# *Candidate Positioning in U.S. Senate Elections: An Empirical Investigation of the Revised Spatial Model*[*]

Michael Peress
Department of Political Science,
University of Rochester
mperess@mail.rochester.edu

July 15, 2009

**Abstract**

Existing empirical investigations of the spatial model of two candidate elections have focused on recovering the utility functions of voters, and testing the validity of the spatial model's characterization of voter behavior. The spatial model also makes strong predictions about candidate positioning. Empirical work testing the implications of the spatial model's predictions for candidate positioning is more limited because simple models are too easy to reject and more realistic models are hard to solve. We develop a 'revised' spatial model that incorporates non-policy factors, uncertainty, and policy-motivated candidates. We develop a structural approach to estimate both voter-specific and candidate-specific parameters using the Simulated Method of Moments. We use this approach to test the validity of the revised spatial model as an account of candidate positioning in U.S. Senate elections. . We apply our framework to decompose the sources of the electoral success in Senate elections between 1988 and 1992. Our results suggest that while the Republican Party recruited more high quality candidates, the Democratic Party's candidates compensated for their disadvantage by taking more moderate positions.

[*] This paper benefited from comments and suggestions from James Adams, Barry Burden, Ken Kollman, Keith Poole, Ming Li, and participants of seminars at UC San Diego, the University of Rochester, the Midwest Political Science Association conference (Chicago, 2008), the Public Choice Society meetings (Las Vegas, 2009), and the American Political Science Association conference (Boston, 2009). I would like to thank Jay Goodliffe for providing some of the data used in this paper.

# 1 - Introduction

Since its' introduction by Downs (1957), the spatial model of politics has proved a valuable tool for analyzing two candidate elections. The Downsian model predicts that, in two candidate elections, competition forces both candidates to adopt the position of the median voter. Shortcomings of the Downsian model as a descriptive theory of candidate positioning were noticed soon after its' introduction. The most obvious shortcoming is the prediction that the candidates take identical positions. This prediction is easily refuted by casual empiricism- Democratic candidates consistently take positions to the left of their Republican counterparts (Ansolabehere, Snyder, and Stewart, 2001; Burden, 2004). This stylized fact – divergence – has motivated theorists to update the spatial model to eliminate this embarrassing prediction, by replacing various Downsian assumptions with more realistic ones.

Theorists have been largely successful in this regard- there is no shortage of spatial models that predict divergence. Yet, while the theoretical literature has shown that the spatial model can generate divergence (a weak test of the model), it has not shown that the spatial model can account for the degree of divergence observed in the data. Moreover, there are a host of other stylized facts that the spatial model can potentially account for, and should be able to account for to be deemed a good theory of candidate positioning.

The spatial model's predictions for voter behavior have been subjected to extensive empirical testing (Rabinowitz and McDonald, 1989; Erikson and Romero, 1990; Merrill and Grofman, 1999; Grynaviski and Corrigan, 2006). And there has surely been some work testing the spatial model's predictions for candidate positioning (Ansolabehere, Snyder, and Stewart, 2001; Burden, 2004; Adams, Merrill, and Grofman, 2005). Certain complications arise in implementing tests of the spatial model. Simple variants of the spatial model are too easy to reject. More realistic versions of the spatial model typically yield ambiguous comparative statics so that generating clear predictions requires choosing values for many free parameters.

Our goals in this paper are threefold. First, we will present a more general version of the spatial model, and develop an approach for computing the equilibrium and estimating the parameters of this model. We will consider a ‘revised’ spatial model that includes non-policy factors, policy-motivated candidates, and uncertainty. While not an exhaustive list of features that spatial modelers have considered important, they represents the least controversial set of features one can introduce into the framework, while still giving the spatial model a fair chance of accounting for patterns in candidate positioning. We will then develop a strategy for computing the equilibrium of this model numerically, and develop an approach for estimating the free parameters of this model. Unlike previous work which has focused on recovering the parameters of the voters’ utility functions from data, our approach will be able to recover the parameters of the voters’ and candidates’ utility functions.

Second, we will provide a comprehensive test of the spatial model’s ability to account for the major patterns of candidate positioning. This is a test which the Downsian model would clearly fail, but ‘modern’ spatial models may fair better. We will demonstrate that endogeneity and simultaneity problems are endemic to the study of candidate positioning and discuss how our approach can deal with these problems. For example, scholars of legislative politics would like to determine whether competitiveness leads to better representation (which has occasionally been hypothesized to be a consequence of spatial competition). Our results show that while competitiveness and representation are negatively correlated in the data, the spatial model is able to fit this pattern.

Finally, we will use our framework to decompose the sources of electoral success in U.S. Senate elections between 1988 and 1992. We find that the political climate favored Republicans in 1990 more so than in 1988 and 1992. As a result, the Republican Party was able to recruit quality challengers, reinforcing their advantage. Overall, the Republican Party was advantaged in terms of non-policy factors in the 1990 elections. Nonetheless, the Republicans won half of the Senate races and garnered less than half of the votes because the Democratic candidates were more effectively able to moderate their policy positions. The Democratic candidates were able to moderate because they had more moderate personal ideal points.

## 2 – Review of the Spatial Model

The spatial modeling literature has been largely motivated by the desire to explain divergence. To this end, spatial modelers have documented various sources of divergence. These include (but are not limited to) policy-motivation and uncertainty (Whitman, 1983; Calvert, 1985; Groseclose, 2001), valence and uncertainty (Schofield, 2003, 2004; Adams, Merrill, and Grofman, 2005; Peress, forthcoming), abstention due to alienation (Adams, Merrill, and Grofman, 2005; Glasear, Ponzetto, and Shapiro, 2005), primaries and uncertainty (Serra, 2008), and fundraising and persuasive advertising (Moon, 2004).[1]

While many explanations of divergence exist, it is unclear if any of these can explain the *degree* of divergence observed in the data. This issue has largely been ignored, but a few studies suggest that the existing theories cannot account for the degree of divergence observed in the data. For example, Adams, Merrill, and Grofman (2005) show that a spatial model with abstention due to alienation and indifference predicts divergence that is on the order of 25% of the divergence observed in the American National Election Studies data, under reasonable assumptions about the electoral environment (pg. 135-139).[2] They find that incorporating uncertainty and policy-motivation can account for about 50% of the divergence observed in the data (pg. 217).

[1] For survey articles listing additional sources of divergence in the spatial model, see Fiorina (1999) and Grofman (2004).

[2] See Peress (2007) for a similar conclusion about abstention due to alienation and indifference.

Besides the ability of the spatial model to account for the degree of divergence, there are a host of other stylized facts that the spatial model can potentially explain. For example, Miller and Stokes (1963) find that legislators respond to their constituents, but do so imperfectly. While it has been argued that the spatial model predicts that competitive elections lead to better representation, the evidence suggests that the effect is weak at best (MacRae, 1952; Fiorina, 1974; Kuklinski, 1977; Bartels, 1991; Gulati, 2004; Griffin, 2006).[3] Downs' spatial model precludes polarization among elected officials in the absence of polarization across electoral districts. We nonetheless observe that elected officials are polarized (Poole and Rosenthal, 1991, 1997) while public opinion is not (Fiorina, 2005).[4]

In this paper, we will test the 'revised' spatial model's ability to account for all of these patterns. In order to give the spatial model a fair chance of passing this test, we will incorporate a number of features popular in modern spatial models. In particular, the 'revised' spatial model will include non-policy factors, policy-motivated candidates, and uncertainty. Each of these features is necessary to explain patterns of voting behavior and candidate positioning.

The incorporation of non-policy factors follows a long line of papers that have empirically investigated the spatial model as an account of voter decision making.[5] Non-policy factors allow us to relax the assumption of pure ideological voting assumed in the Downsian model. They are incorporated in this empirical literature because the data clearly indicate that voters do not always choose the candidate that they are ideologically closer to.

---

[3] Later in the paper, we show that argue that this logic is misleading- spatial competition does not predict that there is a positive correlation between competitiveness and representation.

[4] We qualify the second statement with a particular definition of polarization. We say that public opinion is polarized if it exhibits a bimodal distribution. An alternative definition used in the literature is that public opinion is polarized (by party) if Democratic and Republican identifiers take distinct positions. Under the second definition, public opinion is polarized and this polarization has increased over time (Levendusky, 2005; Jacobson, 2006; Layman and Carsey, 2006). These two different definitions have unfortunately caused quite a bit of confusion in the literature.

[5] See Erikson and Romero (1990), Schofield et al. (1998), Martin et al. (1999), Adams and Merrill (2003), Adams, Merrill, and Grofman (2005), and Peress (2007).

Second, we incorporate uncertainty into the framework. The uncertainty takes the form of valence uncertainty (Adams, Merrill, and Grofman, 2005).[6] This is an alternative to uncertainty about the position of the median voter (Calvert, 1985; Groseclose, 2001). We chose to incorporate valence uncertainty because it is plausible that there is a large degree of such uncertainty. Uncertainty about the position of the median voter could presumably be resolved early on in the election cycle because the median voter moves slowly over time. Voter's overall perceptions of the candidates, however, may change rather quickly, in response to a television advertisement or a scandal. Valence uncertainty can account for shocks that occur well into the election season, when the candidates have already committed to policy positions. In addition to providing a possible source of divergence, incorporating uncertainty is necessary to explain the observation that a relatively small number of races end in near ties.[7]

Third, we incorporate policy-motivated candidates. While a handful of models explain divergence with purely office-motivated candidates (Moon, 2004; Adams, Merrill, and Grofman, 2005), allowing for policy-motivated candidates will greatly increase our ability to explain the degree of divergence observed in the data.[8] By introducing policy-motivated candidates, we now have two features that allow for divergence. Non-policy factors are a generalization of valence advantages and hence provide one source of divergence. Uncertainty provides the second source of divergence in the framework.

We argue that a spatial model with these three features is able to explain the most important aspects of the data. There are two issues we must deal with however. As we move from simple spatial models to more complicated ones, analytical solutions are no longer available. Second, more general models typically have more free parameters. For example, Adams, Merrill, and Grofman (2005) mention the difficulty of choosing such parameters as the ideal points of the

[6] A more appropriate term may be "non-policy uncertainty" since the term "valence" was originally used to connote the voters' evaluations of the candidates attractiveness, but the broader use of valence has become standard in the literature.
[7] See Peress (forthcoming) for a formal justification of this claim.
[8] Adams, Merrill, and Grofman (2005) already provide some evidence that this is the case.

candidates and the variance of the electoral 'shock'. Even the sign of the comparative statics generated by this framework will depend on these free parameters. In our case, the predictions generated by the model will depend on the ideal points of the candidates, the degree of uncertainty, the location of the median voter, the valences of the candidates, etc.

We deal with both of these issues by employing a structural estimation approach. Previous work has employed a structural approach to obtain the parameters of the voter's utility function and the distribution of voter characteristics (Erikson and Romero, 1990; Schofield et al., 1998; Martin et al., 1999; Adams and Merrill, 2003; Adams, Merrill, and Grofman, 2005; Peress, 2007). In this paper, we extend the structural approach to obtain the parameters of the candidates' utility functions and the distribution of candidate characteristics. This step is more involved because it requires repeated numerical computations of the equilibrium.[9] We develop an effective automatic algorithm for computing equilibria in models of spatial competition. We employ the Simulated Method of Moments approach to recover the candidate parameters. Our approach makes use of a recent advancement in the structural estimation literature which drastically reduces the computational burden of the nested fixed point algorithm by employing importance sampling (Ackerberg, 2007).

## 3 - The Model

We assume that there are two candidates competing for office- candidate D and candidate R (the Democratic and Republican candidates). Policy is characterized by the one-dimensional spatial model.[10] Denote the positions of candidate D and candidate R by $y_D \in \mathbb{R}$ and $y_R \in \mathbb{R}$. Voters are completely characterized by their ideal points $v \in \mathbb{R}$ and by their non-policy factors

[9] For applications of the nested fixed point approach to political economy, see Merlo (1997) and Mebane (2000).
[10] While the theory easily extends to the multi-dimensional spatial model, the data we seek to explain is one-dimensional.

$(z_D, z_R) \in \mathbb{R}^2$. There is a continuum of voters represented by the density function $f(v, z_D, z_R)$.[11] The utility a voter characterized by $(v, z_D, z_R)$ receives from voting for a candidate is represented by $u(y_k - v, z_k) = z_k - \rho \mid y_k - v \mid$ where $\rho > 0$.

For expositional convenience, we let $z = z_D - z_R$ and replace the distribution $f(v, z_D, z_R)$ with the distribution $f(v, z)$ since only the difference $z_D - z_R$ matters in terms of partitioning the set of voters among the candidates. Candidates do not know the true distribution of non-policy factors. We assume that $f(v, z; \lambda) = g(v, z - \lambda)$. The candidates know $g$, but they are uncertain about the value of $\lambda$ at the time they take their positions. This uncertainty is fully resolved when the election takes place.

The random variable $\lambda$ can be thought of as a valence 'shock', which shifts the distribution of non-policy factors. A positive shock makes all voters more favorable to the Democratic candidate and a negative shock makes all voters more favorable to the Republican candidate. Candidates may be uncertain about a number of things including whether their characteristics will be viewed positively by the voters, whether their youthful indiscretions will be discovered and how they will play in the media, whether there will be a terrorist attack in the month before the election, etc.

We denote the candidates' beliefs over $\lambda$ using the cumulative distribution function $F_\lambda$, which we assume represents a continuous random variable with full support. The candidates' vote shares are functions of their positions and the realization of the valence shock,

$$s_D(y_D, y_R; \lambda) = \int_{v=-\infty}^{\infty} \int_{z=\rho|y_R - v| - \rho|y_D - v|}^{\infty} g(v, z - \lambda) dz dv$$

$$s_R(y_D, y_R; \lambda) = \int_{v=-\infty}^{\infty} \int_{z=-\infty}^{\rho|y_R - v| - \rho|y_D - v|} g(v, z - \lambda) dz dv$$

[11] We assume that $f(v, z_D, z_R)$ represents the subpopulation of voters rather than the entire population. The restriction of full turnout is without loss of generality so long as the voters' turnout decisions do not depend on the positions that the candidates take.

Define $\lambda^*(y_D, y_R)$ by,

$$s_D(y_D, y_R; \lambda^*(y_D, y_R)) = s_R(y_D, y_R; \lambda^*(y_D, y_R)) = \tfrac{1}{2}$$

Since $s_D(y_D, y_R; \lambda)$ is strictly increasing in $\lambda$, $\lambda^*(y_D, y_R)$ is uniquely defined and we can characterize the probabilities that each candidate wins the election by,

$$\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2}) = 1 - F_\lambda(\lambda^*(y_D, y_R))$$

$$\Pr(s_R(y_D, y_R; \lambda) \geq \tfrac{1}{2}) = F_\lambda(\lambda^*(y_D, y_R))$$

We assume that candidates have mixed motivations- they care about policy outcomes as well as holding office. The benefit they receive from winning office is $\alpha_D$ and $\alpha_R$ and the candidates have ideal points $q_D$ and $q_R$. This gives rise to the following utility functions,[12]

$$V_D(y_D, y_R) = \Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})(\alpha_D - (y_D - q_D)^2) + \Pr(s_D(y_D, y_R; \lambda) < \tfrac{1}{2})(-(y_R - q_D)^2)$$

$$V_R(y_D, y_R) = \Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})(-(y_D - q_R)^2) + \Pr(s_D(y_D, y_R; \lambda) < \tfrac{1}{2})(\alpha_R - (y_R - q_R)^2)$$

A pure-strategy equilibrium is defined as a point $(y_D^*, y_R^*) \in \mathbb{R}^2$ such that $y_D^* \in \arg\max_{y_D \in \mathbb{R}} V_D(y_D, y_R^*)$ and $y_R^* \in \arg\max_{y_R \in \mathbb{R}} V_R(y_D^*, y_R)$. As is often the case in spatial models with policy-motivated candidates, our framework will not admit an analytical characterization of the equilibrium.[13] We rely on numerical methods to compute the equilibrium in this model. Although the sufficient conditions for the existence and uniqueness of pure-strategy equilibrium are not available in our framework, we have found through numerical simulations on randomly generated parameters that a unique equilibrium typically exists.[14] We detail our approach to computing the equilibrium in Appendix I.

---

[12] The restriction of the candidates' policy weight to one is without loss of generality.

[13] See Peress (forthcoming) for characterizations of the equilibrium when uncertainty is not present.

[14] Specifically, we searched for multiple equilibria using randomly generated starting values. It is possible that there exist equilibria that are not reachable using best-response iterations, but these equilibria are unlikely to be learned by Senate candidates. We also note that Peress (forthcoming) provides assumptions under which a unique equilibrium exists in the case where no uncertainty is

# 4 – Estimation Procedure

The key methodological innovation of this paper is to provide a technique that can recover the parameters of both the voter and candidate utility functions. Applications studying voters are common in the literature. Estimating of the candidates' utility functions is more involved because it requires solving for the equilibrium of the model repeatedly. In order to reduce the computational complexity of the estimation procedure, we divide the estimation into a number of stages. Only in the last stage will we have to numerically solve for the equilibrium.

## 4.1 – First Stage Estimation

In the first stage, we will estimate the distribution of voter characteristics, $f(v,z)$, and the policy weight $\rho$. Let $j*(n)$ denote the Senate race that respondent $n$ voted in. We start with the equation,

$$\kappa_n* = -\rho(y_{D,j*(n)} - v_n)^2 + \rho(y_{R,j*(n)} - v_n)^2 + \gamma' X_n + \lambda_{j*(n)} + \varepsilon_n$$

Here, $X_n$ is a vector of voter specific variables and $\lambda_{j*(n)}$ are district specific fixed-effects. The voter is assumed to vote for the Democratic candidate when $\kappa_n* \geq 0$. The voter's non-policy factor is given by $z_n = \gamma' X_n + \varepsilon_n$. Hence, we can estimate $f(v,z)$ using the empirical distribution of $(v_n, z_n)$ where $\varepsilon_n$ is a random draw from a standard normal distribution. Then, we can estimate the mean voter's position using,

---

present. Although we don't formalize the argument here, it is likely that adding uncertainty to this model will preserve uniqueness.

$$\hat{\tau}_j = \frac{\sum_{j=1}^{J} 1\{j^*(n) = j\} v_n}{\sum_{j=1}^{J} 1\{j^*(n) = j\}}$$

where $J$ denotes the number of races.

### 4.2 – Second Stage Estimation

The first stage yields estimates of $\lambda_j$, the realization of the valence shock in the district. In order to introduce uncertainty in the model, we assume that candidates do not know the value of this variable at the time they are choosing their positions. The valence shock is subject to two types of uncertainty. First, there is a race-level unobservable component which accounts for the fact that candidates cannot be sure how the voters will react to their non-policy characteristics. Second, there is an election-level unobservable component which affects the relative strength of the parties in each election cycle.

Candidates can predict the race-level component using incumbency, the quality of the challenger (Lublin, 1994), and the region. This set of variables is motivated by the literature on the incumbency advantage. We note that these variables are known to the candidates well before Election Day, and hence the candidates can use these variables to predict their likely vote shares at the time that they are select their positions. Candidates can predict the election-level component using the party of the sitting president, whether the election is a midterm election, the net approval rating of the president, and macroeconomic variables. This set of variables is motivated by the literature on predicting presidential elections. We treat both sets of variables as observable to both the candidates and the econometrician.

The obtain a measure of the candidates' beliefs, we use the auxiliary regression,

$$\lambda_j = \eta_A ' A_{t^*(j)} + \eta_C ' C_j + \sigma_\xi \xi_{t^*(j)} + \sigma_\omega \omega_j$$

Here, $t^*(j)$ is the year corresponding to race $j$, $A_t$ are election-level predictors of the valence shock, $C_j$ are race-level predictors of the valence shock, $\xi_t$ is an election-level shock (common to all races in a particular election year), and $\omega_j$ is a race-level shock. We assume that $\xi_t$ and $\omega_j$ have the standard normal distribution and that,

$$Cov(\xi_t,\omega_j)=\begin{cases}\rho_{\xi\omega}, & t^*(j)=t\\ 0, & otherwise\end{cases}$$

We can estimate $(\eta_A,\eta_C,\sigma_\xi,\sigma_\omega,\sigma_{\xi\omega})$ using maximum likelihood.[15]

We assume that the candidates know the variables $A_{t^*(j)}$ and $C_j$, but do not know the realization of the shocks $\xi_{t^*(j)}$ and $\omega_j$. Instead, they predict the likely vote share of the Republican candidate, conditional on the positions that both candidates take using,

$$\lambda_j \sim N(\mu_j^\lambda,\sigma_\lambda{}^2)$$

where $\mu_j^\lambda=\eta_A{}'A_{t^*(j)}+\eta_C{}'C_j$, $\sigma_\lambda{}^2=\sigma_\xi{}^2+\sigma_\omega{}^2+2\sigma_{\xi\omega}$, and $\sigma_{\xi\omega}=\sigma_\xi\sigma_\omega\rho_{\xi\omega}$.

### 4.3 – Third Stage Estimation

In the third stage, we estimate the remaining parameters which consist of the parameters of the candidates' utility functions and the distribution of characteristics of the candidates. The

[15] Define $W_j=(A_{t^*(j)},C_j)$, $\beta=(\beta_A,\beta_C)$, and $\boldsymbol{W}_t$ to be the stacked rows of $W_j$ with $t^*(j)=t$. Define $\Omega_t(\sigma_\xi,\sigma_\omega,\sigma_{\xi\omega})$ using,

$$[\Omega_t(\sigma_\xi,\sigma_\omega,\sigma_{\xi\omega})]_{j,k}=\begin{cases}\sigma_\xi{}^2+\sigma_\omega{}^2+2\sigma_{\xi\omega}, & j=k\\ 2\sigma_{\xi\omega}, & j\neq k \text{ and } t^*(j)=t^*(k)\\ 0, & t^*(j)\neq t^*(k)\end{cases}$$

We can estimate $(\beta_A,\beta_C,\sigma_\xi,\sigma_\omega,\sigma_{\xi\omega})$ by maximizing the likelihood,

$$l(\beta_A,\beta_C,\sigma_\xi,\sigma_\omega,\sigma_{\xi\omega})=-\sum_{t=1}^{T}\left[\log\det\Omega_t(\sigma_\xi,\sigma_\omega,\sigma_{\xi\omega})+\tfrac{1}{2}(\lambda_t-\boldsymbol{W}_t\beta)'\Omega_t(\sigma_\xi,\sigma_\omega,\sigma_{\xi\omega})^{-1}(\lambda_t-\boldsymbol{W}_t\beta)\right]$$

In practice, we use a slightly different procedure (which we describe in Supplementary Appendix II) due to data limitations.

game described in section 3 of the paper describes a map between the candidate's characteristics, $(q_D, q_R, \alpha_D, \alpha_R)$, and their equilibrium positions $(y_D, y_R)$. As $(q_D, q_R, \alpha_D, \alpha_R)$ are not observable, our aim is to recover the parameters of the distribution of $(q_D, q_R, \alpha_D, \alpha_R)$ from the distribution of $(y_D, y_R)$. We start by specifying,

$$q_{D,j} = \pi_D + \beta_D{}'\tau_j + \delta_D{}'\lambda_j + \sigma_R \eta_{j,1}$$

$$q_{R,j} = \pi_R + \beta_R{}'\tau_j + \delta_R{}'\lambda_j + \sigma_D \eta_{j,2}$$

$$\alpha_{D,j} = \varphi_D{}^2 \eta_{j,3}{}^2, \qquad \alpha_{R,j} = \varphi_R{}^2 \eta_{j,4}{}^2$$

where $\eta_j \sim N(0, I)$. The specification allows the candidates' ideal points to be correlated with the mean voter's position and the valence advantage of the Democratic candidate.

We will attempt to recover the vector of parameters,

$$\theta = (\pi_D, \pi_R, \beta_D, \beta_R, \delta_D, \delta_R, \sigma_D, \sigma_R, \varphi_D, \varphi_R)$$

Our approach will rely on the Simulated Method of Moments (McFadden, 1989; Pakes and Pollard, 1989). Specifically, given the vector of parameters, we will generate $S$ samples of $J$ observations, $(q_{D,j,s}, q_{R,j,s}, \alpha_{D,j,s}, \alpha_{R,j,s})$. For each $j$ and $s$, we compute predicted equilibrium positions from the theoretical model, $(\tilde{y}_{D,j,s}, \tilde{y}_{R,j,s})$. Our estimation procedure chooses the parameters vector $\theta$ to minimize a quadratic form of the differences between the empirical and predicted moments (Hansen, 1982). We obtain efficient estimates by employing the continuously updated Generalized Method of Moments estimator (Hanson, Heaton, and Yaron, 1996).

Note the computational complexity involved here- we must solve for the equilibrium numerically $J * S$ times each time we evaluate the objective function. Fortunately, we can significantly reduce the computational burden by employing a technique devised by Ackerberg (2006), which we detail in Appendix II. The technique uses importance sampling to reduce the computational burden in structural estimation problems. As a result of this technique, we only have to solve for the equilibrium of the model $J * S$ times before the first iteration.

We are free to choose any set of moments that identify the unknown parameter vector. In our application, we will choose moments based on the means and variances of $(y_{D,j}, y_{R,j})$ and their correlations with $(\tau_j, \lambda_j)$. We will choose more moments than there are free parameters in order to allow for a test of the model based on over-identifying conditions (Hansen, 1982).

## 5 – Data

Our estimation strategy will require us to observe a cross-section of races with individual level data from each of these races. In addition, it is important that we obtain estimates of the candidates' positions and the voters' ideal points *on the same scale*. Congressional and Gubernatorial elections are natural applications for our approach. The requirement that we observe individual level data, however, precludes most surveys that rely on nationwide probability samples because we will not observe a sufficient number of individuals in each state. The Senate Election Study is uniquely suitable for our purposes because sampling is stratified by state, ensuring that we obtain a reasonably large sample of individuals even in small states.

The Senate Election Study provides us with respondent self-placements on a liberal conservative scale, respondent placements of the candidates on the same scale, whether the respondents voted, for whom they voted, the partisan identification of the respondents, and various demographic characteristics.

A limitation of the Senate Election Study is that we only observe three elections of data. Most of the parameters of the model can be estimated with this data, but there is reason to believe that Senate elections are subject to election-level as well as race-level shocks. These election-level shocks may come from changes in the reputation of the major political parties or retrospective economic voting. Furthermore, these shocks are partially predictable. For example, it was clear in early 2008 that the Republican Party would do poorly in the 2008 Senate elections.

To estimate the election-level shock model, we supplement the Senate Election Study with the American National Election Study cumulative file.

Perhaps the biggest drawback of the Senate Election Survey is the quality of the candidate position estimates. A number of anomalies are present in the data. The most striking is that the degree of centrism of the candidate placements is beyond believability. The representation literature has grappled with this problem. Two approaches have been used. The first is to employ measures that do not require voter and candidate ideology to be on the same scale (Miller and Stokes, 1963). The second approach is to rely on voter placements of the candidates, but to correct them in some way.[16] The first approach is not an option for us- without scale comparability, we would lose much of our power to reject the spatial model as an account of candidate positioning.

Powell (1989) notes that while placements overall are not reliable, the placements of knowledgeable voters are quite reliable.[17] This is the approach we will use to estimate candidate positions. We estimate the relationships,[18]

$$\hat{y}_{D,j^*(n)} = y_{D,j^*(n)} + \beta_D '(Q_n - Q^*) + \varepsilon_{D,n}$$

$$\hat{y}_{R,j^*(n)} = y_{R,j^*(n)} + \beta_R '(Q_n - Q^*) + \varepsilon_{R,n}$$

where $Q_n$ is a vector of variables capturing the political knowledge of respondent $n$ and $Q^*$ is the knowledge level of the most knowledgeable set of respondents.[19] If $\beta_D > 0$ and $\beta_R < 0$, then we would except that mean placements would yields estimates of candidate positions that are biased upward for Democrats and biased downward for Republicans. By estimating the candidate positions based on the above regressions, we can obtain corrected estimates of the candidate

---

[16] See Bafumi and Herron (2007) and Masket and Noel (2008) for an alternative approach.

[17] See also Palfrey and Poole (1987).

[18] In principal, we could measure the candidates' positions simply by averaging the placements of the most knowledgeable respondents, but the sample sizes available in the Senate Election Study would not allow for reliable estimates.

[19] Political knowledge is measured using the interviewers rating of the respondent's knowledge on a 1 to 5 scale.

positions. We computed corrected estimates by estimating the above regressions using ordinary least squares.[20]

# 6 – Estimation Results

### 6.1 – First Stage Parameters

The estimates of the parameters of the voters' utility functions are given in Table 1. The estimates here are not surprising- voters respond to spatial proximity, rewarding candidates that are ideologically closer to them. Holding other things constant, Democratic Party identifiers are more likely to vote for Democratic candidates and Republican Party identifiers are more likely to vote for Republican candidates. Black voters are more likely to vote for Democratic candidates. The remaining demographics- gender, age, and education- fail to obtain statistical significance at the 5% level. We note that in the first stage, we have included race-level fixed effects. Hence, we do not include regressors that vary only at the race level in this stage of the estimation. These fixed effects measure the quality difference (broadly construed) between the Republican and Democratic candidates.

### 6.2 – Second Stage Parameters

The results of the second stage estimation procedure are given in Table 2. Here, we regress the fixed effects from the first stage estimation on a number of candidate-specific and

[20] One may hypothesize that more knowledgeable respondents have a lower variance term. We estimated a heteroskedastic linear model by maximum likelihood, and found that this was indeed the case. However, the estimates turned out to be nearly identical, so we relied on the OLS estimates. Our estimator therefore corrects for the incorrect scaling of the candidates by uninformed voters.

election-specific factors that the candidates could reasonably use to form their belief over the likelihood of victory. We find that incumbent candidates expect to perform best (this effect is highly statistically significant). Candidates who have held high elective office are expected to do better than candidates who have not held previous office, though the effect is only statistically significant for Democratic candidates.[21] We also find that Democratic candidates do less well in the South, though this effect is also not statistically significant.

Aggregate level factors have an influence as well. Democratic candidates expect to obtain a lower vote share when the incumbent president is a Democrat, and an even lower vote share in midterm elections when the sitting president is a Democrat.[22] A candidate whose party matches the incumbent president's party expects his own vote share to increase as that the President's net approval rating increases.[23] Overall, the results are consistent with a midterm loss that is moderated by Presidential approval (Erikson, 1988).

All the variables we have included in the second stage procedure are variables that the candidates would know at the time they are selecting their positions. For example, both candidates would know whether one of them is an incumbent, whether they are running in the South, etc. Candidates would then form their beliefs as follows. An incumbent running against a low quality challenger would expect to do well. A challenger in an open race running against a challenger of similar quality would expect a close race. Variables such as the spending advantage of the Democratic candidate may in fact be correlated with $\lambda_j$, but would not be appropriate for inclusion here because the candidates would not know this amount at the time they choose their

[21] Former Senators, Governors, member of the House, and holders of statewide elected offices are said to have held high elected office. State legislators and local elected officials are said to have held low elective office.

[22] This statement may at first seem to be at odds with conventional wisdom, but recall that the conventional wisdom is that the elected President's party gains seats in the Presidential election and that the incumbent president's party loses seats in the midterm election. The results presented hear do not concern gains and loses, but absolute levels.

[23] We also considered economic factors such as GDP growth and inflation, but found that these factors were not statistically significant. This is likely due to the fact that while these factors are important, they operate through their effect on Presidential approval.

positions. To the extent that incumbents can be expected to raise more money, this factor will be accounted for by the inclusion of the incumbent dummy variable in the specification.

### 6.3 – Third Stage Parameters

The estimated parameters from the third stage are reported in Table 3. The coefficients on $\beta_D$ and $\beta_R$ have the expected sign. As the mean voter in a district moves to the right, the expected ideal point of the Democratic candidate moves to the right. This effect is only statistically significant for Republican candidates. We note that $q_{D,j}$ and $q_{R,j}$ represent the true policy preferences of the candidate, not the positions that they take. Consequently, this result suggests that the mean ideology of the party elites that make up likely Senate candidates is related to the distribution of voters in that district, with more conservative districts having more conservatives elites for *both* parties. This result is consistent with the findings of Erikson, Wright, and McIver (1994).

The coefficient on $\delta_D$ suggests that districts that tend to have high quality Democratic candidates also tend to have more conservative Democrats (in their policy-ideal points, but not necessarily in the positions that they take). This effect is not statistically significant however. The coefficient on $\delta_R$ suggests that districts that tend to have high quality Republicans also tend to have more conservative Republicans.

The remaining coefficients are hard to interpret, so we instead report the results as follows. Table 4 reports summary statistics of the data and the model. First, we find that $q_{D,j}$ has a mean of 2.7 with a standard deviation of 2.0, while $q_{R,j}$ has a mean of 8.6 and a variance of 2.8. The average mean voter across districts is located at 4.4, so this indicates that Democratic candidates are on average more moderate in their policy ideal points. We also find that Democratic candidates are more heterogeneous in their policy ideal points, while the Democratic

Party retains a mass of moderate Southern Democrats (or at least it did in the 1988-1992 time period).

## 7 – Model Fit

Here, we will focus mainly on the fit of the model in the third stage since the fit of the spatial model as a description of voter behavior has been studied extensively. We start by reporting the result of an over-identification test (Hansen, 1982). We computed the J-Statistic to be 3.45 with 1 degree of freedom.[24] The p-value is calculated as 0.063, indicating that we fail to reject the model at conventional levels of statistical significance.

While the failure to reject the model is reassuring, we believe our results suggest a more nuanced interpretation of the fit of the spatial model. The spatial model is capable of explaining many patterns observed in the data, but there are some anomalies. The failure to reject the model using the J-test means that these anomalies may be caused by sampling error alone. We will present a number of plots which compare patterns observed in the data with the patterns predicted by the model at the estimated parameter values. In all of these plots, one should keep in mind that because the number of races is relatively small, sampling error will be associated with these quantities.

We start by producing the classic representation plot in Figures 1 and 2. The results indicate that the predictions of the model are largely consistent with the data. Candidates respond to their constituents, but responsiveness is weak. To the extent that responsiveness exists, it operates mostly through the probability of electing a Republican candidate, rather than through

[24] For the J-Test, the null hypothesis is that the model is correctly specified, meaning that all of the population moments are equal to zero when evaluated at the true parameter value. The alternative hypothesis is that at least one of the moments is different from zero for any parameter value.

within party responsiveness.[25] The data indicate that within party responsiveness is positive for Democratic candidates and slightly negative for Republican candidates. The model predicts positive relationships for both parties. This is likely more a reflection on the quality of the Senate Election Study data than a failure of the model- both Ansolabehere, Snyder, and Stewart (2001) and Burden (2004) find that the responsiveness is positive for candidates of either party.

In Figures 3 and 4, we examine polarization among elected officials. Polarization by party and bimodality are present in both the data and the model predictions. The model however predicts more overlap in the distributions and the distribution of positions for Democratic candidates contains too little kurtosis as compared with the data. We suspect that these problems could be addressed by selecting an alternative specification for the random components of $q_{D,j}$ and $q_{R,j}$. More importantly, one might suppose that polarization is inconsistent with spatial competition. Our results indicate that this is not the case- although spatial competition constrains legislators from locating at their ideal points, the spatial model is consistent with observed patterns of polarization. Moreover, we note that our spatial model predicts substantial amounts of divergence. We can determine that the average divergence present in the data is 2.0, while the average divergence predicted by the model is 1.5.

In Figures 5 and 6, we plot the predicted and actual Republican vote shares across districts. The results suggest a degree of inconsistency between the model and the data. The model generally predicts less competitive elections than we observe in the data. This difference, however, can be largely explained by the fact that we purged uncontested races from our dataset, while the model does not explicitly incorporate a mechanism for censoring uncontested races. If one were to purge races with very small vote shares for the losing party from the model generated data, one would obtain a figure matching Figure 5 much more closely.

[25] Bartels (1991) makes a similar point.

### 7.1 - Competitiveness and Representation

A key finding that we will draw from our analysis is that even when the spatial model generates monotone comparative statics, these comparative statics may not appear as correlations in the actual data. The problem is that candidate positioning is plagued by endogeneity problems. Even if one could hold constant observable variables such as the position of the median voter, the valence advantage of the Democratic candidate, etc., one could not hold constant certain unobservable variables, such as the personal ideal points of the candidates, which may in turn be correlated with the observable variables. We illustrate this problem with an application to the marginality hypothesis.

An important issue in the representation literature is the relationship between competitiveness and representation. The marginality hypothesis, first proposed by MacRae (1952), states that legislators who face strong competition will better represent their constituencies than legislators whose are safe from electoral threat. This topic has received considerable attention in the literature.[26]

The difficulty of studying this problem is that there is clearly a simultaneity problem here- races may be marginal because the stronger candidate is poorly representing the constituency. More recent work has addressed this problem with measures not directly related to the current race. For example, Gulati (2004) constructs a measure based on partisan identification in the state and Griffin (2006) constructs a measure based on presidential vote. While these approaches provide an improvement over previous work, both of these measures have been used elsewhere as measures of ideology.[27] The finding that competitive elections lead to better

[26] See Fiorina (1973), Kuklinski (1977), Bartels (1991), Gulati (2004), and Griffin (2006).
[27] District partisanship is used in Erikson, Wright, and McIver (1994) as a measure of ideology. Republican presidential vote share is used widely used as a measure of ideology in the representation literature, and is in fact used by Griffin (2006) for this purpose.

representation may simply indicate that more moderate constituencies lead to better representation. This is an equally reasonable hypothesis, and an equally reasonable interpretation of what these variables measure.

An alternative solution to the simultaneity problem is to rely on more sophisticated theorizing. Groseclose (2001) shows that in a model where candidates are uncertain about the location of the median voter, a 'reverse marginality hypothesis' obtains- giving one candidate a small valence advantage leads that candidate to move towards the expected median voter.[28] Adams, Merrill, and Grofman (2005) consider a model with valence uncertainty, and show that increasing a candidate's expected valence leads that candidate to move away from the median voter. They also argue that given reasonable values of the variance of the median voter position, increasing a candidate's valence would lead the candidate to move away from the expected median voter. Peress (forthcoming) shows that in a spatial model with no uncertainty, the position of the winning candidate exhibits monotone comparative statics in the valence of the candidates, with perfect representation resulting when neither candidate has a valence advantage.

The theoretical work summarized above is generally consistent with the marginality hypothesis, but this does not mean that a positive correlation will be found between competitiveness and policy congruence or valence and policy congruence. Even if we control from such observable variables as the position of the median voter, we will not be able to control for unobservable variables such as the policy ideal points of the candidates, which may themselves be correlated with various observable variables.

We illustrate this problem in Figures 7 and 8. We plot the relationship between Republican vote share and the position of the winning candidate. We see that there is no tendency for candidates in marginal races to take more moderate positions. In fact, Democratic candidates

---

[28] We note that Groseclose's (2001) theoretical result implies that the derivative of the advantaged candidate's position with respect valence indicates that the candidate will move toward the expected median voter for *small values of valence*. Groseclose's computational results suggest that effect reverses for larger values of valence.

facing competitive races actually take more extreme positions. We do not take this as evidence of a 'reverse marginality hypothesis', nor do we take this as evidence of an 'in step-out of office' effect.[29] Instead, the similarity between Figures 7 and 8 indicate that this version of the spatial model is able to fit a surprising pattern in the data, and one that others have (incorrectly) suggested is inconsistent with spatial competition.

Despite findings in the literature indicating a lack of correlation between competitiveness and representation, these results do not suggest that the marginality hypothesis is false. Here, we use the model to determine the comparative static between competitiveness and representation. In this comparative static, we will shift the distribution of non-policy factors in each race so that the distribution will have mean $\phi$, and we will vary $\phi$ between -5 and 5. In Figure 9, we plot the predictions for the average Republican vote share.[30] We see that Republican vote share increases monotonically with valence, indicating the validity of the measures of competitiveness.

Figure 10 plots the average positions of the Democratic and Republican candidates as $\phi$ is varied. The results indicate that the average position of the winning candidate varies little as $\phi$ varies. The positions of the Democratic and Republican candidates vary quite a bit, however. When the Republican candidate has a large valence advantage, he will position far to the right while the Democratic candidate will position near the political center. In the situation where the Democratic candidate has a large valence advantage, the reverse will happen.

The comparative statics of the position of the winning candidate are monotone in the valence advantage of the Democratic candidate, but the effect is largely tempered by candidate positioning. As the valence advantage of the Democratic candidate increases, both candidates move to the left. The Democratic candidate's probability of winning decreases, however, leaving the expected positioning of the winning candidates largely unchanged.

---

[29] Canes-Wrone, Brady, and Cogan (2002) argue that legislators who are "out of step" with their constituencies may soon find themselves "out of office".

[30] Following Achen (1978), we define congruence to be the absolute value of the difference between the position of the winning candidate and the position of the median voter in the district.

## 8 – Electoral Success and Representation

In this section, we use our framework to decompose the sources of the Democratic Party's success (or lack of success) in Senate elections between 1980 and 2004, and in more detail, in Senate elections between 1988 and 1992.

### 8.1 – Political Climate

We start by considering the political climate. Our model assumed that candidates use three variables to predict the election-level shock- the party of the incumbent president, whether the election is a midterm election (coded by the party of the incumbent president) and the president's net approval rating (coded by the party of the incumbent president). Table 5 presents the predicted and realized election-level shocks. Higher shocks mean that Republican candidates expect to do better, holding constant the characteristics of candidates and districts. We note that the shocks are normalized so that the realized value of the 1988 shock is equal to zero.

We can see that 1984, 1994, and 2002 were expected to have a political climate favorable to Republican candidates. Alternatively, 1982 and 1992 were predicted to be favorable towards Democratic candidates. We note here that our results suggest that the Democratic party's 1994 midterm loss was largely predictable based on the fact that the election-level predicted shock changed by 0.36 units over the two year period (which represents nearly the entire range of shocks observed over a 26 year period).

### 8.2 – Electoral Success

Predicted election-level shocks can have a reinforcing effect through the recruitment of quality challengers. We can compute the average challenger quality (using the 0-5 scale as in Lublin, 1994) for both the Democratic and Republican candidates. The correlation between the predicted aggregate shock and the average quality of the challenger is -42% and 58% for the Democratic and Republican parties respectively. When the political climate is expected to favor the Republican Party, the Democratic Party is able to recruit fewer high quality candidates and the Republican Party is able to recruit more high quality candidates. We note that these correlations are larger than the correlations with the realized shock (which are -22% and 33% respectively).

We did not see large differences in the success of the Republican Party over 1988, 1990, and 1992. The Republicans won 14 out of 32 Senate races (and a 46.8% vote share) in 1988, 17 out of 34 Senate races (and a 49.7% vote share) in 1990, and 14 out of 33 Senate races (and a 48.1% vote share) in 1992.[31] Of these years, 1990 was the year that was predicted to favor the Republican Party most, and Republicans were therefore most successful in recruiting quality challengers. We also found that the average valence shock (which includes both election-level and race-level components) was highest 1990. Nonetheless, the Republican Party's success in 1990 was tempered by their inability to take moderate positions, which itself was due to more extreme personal ideal point among Republican candidates.

**8.3 - Responsiveness to the Mean Voter**

The 1990 Senate elections illustrated how positioning by the candidates can mitigate a pre-existing valence advantage. While competitive forces are still active in asymmetric and uncertain elections, they are muted relative to the Downsian benchmark, which features perfect

[31] The 1988 Nevada Senate race is excluded from these calculations because there was no data available in the Senate Election Study.

responsiveness to the median voter. Here, we assess the degree of such responsiveness in a more realistic model of Senate elections.

In section 7, we demonstrated that there was a correlation between the position of the mean voter and the position of the winning candidate. Here, we investigate a different but related issue- if the mean voter shifts one unit to the left (right), how far left (right) will the winning candidate shift? We consider the comparative statics between the positions of the candidates and the position of the mean voter. We shift the distribution of voter ideal points in each district so that the mean is $\varsigma$ and we vary $\varsigma$ between 1 and 7.

The results are reported in Figure 11. We see that as the mean voter becomes more conservative, both the Democratic and Republican candidates become more conservative. The slope of the relationship, however, is less than one, and the average position of the winning candidate and the mean voter is only close when the mean voter is quite conservative. When $\varsigma = 1$, the divergence of the winning candidate from the mean voter is substantial. We can attribute this to the fact that Democratic candidates are more moderate than Republican candidates in terms of their personal ideal points. It also indicates that the spatial model provides an explanation for the weak pattern of responsiveness we observe in the data.

## 9 – Conclusion

As we have argued, the study of candidate positioning is plagued with endogeneity problems. To properly theorize and interpret empirical work in this area, it is essential to have a working theory of candidate positioning. In this paper, we argue that the revised spatial is a good candidate for such a theory. Our goals in this paper were threefold. First, we sought to develop an approach for estimating the parameters of a general spatial model. In particular, we sought to extend existing work by providing a technique for recovering the parameters characterizing both voter and candidates. Our second goal was to empirically test the revised spatial model's ability to explain patterns of candidate positioning. Our third goal was to use the spatial model to decompose the sources of electoral success in U.S. Senate elections.

We believe our results reflect quite favorably on the ability of the spatial model to explain patterns of candidate positioning. The results of this paper suggest while the Downsian model obviously cannot explain the patterns of candidate positioning, a revised spatial model can explain many of the patterns we observe. While the fit of the model is not perfect in every dimension, recent theoretical work has been largely successful in improving the spatial model by making more realistic assumptions about the motivations of the candidates and the electoral environment.

Given the success of the spatial model in explaining patterns of candidate positioning, we believe that it should be more frequently used to guide empirical work in this area. Our application to the marginality hypothesis illustrated an instance where spatial competition does not generate monotone comparative statics. In spatial competition, observed correlations are particularly uninformative about underlying phenomena because endogeneity problems are endemic. We believe we have outlined an approach for dealing with such endogeneity. To outline our approach, we begin with a baseline model, which has fared well against empirical tests to date. We discover some new phenomenon in need of explanation. We check if our baseline model is consistent with this hypothesis. If it is consistent, then there is nothing to new to explain. If it is not consistent, we consider alternative theories, but we keep in mind that the new theories must explain the old facts as well.

Finally, we applied our theory to decompose the sources of electoral success. First, elections are subject to an aggregate shock which derives from the fact that the fortunes of Senate candidates are tied to the performance of the sitting president. This is an important source of variation though not as important as differences in the characteristics of the candidates. Nonetheless, since candidate characteristics do not vary much in aggregate (e.g. most elections feature about the same number of Democratic and Republican incumbents), the election-level shock turns out to be an important determinant of the overall electoral success of a party. Furthermore, some candidate characteristics are likely to be positively influenced by the electoral climate. In particular, a party is able to recruit higher quality challengers when the electoral environment is expected to be favorable towards that party.

Positioning by the candidates acts against the predetermined sources of advantage-weaker candidates take more moderate positions in order to overcome their non-policy advantage. This is not effective in very lop-sided elections because policy voting is only one component of the voters' choice strategies. Moreover, Democratic candidate are able to more effectively moderate because they have more moderate personal ideal points.

The results suggest that Senate elections are far from the Downsian ideal of perfect responsiveness. If the mean voter where to move one unit to the left (right), the position of the winning candidate would be expected to move only about half a unit to the left (right). The presence of valence (or non-policy voting) and uncertainty limit the responsiveness of candidates to the voters' policy preferences, but the basic logic of the Downsian competition still applies-competition encourages candidates to respond to the voters' policy preferences.

## References


[1] Achen, Christopher H. (1978). "Measuring Representation". *American Journal of Political Science* 3:475-510.

[2] Ackerberg, Daniel A. (2001). "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation". Working Paper.

[3] Adams, James, Thomas Brunell, Bernard Grofman, and Samuel Merrill, III (2007). "Move to the Center or Mobilize the Base? Effects of Political Competition, Voter

Turnout, and Partisan Loyalties on the Ideological Divergence of Vote-maximizing Candidates". Working Paper.

[4] Adams, James and Samuel Merrill, III (2003). "Voter Turnout and Candidate Strategies in American Elections". *Journal of Politics* 65:161-189.

[5] Adams, James, Samuel Merrill, III, and Bernard Grofman (2005). *A Unified Theory of Party Competition*. New York: Cambridge University Press.

[6] Aldrich, John H., and Richard D. McKelvey (1977). "A Method for Scaling with Applications to the 1968 and 1972 Presidential Elections". American Political Science Review 71:111-130.

[7] Ansolabehere, Stephen and James M. Snyder, Jr. (2000). "Valence Politics and Equilibrium in Spatial Election Models". *Public Choice* 103:327-336.

[8] Ansolabehere, Stephen, James M. Snyder, Jr., and Charles Stewart, III (2001). "Candidate Positioning in U.S. House Elections". *American Journal of Political Science* 45:136-149.

[9] Bafumi, Joseph, and Michael Herron (2008). "Preference Aggregation, Representation and Elected American Political Institutions". Working Paper.

[10] Banks, Jeffery, and John Duggan (2003). "Probabilistic Voting in the Spatial Model of Elections: The Theory of Office Motivated Candidates". Working Paper.

[11] Bartels, Larry M. (1991). "Constituency Opinion and Congressional Policy Making: The Reagan Defense Build Up". *American Political Science Review* 2:457-474.

[12] Burden, Barry (2004). "Candidate Positioning in U.S. Congressional Elections". *British Journal of Political Science* 34:211-227.

[13] Calvert, Randall (1985). "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence" *American Journal of Political Science* 29:69-95.

[14] Canes-Wrone, Brandice, David W. Brady and John F. Cogan (2002). "Out of Step, Out of Office: Electoral Accountability and House Members' Voting" *American Political Science Review* 96:127-140.

[15] Coughlin, P. J. and S. Nitzan (1981a). "Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Optima". *Journal of Public Economics* 15:113-122.

[16] Coughlin, P. J. and S. Nitzan (1981b). "Directional and Local Electoral Equilibria with Probabilistic Voting". *Journal of Economic Theory* 24:226-240.

[17] Downs, Anthony (1957). *An Economic Theory of Democracy*.

[18] Erikson, Robert S. (1988). "The Puzzle of Midterm Loss." *Journal of Politics* 50: 1011–29.

[19] Erikson, Robert S. and David W. Romero (1990). "Candidate Equilibrium and the Behavioral Model of the Vote". *American Political Science Review* 84:1103-1126.

[20] Erikson, Robert S., Gerald C. Wright, and John P. McIver (1994). *Statehouse Democracy: Public Opinion and the American States*. Cambridge: Cambridge University Press.

[21] Fiorina, Morris P. (1999). "What Ever Happened to the Median Voter?" Working Paper.

[22] Fiorina, Morris (2005). Culture Wars: The Myth of a Polarized American.

[23] Griffin, John D. (2006). "Electoral Competition and Democratic Responsiveness: A Defense of the Marginality Hypothesis". *Journal of Politics*.

[24] Grofman, Bernard (2004). "Downs and Two Party Convergence". *Annual Review of Political Science* 7:25-46.

[25] Groseclose, Tim (2001). "A Model of Candidate Location When One Candidate Has a Valence Advantage." Working Paper.

[26] Grynaviski, Jeffery, and Bryce E. Corrigan (2006). "Specification Issues in Proximity Models of Candidate Evaluation (with Issue Importance)". *Political Analysis* 14:393-420.

[27] Gulati, Girish J. (2004). “Revisiting the Link Between Electoral Competition and Policy Extremism in the U.S. Congress”. *American Politics Research* 32:495-520.

[28] Hansen, Lars (1982). “Large Sample Properties of Generalized Method of Moments Estimators”. *Econometrica* 50:1029-1054.

[29] Hansen, Lars P., John Heaton, and Amir Yaron (1996). “Finite-Sample Properties of Some Alternative GMM Estimators”. *Journal of Business and Economic Statistics* 14:262-280.

[30] Jacobson, Gary C. (2006). *A Divider, not a Uniter: George W. Bush and the American People*. Longman.

[31] Layman, Geoffrey C., and Thomas M. Carsey (2006). “Changing Sides or Changing Minds? Party Identification and Policy Preferences in the American Electorate”. *American Journal of Political Science* 50:464-477.

[32] Levendusky, Matthew (2005). “American the Purple”. Working Paper.

[33] Lublin, David (1994). “Quality, Not Quantity: Strategic Politicians in U.S. Senate Elections, 1952-1990”. *Journal of Politics* 56.

[34] MacRae, D., Jr. (1952). “The Relation Between Roll Call Votes and Constituencies in the Massachusetts House and Representatives”. *American Political Science Review* 46:1046-1055.

[35] Masket, Seth E. and Hans Noel (2008). “How Out of Step? Measuring Differences in Voter and Legislator Preferences”. Working Paper.

[36] McFadden, Daniel (1989). “A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration”. *Econometrica* 57:995-1026.

[37] Mebane, Walter R., Jr. (2000). “Coordination, Moderation, and Institutional Balancing in American Presidential and House Elections”. *American Political Science Review* 94:37-57.

[38] Merlo, Antonio (1997). “Bargaining over Governments in a Stochastic Environment”. *Journal of Political Economy* 105:101-131.

[39] Merrill, Samuel, III, and Bernard Grofman (1999). *A Unified Theory of Voting: Direction and Proximity Spatial Models*. Cambridge, UK: Cambridge University Press.

[40] Moon, Woojin (2004). “Party Activists, Campaign Resources and Candidate Position Taking: Theory, Tests and Applications”. British Journal of Political Science 34:611-633.

[41] Pakes, Ariel, and David Pollard (1989). “Simulation and the Asymptotic of Optimization Estimators”. *Econometrica* 57:1027-1057.

[42] Palfrey, Thomas R., and Keith T. Poole (1987). “The Relationship between Information, Ideology, and Voting Behavior”. *American Journal of Political Science* 31:511-530.

[43] Peress, Michael (2007). “Securing the Base: Electoral Competition under Variable Turnout”. Working Paper.

[44] Peress, Michael (forthcoming). “The Spatial Model with Non-policy Factors: A Theory of Policy Motivated Candidates”. Forthcoming in *Social Choice and Welfare*.

[45] Poole, Keith T. and Howard Rosenthal (1991). “Patterns of Congressional Voting”. *American Journal of Political Science* 35:228-78.

[46] Poole, Keith T. and Howard Rosenthal (1997). *Congress: A Political Economic History of Roll Call Voting*. New York: Oxford University Press.

[47] Quinn, Kevin M., Andrew D. Martin, and Andrew B. Whitford (1999). “Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models”. *American Journal of Political Science* 43:1231-1247.

[48] Rabinowitz, George, and Elaine Stuart McDonald (1989). “A Directional Theory of Issue Voting”. *American Political Science Review* 83:93-121.

[49] Schofield, Norman (2003). “Valence Competition in the Spatial Stochastic Model”. *Journal of Theoretical Politics* 15:371-383.

[50] Schofield, Norman (2004). “Equilibrium in the Spatial ‘Valence’ Model of Politics”. *Journal of Theoretical Politics* 16:447-481.

[51] Schofield, Norman, Andrew D. Martin, Kevin M. Quinn, and Andrew B. Whitford (1998). “Multiparty Electoral Competition in the Netherlands and Germany: A Model Based on Multinomial Probit”. *Public Choice* 97:257-293.

[52] Wittman, Donald (1983). “Candidate Motivation: A Synthesis of Alternatives.” *American Political Science Review* 77:142-157.

# Appendix I – Computing the Equilibrium

Here, we will describe an algorithm that can be used to compute the equilibrium of the game. We have found that an effective approach for solving many continuous games (including this one) is to maximize $V_D(y_D, y_R)$ over $y_D$ to produce a new iterate for $y_D$, and maximize $V_R(y_D, y_R)$ over $y_R$ to produce a new iterate for $y_R$. We refer to this algorithm as Gauss-Seidel Best Response Iterations.

The main difficulty is that both $V_D$ and $V_R$ involve integrals over $(v, z, \lambda)$, and these integrals do not generally have analytical solutions. Theoretically, we could easily compute these integrals using simulation methods. For example, we can compute,

$$\hat{s}_D(y_D, y_R; \lambda) = \tfrac{1}{N}\sum_{n=1}^{N} 1\{\tilde{z}_n > \rho \,|\, y_R - \tilde{v}_n \,| - \rho \,|\, y_D - \tilde{v}_n \,|\}$$

where $(\tilde{v}_n, \tilde{z}_n)$ are independent draws from $g(v, z - \lambda)$. We have found this approach to be quite undesirable because $\hat{s}_D(y_D, y_R; \lambda)$ will be discontinuous function of the candidate's positions in finite samples due to the presence of the indicator function. This not only trips up the computational algorithms, but will often generate an approximate game with no equilibrium, even when the original game had an equilibrium.

Our approach will integrate over $v$ using simulation methods. In particular, we will draw a sample of $N$ voters from $f_v$. Voter $n$'s ideal point will be given by $v_n$. We assume that $G_{z|v}$ can be characterized analytically however. This assumption will indeed be satisfied in our application. We can compute candidate D's vote share as a function of $\lambda$ using,

$$\hat{s}_D(y_D, y_R; \lambda) = 1 - \tfrac{1}{N}\sum_{n=1}^{N} G_{z|v}(\rho \,|\, y_R - \tilde{v}_n \,| - \rho \,|\, y_D - \tilde{v}_n \,| - \lambda \,|\, \tilde{v}_n)$$

Now, we would like to characterize the probability $\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})$. We can characterize this using,

$$\widehat{\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})} = \int_{\lambda=\hat{\lambda}^*(y_D, y_R)}^{\infty} f_\lambda(\lambda) d\lambda = 1 - F_\lambda(\hat{\lambda}^*(y_D, y_R))$$

where $\hat{\lambda}^*(y_L, y_R)$ is the solution to,

$$\hat{s}_D(y_D, y_R; \hat{\lambda}^*(y_D, y_R)) = \tfrac{1}{2}$$

Notice that $\lim_{\lambda \to -\infty} \hat{s}_D(y_L, y_R; \lambda) = 0$, $\lim_{\lambda \to \infty} \hat{s}_D(y_L, y_R; \lambda) = 1$, and $\hat{s}_D(y_L, y_R; \lambda)$ is continuous in $\lambda$. Hence, the intermediate value theorem indicates that a solution to this equation must exist. The solution will be unique because $\hat{s}_D(y_D, y_R; \lambda)$ is strictly increasing in $\lambda$.

Continuity of $\hat{s}_D(y_D, y_R; \hat{\lambda}^*(y_D, y_R))$ in all three arguments means that we can use the implicit function theorem to establish that $\hat{\lambda}^*(y_D, y_R)$ is continuous in $(y_D, y_R)$. Then, continuity of $F_\lambda$ implies that $\widehat{\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})}$ is continuous in $(y_D, y_R)$. Finally, we define,

$$\hat{V}_D(y_D, y_R) = \widehat{\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})}(\alpha_D - (y_D - q_D)^2) + (1 - \widehat{\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})})(-(y_R - q_D)^2)$$

$$\hat{V}_R(y_D, y_R) = \widehat{\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})}(-(y_D - q_R)^2) + (1 - \widehat{\Pr(s_D(y_D, y_R; \lambda) \geq \tfrac{1}{2})})(\alpha_R - (y_R - q_R)^2)$$

These functions are clearly continuous in $(y_D, y_R)$ as well. Because we employ the absolute value utility function for the voters, the approximations $\hat{s}_D(y_D, y_R; \lambda)$ will not be differentiable in $(y_D, y_R)$, despite the fact that $s_D(y_D, y_R; \lambda)$ are themselves differentiable (since the set of discontinuities of $\int_{z=\rho|y_R - v| - \rho|y_D - v|}^{\infty} g(v, z - \lambda) dz$ has measure zero).

## Appendix II – Importance Sampling

Define the equilibrium correspondence by $(y_D, y_R) = m_j(q_D, q_R, \alpha_D, \alpha_R)$. Define $y = (y_D, y_R)$ and $\psi = (q_D, q_R, \alpha_D, \alpha_R)$ so that $y = m_j(\psi)$. The model implies that $\psi \sim f_j(\psi;\theta)$. We can form moments from the observed data, $\frac{1}{J}\sum_{j=1}^{J} h(y_j)$. We would like to match these to their theoretical counterparts, $\frac{1}{J}\sum_{j=1}^{J} E[h(y_j)]$, which are simulated from the model. Following Ackerberg (2006), we can calculate,

$$\frac{1}{J}\sum_{j=1}^{J} E[h(y_j)] = \frac{1}{J}\sum_{j=1}^{J} E[h(m_j(\xi_j))] = \frac{1}{J}\sum_{j=1}^{J}\int_{\psi} h(m_j(\psi)) f_j(\psi;\theta) d\xi = \frac{1}{J}\sum_{j=1}^{J}\int_{\psi} h(m_j(\psi)) \frac{g(\psi)}{f_j(\psi;\theta)} g(\psi) d\xi$$

We employ importance sampling by taking random draws, $\{\tilde{\psi}_{j,s}\}_{j=1,s=1}^{J,S}$ for some fixed $S \geq 1$. We can then approximate the theoretical moments by,

$$\frac{1}{J}\sum_{j=1}^{J} E[h(y_j)] \approx \frac{1}{J}\sum_{j=1}^{J}\sum_{s=1}^{S} \frac{g(\tilde{\psi}_{j,s})}{f_j(\tilde{\psi}_{j,s};\theta)} h(m_j(\tilde{\psi}_{j,s}))$$

To compute $m_j(\tilde{\psi}_{j,s})$, we must numerically solve for the equilibrium of the model, but since the unknown parameter $\theta$ does not enter into the expression for $m_j(.)$, we need only solve the model $J * S$ times, rather than $J * S$ times each time the moments are evaluated, as would be required by naïve sampling approaches. In our application, we draw $S = 50$ samples for each observation.

This approach is effective in drastically reducing the computational burden of nested fixed point problems if the importance sampling ratio, $\frac{g(\tilde{\psi}_{j,s})}{f_j(\tilde{\psi}_{j,s};\theta)}$, is easy to compute, which indeed is the case for the $g$ we select. In particular, we will choose,

$$g(\tilde{\psi}_{j,s}) = g_1(\tilde{q}_{D,j,s}) g_2(\tilde{q}_{R,j,s}) g_3(\tilde{\alpha}_{D,j,s}) g_4(\tilde{q}_{R,j,s})$$

where $g_1 \sim N(1, 2^2)$, $g_2 \sim N(7, 2^2)$, $g_3 \sim 2\chi_1^2$, and $g_4 \sim 2\chi_1^2$.

## Appendix III – Tables and Figures

### Table 1 – Voter Utility Parameters

| | Estimate | Standard Error |
|---|---|---|
| Spatial Distance | 0.227*** | (0.019) |
| Democratic Party Identifier | 0.703*** | (0.069) |
| Republican Party Identifier | -0.695*** | (0.069) |
| Female | 0.031 | (0.055) |
| Black | 0.457** | (0.163) |
| Age: <35 | - | |
| Age: 35-49 | -0.050 | (0.071) |
| Age: 50-64 | 0.032 | (0.082) |
| Age: 65+ | -0.071 | (0.086) |
| Education: No Degree | - | |
| Education: Bachelors Degree | -0.030 | (0.066) |
| Education: Graduate Degree | -0.050 | (0.092) |

## Table 2 – Candidate Belief Parameters[32]

| | Estimate | Standard Error |
|---|---|---|
| Constant | 0.033 | (0.199) |
| South | -0.226 | (0.129) |
| Democrat: Incumbent | 0.775*** | (0.151) |
| Democrat: Held Higher Office | 0.424** | (0.147) |
| Democrat: Held Lower Office | 0.235 | (0.200) |
| Republican: Incumbent | -0.712*** | (0.159) |
| Republican: Held Higher Office | -0.170 | (0.123) |
| Republican: Held Lower Office | -0.049 | (0.134) |
| Democratic Incumbent President | 0.133* | (0.057) |
| Midterm Effect (Coded by Pres. Inc.) | 0.101 | (0.090) |
| Pres. Net Approval (Coded by Pres. Inc.) | -0.519** | (0.195) |
| Sigma | 0.529 | |

[32] Higher office includes former Senators, members of the House, governors, and holders of other statewide elective office. Lower office includes state legislators and local elected offices. Presidential net approval is the average from Gallup polls conducted between February and April of the year.

## Table 3 – Candidate Utility Function Parameters

| | Estimate | Stan. Error |
|---|---|---|
| Democratic Intercept ( $\pi_D$ ) | -6.409 | (7.218) |
| Republican Intercept ( $\pi_R$ ) | -7.148 | (2.233) |
| Dem. Mean Voter Slope ( $\beta_D$ ) | 2.096 | (1.620) |
| Rep. Mean Voter Slope ( $\beta_R$ ) | 3.638 | (0.456) |
| Dem. Valence Slope ( $\delta_D$ ) | 1.729 | (2.235) |
| Rep. Valence Slope ( $\delta_R$ ) | -4.502 | (0.554) |
| Dem. Stan. Dev. ( $\sigma_D$ ) | 1.813 | (0.538) |
| Rep. Stan. Dev. ( $\sigma_R$ ) | 0.199 | (0.199) |
| Office Holding Benefit ( $\varphi_D$ ) | 1.286 | (7.355) |
| Office Holding Benefit ( $\varphi_R$ ) | 5.956 | (22.989) |

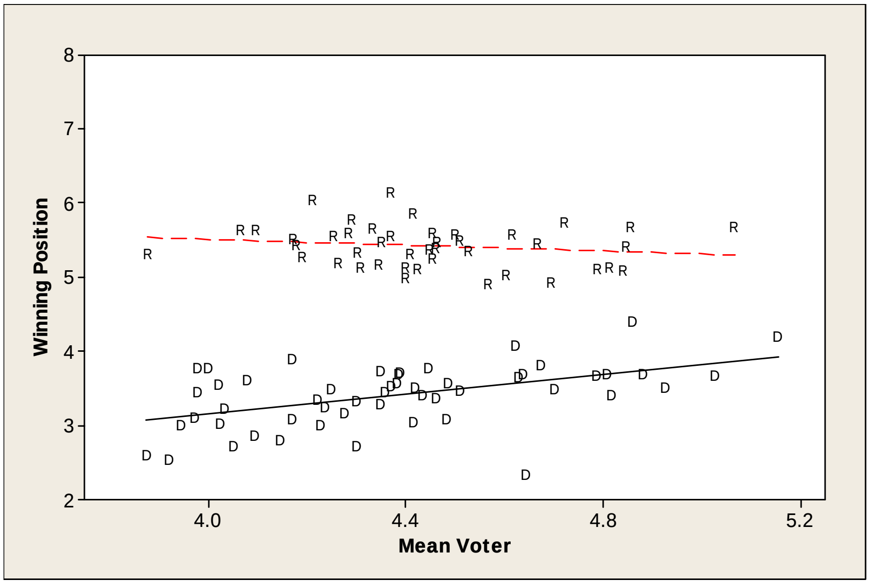
## Table 4 – Summary Statistics (Data and Model)

| | Mean | Stan. Dev. |
|---|---|---|
| ***Data*** | | |
| Mean Voter ( $\tau_j$ ) | 4.410 | [0.284] |
| Valence ( $\lambda_j$ ) | 0.055 | [0.556] |
| Dem. Position ( $y_{D,j}$ ) | 3.253 | [0.416] |
| Rep. Position ( $y_{R,j}$ ) | 5.295 | [0.335] |
| Rep. Vote Share | 0.481 | [0.127] |
| ***Model*** | | |
| Dem. Ideal Point ( $q_{D,j}$ ) | 2.689 | [2.030] |
| Rep. Ideal Point ( $q_{R,j}$ ) | 8.627 | [2.834] |
| Dem. Position ( $y_{D,j}$ ) | 4.333 | [0.722] |
| Rep. Position ( $y_{R,j}$ ) | 5.869 | [0.691] |
| Dem. Office Benefit ( $\alpha_{D,j}$ ) | 0.802 | [0.995] |
| Rep. Office Benefit ( $\alpha_{R,j}$ ) | 4.142 | [4.764] |
| Rep. Vote Share | 0.414 | [0.290] |

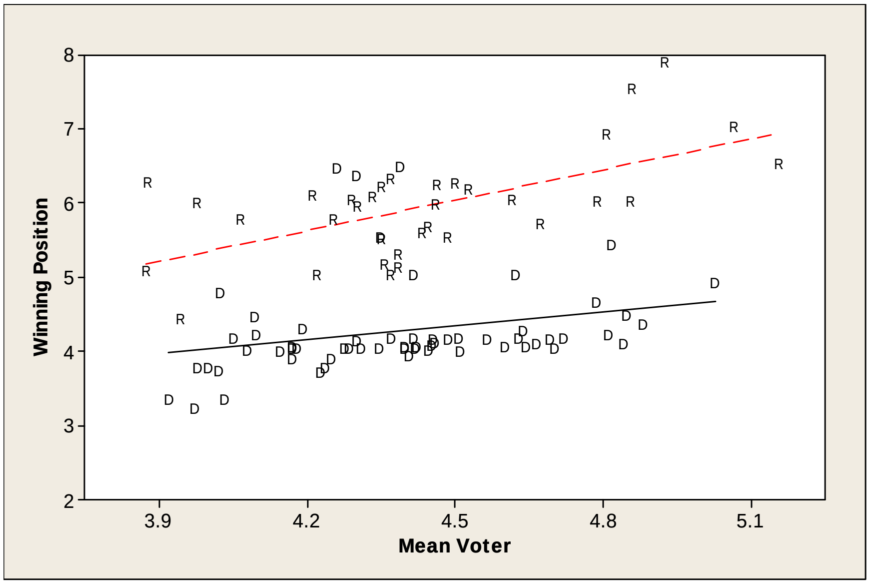
# Table 5 – Election-level Shock Model

| | Incumbent President's Party | Midterm Election | President's Net Approval Rating | Predicted Election-level Shock | Realized Election-level Shock |
|---|---|---|---|---|---|
| 1980 | Dem. | No | 2 | 0.310 | 0.099 |
| 1982 | Rep. | Yes | 0 | -0.050 | -0.201 |
| 1984 | Rep. | No | 17 | 0.143 | 0.265 |
| 1986 | Rep. | Yes | 35 | 0.133 | 0.164 |
| 1988 | Rep. | No | 11 | 0.110 | 0.000 |
| 1990 | Rep. | Yes | 55 | 0.236 | 0.151 |
| 1992 | Rep. | No | -9 | 0.006 | 0.111 |
| 1994 | Dem. | Yes | 10 | 0.368 | 0.410 |
| 1996 | Dem. | No | 16 | 0.236 | 0.390 |
| 1998 | Dem. | Yes | 37 | 0.230 | 0.049 |
| 2000 | Dem. | No | 24 | 0.192 | 0.390 |
| 2002 | Rep. | Yes | 62 | 0.276 | 0.341 |
| 2004 | Rep. | No | 6 | 0.082 | 0.105 |

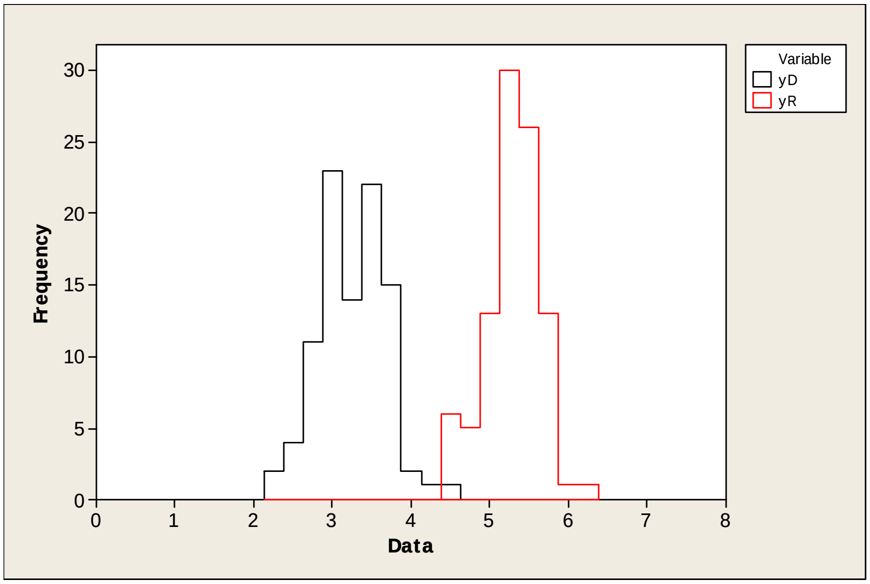
## Figure 1 – Representation Plot (Data)



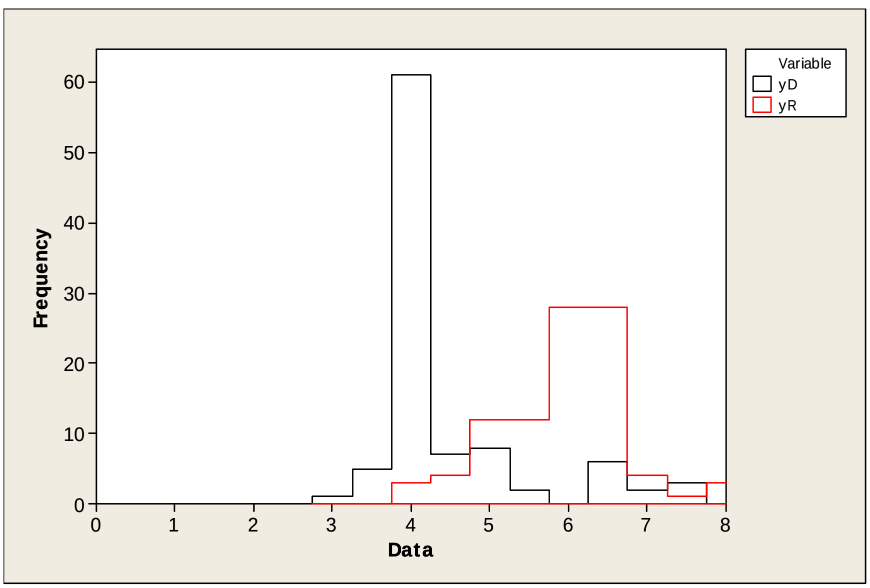

## Figure 2 – Representation Plot (Model)



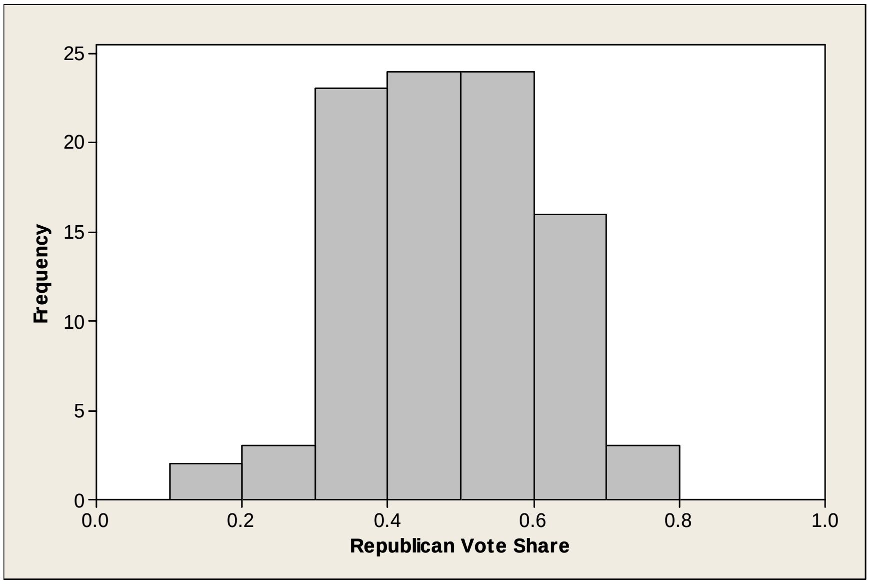
# Figure 3 – Candidate Positions (Data)



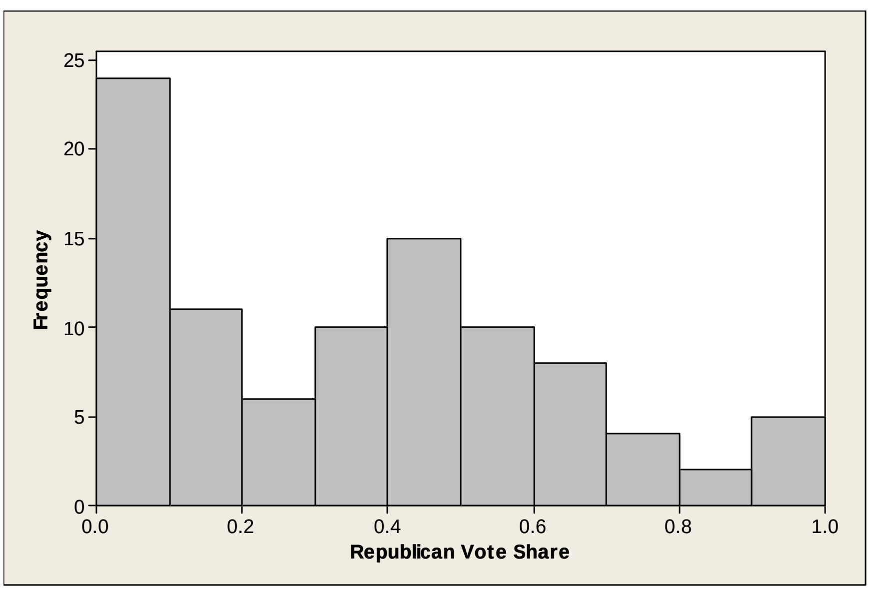

# Figure 4 – Candidate Positions (Model)



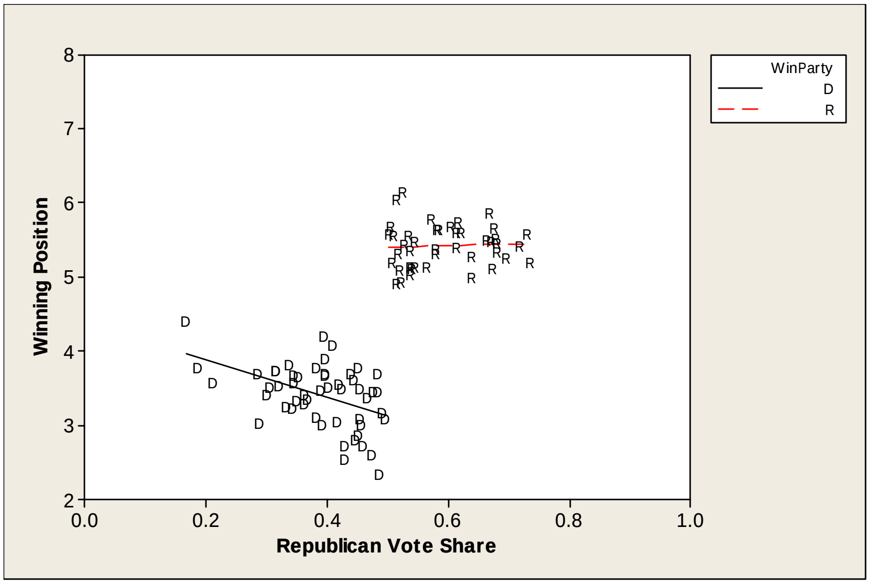
## Figure 5 – Republican Vote Shares (Data)




## Figure 6 – Republican Vote Shares (Model)



## Figure 7 - Competitiveness vs. Winning Position (Data)




## Figure 8 - Competitiveness vs. Winning Position (Model)

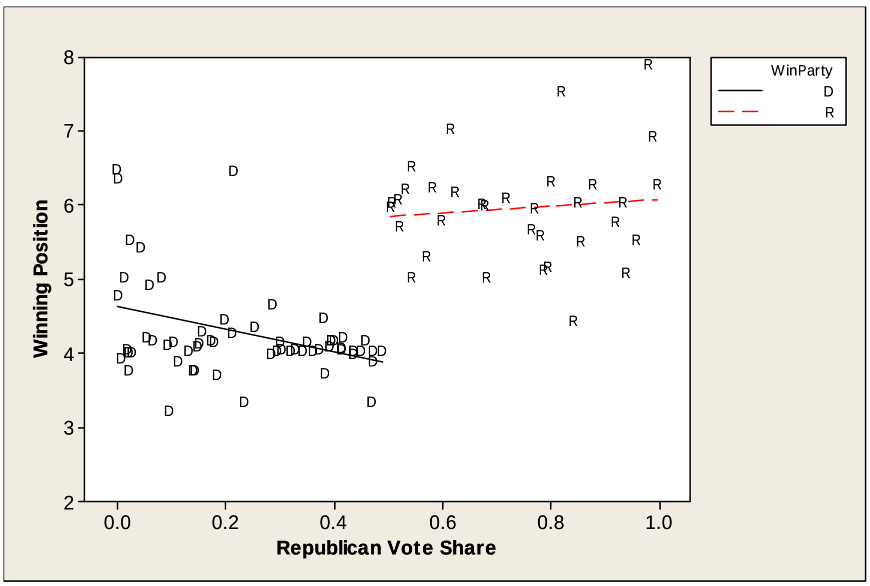


**Figure 9 – Comparative Statics: Republican Vote Share vs. Valence**

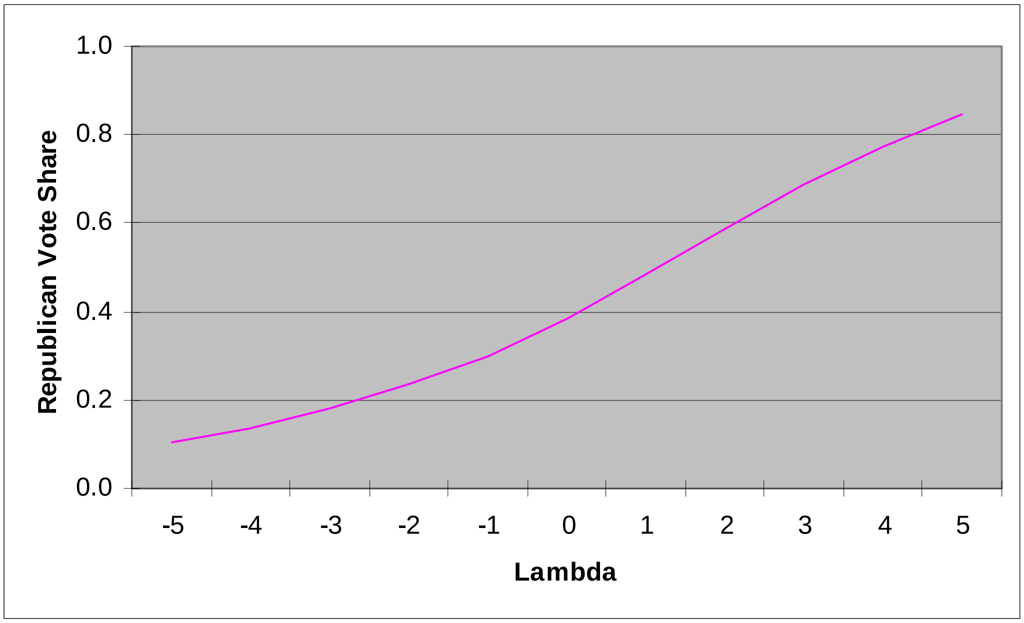



**Figure 10 – Comparative Statics: Candidate Positions vs. Valence**

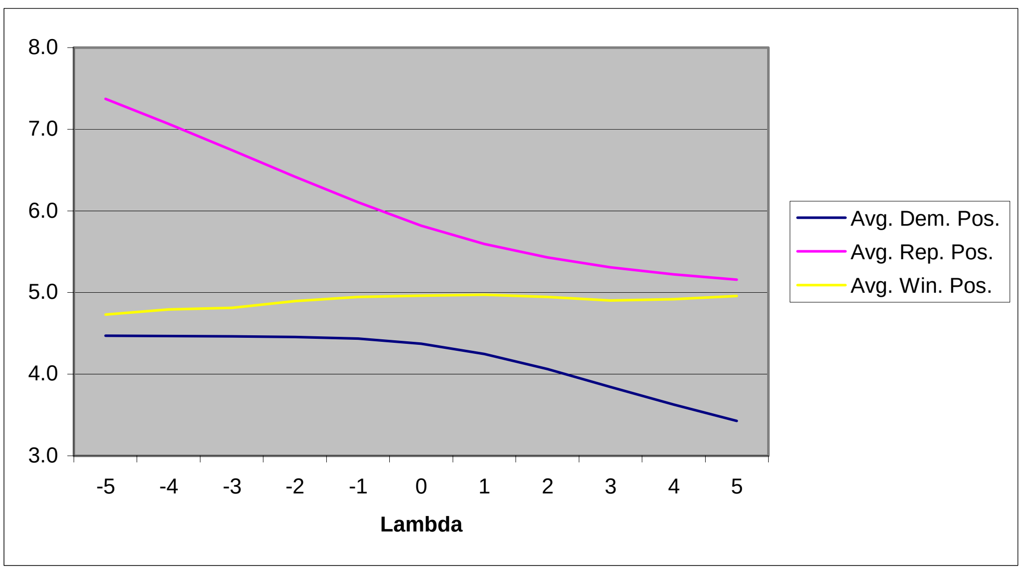


**Figure 11 – Comparative Statics: Candidate Positions vs. Mean Voter Position**