

Bayesian Estimation of Incumbency Advantage and Campaign Spending Effect without Simultaneity Bias *

Kentaro Fukumoto[†]

April 14, 2007

*This is a paper prepared for the annual meeting of the Midwest Political Science Association, Chicago, IL, USA, April 12-15, 2007. Its earlier version was presented at the 23rd Annual Summer Methodology Conference, University of California, Davis, July 20-22, 2006. I appreciate Gary Jacobson for giving me his data and Ken Shotts for his comment on the earlier version of this paper. My thank also goes to Christopher Achen, Robert Erikson, Jonathan Katz, Gary King and Walter Mebane for their comments at the presentation. I express my gratitude to the Japan Society for the Promotion of Science for research grant. This is work in progress. Comments are really welcome.

[†]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>.

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. I show superiority 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, defender’s campaign spending effect is larger and 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.¹ 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

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

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.

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.

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:²

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

where $\mathcal{N}(\mu, \sigma)$ is normal distribution whose mean is μ and standard deviance is σ . 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).

²Since 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.

2.2 Problems

2.2.1 Incumbency Advantage: β_{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}) .³

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

Then, the effect of incumbency status of defender party's candidate, β_{QD} , is their estimate of incumbency advantage and it is estimated by least square.

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 ε_V into analyst's error ε_{VK} , which is known to players but not analysts, and player's error ε_{VU} , which is unknown to players and analysts. I assume that both are independent of each other and jointly follow the bivariate

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

normal distribution.⁴

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

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 ε_{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 ε_{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 ε_{VK} , GK estimator of β_{QD} is as much biased as $E((Q'_D z)^{-1} z' \varepsilon_{VK})$, 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

⁴According to Signorino (2003), ε_V is regressor error and ε_{VU} is agent error.

by omitting the variable ε_{VK} which affects the dependent variable V and a regressor Q_D . Since stochastic error ε_{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 ε_{VK} . We should take ε_{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 β_{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 β_{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 β_{QD} (θ is the parameter set). When we estimate β_{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: β_{QC}

The above argument also holds for high quality challenger's effect on vote (β_{QC}). The challenger is also a strategic player. The smaller ε_{VK} or the larger β_{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. $E(Q'_C \varepsilon_{VK}) < 0$ and β_{QC} is also likely to be overestimated.

2.2.3 Campaign Spending Effect: β_{MP}

Campaign spending effect β_{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 ϵ_{VK} or the larger β_{MP} , the larger V and, therefore, the smaller M_P . Since usually $E(M'_P \epsilon_{VK}) < 0$, β_{MD} tends to be underestimated and β_{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 run 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 ε_{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 (θ) 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 $I(V_{t-1})$. ϵ_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 δ , x does not include any lagged variables.

Eq. (3) also illustrates that the coefficient of error's autoregressive term, δ , 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 ϵ_V into analysts' error ϵ_{VK} and player's error ϵ_{VU} in the same way as ϵ_{VK} and ϵ_{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 δ .

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).⁵ 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 ϵ_{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 (λ_P) multiplied by the probability to win (\mathcal{W}_P), in addition to random utility (ϵ_{UP}) which is independent of M .

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

⁵Mebane (2000) also constructs a game theoretical model of campaign spending and electoral outcomes and test its empirical implication using the U.S. data.

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 Φ is the standard normal cumulative probability function. The probability for the challenger to win is

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

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

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

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

According to 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 φ is the standard normal probability density function.⁶

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

⁶Since Erikson and Palfrey (2000) do not model candidate quality selection, their model does not contain (nor identify) λ_P . As I will show shortly, however, my model makes much of Q_P and can identify λ_P .

M_P^* by \bar{M}_P in the following way;

$$\bar{M}_P = \gamma_P \times \varphi((\bar{V}_{M0} - \alpha_1)\alpha_2)$$

$$\gamma_P = \frac{\lambda_P \beta_{MP}}{2\sqrt{2\pi\kappa_{P2}\sigma_{VU}}} > 0$$

γ_P is a shape parameter proportional to the maximum amount of spending and is estimated instead of λ_P . α_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 and spend the most money when an election seems to be 50 – 50 competition, namely, the vote margin is 0. Thus, we expect $\alpha_1 = 0$. α_2 is a shape parameter to indicate how fast deviance of \bar{V}_{M0} from α_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 ϵ_{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 and \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.⁷ In this subsection, all distributions, functions

⁷As 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.

and values (but not parameters) are conditioned on x and ϵ_{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 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 γ_P instead of λ_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

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

I approximate it by

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

where

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

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

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

Type and Best Response. Note that 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). Unfortunately, we are not sure of whether they leave Congress for electoral reason or not.

I assume that there are two types of P 's high quality candidate, responsive candidate

($T_P = 1$) and irresponsible candidate ($T_P = 0$). The opponent $-P$ and analysts assumes that T_P follows Bernoulli distribution $\mathcal{B}(\tau_P)$. An irresponsible type never runs irrespective of electoral prospect because of death, age and so on: $Q_P(T_P = 0) = 0$. A responsive type runs if its expected utility is positive. Its best response is

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

Thus, conditioned on \bar{Q}_{-P} , the best response probability for P 's responsive type to run is

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

Thus, P 's best response marginal probability to run unconditional on type is

$$\begin{aligned} \bar{Q}_P^*(\bar{Q}_{-P}) &= \tau_P \bar{Q}_P^*(\bar{Q}_{-P}|T_P = 1) + (1 - \tau_P) \bar{Q}_P^*(\bar{Q}_{-P}|T_P = 0) \\ &= \tau_P \Phi(\bar{U}_P(\bar{Q}_{-P})/\sigma_{UP}) \end{aligned}$$

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.

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

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 $x, \epsilon_{VK}, \beta, \delta, \gamma, \alpha, \kappa, \tau, \sigma_{VU}$ and σ_{UP} , where $\gamma = (\gamma_D, \gamma_C, \kappa = (\kappa_{D1}, \kappa_{C1}, \kappa_{D2}, \kappa_{C2}), \tau = (\tau_D, \tau_C)$.

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 and Q . Eq. (4) offers ϵ_{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 ϵ_{VK} 's as if they were observed. In fact, however, they are not. Rather, they are parameters to be estimated. Thus, I sample ϵ_{VK} 's in MCMC. To integrate ϵ_{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 κ , σ_{UP} is assumed to be 1. For computational convenience, this paper assumes $\tau_P = 1$. Thus, the parameter set to be estimated is

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

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

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

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

As for ϵ_{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 ϵ_{VK} , $\mathcal{MVN}(\tilde{\epsilon}_{VK}|\epsilon_{VK}, \Sigma_{\epsilon_{VK}})$, 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. (By mistake, I forgot to multiply lagged variables by $I(v_{t-1})$. Thus, I do not report δ .)

I make data following my own model. Parameters are set as follows: $\beta_0 = 0, \beta_{QD} = \beta_{QC} = 2, \beta_{MD} = \beta_{MC} = 1, \beta_R = 0.5, \delta = 0.7, \alpha = (0, 0.1), \gamma = (20, 15), \kappa = (0, 0, 0.001, 0.001), \sigma_{VU} = 8, \sigma_{VK} = 4, \sigma_{MD} = \sigma_{MC} = 0.1$. 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 19 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.

In MCMC, I discard 5,000 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 five draw (thinning) from 5,000 draws as 1,000 samples from posterior distribution of parameters. 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, medians of sample draws are stored for every data

set and their mean and standard deviation across data sets are reported in Table 1. Root mean squared errors (RMSEs) are calculated for every data set and their average values are shown in Table 1.

	True	My Model Median			Conventional Model MLE		
	Value	Mean	SD	RMSE	Mean	SD	RMSE
β_0	0.000	3.056	2.792	3.609	20.316	5.570	20.492
β_{QD}	2.000	1.991	0.655	0.603	2.089	0.676	1.129
β_{QC}	2.000	2.824	0.631	1.029	1.650	0.693	1.163
β_{MD}	1.000	0.440	0.066	0.595	-0.828	4.299	3.516
β_{MC}	1.000	0.559	0.134	0.485	1.857	7.220	4.217
β_x	0.500	0.528	0.068	0.259	0.343	0.115	0.519

Table 1: Results of Monte Carlo Simulation

When it comes to the conventional model, I calculate maximum likelihood estimates (MLEs) for every data set. Then, their mean and standard deviation 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.

An important result is that the conventional model underestimates defender's campaign spending effect (β_{MD}) and overestimate challenger's (β_{MC}), which also supports the common concern. My estimates of defender's spending effect (β_{MD}) is not only larger than that of the conventional model but also positive. Moreover, it overestimates incumbency advantage (β_{QD}) as suspected. What I did not anticipate is that my estimates of challenger's candidate quality effects (β_{QC}) is larger and more biased than that of the conventional model. For all of them, standard deviances of the conventional model are larger than those of my model. Therefore, the former has smaller RMSEs than the latter.

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.⁸ 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.
- *Constant*.

⁸Gary Jacobson kindly gave me his data. I appreciate him.

6.2 Results

In MCMC, I discard 4,000 draws as burn-in. For each parameter, I adapt jumping width comparing acceptance rate of the last 40 draws against the benchmark of 44% during the whole burn-in period. After that, I use every four draw (thinning) from 4,000 draws as 1,000 samples from posterior distribution of parameters. Computation takes 4 hours and 34 minutes. Unfortunately, convergence does not seem to be achieved. Though, due to time constrain, this paper reports the current results of my study.

6.2.1 Effects on Vote (β , δ , ϵ_{VK} , σ_{VK} and σ_{VU})

To make clear how different my model is from previous ones, Table 2 compares my estimates with those of the conventional model I used in the Monte Carlo section. As point estimates of my model, median of sample draws are reported. As suspected, my estimates of candidate quality effects (β_{QP}) and challenger's spending effect (β_{MC}) are smaller than those of the conventional model. My estimates of defender's spending effect (β_{MD}) is larger than that of the conventional model and positive. Moreover, all standard errors of my model are narrower than the conventional model.

Model	Mine		Conventional		
	Median	SE	MLE	SE	Bias
β_0	1.403	0.137	4.344	0.510	
β_{QD}	4.008	0.059	6.472	0.385	2.239
β_{QC}	2.654	0.065	4.034	0.286	1.444
β_{MD}	16.165	0.670	-22.342	3.656	-16.615
β_{MC}	0.002	0.000	34.353	3.697	17.129
β_x	3.103	0.548	1.546	0.748	
δ	0.794	0.005	0.630	0.015	

Table 2: The Effects of Endogenous Variables on Vote Margin

As I explained in the second section, the conventional estimates should be biased as much as $E((z'z)^{-1}z'\epsilon_{VK})$, where $z = (q, m)$ is the endogenous variable matrix. The fifth column

of bias reports these values, which are average of $(z'z)^{-1}z'\epsilon_{VK}$ over 4,000 sample draws. As for β_{QP} , differences between my model estimates and the conventional estimated is almost equal to this bias value. Thus, we may well consider that most part of simultaneity bias of β_{QP} is caused by stochastic dependence rather than parametric dependence. When it comes to β_{MP} , the direction of bias is positive for β_{MD} and negative for β_{MC} as expected, while the bias size is almost half of discrepancies between the two estimates. Accordingly, I suspect that parametric dependence also bring about simultaneity bias.

The median of σ_{VK} (denoted by $\hat{\sigma}_{VK}$) is 5.7 and $\hat{\sigma}_{VU}$ is 6.1. Therefore, candidates know almost half of what we analysts do not know.

6.2.2 Effects on Campaign Spending (γ and α)

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 all parameters are equal to their median values of my estimates and $\epsilon_{VK} = 0$ and defender is Democrat. 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. γ decides height, α_1 decides horizontal location, and α_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.

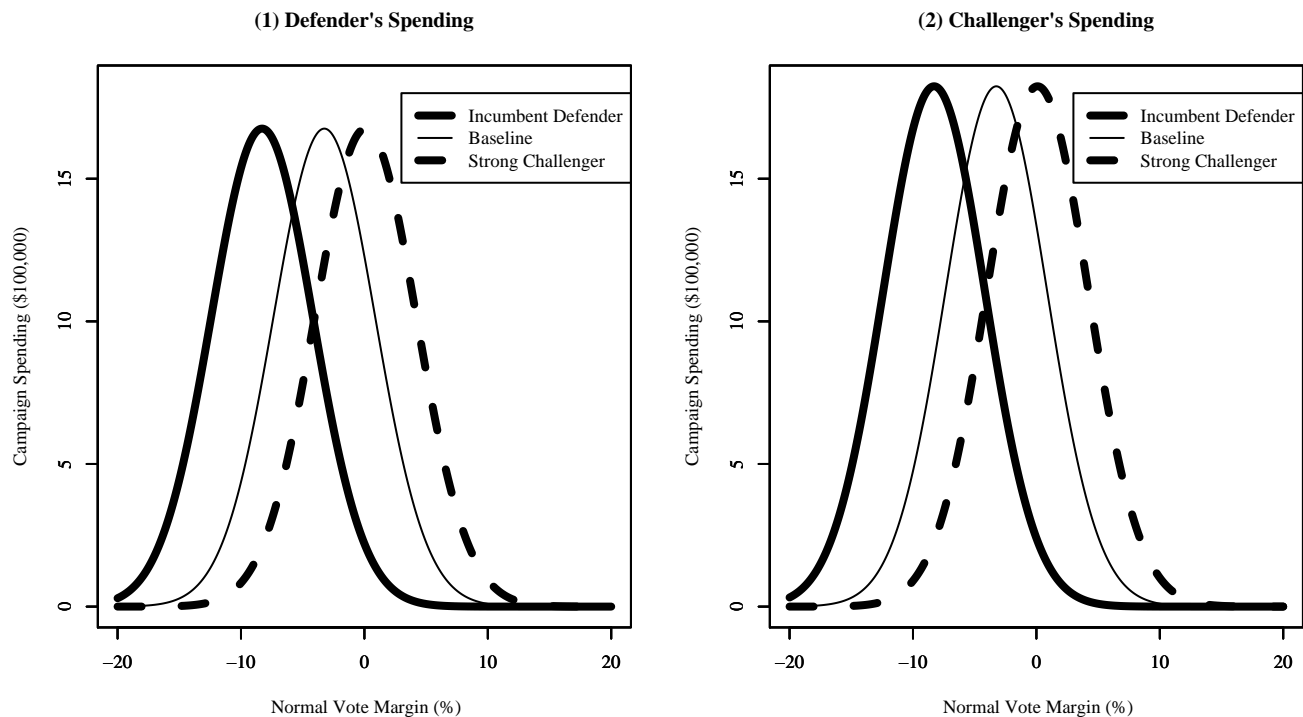


Figure 1: Campaign Spending and Normal Vote

6.2.3 Effects on Candidate Quality (κ)

Figure 2 shows the probabilities for high quality candidate to run depending on normal vote size. κ 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.

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 ϵ_{VK} 's and estimate them by data augmentation in MCMC. Through expected vote

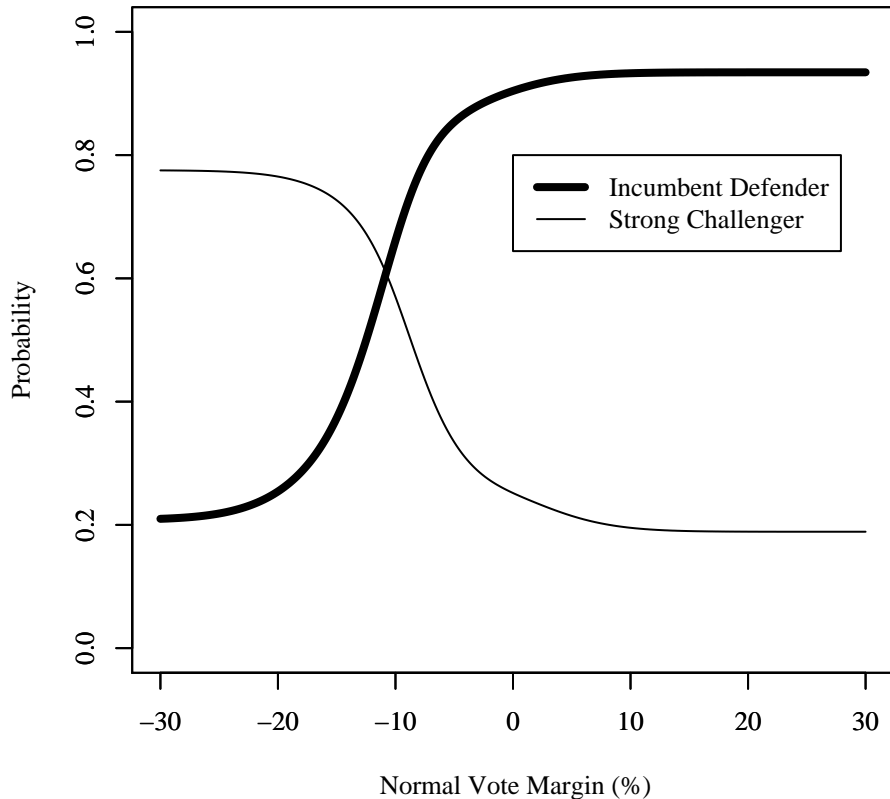


Figure 2: Probabilities for High Quality Candidate to Run

margin $\bar{V}(\epsilon_{VK})$, ϵ_{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. 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 reads 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 no so safe 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

- Ansolahehere, Stephen, James M. Snyder and Jr. Charles Stewart. 2000. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage." *American Journal of Political Science* 44:17–34.
- Banks, Feffrey S. and Roderick Kiewiet. 1989. "Explaining Patterns of Candidate Competition in Congressional Elections." *American Journal of Political Science* 33:997–1015.
- Bianco, William T. 1984. "Strategic Decisions on Candidacy in U.S. Congressional Districts." *Legislative Studies Quarterly* 9:351–64.
- Black, Gordon. 1972. "A Theory of Political Ambition: Career Choices and the Role of Structural Incentives." *American Political Science Review* 66:144–59.
- Bond, Jon R., Cary Covington and Richard Fleisher. 1985. "Explaining Challenger Quality in Congressional Elections." *Journal of Politics* 47:510–529.
- 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* 41:1414–61.
- Brace, Paul. 1984. "Progressive Ambition in the House: A Probabilistic Approach." *Journal of Politics* 46:556–71.
- Brace, Paul. 1985. "A Probabilistic Approach to Retirement from the U.S. Congress." *Legislative Studies Quarterly* 10:107–23.
- 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* 14:549–65.
- Cox, Gary W. and Jonathan N. Katz. 2002. *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. Cambridge: 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* 47:18–32.
- Erikson, Robert S. and Thomas R. Palfrey. 1998. "Campaign Spending and Incumbency: An Alternative Simultaneous Equations Approach." *Journal of Politics* 60:355–373.
- Erikson, Robert S. and Thomas R. Palfrey. 2000. "Equilibria in Campaign Spending Games: Theory and Data." *American Political Science Review* 94:595–609.

- 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* 6:251-73.
- Gelman, Andrew and Gary King. 1990. "Estimating Incumbency Advantage without Bias." *American Journal of Political Science* 34:1142-64.
- 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* 22:125-153.
- Gowrisankaran, Gautam, Matthew F. Mitchell and Andrea Moro. 2004. "Why Do Incumbent Senators Win? Evidence from a Dynamic Selection Model."
- 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* 32:884-907.
- Groseclose, Timothy and Keith Krehbiel. 1994. "Golden Parachutes, Rubber Checks, and Strategic Retirements from the 102d House." *American Journal of Political Science* 38:75-99.
- 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* 89:121-36.
- Heckman, James. 1974. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42:679-694.
- Hibbing, John R. 1982. "Voluntary Retirement from the U.S. House of Representatives: Who Quits?" *American Journal of Political Science* 26:467-84.
- Jacobson, Gary C. 1989. "Strategic Politicians and the Dynamics of U.S. House Elections, 1946-86." *American Political Science Review* 83:773-93.
- Jacobson, Gary C. 1990. "The Effects of Campaign Spending in House Elections: New Evidence for Old Arguments." *American Journal of Political Science* 34:334-362.
- 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* 88:699–703.
- Kiewiet, D. Roderick and Langche Zeng. 1993. "An Analysis of Congressional Career Decisions, 1947-1986." *American Political Science Review* 87:928–41.
- 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* 44:341–55.
- Lazarus, Jeffrey. 2005. "Unintended Consequences: Anticipation of General Election Outcomes and Primary Election Divisiveness." *Legislative Studies Quarterly* 30:435–61.
- 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* 102:777–798.
- Levitt, Steven D. and Catherine D. Wolfram. 1997. "Decomposing the Sources of Incumbency Advantage in the U.S. House." *Legislative Studies Quarterly* 22:45–60.
- 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* 47:186–209.
- Mebane, 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* 23:1–26.
- 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* 93:279–297.

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* 23:419–33.