

# Why Unsafe at Any Margin? Incumbency Advantage and Vulnerability \*

Kentaro Fukumoto<sup>†</sup>

March 27, 2008

## Abstract

In estimating incumbency advantage and campaign spending effect, simultaneity bias is present. In order to solve it, my model explicitly takes into account “analyst’s error” which analysts do not know but players know. Estimation by Markov Chain Monte Carlo, especially data augmentation, enables us to integrate analyst’s error out and employ a non closed-form likelihood function, which is the joint distribution of the five endogenous variables: vote margin, both parties’ campaign spending and candidate quality. I derive equilibrium of my game-theoretical model and plug it into my statistical model. As for incumbency vulnerability, standard deviance of vote margin is explained by redistriction, quality of candidates and time trend. I show superiority

---

\*Paper prepared for the 66th Annual Meeting of the Midwest Political Science Association, Chicago, IL, USA, April 3-6, 2008. Its earlier versions were presented at the 23rd Annual Summer Methodology Conference, University of California, Davis, July 20-22, 2006 and the Annual Meeting of the Midwest Political Science Association, Chicago, IL, USA, April 12-15, 2007. I appreciate Gary Jacobson for giving me his data and Ken Shotts as well as Jeff Gill for their discussion on earlier versions of this paper. My thank also goes to Chris Achen, Ken’ichi Ariga, Robert Erikson, Jonathan Katz, Gary King and Walter Mebane Jr. for their comments at the presentation. I express my gratitude to the Japan Society for the Promotion of Science for research grant. Ryota Natori kindly offers me computational resource. This is work in progress. Comments are really welcome.

<sup>†</sup>Professor, Department of Political Science, Gakushuin University, Tokyo, E-mail: First Name dot Last Name at gakushuin dot ac dot jp, URL: [http://www-cc.gakushuin.ac.jp/~e982440/index\\_e.htm](http://www-cc.gakushuin.ac.jp/~e982440/index_e.htm).

of my model compared to a conventional estimator by Monte Carlo simulation. Empirical application of this model to the recent U.S. House election data demonstrates that, as suspected, incumbency advantage is smaller, redistribution increases variance of vote and incumbency decreases it, defender's campaign spending effect is positive, and challenger's campaign spending effect is smaller than previously shown.

## 1 Introduction

Ordinary Americans take it for granted that incumbents have advantage in the U.S. House election and large campaign spending helps them. If this is true, incumbency advantage and campaign spending effect make representatives less vulnerable to electoral pressure and irresponsible to citizen's voice. Existence of campaign spending effect is a cause of the campaign finance reform.

Though, surprisingly, political scientists have trouble in measuring size of incumbency advantage and campaign spending effect because of "simultaneity bias". The logic is as follows. On one hand, when incumbent legislators foresee its defeat, they do not run for reelection. They are strategic. Only incumbents who expect they will win run. As a result, incumbency advantage is overestimated. On the other hand, those incumbents who have poorer electoral prospect need to and do raise and spend more campaign fund but still end up with not so many votes. Thus, it *seems* as if the more campaign contribution lead to the less votes. In this sense, incumbent's campaign spending effect is underestimated. For both aspects, causal direction between vote and incumbency or money is not only from the latter to the former but also in the opposite way. That is why this is called simultaneity bias. The same argument also holds for the challenger party.

Simultaneity bias arises when part of error term and some parameters in the vote model also affect entry decision of candidates and campaign spending of both parties. I call them

stochastic dependence and parametric dependence, respectively.<sup>1</sup> First, to tackle stochastic dependence, I decompose error term into player's error and analyst's error. Players are blind to the former only, while we analysts know neither. My model take analyst's error into account. Estimation by Markov Chain Monte Carlo (hereafter MCMC), especially data augmentation, enables us to integrate analyst's error out and employ a non closed-form likelihood function. Second, to deal with parametric dependence, I use the joint distribution of the five endogenous variables: vote margin, both parties' campaign spending and candidate quality. In order to do it, I take advantage of theories of electoral politics rigorously, construct a game theoretical model, and plug its equilibrium into my statistical model. In this sense, the present paper aims to show empirical implications of theoretical model.

Moreover, this study also pays attention to incumbency vulnerability. Recently, however many votes incumbents won in the previous election, they are not guaranteed certain reelection. In order to answer "why unsafe at any margin," the model examines what explains variance of vote margin.

This paper is organized as follows. The first section explains the setting of the three-stage game, the simultaneous bias problem, previous solutions and outline of my solution. Next, I derive equilibrium of my game-theoretical model and put it into my statistical model. Third, Monte Carlo simulation is demonstrated. The following section will analyze the recent U.S. House election data, 1972-2004, and show that, as suspected, incumbency advantage is smaller, defender's campaign spending effect is larger and positive, and challenger's campaign spending effect is smaller than previously shown. Finally, I conclude.

---

<sup>1</sup>I borrow the word of "parametric dependence" from King (1989, 190-91)

## 2 Simultaneity Bias: Problems and Solutions

### 2.1 Setting

I outline my three-stage dynamic game and introduce my notation of variables. Players are candidates of the defender party  $D$  and the challenger party  $C$ . Each party has a high quality candidate and a low quality candidate. In order to avoid repeating similar equations for both parties, I mean either of them by  $P \in \{D, C\}$  and let  $-P = C$  if  $P = D$  and  $-P = D$  if  $P = C$ .

At the first stage, players are the high quality candidates of each party. They decide to run ( $Q_P(x) = 1$ ) or not ( $Q_P(x) = 0$ ) in general election based on covariates  $x$  such as national tide (dummy of Democrat in each year) and lagged variables. If they do not run, the low quality candidate runs (Banks and Kiewiet, 1989, I do not suppose uncontested elections). For defender, a high quality candidate is equal to incumbent legislator. Even though the word “incumbent” is usually used for party and candidate, this paper uses it only for candidate but not party and distinguishes defender *party* and incumbent *candidate* for clarification of argument. For candidate quality of the challenger party, the electoral studies almost agree to use prior experience of elective office as its proxy (Bianco, 1984; Cox and Katz, 2002; Jacobson and Kernell, 1983). Though this common notation for both parties is not usual, it makes presentation below simpler.

At the second stage, players are every party’s candidate who runs. Party  $P$ ’s candidate decides how much it spends for campaign,  $M_P(Q_P, Q_{-P}, x)$ , after observing both its own quality  $Q_P$  and that of the opponent  $Q_{-P}$ .

At the last stage, there are no strategic players. The voters return the two-party vote

margin of the defender,  $V(Q_P, Q_{-P}, M_P, M_{-P}, x)$ , in the following way:<sup>2</sup>

$$\begin{aligned} V &= \tilde{V} + \varepsilon_V \\ \tilde{V} &= \beta_0 + \beta_{QD}Q_D - \beta_{QC}Q_C + \beta_{MD}M_D - \beta_{MC}M_C + \beta_x x \\ \varepsilon_V &\sim \mathcal{N}(0, \varsigma_V). \end{aligned} \tag{1}$$

where  $\mathcal{N}(\mu, \sigma)$  is normal distribution whose mean is  $\mu$  and standard deviance is  $\sigma$ . The coefficients of  $Q_C$  and  $M_C$  have minus sign because challenger's candidate quality and campaign spending are reasonably expected to have negative impact on defender's vote and this parameterization makes the following equations simpler.

A large letter refers to a variable (e.g.  $Q_P$ ), while a small letter refers to its observed value (e.g.  $q_P$ ).

## 2.2 Problems

### 2.2.1 Incumbency Advantage: $\beta_{QD}$

Today, the canonical estimator of incumbency advantage is Gelman and King (1990)'s (hereafter, GK estimator). They propose to regress defender's vote on incumbent candidate dummy, Republican defender indicator  $R$  (1 if the defender is Republican and  $-1$  if it is Democrat), and lagged vote margin  $V_{t-1}$  (except for which I suppress time subscript  $t$  for easy presentation). That is, in the Eq. (1), they assume  $\beta_{QC} = \beta_{MD} = \beta_{MC} = 0$  and make  $x$  composed of  $(R, V_{t-1})$ .<sup>3</sup>

$$V = \beta_0 + \beta_{QD}Q_D + \beta_R R + \beta_V V_{t-1} + \varepsilon_V$$

---

<sup>2</sup>Since  $V$  is bounded between -50 and 50, you might well transform it by log odds so that it is unbounded. Though, most scholars do not transform vote, arguing that  $V$  falls between -30 and 30 in reality. In order to make my result comparable to previous studies, I also follow the suit. In addition, I assume that the two-party vote margin is independent of the other parties' vote share.

<sup>3</sup>Their original dependent variable is Democrat's vote margin, not defender's. I arrange their expression so that their model fits my notation.

Then, the effect of incumbency status of defender party's candidate,  $\beta_{Q_D}$ , is their estimate of incumbency advantage and it is estimated by least square method.

GK estimator, however, suffers from simultaneity bias, because an incumbent retires strategically (Cox and Katz, 2002; Jacobson and Kernell, 1983). That is, the more optimistic incumbents are about their prospect of vote margin  $V$ , the more likely they are to run ( $Q_D = 1$ ); Otherwise, they will retire ( $Q_D = 0$ ). Therefore, defender's candidate quality  $Q_D$  is endogenous to vote margin  $V$ . Simultaneity between  $V$  and  $Q_D$  comes from stochastic dependence and parametric dependence between them. Below, I will explain them more formally.

Stochastic Dependence. First,  $V$  and  $Q_D$  are not stochastically independent as GK estimator implicitly assumes. I decompose error term  $\varepsilon_V$  into analyst's error  $\varepsilon_{VK}$ , which is known to players but not analysts, and player's error  $\varepsilon_{VU}$ , which is unknown to players and analysts. I assume that both are independent of each other and jointly follow the bivariate normal distribution.<sup>4</sup>

$$\begin{aligned}\varepsilon_V &= \varepsilon_{VK} + \varepsilon_{VU} \\ &\sim \mathcal{N}(0, \varsigma_V) \\ \begin{pmatrix} \varepsilon_{VK} \\ \varepsilon_{VU} \end{pmatrix} &\sim \mathcal{BVN}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \varsigma_{VK}^2 & 0 \\ 0 & \varsigma_{VU}^2 \end{pmatrix}\right) \\ \therefore \varsigma_V &= \sqrt{\varsigma_{VK}^2 + \varsigma_{VU}^2}.\end{aligned}$$

---

<sup>4</sup>According to Signorino (2003),  $\varepsilon_V$  is regressor error and  $\varepsilon_{VU}$  is agent error.

The vote margin players expect is

$$\begin{aligned}
\bar{V} &= \int V \mathcal{N}(\varepsilon_{VU}) d\varepsilon_{VU} \\
&= \int (\tilde{V} + \varepsilon_{VK} + \varepsilon_{VU}) \mathcal{N}(\varepsilon_{VU}) d\varepsilon_{VU} \\
&= \tilde{V} + \varepsilon_{VK}.
\end{aligned}$$

Note that the vote margin analysts (or GK estimator) expect is

$$\int \int V \mathcal{N}(\varepsilon_{VU}) \mathcal{N}(\varepsilon_{VK}) d\varepsilon_{VU} d\varepsilon_{VK} = \tilde{V}.$$

On one hand, the larger  $\varepsilon_{VK}$ , the larger the player's expected vote margin  $\bar{V}$  and, knowing this, the more likely the incumbent is to run ( $Q_D = 1$ ). On the other hand, this does not hold in the case of  $\varepsilon_{VU}$ , because players do not know its value, either. Thus,  $E(Q'_D \varepsilon_{VK}) > 0$  but  $E(Q'_D \varepsilon_{VU}) = 0$ . Therefore, by omitting  $\varepsilon_{VK}$ , GK estimator of  $\beta_{QD}$  is as much biased as the first element of  $E((z'z)^{-1} z' \varepsilon_{VK})$ , where  $z$  is the matrix of all regressors ( $q_D, r, v_{t-1}$ ). Usually, this bias is positive and inflates GK estimate of incumbency advantage  $\hat{\beta}_{QD}$ . If analysts knew as well as players (i.e.,  $\varepsilon_{VK} = 0$ ), there would be no bias. Unfortunately but usually, this does not hold. This formulation makes it clear that simultaneity bias arises when a model is misspecified by omitting the variable  $\varepsilon_{VK}$  which affects the dependent variable  $V$  and a regressor  $Q_D$ . Since stochastic error  $\varepsilon_{VK}$  of  $V$  in the third stage affects  $Q_D$  in the first stage prospectively, not only the probability of  $V$ ,  $p(v|\theta)$ , but also that of  $Q_D$ ,  $p(q_D|\theta)$ , depends on  $\varepsilon_{VK}$ . We should take  $\varepsilon_{VK}$  into consideration of our model of  $V$  and  $Q_D$ .

Parametric Dependence. Second,  $V$  and  $Q_D$  are not parametrically independent as GK estimator implicitly assumes. The larger incumbency advantage  $\beta_{QD}$ , the wider vote margin  $V$  the defender obtains and the more likely an incumbent is to run for reelection,  $Q_D = 1$ . Since parameters like  $\beta_{QD}$  of  $V$  in the third stage also affects  $Q_D$  in the first stage

prospectively, not only the likelihood of  $v$ ,  $L(v|\theta)$ , but also that of  $q_D$ ,  $L(q_D|\theta)$ , depends on  $\beta_{QD}$  ( $\theta$  is the parameter set). When we estimate  $\beta_{QD}$ , say, by maximizing likelihood or MCMC, we should use likelihood of both  $v$  and  $q_D$ ,  $L(v, q_D|\beta_{QD})$ .

### 2.2.2 Challenger Candidate’s Quality Effect: $\beta_{QC}$

The above argument also holds for high quality challenger’s effect on vote ( $\beta_{QC}$ ). The challenger is also a strategic player. The smaller  $\epsilon_{VK}$  or the larger  $\beta_{QC}$ , the smaller the defender’s vote margin  $V$  (Bond, Covington, and Fleisher, 1985; Green and Krasno, 1988; Jacobson and Kernell, 1983) and, therefore, a strong candidate of the challenger party ( $Q_C = 1$ ) is more likely to run.  $E(Q'_C \epsilon_{VK}) < 0$  and  $\beta_{QC}$  is also likely to be overestimated.

### 2.2.3 Campaign Spending Effect: $\beta_{MP}$

Campaign spending effect  $\beta_{MP}$  is crucial, though its measurement is controversial. Jacobson (1989, 1990) reports that challenger’s campaign spending diminishes defender’s vote  $V$  ( $\beta_{MC} > 0$ ), while defender’s has no effect ( $\beta_{MD} = 0$ ). Since then, a lot of scholars have tried to find that defender’s war chest also matters (Erikson and Palfrey, 1998, 2000; Goidel and Gross, 1994; Green and Krasno, 1988; Kenny and McBurnet, 1994; Levitt, 1994).

The relationship between  $V$  and  $M_P$  is also contaminated with stochastic dependence and parametric dependence, though it is not as straight-forward as that between  $V$  and  $Q_P$ . Suppose that the more money candidates spend, the more votes they receive. Unlike the case of candidate quality, an effect of expected vote on campaign spending depends on not its level but its closeness or competitiveness. On one hand, when they foresee vote margin is nearly 0, they definitely need to expend more. On the other hand, when they are almost sure to win or lose, marginal increase of votes by additional spending is not worth its cost for strategic contributors and candidates (Jacobson and Kernell, 1983). Erikson and Palfrey (2000, 599) formally show that “equilibrium candidate spending should be proportional to

the normal density of the expected incumbent margin of victory.” Accordingly, when  $V > 0$ , the larger  $\epsilon_{VK}$  or the larger  $\beta_{MP}$ , the larger  $V$  and, therefore, the smaller  $M_P$ . Since usually  $E(M_P \epsilon_{VK}) < 0$ ,  $\beta_{MD}$  tends to be underestimated and  $\beta_{MC}$  tends to be overestimated (as many scholars suspect).

Besides, simultaneity also exists between  $Q_P$  and  $M_P$ .

## 2.3 Previous Solutions

So far, scholars have tried to solve stochastic dependence but it is difficult. As I mentioned above, the relation between  $V$  and  $Q_P$  is typical sample selection situation. Heckman (1974)’s sample selection model is, however, unavailable due to exclusion restriction because the same covariates should affect both (Sartori, 2003).

The most common method is to employ instrumental variable (Erikson and Palfrey, 1998; Green and Krasno, 1988; Kenny and McBurnet, 1994). To find appropriate instrumental variable itself is, however, problematic task. Goidel and Gross (1994) model system of four equations ( $V, Q_C, M_P$ ) simultaneously by three-stage least square. A problem of their model is failure to take into consideration expectation of endogenous variables. For example, they do not include expected vote into the equation of candidate quality. Since their equations share some covariates but not parameters, their model implicitly assume parametric independence.

Another way is to utilize natural experiment. Levitt (1994) and Levitt and Wolfram (1997) examine elections where the same two candidates face one another on more than one occasion to control all time invariant district specific features and candidate specific ones, observed or unobserved or unobservable. But this does not control time varying random shocks. Ansolabehere, Snyder, and Stewart (2000) and Desposato and Petrocik (2003) use redistriction as natural experiment. An incumbent should not enjoy personal vote in the area which was not the incumbent’s previous district (the new voters). Difference the vote among the new voters and that among the old ones is an estimate of incumbency advantage.

Their method does not, however, capture the part of incumbency advantage which is not due to personal vote, such as experience in the Capitol Hill. Cox and Katz (2002) pay attention to a non incumbent's vote in such a district where the incumbent fails to run involuntarily (namely, not for electoral reason) because it is a good estimate of the vote the incumbent would receive if it ran as non incumbent. But it is difficult to judge whether the incumbent retires voluntarily or not. Erikson and Palfrey (2000) and Lee (forthcoming) focus on districts where the previous competition nearly 50-50, because candidates are not sure which will win this time and their expectation does not affect decision of running and campaign spending. Though these natural experiment methods are interesting, estimation using limited observations sacrifices efficiency of estimation and may lead to estimate which is different from the average incumbency advantage.

To my knowledge, few works consider parametric dependence.

## 2.4 My Solution

The previous studies try to solve the two problems by erasing them. The present paper considers that they are political mechanisms of interest and should be modeled, not avoided. First, to tackle stochastic dependence, I include the previously excluded variable  $\varepsilon_{VK}$  in my model as if it is observed and integrate it out in estimation process. As I will explain shortly, this can be possible by data augmentation in MCMC. Second, to deal with parametric dependence, I make much of the joint probability function of the five endogenous variables  $(v, m, q)$  instead of their five separate marginal probability functions.

I denote  $m = (m_D, m_C)$  and  $q = (q_D, q_C)$ . The joint probability function of the five

endogenous variables  $(v, m, q)$  conditioned on covariates  $(x)$  and parameters  $(\theta)$  is

$$\begin{aligned} p(v, m, q|x, \theta) &= \int p(v, m, q, \varepsilon_{VK}|x, \theta) d\varepsilon_{VK} \\ &= \int p(v|m, q, x, \varepsilon_{VK}, \theta) p(m|q, x, \varepsilon_{VK}, \theta) p(q|x, \varepsilon_{VK}, \theta) p(\varepsilon_{VK}|x, \theta) d\varepsilon_{VK} \quad (2) \end{aligned}$$

Since the whole three-stage game is dynamic, equilibrium should be subgame perfect and I will consider each stage backward in the next section. Games at the first and second stages will be constructed as static games. I will also use equilibrium of my game theoretic model as conditional expectation values of the five endogenous variables  $Q_P$ 's,  $M_P$ 's and  $V$  in my statistical model. This connection between the game theoretic model and the statistical model will illustrate empirical implications of this theoretical model.

## 3 Model

### 3.1 Vote Margin: $V$

#### 3.1.1 Normal Vote Margin

Analysts usually control “normal vote margin” as baseline, that is, the partisan vote the defender would have in the district if all explanatory variables (including the constant term but excluding the party indicator) had no effect. Which measurement to use as the normal vote margin is, however, a controversial issue. An usual proxy is lagged vote (Cox and Katz, 2002; Gelman and King, 1990); some may use presidential vote or vote for other electoral offices in the same district; others calculate their mean for a decade (Bond, Covington, and Fleisher, 1985; Ansolabehere, Snyder, and Stewart, 2000). I advocate for lagged vote, not just because it well explains the current election, but because the lagged dependent variable conveys unmeasured information.

I assume the sign corrected first order autoregressive (AR(1)) error process:

$$\begin{aligned}\varepsilon_{V,t} &= \delta I(V_{t-1})\varepsilon_{V,t-1} + \epsilon_{V,t} \\ \epsilon_{V,t} &\stackrel{iid}{\sim} \mathcal{N}(0, \sigma_V = \sqrt{1 - \delta}\varsigma_V) \\ I(z) &= \begin{cases} 1 & \text{if } z \geq 0 \\ -1 & \text{if } z < 0. \end{cases}\end{aligned}$$

where  $0 < \delta < 1$ . If a challenger won in the previous election, it becomes a defender in the current election and not  $\varepsilon_{V,t-1}$  but  $-\varepsilon_{V,t-1}$  shows its vote not explained by the model. That is why sign is corrected by  $I(V_{t-1})$ .  $\epsilon_V$  is unmeasured change of district partisan strength at time  $t$  in the district. Examples are scandals, disasters, entry of a third party, redistricting, and so on. I also assume that the current shock  $\epsilon_{V,t}$  is unpredicted from (i.e. independent of) the past shocks  $\epsilon_{V,s < t}$  and their accumulation  $\varepsilon_{V,t-1}$ , but follows the same normal distribution.

Then,

$$\begin{aligned}V_t &= \tilde{V}_t + \varepsilon_{V,t} \\ &= \tilde{V}_t + \delta I(V_{t-1})\varepsilon_{V,t-1} + \epsilon_{V,t} \\ &= \tilde{V}_t + \delta [I(V_{t-1})(V_{t-1} - \tilde{V}_{t-1})] + \epsilon_{V,t}\end{aligned}\tag{3}$$

This expression makes it clear that  $I(V_{t-1})[V_{t-1} - \tilde{V}_{t-1}]$  measures the normal vote margin: “the partisan vote the defender would have if all explanatory variables had no effect”. The previous vote margin which a challenger Democrat won in the previous open election has different meaning from that which an incumbent candidate of (defender) Democratic party won. Even if both are the same value, the former candidate is expected to be stronger than the latter. Thus, it is preferable to subtract covariates’ effect from the previous vote (see

also Gowrisankaran, Mitchell, and Moro, 2004). For purpose of identification of  $\delta$ ,  $x$  does not include any lagged variables.

Eq. (3) also illustrates that the coefficient of error's autoregressive term,  $\delta$ , is equivalent to that of the lagged vote (and the normal vote margin). As always in AR(1) model, normal vote margin is accumulation of past changes of district partisan strength ( $\epsilon_{V,t}$ ) which are discounted (forgotten) at the rate of  $1 - \delta$  ( $0 < \delta < 1$ ) election by election.

$$I(V_{t-1})[V_{t-1} - \tilde{V}_{t-1}] = \sum_{s=1}^{\infty} \delta^{s-1} \left( \prod_{r=1}^s I(V_{t-r}) \right) \epsilon_{V,t-s}$$

### 3.1.2 Player's Error and Analysts' Error

I decompose error term  $\epsilon_V$  into analysts' error  $\epsilon_{VK}$  and player's error  $\epsilon_{VU}$  in the same way as  $\epsilon_{VK}$  and  $\epsilon_{VU}$ .

$$\begin{aligned} \epsilon_V &= \epsilon_{VK} + \epsilon_{VU} \\ \begin{pmatrix} \epsilon_{VK} \\ \epsilon_{VU} \end{pmatrix} &\sim \mathcal{BVN} \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{VK}^2 & 0 \\ 0 & \sigma_{VU}^2 \end{pmatrix} \right) \end{aligned} \quad (4)$$

The vote margin players (not analysts) expect is

$$\begin{aligned} \bar{V}_t &= \int V_t \mathcal{N}(\epsilon_{VU,t}) d\epsilon_{VU,t} \\ &= \tilde{V}_t + \delta [I(V_{t-1})(V_{t-1} - \tilde{V}_{t-1})] + \epsilon_{VK,t}. \end{aligned}$$

Finally, the conditional probability of  $V$  is (time subscript  $t$  and  $t - 1$  is suppressed for simplicity)

$$V \sim \mathcal{N}(v|\bar{V}, \sigma_{VU}). \quad (5)$$

where  $\bar{V}$  depends on  $m, q, x, \epsilon_{VK}, \beta$  and  $\delta$ .

### 3.1.3 Modeling Variance

This model explains variance of vote margin,  $\sigma_V^2 = \sigma_{VK}^2 + \sigma_{VU}^2$ , in the following way:<sup>5</sup>

$$\sigma_{VK} = \exp(z_{VK}\omega_{VK})$$

$$\sigma_{VU} = \exp(z_{VU}\omega_{VU})$$

What are covariates,  $z_{VU}$  and  $z_{VK}$ , other than constant term? First, Jacobson points out that variance of vote margin year by year. Thus, simply, the calendar year variable (minus 1972) is included in both covariates and their coefficients are supposed to be positive.

Second, redistriction brings new voters to the districts (Ansolabehere, Snyder, and Stewart, 2000). Since  $\epsilon_{VK}$  and  $\epsilon_{VU}$  contains information about new voters,  $\sigma_{VK}^2$  and  $\sigma_{VU}^2$  should be larger in redistriction years than in usual election. Though the data drops redistriction year observations (which ends with 2), the next elections (whose year ends with 4) may still suffer volatility due to redistriction. Therefore, the lag redistriction year dummy variable is used and their coefficients are expected to be positive. This is weak test; if the coefficient of lag redistriction is significantly larger than zero, non lag redistriction year dummy should have stronger effect.

Third, candidate quality of both parties,  $q_D$  and  $q_C$ , are employed in  $z_{VK}$  only but not  $z_{VU}$  because parties do not know  $\sigma_{VU}$ . When the defender party field a new candidate ( $q_D = 0$ ), normal vote delivers insufficient information. It results in negative coefficient of  $q_D$ . On the other hand, coefficient of  $q_C$  will be positive. Low quality challengers ( $q_C = 0$ ) tend to be homogeneously weak, while high quality challengers' ( $q_C = 1$ ) strength is heterogeneous.

---

<sup>5</sup>Nowadays, variance becomes quantity of more interest to political scientists. For review, see Braumoeller (2006).

## 3.2 Campaign Spending: $M_P$

### 3.2.1 Game Theoretical Model

At the second stage, both party candidates decide simultaneously how much they spend for campaign,  $M$ . Since we can not fix the order of their decision, this is a static game and I will take advantage of the Nash equilibrium derived by Erikson and Palfrey (2000).<sup>6</sup> Moreover, since they have already decided their own candidate's quality  $Q_P$  and found the opponent's  $Q_{-P}$  in the first stage, there is neither incomplete nor imperfect information and all distributions, functions and values in this subsection (but not parameters) are conditioned on  $Q$ ,  $x$  and  $\epsilon_{VK}$  and suppressed for notational simplicity.

We obtain party  $P$ 's candidate utility ( $\mathcal{U}_P$ ) by subtracting electoral cost ( $\mathcal{K}_P$ ) from expected benefit of seat, which is benefit of seat ( $\lambda_P$ ) multiplied by the probability to win ( $\mathcal{W}_P$ ), in addition to random utility ( $\epsilon_{UP}$ ) which is independent of  $M$ .

$$\mathcal{U}_P(M) = \mathcal{W}_P(M)\lambda_P - \mathcal{K}_P(M_P) + \epsilon_{UP}$$

The probability for the defender to win is

$$\begin{aligned}\mathcal{W}_D(M) &= Pr(V > 0|M) \\ &= \int_0^\infty \mathcal{N}(v|\bar{V}(M), \sigma_{VU})dv \\ &= \Phi(\bar{V}(M)/\sigma_{VU}).\end{aligned}$$

where  $\Phi$  is the standard normal cumulative probability function. The probability for the challenger to win is

$$\mathcal{W}_C(M) = 1 - \mathcal{W}_D(M).$$

---

<sup>6</sup>Mebane (2000) also constructs a game theoretical model of campaign spending and electoral outcomes and test its empirical implication using the U.S. data.

I suppose that electoral cost is constant value plus quadratic of campaign spending:

$$\mathcal{K}_P(M_P) = \kappa_{P1} + \kappa_{P2}M_P^2.$$

$\kappa_{P2}$  is expected to be positive but is not restricted as such so that we can check whether my estimator works well.

According to the first condition to maximize  $\mathcal{U}_P(M)$  (Erikson and Palfrey, 2000), the Nash equilibrium  $M^*$  should meet the following equation;

$$M_P^* = \frac{\lambda_P \beta_{MP}}{2\sqrt{2\pi\kappa_{P2}\sigma_{VU}}} \varphi(\bar{V}(M^*)/\sigma_{VU}) \quad (6)$$

where  $\varphi$  is the standard normal probability density function.<sup>7</sup>

### 3.2.2 Statistical Model

Since it is probably impossible to solve Eq. (6) for  $M^*$  analytically, I approximate scaled expected vote margin given equilibrium spending  $\bar{V}(M^*)/\sigma_{VU}$  by linear function of pre-spending expected vote margin  $\bar{V}_{M0} = \bar{V}(M = (0, 0))$  and approximate equilibrium spending  $M_P^*$  by  $\bar{M}_P$  in the following way;

$$\begin{aligned} \bar{M}_P &= \gamma_P \times \varphi((\bar{V}_{M0} - \alpha_1)\alpha_2) \\ \gamma_P &= \frac{\lambda_P \beta_{MP}}{2\sqrt{2\pi\kappa_{P2}\sigma_{VU}}} > 0 \end{aligned}$$

$\gamma_P$  is a shape parameter proportional to the maximum amount of spending and is estimated instead of  $\lambda_P$ .  $\alpha_1$  is a scale parameter of  $V$  to indicate which value of pre-spending expected vote margin  $\bar{V}_{M0}$  necessitates campaign spending  $\bar{M}_P$  most. The literature on campaign spending effect almost agrees that a defender and a challenger collect

---

<sup>7</sup>Since Erikson and Palfrey (2000) do not model candidate quality selection, their model does not contain (nor identify)  $\lambda_P$ . As I will show shortly, however, my model makes much of  $Q_P$  and can identify  $\lambda_P$ .

and spend the most money when an election seems to be 50 – 50 competition, namely, the vote margin is 0. Thus, we expect  $\alpha_1 = 0$ .  $\alpha_2$  is a shape parameter to indicate how fast deviance of  $\bar{V}_{M0}$  from  $\alpha_1$  decrease  $\bar{M}_P$ . Since  $\varphi(z) = \varphi(-z)$ , I assume that  $\alpha_2 > 0$  for identification. The above reparameterization makes estimation more efficient. I also assume that we observe the approximate equilibrium spending  $\bar{M}_P$  plus normally distributed error  $\epsilon_{MP}$  as  $M_P$ . Therefore, the conditional probability of  $M_P$  is

$$M_P \sim \mathcal{N}(m_P | \bar{M}_P, \sigma_{MP}). \quad (7)$$

where  $\bar{M}_P$  depends on  $q, x, \epsilon_{VK}, \beta, \delta, \gamma_P$  and  $\alpha = (\alpha_1, \alpha_2)$ .

### 3.3 Quality of Candidate: $Q_P$

#### 3.3.1 Game Theoretical Model

I assume that, at the first stage, the high quality candidates of both parties have random utility and decide simultaneously whether they run ( $Q_P = 1$ ) or not ( $Q_P = 0$ ). Thus, quantal response equilibrium will be derived.<sup>8</sup> In this subsection, all distributions, functions and values (but not parameters) are conditioned on  $x$  and  $\epsilon_{VK}$  and suppressed for notational simplicity.

Static Game. Some researchers formulate choice of candidate as a dynamic game. Banks and Kiewiet (1989) suppose the defender is the first mover, while Carson (2003) assumes that the challenger is the first. But this disagreement about the order of player's turn in the literature shows that it is inappropriate to model the situation as a dynamic game. Moreover, for instance, even if the weak first mover makes a bluff and fields a high quality candidate, it may want to take the would-be third move and back down after the second mover defies the

---

<sup>8</sup>As for quantal response equilibrium, see McKelvey and Palfrey (1995, 1996), and Signorino (1999). (Carson, 2003) apply it to candidate entry game but his game is dynamic, not static.

threat and a high quality candidate runs. Or, the first mover might pick up a low quality candidate but reconsider it if the second mover also chooses a low quality candidate. They may not predict which candidate of the opponent party wins its primary. The bottom line is this: from the previous election to the next, both parties are always changing their minds, expecting the opponent's behavior, namely, strategically. Therefore, I suppose that the first stage is a static game (cf. Lazarus, 2005).

Random Utility. Using  $\gamma_P$  instead of  $\lambda_P$ ,  $P$ 's candidate utility is reparameterized as

$$\mathcal{U}_P(Q, M) = (2\sqrt{2\pi}\kappa_{P2}\sigma_{VU}\gamma_P/\beta_{MP})\mathcal{W}_P(Q, M) - \mathcal{K}_P(M_P(Q)) + \epsilon_{UP}.$$

If  $\beta_{MP} = 0$ , however, we can not evaluate this. Even if not, a computer may not calculate utility numerically in the case of  $\beta_{MP} \simeq 0$ . For fear of that, I rescale  $P$ 's utility as

$$\begin{aligned}\tilde{\mathcal{U}}_P(Q, M) &= \lim_{b \rightarrow +|\beta_{MP}|} b \times \mathcal{U}_P(Q, M) \\ &= (I(\beta_{MP})2\sqrt{2\pi}\sigma_{VU}\gamma_P\mathcal{W}_P(Q, M) - |\beta_{MP}|M_P^2(Q))\kappa_{P2} - |\beta_{MP}|\kappa_{P1} + \tilde{\epsilon}_{UP} \\ \tilde{\epsilon}_{UP} &= \lim_{b \rightarrow +|\beta_{MP}|} b \times \epsilon_{UP}.\end{aligned}$$

Given  $P$ 's opponent  $-P$ 's quality  $Q_{-P}$ , utility of  $P$ 's high quality candidate expects is

$$\int \tilde{\mathcal{U}}_P(Q_P = 1, Q_{-P}, M_P = \bar{M}_P(Q_P = 1, Q_{-P}) + \epsilon_{MP})d\epsilon_{MP}$$

I approximate it by

$$\hat{\mathcal{U}}_P(Q_{-P}) + \hat{\epsilon}_{UP}$$

where

$$\begin{aligned}\hat{U}_P(Q_{-P}) &= \tilde{U}_P(Q_P = 1, Q_{-P}, M = \bar{M}(Q_P = 1, Q_{-P})) \\ \hat{\epsilon}_{UP} &\stackrel{iid}{\sim} \mathcal{N}(0, \sigma_{UP}).\end{aligned}$$

In a static game,  $P$  does not know  $Q_{-P}$ . Thus, conditioned on the probability for the opponent to field a high quality candidate,  $\bar{Q}_{-P}$ , utility of  $P$ 's high quality candidate is

$$\bar{U}_P(\bar{Q}_{-P}) + \hat{\epsilon}_{UP}$$

with

$$\bar{U}_P(\bar{Q}_{-P}) = \bar{Q}_{-P}\hat{U}_P(Q_{-P} = 1) + (1 - \bar{Q}_{-P})\hat{U}_P(Q_{-P} = 0)$$

Best Response.  $P$ 's high quality candidate runs if its expected utility is positive.<sup>9</sup> Thus, its best response is

$$Q_P^*(\bar{Q}_{-P}) = \begin{cases} 1 & \text{if } \bar{U}_P(\bar{Q}_{-P}) + \hat{\epsilon}_{UP} > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Thus, conditioned on  $\bar{Q}_{-P}$ , the best response probability for  $P$  to field high quality candidate is

$$\begin{aligned}\bar{Q}_P^*(\bar{Q}_{-P}) &= Pr(Q_P = 1) \\ &= Pr(\bar{U}_P(\bar{Q}_{-P}) + \hat{\epsilon}_{UP} > 0) \\ &= \Phi(\bar{U}_P(\bar{Q}_{-P})/\sigma_{UP})\end{aligned}$$

---

<sup>9</sup>Admittedly, not all incumbent lawmakers leave House for electoral reasons (Box-Steffensmeier and Jones, 1997; Frantzich, 1978; Kiewiet and Zeng, 1993). Some have ambition for other offices such as senator or governor (Black, 1972; Brace, 1984; Copeland, 1989; Rohde, 1979). Some die. Others retire because they are too old, lose fun, or do not expect be promoted to the leadership (Brace, 1985; Groseclose and Krehbiel, 1994; Hall and Houweling, 1995; Hibbing, 1982; Theriault, 1998). These non electoral reasons are incorporated into random error term.

Quantal Response Equilibrium. When the following equation holds for both  $P = D$  and  $P = C$ , the pair  $(\bar{Q}_D^*, \bar{Q}_C^*)$  is the quantal response equilibrium.

$$\bar{Q}_P^* = \Phi(\bar{U}_P(\bar{Q}_{-P}^*)/\sigma_{UP})$$

When  $\bar{U}_P(\bar{Q}_{-P} = 1) < \bar{U}_P(\bar{Q}_{-P} = 0)$ ,

$$\frac{\partial \bar{Q}_P^*}{\partial \bar{Q}_{-P}} < 0 \quad \text{and} \quad 0 \leq \bar{Q}_P^*(\bar{Q}_{-P} = 1) < \bar{Q}_P^*(0 < \bar{Q}_{-P} < 1) < \bar{Q}_P^*(\bar{Q}_{-P} = 0) < 1$$

when  $\bar{U}_P(\bar{Q}_{-P} = 1) > \bar{U}_P(\bar{Q}_{-P} = 0)$ ,

$$\frac{\partial \bar{Q}_P^*}{\partial \bar{Q}_{-P}} > 0 \quad \text{and} \quad 1 > \bar{Q}_P^*(\bar{Q}_{-P} = 1) > \bar{Q}_P^*(1 > \bar{Q}_{-P} > 0) > \bar{Q}_P^*(\bar{Q}_{-P} = 0) \geq 0$$

Therefore, this equilibrium must exist and be unique.

### 3.3.2 Statistical Model

It is probably impossible to solve these equations for  $\bar{Q}_P^*$ 's analytically. Thus, I approximate it by  $\bar{Q}_P^{**}$  which is a linear function of  $\bar{Q}_{-P}^{**}$ :

$$\bar{Q}_P^{**} = \bar{Q}_P^*(0) - (\bar{Q}_P^*(0) - \bar{Q}_P^*(1))\bar{Q}_{-P}^{**}$$

When one solves the system of this equation for  $P = D$  and that for  $P = C$ , one obtains

$$\bar{Q}_P^{**} = \frac{\bar{Q}_P^*(0) - (\bar{Q}_P^*(0) - \bar{Q}_P^*(1))\bar{Q}_{-P}^*(0)}{1 - (\bar{Q}_P^*(0) - \bar{Q}_P^*(1))(\bar{Q}_{-P}^*(0) - \bar{Q}_{-P}^*(1))}$$

For numerical reason, if  $\bar{Q}_P^{**} < 0.01$ , I coerce  $\bar{Q}_P^{**} = 0.01$ . Similarly, if  $\bar{Q}_P^{**} > 0.99$ , I redefine  $\bar{Q}_P^{**} = 0.99$ . From above, the conditional probability of  $Q_P$  is the following Bernoulli

distribution:

$$Q_P \sim \mathcal{B}(q_P | \bar{Q}_P^{**}). \quad (8)$$

where and  $\bar{Q}_P^{**}$  depends on  $x, \epsilon_{VK}, \beta, \delta, \gamma, \alpha, \kappa, \sigma_{VU}$  and  $\sigma_{UP}$ , where  $\gamma = (\gamma_D, \gamma_C), \kappa = (\kappa_{D1}, \kappa_{C1}, \kappa_{D2}, \kappa_{C2})$ .

## 4 Estimation

Eqs. (5), (7) and (8) at the end of each subsection of the previous section give conditional probabilities of the five endogenous variables  $V, M_D, M_C, Q_D$  and  $Q_C$ . Eq. (4) offers  $\epsilon_{VK}$ 's probability. These compose their joint probability in Eq. (2), which does not have closed-form and is difficult to maximize. Thus, I employ MCMC.

So far, I treat  $\epsilon_{VK}$ 's as if they were observed. In fact, however, they are not. Rather, they are parameters to be estimated. Thus, I sample  $\epsilon_{VK}$ 's in MCMC. To integrate  $\epsilon_{VK}$  out, I just ignore their draws. This method is called data augmentation.

I reparameterize some parameters. I estimate logarithm of parameters which are positive values (denoted by, say,  $\dot{\sigma} = \log(\sigma)$ ) and log odds of parameters which range between 0 and 1 (denoted by, e.g.,  $\ddot{\delta} = \log(\delta/(1 - \delta))$ ) so that their parameter space is unbounded and it is easy to propose candidate values by symmetric proposal (normal) distribution. In order to identify  $\kappa$ ,  $\sigma_{UP}$  is assumed to be 1. Thus, the parameter set to be estimated is

$$\theta = (\beta, \ddot{\delta}, \acute{\alpha}, \acute{\gamma}, \kappa, \dot{\sigma}_{VU}, \dot{\sigma}_{VK}, \dot{\sigma}_{MD}, \dot{\sigma}_{MC})$$

where  $\beta = (\beta_0, \beta_{QD}, \beta_{QC}, \beta_{MD}, \beta_{MC}, \beta_x), \acute{\alpha} = (\alpha_1, \alpha_2)$ .

According to Bayes theorem, the posterior distribution is

$$p(\theta|v, m, q, x) \propto p(v, m, q|x, \theta)p(\theta)$$

As already noted, likelihood function  $p(v, m, q|x, \theta)$  is given by Eq. (2) which is calculated using Eqs. (4), (5), (7) and (8). Prior probability of each parameter is *a priori* independent of each other. Their joint distribution  $p(\theta)$  is

$$\begin{aligned}
p(\theta) = & \mathcal{M}\mathcal{V}\mathcal{N}(\beta|b, B) \times \mathcal{N}(\ddot{\delta}|d, D) \\
& \times \mathcal{N}(\alpha_1|a_1, A_1) \times \mathcal{N}(\dot{\alpha}_2|a_2, A_2) \\
& \times \mathcal{N}(\dot{\gamma}_D|g_D, G_D) \times \mathcal{N}(\dot{\gamma}_C|g_C, G_C) \\
& \times \mathcal{N}(\kappa_{1D}|k_{1D}, K_{1D}) \times \mathcal{N}(\kappa_{1C}|k_{1C}, K_{1C}) \\
& \times \mathcal{N}(\dot{\kappa}_{2D}|k_{2D}, K_{2D}) \times \mathcal{N}(\dot{\kappa}_{2C}|k_{2C}, K_{2C}) \\
& \times \mathcal{N}(\dot{\sigma}_{VU}|s_{VU}, \Sigma_{VU}) \times \mathcal{N}(\dot{\sigma}_{VK}|s_{VK}, \Sigma_{VK}) \\
& \times \mathcal{N}(\dot{\sigma}_{MD}|s_{MD}, \Sigma_{MD}) \times \mathcal{N}(\dot{\sigma}_{MC}|s_{MC}, \Sigma_{MC})
\end{aligned}$$

For every single parameter, I derive its full conditional probability density (Gibbs sampling) and sample values from it by Metropolis-Hastings sampling.

As for  $\epsilon_{VK}$ 's, one computational note is in order. I index each observation by subscript  $i$ . Its full conditional probability density is

$$\begin{aligned}
& p(\epsilon_{VK, i}|\theta) \times p(v_i|m_i, q_i, x_i, \epsilon_{VK, i}, \theta) \\
& \times p(m_{D, i}|q_i, x_i, \epsilon_{VK, i}, \theta) \times p(m_{C, i}|q_i, x_i, \epsilon_{VK, i}, \theta) \\
& \times p(q_{D, i}|x_i, \epsilon_{VK, i}, \theta) \times p(q_{C, i}|x_i, \epsilon_{VK, i}, \theta)
\end{aligned}$$

This does not depend on the current values of the other  $\epsilon_{VK, -i}$ 's. Thus, I decline to sample each candidate *scalar*  $\tilde{\epsilon}_{VK, i}$   $N$  (the number of observations) times. Instead, as a more efficient method, I sample a candidate *vector*  $\tilde{\epsilon}_{VK}$  once from the multivariate normal proposal density based on the current vector  $\epsilon_{VK}$ ,  $\mathcal{M}\mathcal{V}\mathcal{N}(\tilde{\epsilon}_{VK}|\epsilon_{VK}, \Sigma_{\epsilon_{VK}})$ , where the  $i$ th element of  $\Sigma_{\epsilon_{VK}}$ 's diagonal is the  $i$ th jumping width and all off diagonals are equal to 0. Then, I decide

to accept or reject each  $\tilde{\epsilon}_{VK, i}$  *separately*. Though it does not change densities analytically, this trick saves frequency of sampling.

## 5 Monte Carlo Simulation

I perform Monte Carlo simulation to study how much simultaneity bias contaminates a conventional estimator. I use the following linear model with all independent variables and their sign corrected lags as a conventional model and estimate parameters by maximum likelihood (this is better than no lag model).

$$\begin{aligned} V_t = & \beta_0 + \beta_{QD}Q_{D, t} - \beta_{QC}Q_{C, t} + \beta_{MD}M_{D, t} - \beta_{MC}M_{C, t} + \beta_{x1}x_t \\ & + \delta I(v_{t-1})V_{t-1} + \beta_{QDL}I(v_{t-1})Q_{D,t-1} - \beta_{QCL}I(v_{t-1})Q_{C,t-1} \\ & + \beta_{MDL}I(v_{t-1})M_{D,t-1} - \beta_{MCL}I(v_{t-1})M_{C,t-1} + \beta_{x1}I(v_{t-1})x_t + \epsilon_{V, t} \end{aligned}$$

Note that the coefficients of  $Q_C$  and  $M_C$  have minus sign so that this model is comparable to my model. I make data following my own model. Parameters are set as follows:  $\beta_0 = 10$ ,  $\beta_{QD} = \beta_{QC} = 2$ ,  $\beta_{MD} = \beta_{MC} = 1$ ,  $\beta_R = 0.5$ ,  $\delta = 0.7$ ,  $\alpha = (0, 0.1)$ ,  $\gamma = (20, 15)$ ,  $\kappa = (0.1, 0.2, 0.002, 0.002)$ ,  $\omega_{VU} = (1.4, 0.1)$ ,  $\omega_{VK} = (2, -0.5, -0.25, 0.05)$ ,  $\sigma_{MD} = \sigma_{MC} = 0.1$ . Once, I randomly produce 500 observations of  $q_{P, t-1}$  from binomial distribution,  $m_{P, t-1}$  from gamma distribution and  $x_t$  (one variable) and  $x_{t-1}$  from standard normal distribution. Using them, I calculate  $\bar{V}_{t-1}$  and sample  $V_{t-1}$  once. Then, I make 178 sets of the five endogenous variables  $(V_t, Q_{P, t}, M_{P, t})$ . For every data set, I estimate parameters by my model and conventional model.

When it comes to the conventional model, I calculate maximum likelihood estimates (MLEs) for every data set. Then, their mean and standard deviance across data sets are reported in Table 1. RMSEs are calculated for every data set by squared difference between the true values and MLEs plus squared standard error. Their average values are shown in

Table 1. How often the true values are within 95% confidence interval is indicated in the column of 95% coverage.

	True	Conventional Model MLE				MCMC Mean			
		Mean	SD	RMSE	Cover	Mean	SD	RMSE	Cover
Const.	10.000	21.462	2.021	140.167	0.000	16.667	15.194	8.953	0.079
$\beta_{QD}$	2.000	0.202	0.676	3.978	0.139	-0.460	3.543	3.029	0.180
$\beta_{QC}$	2.000	0.061	0.445	4.152	0.006	1.361	1.962	1.567	0.219
$\beta_{MD}$	1.000	-0.518	1.198	5.352	0.806	0.254	2.552	1.635	0.213
$\beta_{MC}$	1.000	2.223	1.595	6.894	0.911	1.742	3.272	1.879	0.213
$\beta_x$	0.500	0.012	0.232	0.340	0.394	-1.118	23.816	2.622	0.410
$\delta$	0.700	0.023	0.116	0.485	0.000	0.023	0.126	0.684	0.006

Table 1: Results of Monte Carlo Simulation

In MCMC, I discard 2,500,000 draws as burn-in. For each parameter, I adapt jumping width comparing acceptance rate of the last 100 draws against the benchmark of 44% during the whole burn-in period. After that, I use every ten draw (thinning) from the last 500,000 draws as 50,000 samples from posterior distribution of parameters.<sup>10</sup> Unfortunately, convergence does not seem to be achieved. Though, due to time constrain, this paper reports the current results of my study. As point estimates of my model, mean of sample draws are stored for every data set and their mean and standard deviance across data sets are shown in Table 1. Root mean squared errors (RMSEs) are calculated for every data set and their average values are shown in Table 1. In addition, 95% coverage is indicated.

An important result is that the conventional model underestimates defender’s campaign spending effect ( $\beta_{MD}$ ) and overestimate challenger’s ( $\beta_{MC}$ ), which also supports the common concern. My estimates of defender’s spending effect ( $\beta_{MD}$ ) is not only larger than that of the conventional model but also positive. For five of seven coefficients reported, my model has smaller RMSEs than conventional model.

<sup>10</sup>For 53 datasets, after burn-in period of 1,000,000 draws, every two draw from the last 500,000 draws is used as 250,000 samples. For 37 datasets, after burn-in period of 500,000 draws, every draw from the last 500,000 draws is used.

Since the data is generated according to my model, it is no wonder if my estimator works better than the conventional model. The purpose of this comparison is to show how much of simultaneity bias the conventional estimator produces when stochastic and parametric dependence exists among endogenous variables but they are not taken into account.

## 6 Empirical Analysis of the U.S. Data

### 6.1 Data

I use the U.S. House election data, 1972 to 2004, made by Gary Jacobson.<sup>11</sup> I delete observations which measures elections just after redistriction or in the year ending in 2, contain any missing value or do not have one major party defender candidate and one challenger candidate. The number of observations is 3928.

Endogenous variables are:

- *Vote* ( $V$ ): The defender's two-party vote share in percentage terms.
- *Defender's Quality* ( $Q_D$ ): A dummy variable of incumbent candidate.
- *Challenger's Quality* ( $Q_C$ ): A dummy variable which indicates whether the candidate has held elective office or not.
- *Defender's Spending* ( $M_D$ ): Defender's expenditures. The unit is \$10,000,000.
- *Challenger's Spending* ( $M_C$ ): Challenger's expenditures. The unit is \$10,000,000.

Exogenous Variables ( $x$ ) are:

- *Democrat*: A dummy variable which indicates whether the defender party is Democrat or not.

---

<sup>11</sup>Gary Jacobson kindly gave me his data. I appreciate him.

- *Constant.*

Variance model covariates ( $z$ ) are ( $z_{VU}$  and  $z_{VU}$ ):

- *Lag Redistriction:* A dummy variable which indicates whether the election is the second one since redistriction (in the year ending in 4) or not.
- *Year:* Calendar year number minus 1972.
- *Defender's Quality ( $Q_D$ ).*
- *Challenger's Quality ( $Q_C$ ).*
- *Constant.*

## 6.2 Results

In MCMC, I discard 7,500 draws as burn-in. For each parameter, I adapt jumping width comparing acceptance rate of the last 100 draws against the benchmark of 44% during the whole burn-in period. After that, I store 7,500 samples from posterior distribution of parameters.

### 6.2.1 Effects on Vote ( $\beta$ , $\delta$ , $\epsilon_{VK}$ , $\sigma_{VK}$ and $\sigma_{VU}$ )

To make clear how different my model is from previous ones, Table 2 compares my estimates (the third and fourth columns) with those of the conventional model I used in the Monte Carlo section (the first two columns). As point estimates of my model, mean of sample draws are reported. The last four columns demonstrate diagnosis statics of MCMC convergence: Geweke's Z score (preferable if less than 1.96), Heidel's stationary test p-value (preferable if more than 0.05), autocorrelation of 50th lag and effective size of chains. Though it is not yet confirmed that MCMC chains converge, this paper reports the result as tentative analysis.

As suspected, my estimates of candidate quality effects ( $\beta_{QP}$ ) and challenger's spending effect ( $\beta_{MC}$ ) are smaller than those of the conventional model. My estimate of defender's spending effect ( $\beta_{MD}$ ) is positive (which is reasonable), while that of the conventional model is negative. Moreover, most of standard errors of my model are narrower than the conventional model. Since my MCMC chain of  $\beta_{MC}$  is stacked, its result is not reliable yet.

Model	Conventional		MCMC		Geweke	Heidel	Autocorr.	Eff. Size
	MLE	SE	Mean	SE				
$\beta_0$	5.922	0.539	0.890	0.353		0.012	0.890	8.067
$\beta_{QD}$	6.498	0.363	4.872	0.295	-1.379	0.004	0.911	7.444
$\beta_{QC}$	3.435	0.270	3.262	0.254		0.012	0.904	7.275
$\beta_{MD}$	-32.486	3.723	16.511	1.506	-4.891	0.064	0.539	40.673
$\beta_{MC}$	38.843	3.876	-0.001	0.000	-1.357	0.210	0.651	0.000
Democrat	6.656	0.474	5.719	0.585	-13.711	0.020	0.764	22.949
$\delta$	0.573	0.013	0.832	0.011	-2.424	0.142	0.460	60.837

Table 2: The Effects of Endogenous Variables on Vote Margin

### 6.2.2 Effects on Standard Deviance of Vote ( $\omega$ )

Table 3 shows what affects standard error of player's error ( $\sigma_{VK}$ ) and that of analysts' error ( $\sigma_{VU}$ ). The first column is mean of chains and the second is standard error. Since these are difficult to understand, the third column calculates first difference. For example, in an observation in 1972 *without* redistriction,  $\sigma_{VK} = \exp(1.964) = 7.124$ . On the other hand, in an observation in 1972 *with* redistriction,  $\sigma_{VK} = \exp(1.964 + 0.099 \times 1) = 7.865$ . Thus, first difference is  $7.865 - 7.124 = 0.741$ . That is, redistriction increase player's error 0.741 percentage point. The fourth to seventh columns display the same diagnosis statistics of MCMC convergence as in the previous sub-subsection. Though chains are not yet converged, convergence of  $\omega$ 's is better than that of  $\beta$ 's.

As anticipated, redistriction deteriorates both standard errors. Unexpectedly, though, time trend ameliorates them. An incumbent candidate decreases standard error of player's

	Est.	St.Err.	1st Dif.	Geweke	Heidel	Autocorr.	Eff. Size
$\omega_{VU}$							
Constant	1.964	0.039		1.580	0.119	0.434	83.716
Redistrict (Lag)	0.099	0.041	0.741	0.018	0.640	0.230	154.972
Year	-0.020	0.002	-0.143	-3.480	0.062	0.468	73.157
$\omega_{VK}$							
Constant	2.202	0.054		-2.351	0.014	0.691	30.843
$Q_D$	-0.497	0.052	-3.543	4.166	0.077	0.701	29.354
$Q_C$	0.130	0.041	1.252	-3.212	0.084	0.233	185.534
Redistrict (Lag)	0.061	0.028	0.567	-2.042	0.103	0.219	270.431
Year	-0.008	0.001	-0.070	-0.619	0.067	0.304	127.058

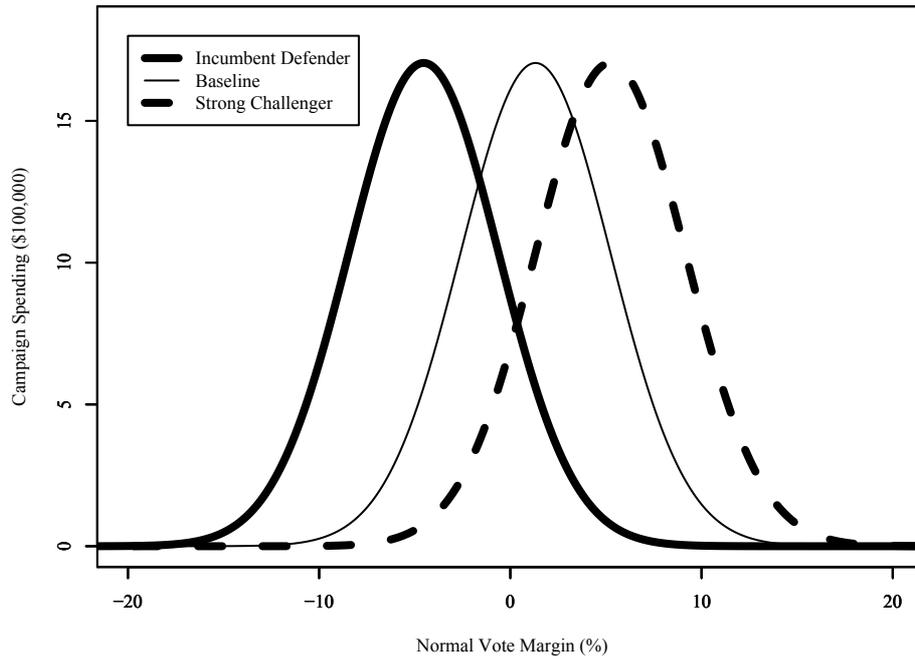
Table 3: The Effects on Standard Deviance of Vote Margin

error, while high quality challenger increases it. Constant terms imply that, when all independent variables are zero,  $\sigma_{VU}^2$  is 50.8 and  $\sigma_{VK}^2$  is 81.7. Therefore, candidates know more than half (61.5%) of what we analysts do not know.

### 6.2.3 Effects on Campaign Spending ( $\gamma$ and $\alpha$ )

Since estimates of parameters themselves are difficult to interpret, I demonstrate their effects by simulation (King, Tomz, and Wittenberg, 2000). Figure 1 displays the relationship between normal vote and both parties' campaign spending when defender is Democrat, lag  $q$  and  $m$  are zeros and  $\epsilon_{VK} = 0$ . In this figure, unit of spending is \$ 10,000. Baseline is the case where both parties field low quality candidates ( $\beta_{QD} = \beta_{QC} = 0$ ). The lines are bell shaped by construction. The more competitive the normal vote margin, the more campaign money each candidate spend.  $\gamma$  decides height,  $\alpha_1$  decides horizontal location, and  $\alpha_2$  decides width. On one hand, bold lines illustrate the case of incumbent against weak challenger ( $\beta_{QD} = 1, \beta_{QC} = 0$ ). Reasonably, this case compensates normal vote margin and the lines move leftward. On the other hand, dotted lines show the case of non incumbent versus strong challenger ( $\beta_{QD} = 0, \beta_{QC} = 1$ ), where normal vote margin is sacrificed and the lines move rightward. All these results are as expected.

(1) Defender's Spending



(2) Challenger's Spending

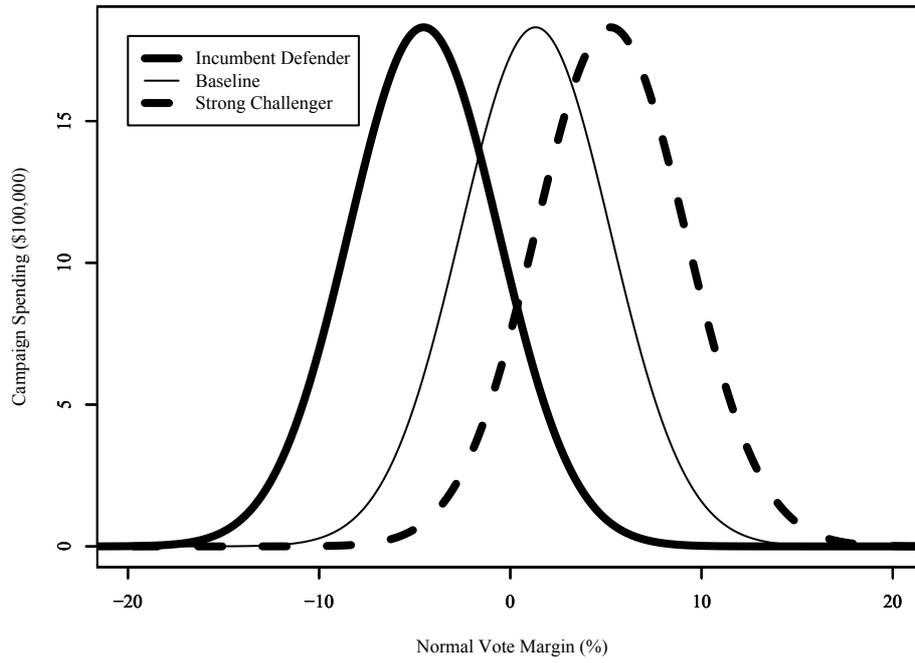


Figure 1: Campaign Spending and Normal Vote

### 6.2.4 Effects on Candidate Quality ( $\kappa$ )

Figure 2 shows the probabilities for high quality candidate to run depending on normal vote size.  $\kappa$  affects the shape of the curve lines. It is clear that, as normal vote becomes smaller, an incumbent hesitates to enter the race and a strong challenger candidate is more willing to run. This is why simultaneity bias occurs.

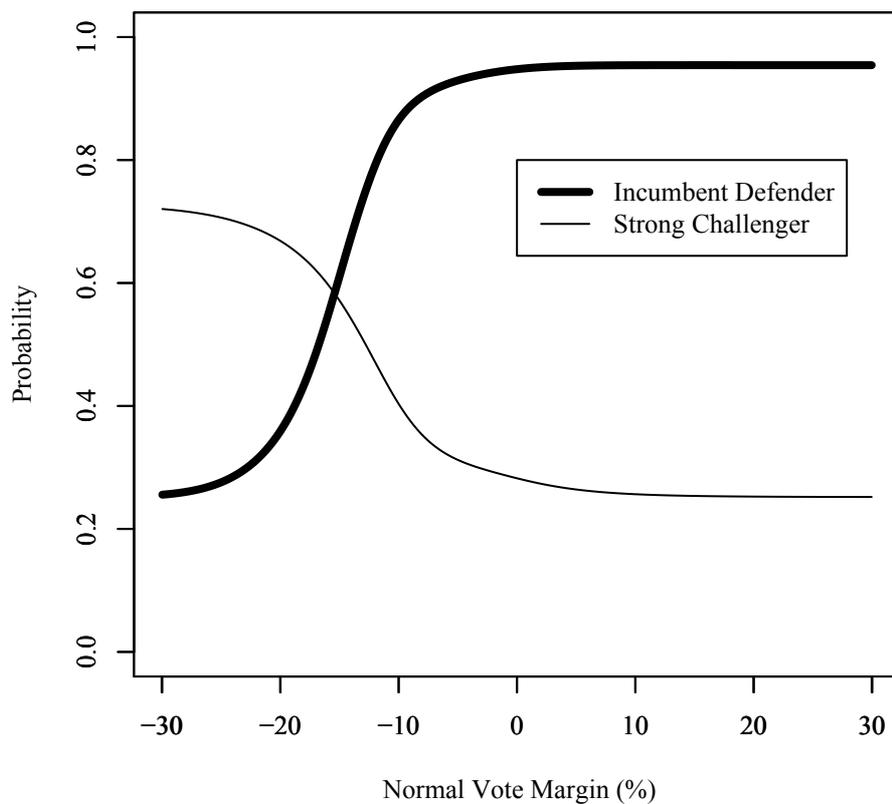


Figure 2: Probabilities for High Quality Candidate to Run

## 7 Conclusion

This paper proposes a solution to simultaneity bias of incumbency advantage and campaign spending. In order to take into account stochastic dependence, I explicitly model analyst's error  $\epsilon_{VK}$ 's and estimate them by data augmentation in MCMC. Through expected vote margin  $\bar{V}(\epsilon_{VK})$ ,  $\epsilon_{VK}$  affects probability of high quality candidate  $\bar{Q}^{**}$  and mean campaign spending  $\bar{M}$ . In order to deal with parametric dependence, I use the joint distribution of all the endogenous variables. I derive equilibrium of my game-theoretical model and plug it into my statistical model. As for incumbency vulnerability, standard deviance of vote margin is explained by redistriction and quality of candidates. I show superiority of my model compared to the conventional estimators by Monte Carlo simulation. Empirical application of this model to the recent U.S. House election data demonstrates that incumbency advantage is smaller than previously shown and that entry of incumbent and strong challenger is motivated by electoral prospect.

Practically speaking, the result of the paper gives readers both hope and concern about American democracy. On one hand, incumbency advantage is smaller and challenger's campaign spending effect is smaller than previously shown. Election is "why unsafe at any margin" even to incumbent and money can not buy sufficient votes. Thus, citizens seem to be powerful enough to make their voice be heard. On the other hand, defender's campaign spending effect is larger and positive. Necessity of campaign finance reform still remains.

I also intend to contribute to electoral studies by redefining the normal vote. My model subtracts effects of lagged variables from the lagged vote to obtain the normal vote margin, because substantial meaning of lagged vote differs depending on how it was fought.

It goes without saying that my model can be applied to any single member district election fought by the two major parties beyond the U.S. Moreover, you can use it in analyzing mixed proportional representation (PR) electoral system. Ferrara, Herron, and Nishikawa (2005)

argue that a party which fields a candidate in a single member district (SMD) has bonus votes in PR tier in that SMD. If you take  $Q_P$  as a dummy of SMD candidate and  $V$  as PR vote share and collapse parties into two major blocs, you can use my model.

This paper assumes incumbency advantage is constant, though it is promising to make it varying, especially with some covariates such as year when the election was held (Gelman and King, 1990) and partisanship (party registration rate, Desposato and Petrocik, 2003). Gelman and Huang (forthcoming) estimate individual incumbency advantage thanks to hierarchical model. Moon (2006) argues that safe incumbent spending is less effective than marginal incumbent spending and campaign spending effect varies with the previous vote margin because the former has fewer votes to buy. These are future agendas to be solved.

## References

- Ansolabehere, Stephen, James M. Snyder, and Jr. Stewart, Charles. 2000. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage." *American Journal of Political Science* 44 (1): 17-34.
- Banks, Feffrey S., and Roderick Kiewiet. 1989. "Explaining Patterns of Candidate Competition in Congressional Elections." *American Journal of Political Science* 33 (4): 997-1015.
- Bianco, William T. 1984. "Strategic Decisions on Candidacy in U.S. Congressional Districts." *Legislative Studies Quarterly* 9 (2): 351-64.
- Black, Gordon. 1972. "A Theory of Political Ambition: Career Choices and the Role of Structural Incentives." *American Political Science Review* 66 (1): 144-59.
- Bond, Jon R., Cary Covington, and Richard Fleisher. 1985. "Explaining Challenger Quality in Congressional Elections." *Journal of Politics* 47 (2): 510-29.
- Box-Steffensmeier, Janet M., and Bradford S. Jones. 1997. "Time Is of the Essence: Event History Models in Political Science." *American Journal of Political Science* 41 (4): 1414-61.
- Brace, Paul. 1984. "Progressive Ambition in the House: A Probabilistic Approach." *Journal of Politics* 46 (2): 556-71.
- Brace, Paul. 1985. "A Probabilistic Approach to Retirement from the U.S. Congress." *Legislative Studies Quarterly* 10 (1): 107-23.
- Braumoeller, Bear F. 2006. "Explaining Variance; Or, Stuck in a Moment We Can't Get Out Of." *Political Analysis* 14 (3): 268-90.
- Carson, Jamie L. 2003. "Strategic Interaction and Candidate Competition in U.S. House Elections: Empirical Applications of Probit and Strategic Probit Models." *Political Analysis* 11: 368-380.
- Copeland, Gary W. 1989. "Choosing to Run: Why House Members Seek Election to the Senate." *Legislative Studies Quarterly* 14 (4): 549-65.
- Cox, Gary W., and Jonathan N. Katz. 2002. *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. Cambridge, UK: Cambridge University Press.
- Desposato, Scott W., and John R. Petrocik. 2003. "The Variable Incumbency Advantage: New Voters, Redistricting, and the Personal Vote." *American Journal of Political Science* 47 (1): 18-32.

- Erikson, Robert S., and Thomas R. Palfrey. 1998. "Campaign Spending and Incumbency: An Alternative Simultaneous Equations Approach." *Journal of Politics* 60 (2): 355-373.
- Erikson, Robert S., and Thomas R. Palfrey. 2000. "Equilibria in Campaign Spending Games: Theory and Data." *American Political Science Review* 94 (3): 595-609.
- Ferrara, Federico, Erik S. Herron, and Misa Nishikawa. 2005. *Mixed Electoral Systems: Contamination and Its Consequences*. New York: Palgrave Macmillan.
- Frantzich, Stephen E. 1978. "Opting Out: Retirement from the House of Representatives, 1966-1974." *American Politics Quarterly* 6 (3): 251-73.
- Gelman, Andrew, and Gary King. 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34 (4): 1142-64.
- Gelman, Andrew, and Zaiying Huang. forthcoming. "Estimating Incumbency Advantage and Its Variation, as an Example of a before-after Study." *Journal of American Statistical Association*.
- Goidel, Robert K., and Donald A. Gross. 1994. "A Systems Approach to Campaign Finance in U.S. House Elections." *American Politics Quarterly* 22 (2): 125-153.
- Gowrisankaran, Gautam, Matthew F. Mitchell, and Andrea Moro. 2004. "Why Do Incumbent Senators Win? Evidence from a Dynamic Selection Model." p. Working Paper 10748. National Bureau of Economic Research.
- Green, Donald Phillip, and Jonathan S. Krasno. 1988. "Salvation for the Spend Thrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections." *American Journal of Political Science* 32 (4): 884-907.
- Groseclose, Timothy, and Keith Krehbiel. 1994. "Golden Parachutes, Rubber Checks, and Strategic Retirements from the 102d House." *American Journal of Political Science* 38 (1): 75-99.
- Hall, Richard L., and Robert P. van Houweling. 1995. "Avarice and Ambition in Congress: Representatives' Decisions to Run or Retire from the U.S. House." *American Political Science Review* 89 (1): 121-36.
- Heckman, James. 1974. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42 (4): 679-694.
- Hibbing, John R. 1982. "Voluntary Retirement from the U.S. House of Representatives: Who Quits?" *American Journal of Political Science* 26 (3): 467-84.
- Jacobson, Gary C. 1989. "Strategic Politicians and the Dynamics of U.S. House Elections, 1946-86." *American Political Science Review* 83 (3): 773-93.

- Jacobson, Gary C. 1990. "The Effects of Campaign Spending in House Elections: New Evidence for Old Arguments." *American Journal of Political Science* 34 (2): 334-362.
- Jacobson, Gary C., and Samuel Kernell. 1983. *Strategy and Choice in Congressional Elections*. 2nd ed. New Haven: Yale University Press.
- Kenny, Christopher, and Michael McBurnet. 1994. "An Individual-Level Multiequation Model of Expenditure Effects in Contested House Elections." *American Journal of Political Science* 88 (3): 699-703.
- Kiewiet, D. Roderick, and Langche Zeng. 1993. "An Analysis of Congressional Career Decisions, 1947-1986." *American Political Science Review* 87 (4): 928-41.
- King, Gary. 1989. *Unifying Political Methodology*. Cambridge: Cambridge University Press.
- King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44 (2): 341-55.
- Lazarus, Jeffrey. 2005. "Unintended Consequences: Anticipation of General Election Outcomes and Primary Election Divisiveness." *Legislative Studies Quarterly* 30 (3): 435-61.
- Lee, David S. forthcoming. "Randomized Experiments from Non-random Selection in U.S. House Elections." *Journal of Econometrics*.
- Levitt, Steven D. 1994. "Using Repeat Challengers to Estimate the Effect of Campaign Spending on Election Outcomes in the U.S. House." *Journal of Political Economy* 102 (4): 777-798.
- Levitt, Steven D., and Catherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U.S. House." *Legislative Studies Quarterly* 22 (1): 45-60.
- McKelvey, Richard D., and Thomas R. Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior* 10: 6-38.
- McKelvey, Richard D., and Thomas R. Palfrey. 1996. "A Statistical Theory of Equilibrium in Games." *The Japanese Economic Review* 47 (2): 186-209.
- Mebane, Jr., Walter R. 2000. "Congressional Campaign Contributions, District Service, and Electoral Outcomes in the United States: Statistical Tests of a Formal Game Model with Nonlinear Dynamics." In *Political Complexity*, ed. Diana Richards. Ann Arbor: University of Michigan Press.
- Moon, Woojin. 2006. "The Paradox of Less Effective Incumbent Spending: Theory and Tests." *British Journal of Political Science* 36: 705-721.

- Rohde, David. 1979. "Risk-Bearing and Progressive Ambition: The Case of Members of the United States House of Representatives." *American Journal of Political Science* 23 (1): 1-26.
- Sartori, Anne E. 2003. "An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restriction." *Political Analysis* 11: 111-138.
- Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *The American Political Science Review* 93 (2): 279-297.
- Signorino, Curtis S. 2003. "Structure and Uncertainty in Discrete Choice Models." *Political Analysis* 11: 316-344.
- Theriault, Sean M. 1998. "Moving Up or Moving Out: Career Ceiling and Congressional Retirement." *Legislative Studies Quarterly* 23 (3): 419-33.