats when the election is closer. This is plausible, but there is not enough evidence to support it because differences in correlations between the two goals are too small in election years other than 2008.[53]

Due to these results, I will now focus on total spending by PACs, party committees and individuals unless otherwise specified. According to Stromberg's model, if the effect of spending on vote share is of the log form, then both parties should spend the same proportion of their funds in each district. An examination of spending patterns by parties across districts is consistent with this notion. For instance, the simple correlation between spending on Democrats and spending on Republicans is 0.819. Moreover, Figure C.8 in the appendix shows very little differences in the correlations between the Q_i's and spending by party, although Q_i^{seats} is slightly more correlated with spending on Republicans than spending on Democrats.

The finding that campaign contributions are higher in close elections is a common result in the political science and economics literatures. As mentioned earlier, one drawback of these studies is that they use the ex-post electoral margin as a measure of the closeness of the election and assume a linear relationship between this margin and spending.[54] Do the values of Q_i calculated using Stromberg's probabilistic voting model fit the data better?

Figure 9 takes a look at this question by plotting the correlation between various mea-

[53]The correlation between actual spending and Q_i^{maj} is higher (by about 0.1) than the correlation between spending and Q_i^{seats} for all groups in 2000 and 2006. But in 2004 (when the predicted seat share is nearly identical to 2006) and 2010, the correlation with spending is actually slightly higher for Q_i^{seats} than Q_i^{maj}.

[54]Some studies used the electoral margin from the previous election as a measure of closeness.

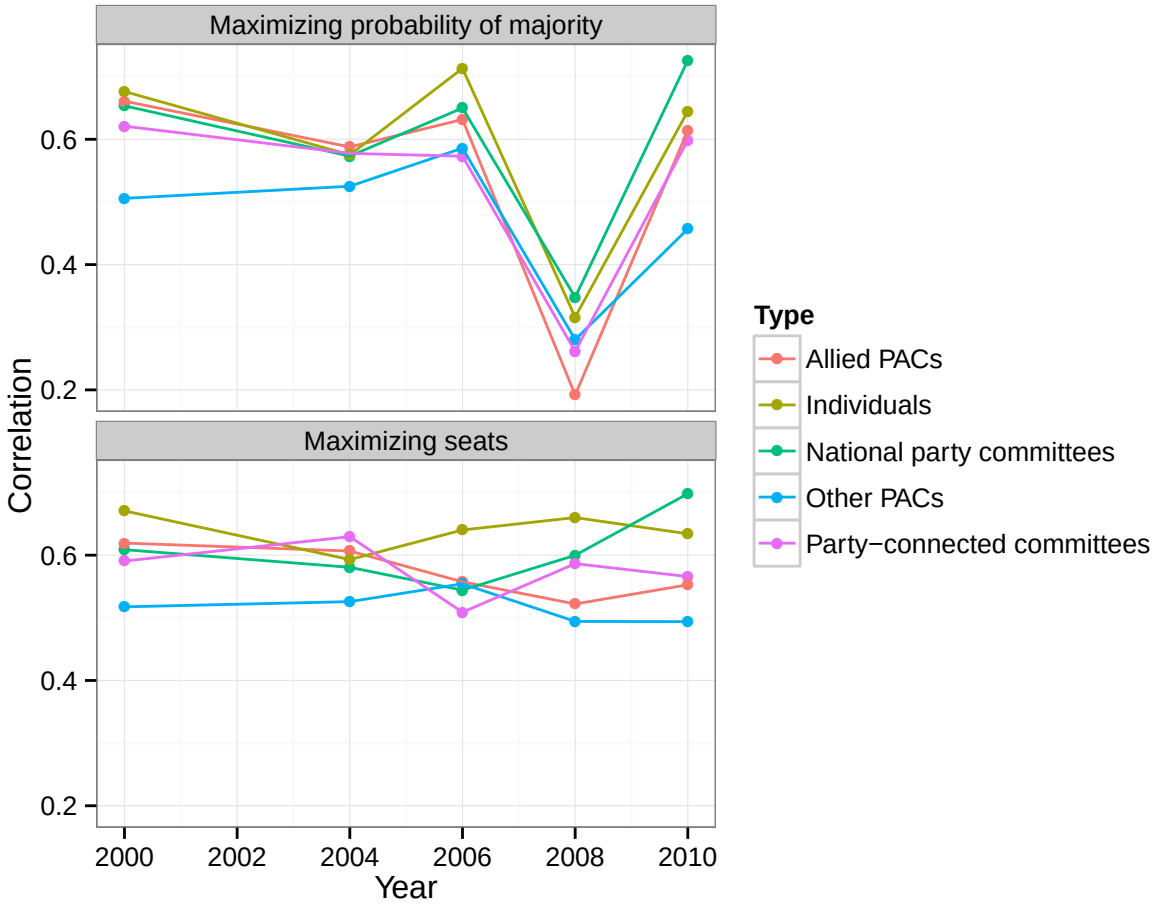

Figure 8: Correlation Between Q_i and Spending, by Contributor Type

Notes: Spending is the sum of spending for Republicans and Democratics in a district by a given contributor type. The top panel shows the correlation between Q_i^{maj} and spending while the bottom panel shows the correlation between Q_i^{seats} and spending.

sure of electoral competitiveness. Based on Figure 8, I use Q_i^{seats} as the measure of electoral competitiveness derived from the probabilistic voting model. The electoral margin is measured as the absolute value of the Democratic vote share less 0.5.[55] Q_i^{seats} fits the data considerably better than the electoral margin for all five elections in the sample. There is also little difference between calculating electoral margin with the actual election outcome or a forecasted value, implying that Q_i^{seats} accounts for observed spending better because it

[55]Uncontested seats are treated as having an electoral margin equal to 0.5 to maintain consistency with the probabilistic voting model.

depends on forecast uncertainty and that the normality assumptions on the error terms are reasonable.

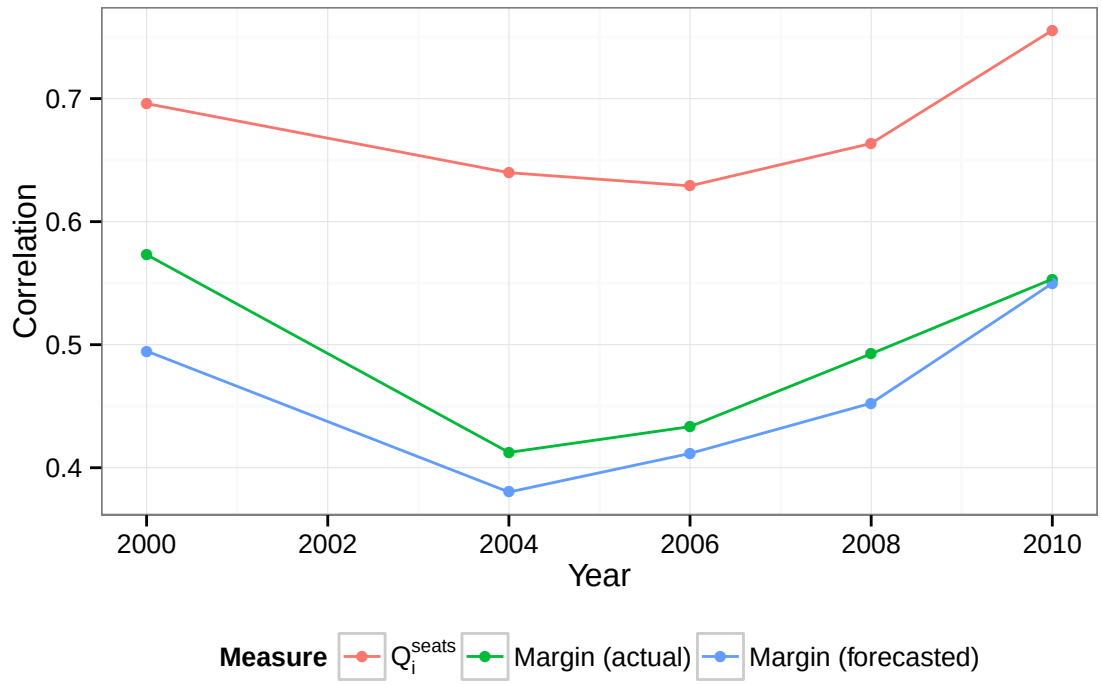

Figure 9: Correlation Between Measures of Electoral Competitiveness and Campaign Spending

Notes: Margin is the absolute value of the actual (forecasted) two-party vote share in a district minus 0.5. Campaign spending is the sum of all spending by PACs, party committees and individuals contributing over $200 in a given district and year.

The correlations presented thus far use the hierarchical model to calculate Q_i. One would expect that the DLM could improve the fit between the model and the observed data. Figure 10 examines this by comparing the correlation between Q_i^{seats} calculated using three different models and spending at different dates during the 2010 campaign.[56] Spending is the sum of all spending in a district between election day and the the date the forecast is made. The three models are the hierarchical model, the prior informed DLM and the non-prior

[56]The figure omits correlations made just prior to the election (late October) because the correlations are significantly lower and distort the figure: the correlations using the prior-informed DLM, no-prior DLM and hierarchical model are 0.422, 0.383 and 0.349 respectively. This likely occurs there are only a couple of days between the final forecast date and election day.

informed DLM.

In early September, or two months prior to the election, forecasts are only available from the hierarchical model so I assign all the models forecast values equal to the forecast from the hierarchical model. The figure shows that the Q_i^{seats} calculated using the DLM's match actual spending better than the Q_i^{seats} calculated using the hierarchical model, although the Q_i^{seats} from the hierarchical model are still pretty accurate. The correlations between the Q_i^{seats} estimated using the hierarchical model and spending decrease in a linear fashion over time while the correlations based on the DLM forecasts reach a peak of around 0.8 one month before the election. There is little difference between the performance of the "prior" and "no prior" DLM's, although the correlations between the "no prior" DLM and spending are somewhat more variable. Overall, this figure suggests that campaign donors use polls to evaluate the competitiveness of a district and update these beliefs when new polls become available.

7.2 Alternative Predictors of District Spending

I now move beyond simple correlations to analyses using regression equation 28. Table 5 reports regression estimates of γ and β when analyzing spending on Democratic candidates (panel A) and Republican candidates (panel B) separately. Equilibrium spending, $Q_{iy}^{seats} / \sum_i Q_{iy}^{seats}$, is estimated using the hierarchical model. Spending shares and equilibrium spending are multiplied by 100 so that they can be interpreted in percentage terms. Mean and median spending shares for Democratic and Republican candidates are 0.23% and 0.085% respectively for both parties.

The first column reports an estimate of a simple linear regression (with an intercept) with equilibrium spending as the only explanatory variable. The equilibrium spending coefficients in both panels A and B are strongly positively associated with actual spending and significantly different than 0. The coefficients are somewhat inconsistent with the equilibrium spending conditions in equation 15 though because they are significantly different than 1.

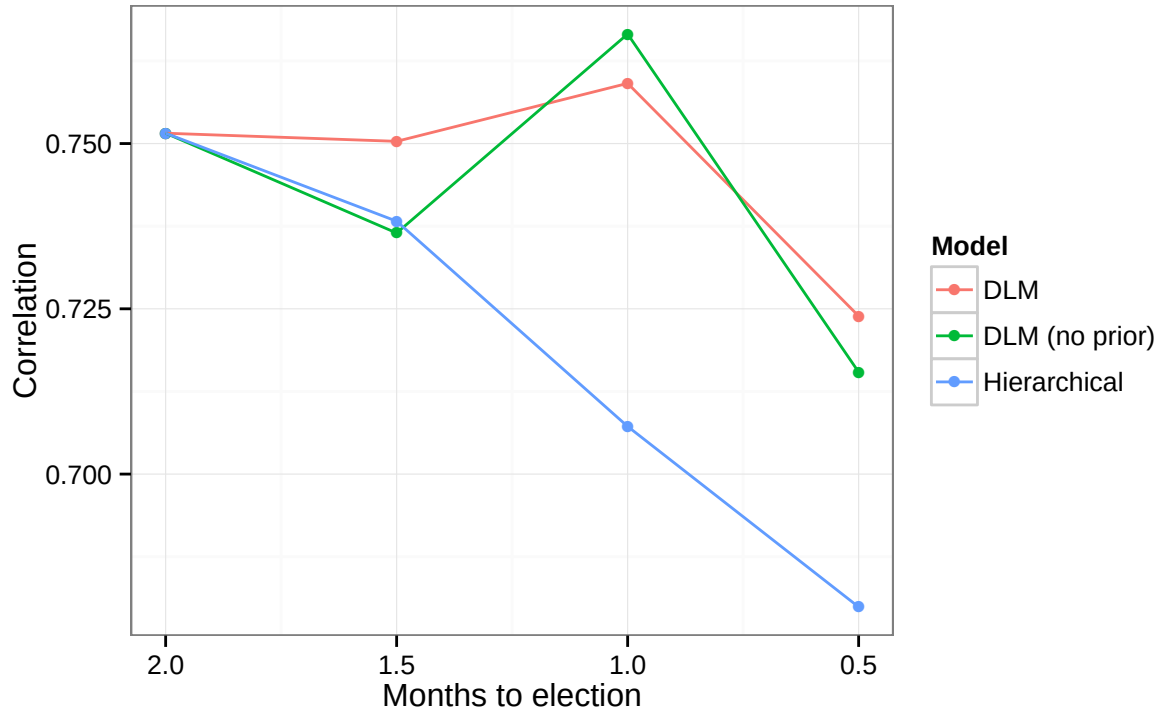

Figure 10: Correlation Between Q_i and PAC Spending by Date and Model, 2010 House Election

Notes: Spending refers to all spending by PACs, party committees and individuals contributing over $200 between a given date and election day.

That said, the coefficients are fairly close to 1 and the R^2 values are high.

The second column adds two dummy variables indicating that a candidate is an incumbent or running in an open seat (the omitted category is a challenger in an incumbent district). Each variable is statistically significant and positively associated with spending in both panels. The R^2 in column 2 improves considerably in the Democratic specification and marginally in the Republican one.

Columns 3 and 4 add influence-motivated variables that campaign donors might consider when deciding which campaigns to contribute to. The influence variables in column 3 are three indicator variables for whether candidates are member of the Ways and Means Committee, party leaders or committee chairmen. In column 4, I include an estimate of the probability that a candidate will win the election (0 to 1 scale) since campaign donors

Table 5: OLS Regressions on Candidate Spending Shares (%)

	(1)	(2)	(3)	(4)	(5)
Panel A. Democrats					
Q share (maximizing seats)	0.753	0.752	0.752	0.750	0.618
	(0.042)	(0.041)	(0.042)	(0.041)	(0.054)
Open seat		0.216	0.216	0.100	0.133
		(0.048)	(0.048)	(0.063)	(0.071)
Incumbent		0.165	0.165	−0.101	−0.099
		(0.012)	(0.013)	(0.084)	(0.094)
Ways and Means Committee			−0.006	−0.012	0.010
			(0.018)	(0.017)	(0.039)
Party Leadership			0.077	0.071	0.069
			(0.023)	(0.022)	(0.015)
Committee Chair			−0.029	−0.033	−0.008
			(0.018)	(0.018)	(0.030)
Probability of Victory				0.295	0.240
				(0.085)	(0.132)
Observations	2,175	2,175	2,175	2,175	1,740
Adjusted R-squared	0.363	0.406	0.406	0.416	0.499
District fixed effects?	No	No	No	No	Yes
Panel B. Republicans					
Q share (maximizing seats)	0.824	0.771	0.773	0.773	0.660
	(0.044)	(0.043)	(0.043)	(0.043)	(0.054)
Open seat		0.161	0.160	0.164	0.156
		(0.045)	(0.045)	(0.058)	(0.064)
Incumbent		0.081	0.070	0.077	0.083
		(0.012)	(0.013)	(0.085)	(0.098)
Ways and Means Committee			0.089	0.089	−0.080
			(0.050)	(0.050)	(0.116)
Party Leadership			0.313	0.313	0.131
			(0.067)	(0.067)	(0.143)
Committee Chair			−0.012	−0.012	−0.115
			(0.041)	(0.041)	(0.075)
Probability of Victory				0.008	−0.082
				(0.087)	(0.137)
Observations	2,175	2,175	2,175	2,175	1,740
Adjusted R-squared	0.415	0.428	0.431	0.431	0.510
District fixed effects?	No	No	No	No	Yes

Notes: The dependent variable is the share of yearly spending by all PACs in each district. The unit of analysis is a candidate-district-year. Robust standard errors are in parentheses. *Open seat*, *Incumbent*, *Ways and Means Committee*, *Ways and Means Committee*, and *Party Leadership* are indicator variables equal to 1 if candidates are running in open seats, incumbents, members of the Ways and Means Committee, chairs of a major committee or House party leaders. Party leaders include the speaker of the House, the majority/minority leaders and the majority/minority whip. *Probability of victory* is the probability that a candidate wins the election calculated using the posterior predictive distribution from the hierarchical model.

concerned with establishing a relationship with a candidate should prefer to donate to a candidate who will be in office (see for instance Snyder Jr 1990). The influence related variables have a small impact on the R^2 and are often not statistically different than zero. Being a party leader is the most important influence variable and is positive and statistically

significant in both panel A and panel B.

To help alleviate concerns that the regression estimates might be biased, column 6 adds district fixed effect that control for time invariant district specific characteristics. The number of observations is smaller in this specification because the 2000 election is dropped since district lines changed in 2002. Adding the fixed effects improves the model fit and only has a small impact on the coefficients. Importantly, equilibrium spending remains a highly significant predictor of actual spending.

As a whole, the regression estimates reported in Table 6 show that equilibrium spending explains a large amount of the variation in actual spending and that the relationship between equilibrium and actual spending is robust to district fixed effects. A few other variables such as whether the candidate is an incumbent or the seat is open help explain the data a little bit better (especially for Democrats) but little is gained in terms of fit from adding influence related variables to the regression.

While spending in a district may be largely explained by equilibrium spending, this ignores some important heterogeneity across campaign donors. Organizations whose welfare depends heavily on policy decisions should have the largest incentives to influence policy. Table 6 looks at whether this is the case by analyzing the spending decisions of the financial industry, a group whose campaign goals might differ markedly from other organizations. Campaign contributions from the financial industry are defined as those coming from PACs classified as a member of the Finance, Insurance, and Real Estate industry by the CRP or from individuals employed in the same industry. The regression analysis examines the impact that being a member of the House Committee on Financial Services Committee has on contributions from these firms. The unit of analysis for the regressions is an incumbent candidate since challengers cannot serve on a congressional committee. The dependent variable is the district share of yearly spending by the financial industry on incumbents.

Column 1 shows that equilibrium spending is positive and statistically significant. Column 2 includes an indicator variables equal to 1 if an incumbent is a member of the Financial

Table 6: Spending by the Financial Industry, Incumbent Races 2000 - 2010

Variable	(1)	(2)	(3)
Q share (maximizing seats)	0.553	0.533	0.295
	(0.084)	(0.084)	(0.088)
On Financial Services Committee		0.249	0.201
		(0.048)	(0.082)
Chair/ranking minority member of Finance Committee		0.574	0.127
		(0.212)	(0.198)
Observations	1,992	1,992	1,592
Adjusted R-squared	0.098	0.137	0.327
District fixed effects?	No	No	Yes

Notes: The dependent variable is the district share of yearly spending by the Finance, Insurance, and Real Estate industry on incumbents. Industry spending covers contributions from PACs and individuals (sorted by employer) donating at least $200. The unit of analysis a district-year. Open seat districts are excluded from the analysis. Robust standard errors are in parentheses.

Services Committee and 0 otherwise. Since the chair and ranking minority members of the Financial Services Committee are likely to hold the most influence, I also include a dummy variable indicating such status. Both variables are statistically significant and have large positive coefficients. In column 3, which controls for district fixed effects, equilibrium spending and committee status remain statistically different than zero but being a chair or ranking minority member does not. This should not be seen as an evidence that a leadership position on the committee is unimportant, but rather a consequence of a fixed effect regression without enough variation in an explanatory variable.[57] In summary, the financial industry appears to pursue both election motivated and influence motivated spending strategies. The overall fit of the financial industry model is considerably worse than the fit reported in the previous table, but this is likely a byproduct of focusing on a single industry.

8 Discussion and Summary

This paper uses Stromberg's 2008 probabilistic voting model to quantify the amount that political parties should spend on districts in House elections. The calculations are made using two assumptions about goals: first, parties maximize the expected number of seats and second, parties maximize the probability of winning a majority of seats. The model is

[57] Members of Congress usually hold committee leadership positions for long periods of time.

estimated using two types of Bayesian forecasting models, one of which is able to update forecasts in real time as new polls become available.

The assumptions regarding the parties' goals have important implications for the optimal allocation of resources. Under the first assumption, parties should spend the most on close districts. On the other hand, when the second assumption is true, parties should spend the most on decisive swing districts; that is, districts that are close when they are pivotal in whether a party wins or loses a majority of seats. In both cases, spending should be more concentrated when the forecasts are more precise.

The empirical results support the first assumption but not the second one. For instance, the correlation between district spending and the amount that should have been spent if parties were maximizing the expected number of seats ranges from 0.5 to 0.8 depending on the date, party of the candidate and forecasting method. Conversely, during the 2008 election in which optimal strategies differed greatly because the Democrats were predicted to win a huge majority, observed spending is highly correlated with a spending strategy based on maximizing the expected number of seats but not with with a spending strategy based on maximizing the probability of winning a majority of seats. That said, the results here cannot rule out the possibility that the parties maximize total seats in lopsided elections in order to increase the probability of majority control in the future.

The empirical results also suggest that most political giving is done to affect elections rather than for other reasons like protecting incumbents or influencing candidates. That said, some political actors, like the financial industry, place close to equal weight on influencing elections and gaining access to politicians.

One curious finding that warrants discussion is that the correlation between equilibrium spending under the model and actual spending is nearly identical for political parties, individuals and PACs. There are two competing explanations for this. First, it could suggest that the national parties are able to coordinate campaign efforts with individuals and groups that have similar political goals. This explanation is consistent with Herrnson's 2009 view

of political parties as multilayered coalitions.

However, a second explanation that cannot be ruled out is that candidates have larger incentives to raise funds in tight races. This interpretation is consistent with the game theoretic model in Erikson and Palfrey (2000). Intuitively, this occurs because an additional dollar of spending has a larger impact on the probability that a candidate will win an election when the projected vote share is closest to 1/2. There is some empirical support for this as well. For example, Stein and Bickers (1994) show that the most vulnerable candidates for reelection (as measured by their vote share in the previous election) are the most likely to obtain new grant money for their constituents. Similarly, Bickers and Stein (1996) find that vulnerable incumbents use the grant system to deter quality challengers from opposing them.

The analyses in this paper suggest a number of fruitful areas for future research. First, researchers should try to separate the competing theories of why close elections receive more campaign contributions. Interviews and surveys of party leaders, interest group leaders, individuals, political consultants and political candidates could be especially useful here. It is of course possible, and perhaps likely, that some campaign donors, like political parties, have the party's collective interest in mind, while others, like individuals, are influenced more by the fundraising efforts of candidates. Second, it would be beneficial to separate campaign spending into separate categories. An analysis of TV and radio advertisements might be particularly interesting, although it would require more detailed information about advertising prices. These ideas could help provide a better understanding of the role that parties play in influencing elections.

Appendices

A Stan Code for Hierarchical Model

```
data {
  int<lower=0> N; // number of observations
  int<lower=0> K; // number of predictors
  int<lower=0> T; // number of years
  int<lower=1,upper=T> year[N]; // year for district
  matrix[N,K] x; // district predictors
  vector[N] y; // outcomes
}
parameters {
  vector[T] delta; // intercept by year
  vector[K] beta; // district coefs
  real<lower=0,upper=1> sigma_delta; // year error
  real<lower=0,upper =1> sigma; // district error
}
model {
  beta ~ normal(0, 1);
  delta ~ normal(0, sigma_delta);
  for (n in 1:N)
    y[n] ~ normal(delta[year[n]] + x[n] * beta, sigma);
}
```

B Gibbs Sampler for DLM

B.1 DLM Form

B.1.1 National Model

Let K_t be the numbers of polls published on day t and J be the number of polling firms. Then, equations 22 and 23 can be written in the form,

$$\underset{(K_t\times 1)}{Y_t} = \underset{(K_t\times 1)}{F_t}\underset{(1\times 1)}{\theta_t} + \underset{(K_t\times J)}{A_t}\underset{(J\times 1)}{\Lambda} + \underset{(K_t\times 1)}{v_t}\ ,\ v_t \sim N(0, V_t) \tag{29}$$

$$\underset{(1\times 1)}{\theta_t} = \underset{(1\times 1)}{G}\underset{(1\times 1)}{\theta_{t-1}} + \underset{(1\times 1)}{w_t}\ ,\ w_t \sim N(0, W), \tag{30}$$

where Y_t is a column vector of all polls published at time t, F_t is a column vector of 1's, A_t is a matrix of binary indicators equal to 1 if polling firm j published the k'th poll and 0 otherwise, $G = 1$, $\theta_t = \mu_t$, $\Lambda = [\lambda_1\ \lambda_2 \cdots \lambda_J]^T$, V_t is a $K_t \times K_t$ diagonal matrix with σ_k^2 along the diagonal and $W = \sigma_\mu^2$.

B.1.2 District Model

Let m be the number of districts with polling data and L_t be the numbers of polls published on day t. Then, equations 24 and 25 can be written in the form,

$$\underset{(L_t\times 1)}{Y_t} = \underset{(L_t\times m)}{F_t}\underset{(m\times 1)}{\theta_t} + \underset{(L_t\times 1)}{v_t}\ ,\ v_t \sim N(0, V_t) \tag{31}$$

$$\underset{(m\times 1)}{\theta_t} = \underset{(m\times m)}{G}\underset{(m\times 1)}{\theta_{t-1}} + \underset{(m\times 1)}{w_t}\ ,\ w_t \sim N(0, W), \tag{32}$$

where Y_t is a column vector of all district polls less national opinion at time t, F_t is a matrix of binary indicators equal to 1 if poll l is from district m, $G = I_m$, $\theta_t = [\xi_{1t}\ \xi_{2t} \cdots \xi_{mt}]^T$, V_t is an $L_t \times L_t$ diagonal matrix with σ_l^2 along the diagonal and $W = \sigma_\xi^2$.

B.2 Gibbs Sampler

B.2.1 National Model

Using the DLM form, the joint posterior density of the states and variance parameters is,

$$p(\theta_{1:T}, \sigma_\mu^2, \Lambda | Y_{1:T}, A_t, V_{1:T}) = \prod_{t=1}^{T} p(Y_t|\theta_t, \Lambda, A_t, V_t) \prod_{t=1}^{T} p(\theta_t|\theta_{t-1}, \sigma_\mu^2) \times p(\theta_0)p(\sigma_\mu^2)p(\Lambda). \tag{33}$$

The variance parameter, σ_μ^2, is given an inverse gamma prior, $p(\sigma_\mu^2) \sim \text{IG}(\alpha_\mu, \beta_\mu)$ and the house effects, Λ, are given independent normal priors, $p(\lambda_j) \sim N(0, \sigma_\lambda^2)$. Using these priors and recalling that $\theta_t = \mu_t$ and $G_t = 1$, the Gibbs sampler then proceeds as follows by drawing from the conditional distributions of the unknown parameters.

1. σ_μ^2. The conditional distribution is,

$$p(\sigma_\mu^2|\ldots) \propto p(\sigma_\mu^2) \prod_{t=1}^{T} p(\theta_t|\theta_{t-1}, \sigma_\mu^2) \tag{34}$$

$$\propto (\sigma_\mu^2)^{(\alpha_\mu - 1)} \exp\left[\frac{\beta_\mu}{\sigma_\mu^2}\right] (\sigma_\mu^2)^{T/2} \prod_{t=1}^{T} \exp\left[\frac{(\mu_t - \mu_{t-1})^2}{2\sigma_\mu^2}\right] \tag{35}$$

$$\propto \sigma_\mu^{2(\alpha_\mu + T/2 - 1)} \exp\left\{\frac{1}{\sigma_\mu^2}\left[\beta_\mu + \frac{1}{2}\sum_{t=1}^{T}(\mu_t - \mu_{t-1})^2\right]\right\}, \tag{36}$$

   which is an inverse gamma distribution,

$$p(\sigma_\mu^2|\ldots) \sim \mathcal{IG}\left(\alpha_\mu + T/2, \beta_\mu + \frac{1}{2}\sum_{t=1}^{T}(\mu_t - \mu_{t-1})^2\right). \tag{37}$$

2. λ_j. The conditional distributions for the House effects are normally distributed since both the prior and the likelihood are normal. Letting $\mathcal{P}_j$ refer to the set of polls published by firm j during the campaign and using standard results for conjugate

normal models, λ_j is normal with mean,

$$\left[\sum_{k\in\mathcal{P}_j}\frac{y_k-\mu_{t[k]}}{\sigma_k^2}\right]\left[\sum_{k\in\mathcal{P}_j}\frac{1}{\sigma_k^2}+\frac{1}{\sigma_\lambda^2}\right]^{-1} \tag{38}$$

and variance,

$$\left[\sum_{k\in\mathcal{P}_j}\frac{1}{\sigma_k^2}+\frac{1}{\sigma_\lambda^2}\right]^{-1}. \tag{39}$$

3. $\theta_{1:T}$. The states can be sampled from their conditional distribution,

$$p(\theta_{1:T}|\ldots)=p(\theta_T|Y_{1:T},A_t,\Lambda,\sigma_\mu^2)\prod_{t=0}^{T-1}p(\theta_t|\theta_{t+1},Y_{1:t},A_t,\Lambda,\sigma_\mu^2). \tag{40}$$

Equation 40 suggests that if we could sample θ_T then we could recursively backwards sample through the rest of the states, $\theta_{1:T-1}$. The *forward filtering backwards sampling* (FFBS) is the standard method for generating this sample. The filtering portion of the algorithm is the Kalman filter discussed in the text and works by starting from an initial value θ_0 and then sequentially updating each θ_t as new information becomes available. The backwards sampling part then smooths the values of the states by first sampling from θ_T, and then from the conditional distribution of θ_{T-1}, and so on all the way through all of the remaining states $\theta_{1:T-2}$. To summarize, the FFBS algorithm proceeds in the following manner:

(a) Run Kalman filter.

(b) Draw $\theta_T\sim N(m_T,C_T)$ where $m_T=E(\theta_T|Y_{1:T},A_t,\Lambda,\sigma_\mu^2)$ and $C_T=\mathrm{Var}(\theta_T|Y_{1:T},A_t,\Lambda,\sigma_\mu^2)$.

(c) For $T-1,\ldots,0$ draw $\theta_t\sim N(h_t,H_T)$ where $h_t=E(\theta_t|\theta_{t+1},Y_{1:t},A_t,\Lambda,\sigma_\mu^2)$ and $H_t=\mathrm{Var}(\theta_t|\theta_{t+1},Y_{1:t},A_t,\Lambda,\sigma_\mu^2)$.

In this case, letting $\mathcal{P}_t$ refers to the set of all polls published on day t we have,

$$m_t = \left[\sum_{k\in\mathcal{P}_t} \frac{y_k - \lambda_{j[k]}}{\sigma_k^2} + \frac{M_{t-1}}{C_{t-1}+\sigma_\mu^2}\right] \cdot C_T, \tag{41}$$

$$C_t = \left[\sum_{k\in\mathcal{P}_t} \frac{1}{\sigma_k^2} + \frac{1}{C_{t-1}+\sigma_\mu^2}\right]^{-1}, \tag{42}$$

and,

$$h_t = \left[\frac{\theta_{t+1}}{\sigma_\mu^2} + \frac{m_t}{C_t}\right] H_t, \tag{43}$$

$$H_t = \left[\frac{1}{\sigma_\mu^2} + \frac{1}{C_t}\right]^{-1}. \tag{44}$$

For additional details on the Kalman filter and the smoothing algorithm see a text such as West and Harrison (1997).

B.2.2 District Model

Inference for the district model proceeds in a similar manner to the national model using the DLM form from equations 31 and 32. The posterior density is,

$$p(\theta_{1:T}, \sigma_\xi^2 | Y_{1:t}) = \prod_{t=1}^{T} p(Y_t|\theta_t, V_t) \prod_{t=1}^{T} p(\theta_t|\theta_{t-1}, \sigma_\xi^2) p(\theta_0) p(\sigma_\xi^2). \tag{45}$$

Again using an inverse gamma distribution for the variance parameter, $\sigma_\psi^2 \sim \text{IG}(\alpha_\xi, \beta_\xi)$, the Gibbs sampler is as follows.

1. σ_ξ^2. Draw from the conditional distribution,

$$p(\sigma_\xi^2|\ldots) \propto p(\sigma_\xi^2) \prod_{t=1}^{T} p(\theta_t|\theta_{t-1}, \sigma_\xi^2) \tag{46}$$

$$\propto (\sigma_\xi^2)^{(\alpha_\xi - 1)} \exp\left[\frac{\beta_\xi}{\sigma_\xi^2}\right] (\sigma_\xi^2)^{Tm/2} \prod_{t=1}^{T}\prod_{i=1}^{m} \exp\left[\frac{(\xi_{it} - \xi_{i,t-1})^2}{2\sigma_\xi^2}\right] \tag{47}$$

$$\propto \sigma_\xi^{2(\alpha_\xi + Tm/2 - 1)} \exp\left\{ \frac{1}{\sigma_\xi^2} \left[\beta_\xi + \frac{1}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} (\xi_{it} - \xi_{i,t-1})^2 \right] \right\}, \tag{48}$$

which is an inverse gamma distribution,

$$p(\sigma_\xi^2 | \ldots) \sim \mathcal{IG}\left(\alpha_\xi + Tm/2, \beta_\xi + \frac{1}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} (\xi_{it} - \xi_{i,t-1})^2 \right). \tag{49}$$

2. $\theta_{1:T}$. The states ξ_{it} are drawn using the FFBS algorithm.

C Additional Figures

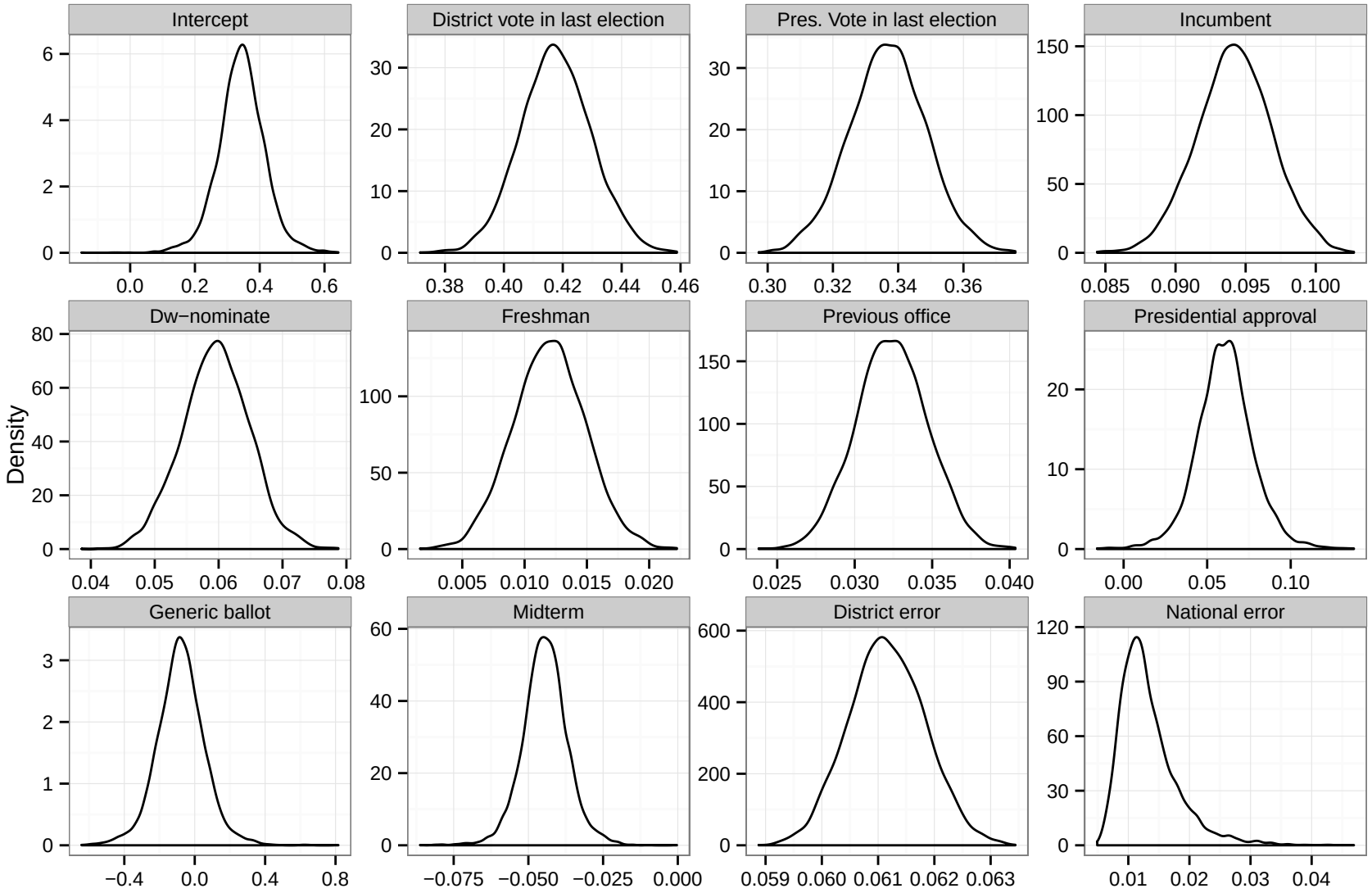

Figure C.1: Posterior Densities for Parameters in Forecasting Model

Notes: Variables are the same as in **??** but names have been shortened for graphical purposes.

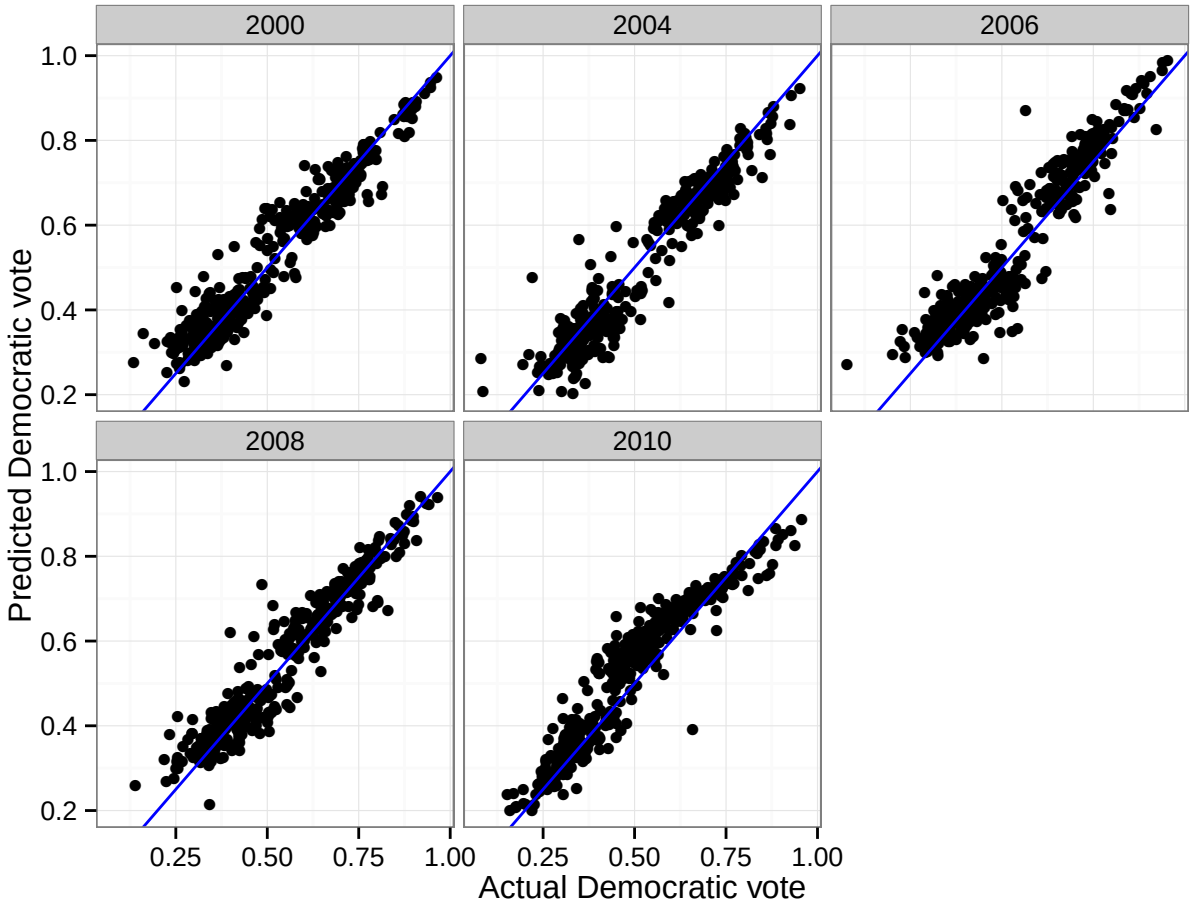

Figure C.2: Out-of-Sample Predicted Versus Actual Plot, House Elections 1990 - 2010

Notes: Predicted value is the mean of the out-of-sample posterior predictive distribution. Model parameters are estimated using data from 1980 up until the year the prediction is being made (i.e. 1980 - 2008 when predicting the 2010 election). The blue line is a 45 degree line.

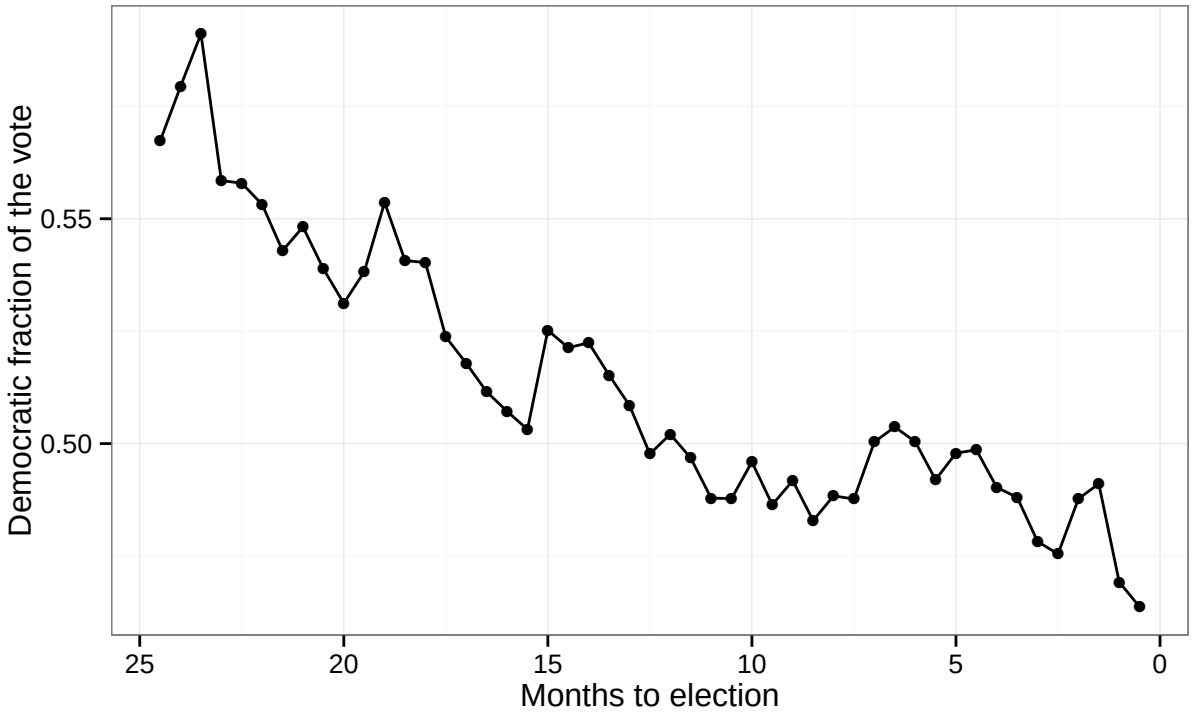

Figure C.3: Estimates of National Opinion During the Election Campaign

Notes: Estimates are the mean of the posterior distribution of the states at each date from a DLM using 2010 generic ballot polling data up until election day.

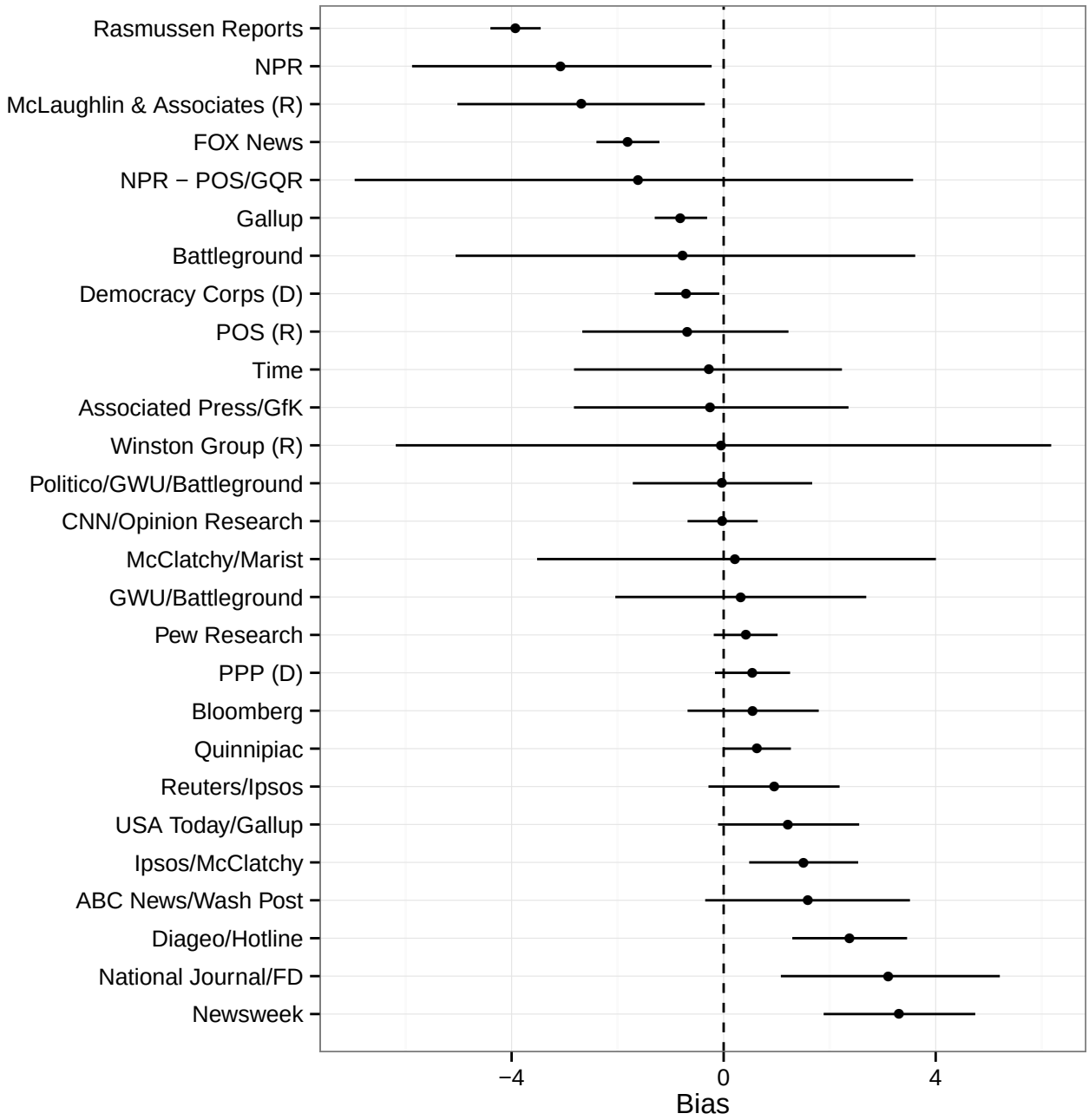

Figure C.4: Estimates of House Effects

Notes: House effects are estimated using the DLM with 2010 generic ballot polling data up until election day. The bias across pollsters is assumed to sum to zero for identification purposes. The reported biases are from the late-October model just prior to election day.

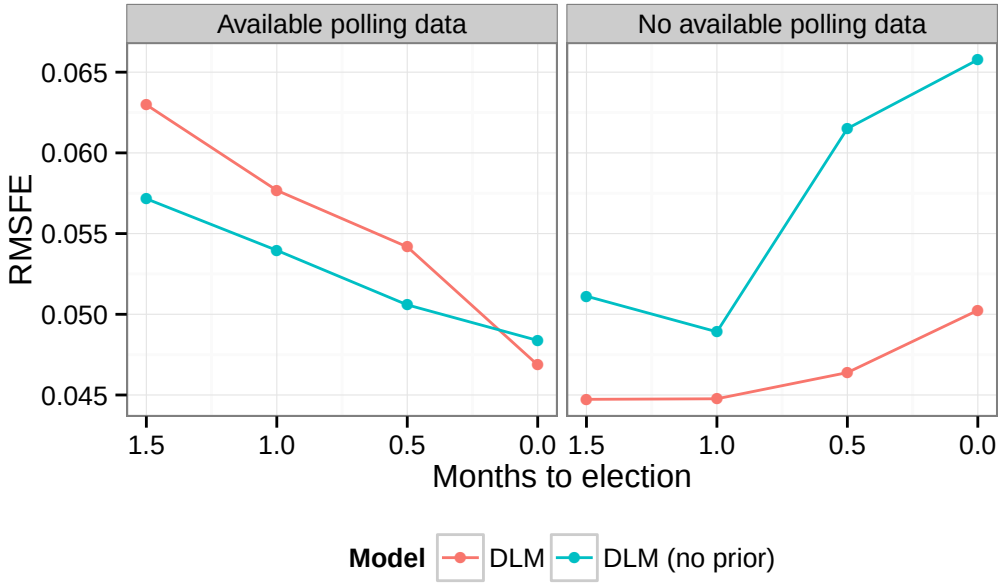

Figure C.5: RMSFE By Polling Availability and Forecast Date

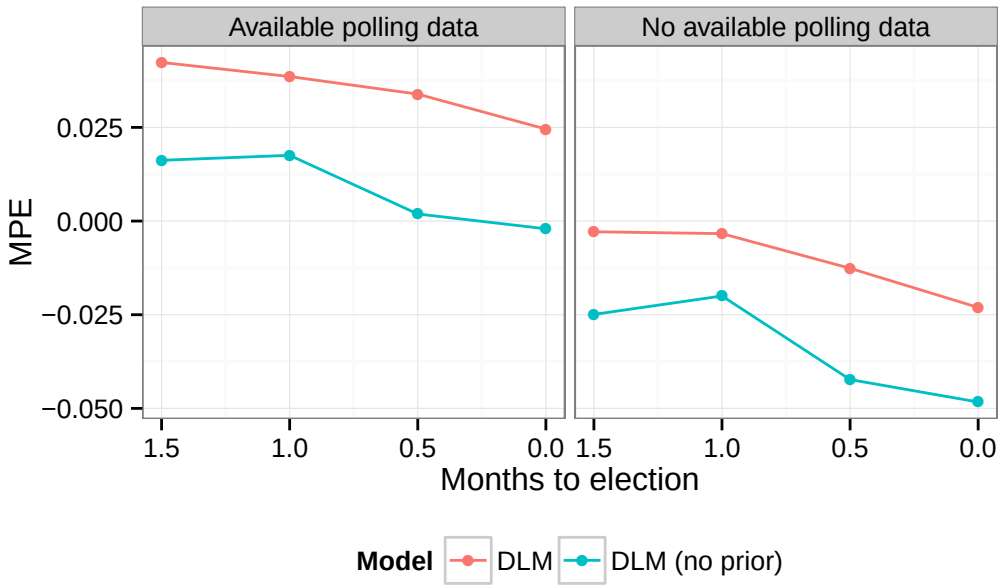

Figure C.6: MPE By Polling Availability and Forecast Date

Notes: The MPE is the mean of the forecasted Democratic share of the two-party vote minus the Democratic share of the two-party vote in the actual election. Negative value reflect a bias toward the Republican party.

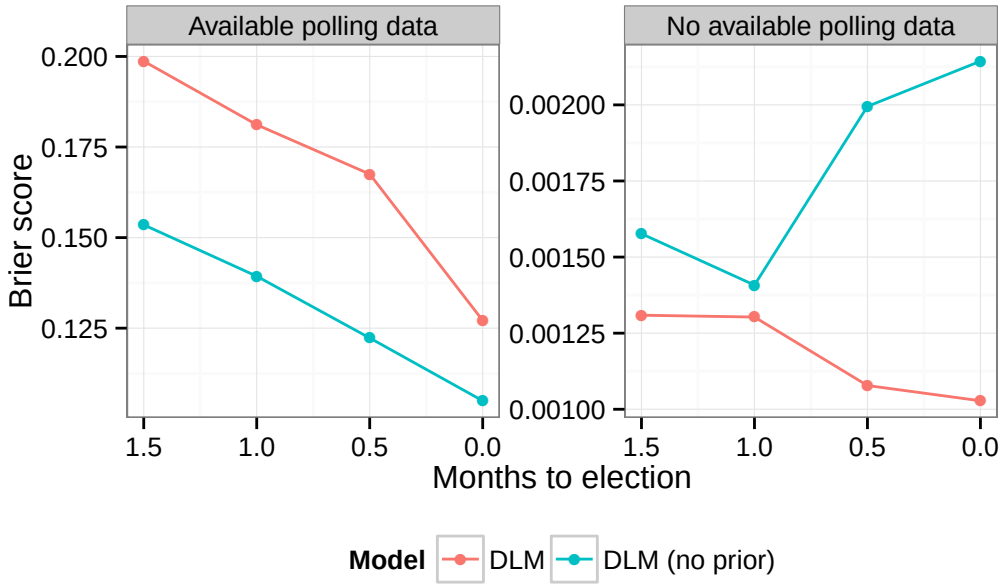

Figure C.7: Brier Score By Polling Availability and Forecast Date

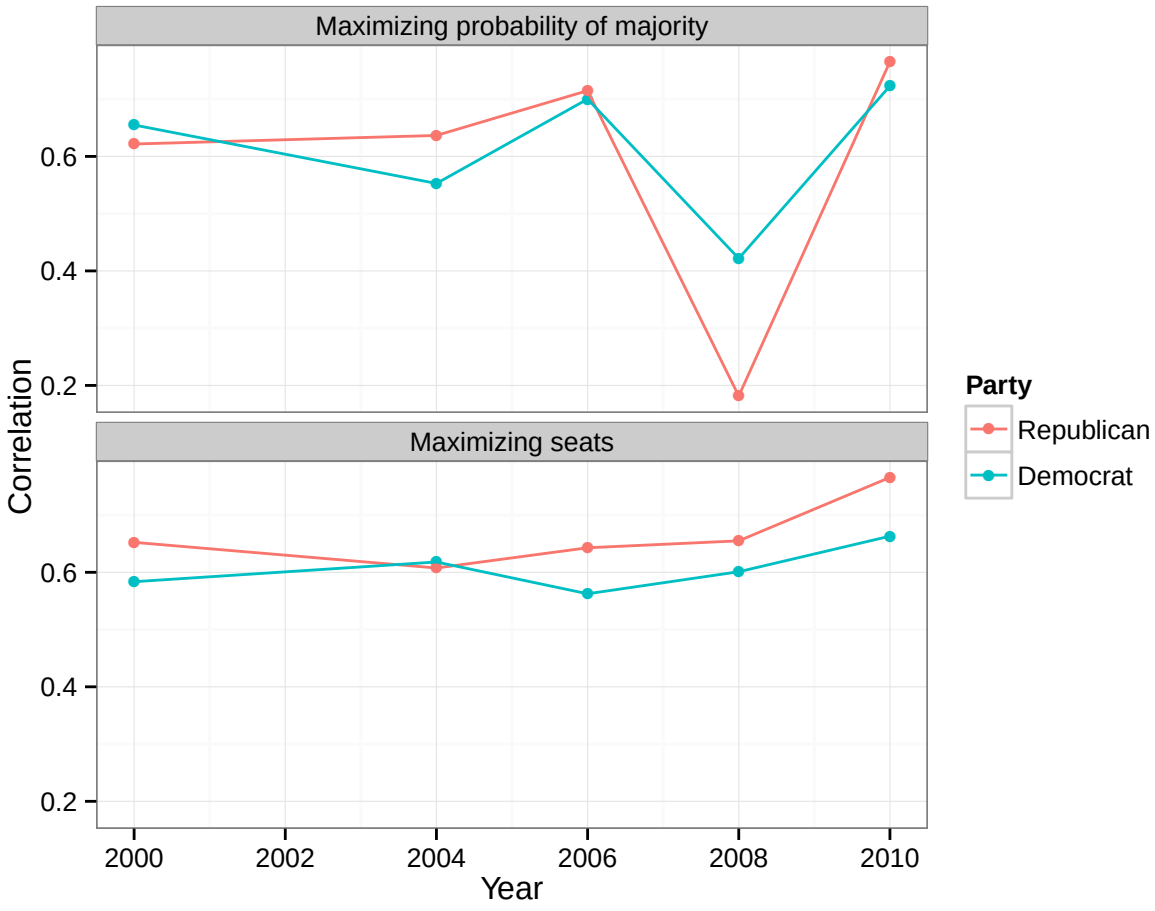

Figure C.8: Correlation Between Q and PAC Spending, by Party

Notes: PAC spending for a party consists of a direct contribution by a PAC to that party or an indirect contribution by a PAC against the opposing party. The top panel shows the correlation between Q_i^{maj} and spending while the bottom panel shows the correlation between Q_i^{seats} and spending.

D Data Appendix

D.1 Gary Jacobson Data

Data were generously provided by Gary Jacobson.

D.2 DW-Nominate Score

The data is available for download at http://www.voteview.com/dwnominate.asp.

D.3 National Variables

Second quarter GDP: The percent change in U.S. GDP from one quarter to the next in chained 2009 dollars was considered but not used in the final hierarchical model. It is available for download from the Bureau of Economic Analysis at http://www.bea.gov/national/index.htm.

Presidential net approval rating: Approval ratings were scraped from the president specific webpages with links available at http://www.ropercenter.uconn.edu/CFIDE/roper/presidential/webroot/presidential_rating.cfm.

Generic Congressional Ballot: Gallup polls were searched for and downloaded using the iPOLL Databank at(http://www.ropercenter.uconn.edu/data_access/ipoll/ipoll.html. There were two main survey questions of interest that were typically worded as "If the elections for Congress were being held today, which party's candidate would you vote for in your Congressional district" and "As of today, do you lean more toward". I consider a respondent to favor a particular party if they say they would vote for that party or they lean more toward that party. Although a sample of likely voters is preferred, Gallup does not include the data used to estimate a likely voter sample, so I use a registered voter sample instead.

D.4 Spending Data

Data on campaign spending are from the campaign finance data provided by the Center for Responsive Politics (CRP) (https://www.opensecrets.org/). CRP obtains the campaign

finance data from the Federal Election Commission and adds value to it by cleaning and categorizing the data. Spending to specific candidates is from the "PAC" table and the "indivs" table which contain itemized spending data for PACs and individuals contributing over $200 respectively. Information on the districts of specific candidates is linked to these tables with the "Cands" table, which provides information on each candidate.

D.5 House Polls

Polls were scraped from the website http://www.realclearpolitics.com/epolls/other/2010_generic_congressional_vote-2171.html and the 2010 New York times House election forecasts (http://elections.nytimes.com/2010/forecasts/house.

References

Ansolabehere, S., J. Snyder and J. de Figueiredo. 2003. "Why Is There So Little Money In US Elections?".

Ansolabehere, Stephen and James M Snyder Jr. 2002. "The incumbency advantage in US elections: An analysis of state and federal offices, 1942-2000." *Election Law Journal* 1(3):315–338.

Aranson, Peter H, Melvin J Hinich and Peter C Ordeshook. 1974. "Election goals and strategies: Equivalent and nonequivalent candidate objectives." *American Political Science Review* 68(01):135–152.

Bickers, Kenneth N and Robert M Stein. 1996. "The electoral dynamics of the federal pork barrel." *American Journal of Political Science* pp. 1300–1326.

Brams, Steven J and Morton D Davis. 1973. "Resource-Allocation Models in Presidential Campaigns: Implications for Democratic Representation." *Annals of the New York Academy of Sciences* 219(1):105–123.

Brams, Steven J and Morton D Davis. 1974. "The 3/2's rule in presidential campaigning." *American Political Science Review* 68(01):113–134.

Campbell, James E. 1992. "Forecasting the presidential vote in the states." *American Journal of Political Science* pp. 386–407.

Colantoni, Claude S, Terrence J Levesque and Peter C Ordeshook. 1975. "Campaign resource allocations under the Electoral College." *American Political Science Review* 69(01):141–154.

Dixit, Avinash and John Londregan. 1996. "The determinants of success of special interests in redistributive politics." *Journal of politics* 58:1132–1155.

Erikson, Robert S. 1988. "The puzzle of midterm loss." *The Journal of Politics* 50(04):1011–1029.

Erikson, Robert S and Thomas R Palfrey. 2000. "Equilibria in campaign spending games: Theory and data." *American Political Science Review* 94(03):595–609.

Gelman, Andrew and Gary King. 1990. "Estimating incumbency advantage without bias." *American Journal of Political Science* pp. 1142–1164.

Gelman, Andrew and Gary King. 1993. "Why are American presidential election campaign polls so variable when votes are so predictable?" *British Journal of Political Science* 23(04):409–451.

Gelman, Andrew and Gary King. 1994. "A unified method of evaluating electoral systems and redistricting plans." *American Journal of Political Science* pp. 514–554.

Gelman, Andrew and Zaiying Huang. 2008. "Estimating incumbency advantage and its variation, as an example of a before–after study." *Journal of the American Statistical Association* 103(482):437–446.

Gerber, Alan. 1998. "Estimating the effect of campaign spending on senate election outcomes using instrumental variables." *American Political Science Review* 92(02):401–411.

Green, Donald Philip and Jonathan S Krasno. 1988. "Salvation for the spendthrift incumbent: Reestimating the effects of campaign spending in House elections." *American Journal of Political Science* pp. 884–907.

Grier, Kevin B and Michael C Munger. 1991. "Committee assignments, constituent preferences, and campaign contributions." *Economic Inquiry* 29(1):24–43.

Herrnson, Paul S. 2009. "The roles of party organizations, party-connected committees, and party allies in elections." *The Journal of Politics* 71(04):1207–1224.

Jackman, Simon. 2005. "Pooling the polls over an election campaign." *Australian Journal of Political Science* 40(4):499–517.

Jackman, Simon. 2009. *Bayesian analysis for the social sciences*. Vol. 846 John Wiley & Sons.

Jackman, Simon. 2012. "Pollster Predictions: 91.4% Chance Obama Wins, 303 or 332 EVs." http://www.huffingtonpost.com/simon-jackman/pollster-predictions_b_2081013.html.

Jacobson, Gary C. 1978. "The effects of campaign spending in congressional elections." *American Political Science Review* 72(02):469–491.

Jacobson, Gary C. 1980. *Money in congressional elections*. Yale University Press.

Jacobson, Gary C. 1985*a*. "Money and votes reconsidered: Congressional elections, 1972–1982." *Public choice* 47(1):7–62.

Jacobson, Gary C. 1985*b*. "Party organization and distribution of campaign resources: Republicans and Democrats in 1982." *Political Science Quarterly* pp. 603–625.

Johansson, Eva. 2003. "Intergovernmental grants as a tactical instrument: empirical evidence from Swedish municipalities." *Journal of Public Economics* 87(5):883–915.

Kastellec, Jonathan P, Andrew Gelman and Jamie P Chandler. 2008. "Predicting and dissecting the seats-votes curve in the 2006 US House election." *PS: Political Science & Politics* 41(01):139–145.

Kau, James B, Donald Keenan and Paul H Rubin. 1982. "A general equilibrium model of congressional voting." *The Quarterly Journal of Economics* pp. 271–293.

King, Gary and Andrew Gelman. 1991. "Systemic consequences of incumbency advantage in US House elections." *American Journal of Political Science* pp. 110–138.

Kroszner, Randall S and Thomas Stratmann. 1998. "Interest-group competition and the organization of congress: theory and evidence from financial services' political action committees." *American Economic Review* pp. 1163–1187.

Lee, David S. 2001. The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the US.. Technical report National bureau of economic research.

Levitt, Steven D. 1994. "Using repeat challengers to estimate the effect of campaign spending on election outcomes in the US House." *Journal of Political Economy* pp. 777–798.

Lindbeck, Assar and Jörgen W Weibull. 1987. "Balanced-budget redistribution as the outcome of political competition." *Public choice* 52(3):273–297.

Linzer, Drew A. 2013. "Dynamic Bayesian Forecasting of Presidential Elections in the States." *Journal of the American Statistical Association* 108(501):124–134.

Lock, Kari and Andrew Gelman. 2010. "Bayesian combination of state polls and election forecasts." *Political Analysis* 18(3):337–348.

Poole, Keith T and Howard L Rosenthal. 2011. *Ideology and congress*. Vol. 1 Transaction Publishers.

Poole, Keith T and Howard Rosenthal. 1997. *Congress: A political-economic history of roll call voting*. Oxford University Press.

Poole, Keith T and Thomas Romer. 1985. "Patterns of political action committee contributions to the 1980 campaigns for the United States House of Representatives." *Public Choice* 47(1):63–111.

Romer, Thomas and James M Snyder Jr. 1994. "An empirical investigation of the dynamics of PAC contributions." *American journal of political science* pp. 745–769.

Snyder, James M. 1989. "Election goals and the allocation of campaign resources." *Econometrica: Journal of the Econometric Society* pp. 637–660.

Snyder Jr, James M. 1990. "Campaign contributions as investments: The US House of Representatives, 1980-1986." *Journal of Political Economy* pp. 1195–1227.

Stein, Robert M and Kenneth N Bickers. 1994. "Congressional elections and the pork barrel." *The Journal of Politics* 56(02):377–399.

Stewart III, Charles and Jonathan Woon. 2015. "Congressional Committee Assignments, 103rd to 112th Congresses, 1993–2011.".

Stratmann, Thomas. 1991. "What do campaign contributions buy? Deciphering causal effects of money and votes." *Southern Economic Journal* pp. 606–620.

Stratmann, Thomas. 2005. "Some Talk: Money in Politics. A (Partial) Review of the Literature." *Public Choice* pp. 135–156.

Strauss, Aaron. 2007. Florida or Ohio? Forecasting presidential state outcomes using reverse random walks. In *Princeton University Political Methodology Seminar*.

Strömberg, David. 2008. "How the Electoral College influences campaigns and policy: the probability of being Florida." *The American Economic Review* 98(3):769–807.

Tufte, Edward R. 1973. "The relationship between seats and votes in two-party systems." *American Political Science Review* 67(02):540–554.

West, Mike and Jeff Harrison. 1997. *Bayesian forecasting and dynamic models*. Vol. 18 Springer New York.