Campaign Finance in U.S. House Elections

Kei Kawai* Takeaki Sunada†
U.C. Berkeley University of Pennsylvania

September 4, 2015

Abstract

This paper structurally estimates a dynamic model of fund-raising, campaign spending, and accumulation of war chest with unobserved candidate quality. We present an identification strategy similar to the one developed in the context of production function estimation that allows us to recover the quality (vote-getting ability) of the candidates. In our counterfactual experiment, we consider the equilibrium effects of government subsidies to challengers. We find that the subsidies substantially crowd out the challenger's fund-raising and increase the incumbent's fund-raising. Analysis that ignores these equilibrium effects substantially overestimates the effect of the policy.

*Department of Economics, U.C. Berkeley Email: kei@berkeley.edu
†Department of Economics, University of Pennsylvania Email: takeaki@sas.upenn.edu

1 Introduction

How campaign financing shapes electoral competition is a question that has attracted much attention among voters, policy makers and the media. This is natural given the fundamental importance of competitive elections in a representative democracy. To the extent that campaign spending and fund-raising practices impede (or facilitate) electoral competition, they have important implications for democracy. Understanding how money affects electoral competition is also the starting point for an informed debate on campaign finance reform.[1]

In this paper, we study how money affects election outcomes and examine the effects of several campaign finance reforms. While there is a large empirical literature on this topic, previous work has focused almost exclusively on the issue of measurement by estimating the extent to which campaign finance variables affect election outcomes. This paper takes the natural next step: We build an estimable model of campaign spending and campaign war chest that allows us to conduct counterfactual policy experiments. The paper also proposes a new identification and estimation strategy that allows us to back out the "quality" of the candidates, rather than treating them just as nuisance parameters. Our strategy, which is similar to those used in estimating production functions, allows us to address the important empirical challenges that complicate the task of measurement as well.

There are two main empirical challenges that we confront when estimating how money affects election outcomes. The first is the endogeneity of spending and the second is sample selection. Endogeneity of spending refers to the well-known problem that incumbent spending is correlated with unobserved challenger quality because incumbents engage in heavy spending whenever they face strong challengers. This issue was first pointed out in Jacobson (1978), which reports a *negative* raw correlation between incumbent spending and incumbent vote share.

The second empirical challenge is to account for sample selection. Sample selection refers to the fact that contested elections are not randomly selected, but,

[1] For example, Congress has considered legislation for public financing of congressional campaigns in almost every session since 1956. See, e.g., Campaign Finance Reform: Early Experiences of Two States that Offer Full Public Funding for Political Candidates, United States General Accounting Office (May 2003).

rather, occur as a result of challengers' endogenous entry decisions. This means that any variable that affects the vote share of the candidates (including those that seem exogenous – e.g., the state of the economy) will be correlated with the challengers' unobserved quality to the extent that challengers make their entry decisions based on those variables.[2] Hence, almost all control variables in the vote share regression are correlated with the challengers' quality. In contrast to the endogeneity of spending, this issue seems to have received no attention in the previous literature.

Our model of campaign spending and war chest builds on the work of Erikson and Palfrey (2000), in which the authors study a static model of campaign spending and fund-raising with unobserved candidate-specific heterogeneity in quality. While their model captures the key static trade-off in determining spending and fund-raising, they abstract from modeling the candidates' savings decision. An important feature of campaign financing in House elections is that incumbents (and to a lesser extent, challengers) can carry over unspent money from past elections for use in future elections. These savings, called war chests, in addition to providing incumbents with a source of funds for tough future elections, may influence the entry decisions of potential challengers (about 15% of U.S. House elections go uncontested). In order to capture possible dynamic behavior both in the estimation and in the counterfactual, we consider a dynamic extension of the Erikson and Palfrey (2000) model, in the spirit of Ericson and Pakes (1995), by introducing an entry decision and a savings decision while also allowing for persistent unobserved heterogeneity in candidate quality.

In terms of identification, our empirical strategy applies ideas from the production function literature. Our strategy is based on the observation that there is a natural parallel between identification of a vote share equation involving endogenous spending and unobserved candidate quality and identification of a production function involving endogenous inputs and unobserved firm-specific productivity. As in the production function literature, we use a nonparametric function of observable characteristics and actions (see Olley and Pakes, 1996; hereafter OP), as

[2] Imagine a story where only the relatively low-quality challengers find it worthwhile to enter when the economy is good, while all challengers find it worthwhile to enter when the economy is bad. Then the state of the economy will be negatively correlated with challenger quality.

a control function to account for unobserved heterogeneity among candidates. We handle the issue of selection bias – which is important in our setting, as well as in the production function literature – by conditioning on nonparametric estimates of the entry probability and the expected number of Primary challengers. We obtain identification of other primitives of the model, such as the cost of raising money, by exploiting the candidates' first-order conditions, which equate the marginal cost of raising money with the marginal benefit of spending and the marginal benefit of saving.

In addition to being able to handle endogeneity and sample selection, a major advantage of the control function approach is that we can recover the quality – vote-getting ability – of the candidates. Our estimates of candidate quality are counterparts of firm productivity measures in the production function literature. The quality measures we recover have an intuitive meaning: Given two candidates with quality q_1 and q_2, candidate 1 obtains $(q_1 - q_2)$ higher vote share, on average, than candidate 2, holding everything else constant. Given that controlling for candidate quality is an issue that arises in many contexts, we believe that our identification strategy can prove useful for studying other topics in Congressional elections, such as the source of incumbency advantage.

For our estimation, we follow a two-step procedure. In the first step, using the control function, we estimate how spending and other control variables translate to votes. In the second step, we estimate other primitives of the model, such as the cost of raising money. The second step uses moment conditions that are generated by the first-order conditions associated with the optimal choices of spending and saving.

Our estimation results find that a one standard deviation increase in the incumbent's spending yields about a 0.4% increase in her vote share, while a standard deviation increase in the challenger's spending yields about a 3.9% increase in her vote share. Our estimates of candidate quality suggest that heterogeneity among incumbents is quite modest, while heterogeneity among challengers is substantial. We estimate that an incumbent at the 75% quantile of the quality distribution wins a 2.1% higher vote share than does an incumbent at the 25% quantile, everything else equal. For challengers, we estimate that a challenger at the 75% quantile receives about a 7.01% higher vote share, on average, than does a challenger at the 25% quantile. Comparing incumbents and challengers, we find that only the top few percent of the challengers have quality comparable to the incumbents' quality. We find that the quality differential between the median challenger and the median incumbent amounts to a more than 30% vote share differential.

In our counterfactual experiment, we consider the effect of public financing of campaigns. In particular, we consider an introduction of a matching fund for challengers.[3] We find that a policy of matching every dollar that a challenger raises with one dollar of public money has a negligible effect on election outcomes, decreasing the average incumbent vote share by only about 0.13%. This is because the matching fund substantially crowds-out the challengers' fund-raising while modestly increasing the incumbents' fund-raising. Absent these equilibrium responses, however, the effect would be much bigger. We predict that the policy would decrease the incumbents' vote share by about 1.65%, on average, and decrease the probability that incumbents win by about 0.68%, on average. That is, if we do not take into account the candidates' equilibrium response, we would greatly overestimate the efficacy of this policy.

Overall, our estimates of the marginal effect of campaign spending are in line with the results reported in the previous literature.[4] The important difference, however, is that our counterfactual predictions account for the equilibrium responses of the candidates, while the previous work does not.

Literature A large body of empirical work studies the effect of campaign spending on the vote share in the context of Congressional elections. Since the pioneering work of Jacobson (1978), in which he points out the endogeneity of candidate spending and the importance of controlling for candidate quality, much of the subsequent work has revolved around overcoming the endogeneity of spending. The most popular line of attack has been to find appropriate instruments for the

[3] Most of the proposed public financing reforms include a spending cap in exchange for a subsidy. The likely outcome of such a policy is that incumbents opt out and only the challengers opt in. The counterfactual policy we consider approximates this type of policy proposal.

[4] The estimated effects are slightly larger in absolute value than the results reported in Levitt (1994) while they are smaller than Green and Krasno (1988).

spending variables. Some of the instruments that have been used in the literature include challenger wealth (Gerber, 1998), previous incumbent spending (Green and Krasno, 1988), and demographic characteristics in multi-member constituencies (Cox and Thies, 2000).[5]

In contrast to the effort devoted to the issue of endogeneity, however, previous work has not addressed the issue of sample selection. This is an important omission given that whether or not to enter is an endogenous decision by the challenger, which means that the quality of the challengers that enter should depend on variables such as district demographic characteristics and incumbent characteristics – which are often used as exogenous control variables in the vote share regression. Identification of the vote share regression via IV would require finding instruments for all control variables – not just for campaign spending – to the extent that challenger entry decisions depend on them. It appears that the prior literature has not addressed this point. Rather than attempting to find instruments for all independent variables in the vote share regression, our paper proposes a control function approach that explicitly accounts for both endogeneity and selection.

In addition to the large body of previous work that relies on instruments, there are two papers that do not take the IV approach and are particularly relevant to our paper: The first paper is Levitt (1994), which looks at repeat challengers and uses pairs of elections in which the same incumbent-challenger pair ran against each other to difference out candidate fixed effects.[6] While our identification strategy is different from Levitt's, we include unobserved heterogeneity of candidates in the vote share function in the same way and extensively exploit the panel structure of the data for identification.

The second related paper is Erikson and Palfrey (2000), which models the candidates' fund-raising and spending and identifies instances in which the endogeneity problem is least problematic. Our theoretical model of electoral competition builds on their model. While they use their model primarily to determine the sample of elections on which to run regressions, we take our model directly to data and struc-

[5]There are many other papers that use IV to estimate the marginal effect of spending. See Stratmann (2005) for a brief survey.

[6]See, also, da Silveira and de Mello (2011) for applying a similar identification strategy in the context of gubernatorial elections in Brazil.

turally estimate it. This allows us to conduct counterfactual policy experiments.

2 Model

Overview We consider an extension of the campaign spending game of Erikson and Palfrey (2000) by adding savings, entry and exit to their model. In each period, the candidates play a stage game in which potential challengers from the out-party (i.e., not the incumbent's party) decide whether or not to enter, and conditional on challenger entry, the incumbent and the challenger simultaneously make spending, saving and fund-raising decisions.[7] The vote share is then realized as a function of the spending of the candidates, the quality of the candidates and a random shock. The winner of the election becomes the new incumbent next period.

Timeline At each period, $t = 1, 2, ...\infty$, an election takes place (We often omit the dependence on t for simplicity). The time between the periods is two years, as Congressional elections take place every two years. The stage game consists of the following three stages:

1. Nature draws the quality of N potential challengers from the out-party, ($q_{C,1}$, $q_{C,2}$, ..., $q_{C,N}$), independently from F_{q_C}. We do not consider entry for the in-party.[8] The number of potential challengers, N, is random. Each potential challenger decides whether or not to enter by comparing the value of entering and staying out. Upon entry, challengers pay an entry cost κ.

2(a) If exactly one challenger decides to enter,

- the incumbent observes the quality of the challenger (and vice versa). The incumbent and the challenger then simultaneously make decisions regarding how much money to spend, raise and save.
- the vote shares of the candidates are determined as a function of spending, candidate quality, other characteristics of the candidates and a random shock.

[7] About 15% of U.S. House elections are uncontested.

[8] We do not model entry for the in-party because virtually all incumbents become the party nominee, barring a major scandal. See Online Appendix for the set of elections we dropped because of scandals.

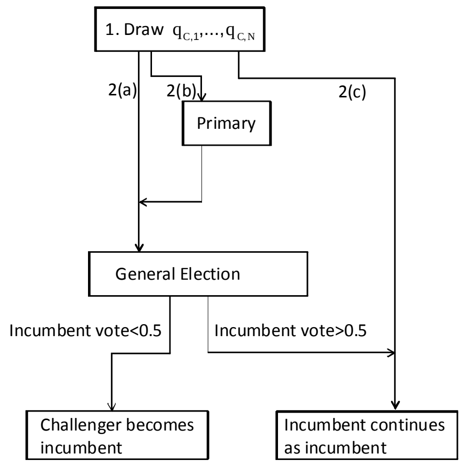

Figure 1: Timeline of the stage game.

2(b) If M $(1 < M \leq N)$ challengers decide to enter,

- a Primary takes place that determines which challenger becomes the Party nominee. Challenger n with quality $q_{C,n}$ becomes the winner of the Primary with probability $\pi(q_{C,n}, \mathbf{q}_{C,-n})$, where $\mathbf{q}_{C,-n} = (q_{C,1},\ ...,\ q_{C,n-1},\ q_{C,n+1},\ ...,\ q_{C,M})$. The winner of the Primary and the incumbent play the game described in 2(a).

2(c) If no challenger decides to enter,

- the incumbent decides how much to spend, raise and save, and the incumbent becomes the winner with probability 1.

3. The winner receives benefit B from being in office. State variables (which we describe below) evolve from current values to the next. Before the start of the next period, the winner decides to retire or run for reelection. Conditional on running for reelection, the winner becomes the incumbent next period and starts out the next period with the amount of money she saved from the previous period.

Figure 1 illustrates the sequence of events. Our model is a dynamic extension of Erikson and Palfrey's (2000) model of campaign spending. In fact, their model corresponds to stage game 2(a) of our model. As in their model, we assume that candidates observe each other's quality before they decide how much to raise, spend and save. This is important, since this is exactly what gives rise to the endogeneity of spending, which the literature has shown to be important.[9] The quality of the candidates are unobserved to the econometrician, however.

State Variables The vector of state variables at the beginning of the stage game is $\mathbf{s}_t = \{q_I, w_{I,t}, ten_{I,t}, D_I, X_t\}$, where q_I is incumbent quality (valance); $w_{I,t}$ is the incumbent war chest at the beginning of the period; $ten_{I,t}$ is the tenure of the incumbent; D_I is a dummy variable that takes the value 1 if the incumbent is a Democrat and -1 if the incumbent is a Republican; and X_t is a vector of demographic characteristics of the Congressional District. The tenure of the incumbent, ten_I, evolves deterministically as $ten_{I,t} = ten_{I,t-1} + 1$. The incumbent war chest is determined by how much the incumbent has saved from the past period. The transition of the demographic characteristics, X_t, is assumed to follow an exogenous first-order Markov process.

The incumbent quality, q_I, is a variable that captures the incumbent's ability to attract votes – i.e., electoral strength.[10] We assume that q_I is constant over time. While this is certainly restrictive, note that we are *not* ruling out the possibility of a deterministic trend in q_I. Any deterministic trend will be accounted for through ten_I: Hence, our definition of q_I can be interpreted as detrended quality. Also, the main reason for restricting q_I to be fixed over time is because of limitations of the data. While it is conceptually straightforward to allow for q_I to evolve stochastically, the estimation of such a model becomes very data-intensive. We will discuss this point in detail at the end of Section 3.1.

Incumbent payoffs and the value function in 2(a) We now specify the primitives of the model in greater detail. We begin with 2(a), the stage game with a contested incumbent. Let q_I and q_C be the incumbent's and challenger's quality,

[9] See the discussion in Jacobson (1978), Green and Krasno (1988), and Gerber (1998), for example.

[10] We are agnostic about how electoral strength is correlated with the quality of the candidate as a legislator.

respectively. We specify u_I, the period utility of the incumbent, and $vote_I$, the incumbent vote share (dropping the time subscript for notational simplicity), as follows:

$$\begin{aligned} u_I &= B \cdot \Pr(vote_I > 0.5) - C_I(w_I' + d_I - w_I; q_I) + H(d_I) \quad (1) \\ vote_I &= \beta_I \ln d_I + \beta_C \ln d_C + q_I - q_C + \beta_{ten} ten_I + D_I \beta_X X + \varepsilon, \quad (2) \end{aligned}$$

B is the benefit from holding office and it is multiplied by the probability of winning – i.e., the probability that the incumbent vote share is greater than 50% of the two-party vote. We specify the incumbent vote share in equation (2) as a linear function of the log spending of the incumbent ($\ln d_I$), the log spending of the challenger ($\ln d_C$), as well as q_I, q_C, ten_I, D_I interacted with X, and a random shock (ε). We assume that ε is an error term that is assumed to follow a Normal distribution with mean equal to 0.5 and variance equal to σ_ε^2. The vector X includes demographic characteristics such as the unemployment rate, and we assume that the sign of X depends on D_I. The linear specification of the vote share function has been used extensively in the previous literature, and we adopt the same specification for comparability. Given our distributional assumption on ε, $\Pr(vote_I > 0.5)$ can also be written as follows:

$$\Pr(vote_I > 0.5) = \Phi\left(\frac{1}{\sigma_\varepsilon}(\beta_I \ln d_I + \beta_C \ln d_C + q_I - q_C + \beta_{ten} ten_I + D_I \beta_X X)\right),$$

where Φ is the c.d.f. of a standard Normal.

The second term of equation (1), $C_I(\cdot)$, captures the costs that the incumbent incurs from raising money. The amount raised by the incumbent is just the sum of the savings (w_I') and spending (d_I) less the war chest (w_I), or $(w_I' + d_I - w_I)$. We let $C_I(\cdot)$ depend on q_I and assume that C_I is decreasing in q_I: That is, higher-quality candidates have a lower cost of raising money. The last term, $H(\cdot)$, captures the consumption value of spending. Campaign spending is sometimes used in ways that seem to benefit the candidates directly, such as for the purchase of personal articles.[11] $H(\cdot)$ is included to capture this.

[11] $H(\cdot)$ explains why there is incumbent spending even in periods when the incumbent seems

We now consider the incumbent's value function. The state variables of the value function are $\{\mathbf{s}, q_C\} = \{q_I, w_I, ten_I, D_I, X, q_C\}$. For a given strategy of the challenger, the value function of the contested incumbent, $v_I^{2(a)}$, can be expressed as follows:

$$v_I^{2(a)}(\mathbf{s}, q_C) = \max_{w_I' \geq 0, d_I \geq 0} u_I + \Pr(vote_I > 0.5)\delta E_{\mathbf{s}'}[V_I(\mathbf{s}')]. \tag{3}$$

The first term on the right-hand side corresponds to the period utility described above, and the second term of the expression corresponds to the continuation value. The continuation value is the expectation of the next period's (ex-ante) value function ($E_{\mathbf{s}'}[V_I(\mathbf{s}')]$) multiplied by the winning probability and discounted by δ. The expectation of the next period's value function is taken with respect to X', the realization of X next period. The incumbent's control variables are the amount of savings w_I' and the amount of spending d_I. The control variables will depend on q_I and q_C because the candidates observe q_I and q_C when they choose d_I and w_I'.

Incumbent payoffs and the value function in 2(c) Let us now consider the incumbent's problem when she is uncontested (i.e., 2(c) in the timeline). The value function of the uncontested incumbent is as follows:

$$v_I^{2(c)}(\mathbf{s}) = \max_{w_I' \geq 0, d_I \geq 0} \tilde{u}_I + \delta E_{\mathbf{s}'}[V_I(\mathbf{s}')],$$

$$\text{where, } \tilde{u}_I = B - \tilde{C}_I(w_I' + d_I - w_I; q_I) + H(d_I).$$

Note that $\tilde{u}_I$ is the period utility of the incumbent when she is uncontested, and the expression is obtained by replacing $\Pr(vote_I > 0.5)$ with 1, and C_I with $\tilde{C}_I$ in equation (1). $\tilde{C}_I$ captures the cost of raising money in uncontested periods, which may be different from C_I. $\Pr(vote_I > 0.5)$ is replaced with 1 because the uncontested incumbent wins with probability one.

(Ex-ante) Value function of the incumbent in stage 1 We now discuss the ex-ante value function of the incumbent, $V_I(\mathbf{s})$. The expression for $V_I(\mathbf{s})$ is as

almost certain to win. In the estimation, we allow for the possibility that $H(\cdot) = 0$, however.

follows:

$$V_I(\mathbf{s}) = (1-\lambda(\mathbf{s}))(1-P_e(\mathbf{s}))v_I^{2(c)}(\mathbf{s}) + (1-\lambda(\mathbf{s}))P_e(\mathbf{s}) \int_{qc} v_I^{2(a)}(\mathbf{s}, q_C) F_{q_C}(q_C | \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}),$$

where $P_e(\mathbf{s})$ is the probability that a challenger enters, and $\lambda(\mathbf{s})$ is the probability that the incumbent retires. The first term corresponds to the value of the incumbent when she is uncontested, an event that occurs with probability $(1-\lambda(\mathbf{s}))(1-P_e(\mathbf{s}))$. The second term corresponds to the value of the incumbent when she is contested, an event which occurs with probability $(1-\lambda(\mathbf{s}))P_e(\mathbf{s})$. Note that in the second term of the expression, we take the expectation of $v_I^{2(a)}$ with respect to the quality distribution of the general election challenger. The conditional distribution is denoted as $F_{q_C}(\cdot | \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, where $\boldsymbol{\chi}$ is a vector of indicator variables, each element of which corresponds to a challenger's entry decision. The expression, $\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}$, denotes the set of circumstances under which there is at least one entrant.

Note that $P_e(\mathbf{s})$ and $F_{q_C}(\cdot | \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, are equilibrium objects. $P_e(\mathbf{s})$ and $F_{q_C}(\cdot | \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$ are both endogenously determined by the behavior of the challengers and, therefore, functions of variables such as q_I and w_I, which potential entrants take into account when they make their entry decision. On the other hand, we let $\lambda(\mathbf{s})$ depend on s but take it as exogenous, an assumption that we will come back to at the end of this section.

Challenger payoffs and value function in 2(a) We now describe the model of the challenger once the challenger has become the Party nominee (corresponds to 2(a) in the timeline). The value function of the challenger in the general election is as follows:

$$\begin{aligned} v_C^{2(a)}(\mathbf{s}, q_C) = \max_{w'_C \geq 0, d_C \geq 0} & B \cdot \Pr(vote_I < 0.5) - C_C(w'_C + d_C, q_C) \\ & + H(d_C) + \delta \Pr(vote_I < 0.5) E_{\mathbf{s}'}[V_I(\mathbf{s}')], \end{aligned}$$

where the choice variables are the amount of spending, d_C, and the amount of savings, w'_C. Note that the challenger's winning probability is given by $\Pr(vote_I < 0.5)$ and that the continuation value is given by V_I, which is the same continuation value we defined in the previous subsection for the incumbent. The challenger's continuation value is the same as the incumbent's because the challenger becomes the incumbent next period if she wins. The argument of the continuation value, $V_I(\cdot)$, consists of the quality of the challenger, q_C, the war chest, w'_C, the tenure $(= 1)$, the Party of the challenger, $-D_I$, and X' – i.e., $\mathbf{s}' = \{q_C,\ w'_C,\ 1,\ -D_I,\ X'\}$. On the other hand, the first argument of $v_C^{2(a)}(\cdot, q_C)$ consists of the incumbent war chest, w_I, and incumbent tenure, ten_I, but it does not include the challenger's war chest or tenure, as the challenger typically starts out without any money from previous elections or any past experience in the U.S. House. Finally, $C_C(\cdot)$ captures the cost of raising money for the challenger.

Challenger's value function in 2(b) Now consider the challenger's value function after paying the entry cost but before winning the Primary (corresponds to 2(b) in the timeline). Suppose that there are M challengers in the Primary with quality given by $q_{C,1}$,$q_{C,2}$,...,$q_{C,M}$. Then, the value function of challenger n with quality $q_{C,n} = q_C$ is simply $\pi(q_C, \mathbf{q}_{C,-n})v_C^{2(a)}(\mathbf{s}, q_C)$, where π denotes the probability that challenger n wins the Primary among M contenders.

Challenger's value function in stage game 1 We can now define the value function of the challenger at the beginning of the stage game (corresponds to 1 in the timeline). Assuming that each challenger must make an entry decision before learning who the other challengers are (i.e., before learning N, M or $\mathbf{q}_{C,-n}$), the value function of the challenger with quality $q_{C,n} = q_C$ at the start of the period is

$$V_C(\mathbf{s}, q_C) = \max\{p(q_C, \mathbf{s})v_C^{2(a)}(\mathbf{s}, q_C) - \kappa, R\}, \tag{4}$$

where κ is the fixed cost of entry, R is the reservation value, and $p(q_C, \mathbf{s})$ is the probability that a challenger with quality q_C becomes the contender in the general election. The expression for $p(q_C, \mathbf{s})$ is

$$p(q_C, \mathbf{s}) = E_M\left[\int \pi(q_C, \mathbf{q}_{C,-n})dF_{\mathbf{q}_{C,-n}}(\mathbf{q}_{C,-n}|M, \mathbf{s})\middle|\,\mathbf{s}\right], \tag{5}$$

where we take the integral of $\pi(q_C, \mathbf{q}_{C,-n})$ with respect to M and $F_{\mathbf{q}_{C,-n}}(\mathbf{q}_{C,-n}|M, \mathbf{s})$. Recall that M is the number of potential challengers that enter (i.e., pay κ), and it is a random variable. $F_{\mathbf{q}_{C,-n}}$ is the quality distribution of the $(M-1)$ other entrants given state $\mathbf{s}$, and M. When there is only one challenger (i.e., when $M = 1$) we assume that $\pi(q_C, \phi) = 1$.

The entry decision of a potential challenger, which we denote by χ, is as follows:

$$\chi(q_C, \mathbf{s}) = \begin{cases} 1\text{: if } p(q_C, \mathbf{s})v_C^{2(a)} - \kappa > R \\ [0,1]\text{: if } p(q_C, \mathbf{s})v_C^{2(a)} - \kappa = R \\ 0\text{: if } p(q_C, \mathbf{s})v_C^{2(a)} - \kappa < R \end{cases}.$$

The number of entrants, M, can be expressed as the sum of the entry decisions of each potential challenger; i.e., $M = \sum_i^N \chi(q_{C,i}, \mathbf{s})$. Note, also, that as long as the value of entry, $v_C^{2(a)}$, is increasing in q_C, the entry decision takes a cut-off form. The entry decision can, therefore, be expressed alternatively as follows:

$$\chi(q_C, \mathbf{s}) = \begin{cases} 1\text{: if } q_C > \bar{q}_C(\mathbf{s}) \\ [0,1]\text{: if } q_C = \bar{q}_C(\mathbf{s}) \\ 0\text{: if } q_C < \bar{q}_C(\mathbf{s}) \end{cases},$$

where $\bar{q}_C(\mathbf{s})$ is defined implicitly as the solution to $p(x, \mathbf{s})v_C^{2(a)}(\mathbf{s}, x) - \kappa = R$. $\bar{q}_C$ is the type of challenger that is indifferent between entering and not entering. As long as $v_C^{2(a)}$ is increasing in q_C, there exists a symmetric pure strategy Nash equilibrium in cut-off strategies.[12] When we consider the entry game of the challengers, we will always focus on this Nash equilibrium.

Entry Probability and the Conditional Distribution of q_C Now, we close the model by describing how $F_{\mathbf{q}_{C,-n}}(\cdot|M, \mathbf{s})$, $P_e(\mathbf{s})$ and $F_{q_C}(\cdot|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$ are determined. We start with $F_{\mathbf{q}_{C,-n}}(\cdot|M, \mathbf{s})$, the quality distribution of the $(M-1)$ other entrants. Recall that each potential entrant decides whether or not to enter by comparing her own quality q_C with the threshold $\bar{q}_C(\mathbf{s})$, which implies that q_C

[12] While it seems natural to require $v_C^{2(a)}$ to be increasing in q_C, we do not have a theoretical proof that it is. What we do in practice is to estimate the parameters of the model assuming that $v_C^{2(a)}$ is increasing in q_C and check ex-post that $v_C^{2(a)}$ is actually increasing at the estimated parameters.

is distributed between $[\bar{q}_C(\mathbf{s}), \infty]$, conditional on entry. Hence, if $q_{C,1}, \cdots, q_{C,N}$ are drawn independently from F_{q_C}, then $F_{\mathbf{q}_{C,-n}}(\cdot|M, \mathbf{s})$ is just the measure obtained by restricting $F_{q_C} \times \cdots \times F_{q_C}$, to $[\bar{q}_C(\mathbf{s}), \infty]^{M-1}$.

Next, consider the expression for $P_e(\mathbf{s})$. Given N potential challengers, the probability that exactly M ($\leq N$) challengers decide to enter is given by $Bin(N, M; 1 - F_{q_C}(\bar{q}_C))$, where $Bin(n_1, n_2; p)$ is the probability that we have n_2 successes out of n_1 trials with success rate p. This follows from the cut-off strategy for challenger entry, which we described above. Then, the probability that there is at least one entrant, P_e, can be expressed as follows:

$$\begin{aligned} P_e(\mathbf{s}) &= \mathbf{E}_N \left[1 - Bin(N, 0; 1 - F_{q_C}(\bar{q}_C(\mathbf{s}))) \middle| \mathbf{s} \right] \\ &= \mathbf{E}_N \left[1 - F_{q_C}(\bar{q}_C(\mathbf{s}))^N \middle| \mathbf{s} \right]. \end{aligned} \tag{6}$$

Now, consider $p(q_C, \mathbf{s})$, the probability that a challenger with quality q_C wins the Primary:

$$\begin{aligned} p(q_C, \mathbf{s}) &= \mathbf{E}_M \left[\int \pi(q_C, \mathbf{q}_{C,-n}) dF_{\mathbf{q}_{C,-n}}(\mathbf{q}_{C,-n}|M, \mathbf{s}) \middle| \mathbf{s} \right] \\ &= \mathbf{E}_M \left[\frac{\int_{\bar{q}_C}^{+\infty} \cdots \int_{\bar{q}_C}^{+\infty} \pi(q_C, \mathbf{q}_{C,-n}) (dF_{q_C})^{M-1}}{(1 - F_{q_C}(\bar{q}_C))^{M-1}} \middle| \mathbf{s} \right] \\ &= \mathbf{E}_N \left[\mathbf{E}_M \left[\frac{\int_{\bar{q}_C}^{+\infty} \cdots \int_{\bar{q}_C}^{+\infty} \pi(q_C, \mathbf{q}_{C,-n}) (dF_{q_C})^{M-1}}{(1 - F_{q_C}(\bar{q}_C))^{M-1}} \middle| N, \mathbf{s} \right] \middle| \mathbf{s} \right] \\ &= \mathbf{E}_N \left[\sum_{M=1}^{N} Bin(N, M; 1 - F_{q_C}(\bar{q}_C)) \frac{\int_{\bar{q}_C}^{+\infty} \cdots \int_{\bar{q}_C}^{+\infty} \pi(q_{C,n}, \mathbf{q}_{C,-n}) (dF_{q_C})^{M-1}}{(1 - F_{q_C}(\bar{q}_C))^{M-1}} \middle| \mathbf{s} \right]. \end{aligned} \tag{7}$$

The first line of the expression comes from (5), and the second line replaces $F_{\mathbf{q}_{C,-n}}(\cdot|M, \mathbf{s})$. The third line uses the law of iterated expectations, and the fourth line uses the fact that the probability of observing M challengers is $Bin(N, M; 1 - F_{q_C}(\bar{q}_C))$. Similarly, the expression for the quality distribution of challengers who win the Primary,

$F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, is as follows:

$$F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}) = \mathbf{E}_N \left[\mathbf{E}_M \left[\Pr(q_C \leq t|M, \mathbf{s})|N, \mathbf{s}\right]|\, \mathbf{s}\right]$$
$$= \mathbf{E}_N \left[\sum_{M=1}^{N} Bin(N, M; 1 - F_{q_C}(\bar{q}_C)) \frac{\int_{\bar{q}_C}^{t} \int_{\bar{q}_C}^{+\infty} \cdots \int_{\bar{q}_C}^{+\infty} M\pi(q_{C,n}, \mathbf{q}_{C,-n})(dF_{q_C})^M}{(1 - F_{q_C}(\bar{q}_C))^M} \middle| \mathbf{s} \right] \quad (8)$$

While we have suppressed the dependence of $\bar{q}_C$ on $\mathbf{s}$ in equations (7) and (8), $\bar{q}_C$ is a function of $\mathbf{s}$, as $\bar{q}_C(\mathbf{s})$.

Sufficient Statistics Property Now we discuss the sufficient statistics property that we will exploit heavily in our identification. For exposition, suppose that N is a constant that does not depend on $\mathbf{s}$. In this case, it is easy to see that the expressions for $p(q_C, \mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$ in equations (7) and (8) depend on the state variables $\mathbf{s}$ only through $\bar{q}_C$. In this sense, $\bar{q}_C$ is a sufficient statistic for $p(q_C, \mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$.[13] Moreover, given that $P_e(\mathbf{s})$ is a monotone function of $\bar{q}_C(\mathbf{s})$ when N is a constant (see expression (6)), $P_e(\mathbf{s})$ is also a sufficient statistic for $p(q_C, \mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$. That is, knowledge of $P_e(\mathbf{s})$ uniquely determines $p(q_C, \mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$.

The sufficient statistics property is important because it allows us to control for challenger quality when identifying the effect of campaign spending on the vote share. To the extent that $P_e(\mathbf{s})$ is a sufficient statistic for $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, we can pin down the distribution of challenger quality once we condition on $P_e(\mathbf{s})$; i.e., for any two state variables $\mathbf{s}$ and $\mathbf{s}'$ such that $P_e(\mathbf{s}) = P_e(\mathbf{s}')$, the distribution of q_C is the same. This means that we can exploit variation in $\mathbf{s}$ such that $E[d_I|\mathbf{s}]$ and $E[d_C|\mathbf{s}]$ change, but $P_e(\mathbf{s})$ is held constant to identify the effect of spending on the vote share without having to worry about changes in $E[q_C|\mathbf{s}]$.

In the Online Appendix, we show that a similar sufficient statistics property holds even when N is a random variable whose distribution depends on $\mathbf{s}$. As long as we are willing to use a large set of sufficient statistics, the sufficient statistics property holds with fairly weak assumptions. In our estimation, we specify $F_N(\cdot|\mathbf{s})$ to follow a negative binomial distribution, $NB(r(\mathbf{s}), p)$, where $r(\mathbf{s})$ can be an ar-

[13] More precisely, we say that $h = h(\mathbf{s})$ is a sufficient statistic for $f(\mathbf{s})$ if $h(\mathbf{s}') = h(\mathbf{s}'') \Rightarrow f(\mathbf{s}') = f(\mathbf{s}'')$.

bitrary function of $\mathbf{s}$. Under this specification, we show that $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are sufficient statistics for $p(q_C, \mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, where $E[M|\mathbf{s}]$ is the expected number of entrants; i.e.,

$$E[M|\mathbf{s}] = \mathbf{E}_N \left[\sum_{M=1}^{N} M \cdot Bin(N, M; 1 - F_{q_C}(\bar{q}_C)) \middle| \mathbf{s} \right] = \mathbf{E}_N \left[N(1 - F_{q_C}(\bar{q}_C)) | \, \mathbf{s} \right]. \tag{9}$$

Remarks on the Model Before we end our model section, we make a few remarks: The first remark is on the incumbent's retirement decision. In our model, we take retirement as an exogenous event. Whether retirement is taken as an exogenous event or an endogenous choice variable actually matters little for the estimation of our model. This is because we estimate $\lambda(\cdot)$ directly as a function of the state variables, $\mathbf{s}$. Depending on whether the retirement is a choice variable or an exogenous event, $\lambda(\mathbf{s})$ can be interpreted either as a policy function or as an exogenous probability. Of course, this distinction can matter in the counterfactual analysis: By taking retirement as an exogenous event, we do not allow the the incumbents to optimize over the retirement decision.

The reason we opt to treat retirement as exogenous is because our estimate of $\lambda(\cdot)$ turns out to be quite flat as a function of $\mathbf{s}$ with the exception of ten_I. That is, ten_I affects the retirement probability, but the effect of other variables, such as q_I and w_I, are not statistically different from zero. This seems to suggest that retirement decisions are *not* made strategically, which is consistent with Ansolabehere and Snyder's (2004) finding of no evidence of strategic retirement.[14] Given these findings, we treat $\lambda(\cdot)$ as exogenous and let $\lambda(\cdot)$ be a function only of ten_I in our estimation.[15]

Our second remark is on the model of the Primary. Given that the focus of this

[14]They state, for example: "In fact, the analysis contradicts the strategic retirement hypothesis directly: instrumenting for incumbency produces effects that are somewhat *higher* than the simple OLS estimators." (p.488)

[15]For a paper that studies endogenous exit decisions, see Diermeier Kean and Merlo (2008). They study the career concerns of House Representatives taking exit as an endogenous choice. They also find that "... the selectivity bias induced by politicians' decisions whether to run for reelection is actually rather modest." (p.349)

paper is on the general election rather than on the Primary, we do not attempt to estimate all of the primitives of the model of the Primary – $\pi(\cdot)$, F_N, R, and κ. In fact, the only data we use from the Primary elections are data on M (the number of candidates that run in the Primary), which implies that we can estimate only reduced form objects, such as $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$. While this may seem like a shortcoming, knowledge of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ is sufficient to identify and estimate the primitives of the general election (e.g., the coefficients of the vote share function, the cost functions, etc.). Moreover, to the extent that $\pi(\cdot)$, F_N, R, and κ are structural to our counterfactual experiments, we can also compute counterfactuals. Our model of the Primary disciplines the estimation of the primitives of the general election, and it is useful to that extent. The main focus of this paper will be on the estimation of the general election and not on the model of the Primary.

Our last remark is about the challengers' entry decision. One of the limitations of our model is that we are not allowing the challengers to have the option value of waiting when they make their entry decision. While this is an important issue, it is hard to incorporate this feature given that we lack data on the set of potential challengers.[16]

Equilibrium Formally, the players of the game are the incumbent and an infinite sequence of potential challengers. The strategies of the game are how much to spend, save, and raise for both the incumbent and the challengers, as well as the challengers' entry decisions. The solution concept we use is stationary Markov Perfect Equilibria (see Maskin and Tirole, 1988). In equilibrium, each player's policy function is the best response to the policy functions of the other players. Under a technical assumption on the size of the state space,[17] there exists an equilibrium (possibly in mixed strategies) of this game.[18]

In general, equilibrium exists in mixed strategies but not necessarily in pure strategies in our setting. One thing to note, however, is that there always exists an

[16]The same issue comes up in much of the empirical work on industry dynamics with entry and exit.

[17]The state space needs to be countable. As long as the state variables are discretized, the model admits an equilibrium.

[18]See Whitt (1980).

equilibrium in which the incumbent plays pure strategies in uncontested races.[19] Henceforth, we assume that the incumbent plays pure strategies in uncontested periods. This is useful because we will use the incumbent's policy function in uncontested races to invert out q_I. Aside from this, however, we do not have to assume that players play pure strategies for identification and estimation.[20]

3 Identification and Estimation

In this section, we discuss identification and estimation of the key components of our model. We first discuss identification and estimation of the vote share equation. We then show how the other primitives of the model, such as C and H etc., can be identified using the first-order conditions that equate the marginal benefit/cost of spending, saving and fund-raising.

3.1 First Step: Vote production function

3.1.1 Overview

The two main challenges in identifying the vote share equation are endogeneity of spending and sample selection. In terms of model notation, the endogeneity of spending simply means that unobserved candidate quality (q_I, q_C) are correlated with d_I and d_C in equation (2). Sample selection refers to the issue that control variables in equation (2), such as ten_I and X, are correlated with q_C, because the distribution of challengers conditional on entry, $F_{q_C}(\cdot|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, depends on $\mathbf{s}$. This is a natural consequence of the fact that contested elections are outcomes of potential entrants' entry decisions rather than a randomly selected sample.

[19]Suppose we take an equilibrium in which the incumbent plays a mixed strategy in uncontested periods. If we then replace that strategy with an alternative one in which the incumbent plays a particular action (in the support of the mixed strategy) with probability one (and keep everything else the same), the resulting profile of strategies will still remain an equilibrium.

[20]Our identification and estimation results are valid under both mixed strategy and pure strategy equilibria. We follow an approach that is similar to that of Bajari, Benkard and Levin (2007) in estimating some of the parameters, but our approach does not require the estimation of the policy function itself: It requires only the distribution of the outcome variables conditional on the observable states.

3.1.2 Sample Selection

We discuss how to deal with sample selection first. From the previous section, recall that the distribution of q_C conditional on entry, $F_{q_C}(\cdot|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, depends on $\mathbf{s} = (q_I, w_I, ten_I, D_I, X)$ only through certain sufficient statistics such as $P_e(\mathbf{s})$ (if N is a constant which does not depend on s) or $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ (if N is a negative binomial with $NB(r(\mathbf{s}), p)$). In what follows, we proceed with the assumption that $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are sufficient statistics, which is consistent with our specification in the estimation. In this case, we can express the conditional mean of q_C as a function of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ as follows:

$$\begin{aligned} E[q_C|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}] &= E[q_C|P_e(\mathbf{s}), E[M|\mathbf{s}]] \\ &\equiv g(P_e(\mathbf{s}), E[M|\mathbf{s}]). \end{aligned}$$

The function g is a complicated equilibrium object the exact shape of which depends on unknown model parameters. From our perspective, however, the important point is that the conditional expectation of q_C, given entry, can be expressed *only* as functions of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ (as opposed to the whole vector of state variables, s). This implies that, given knowledge of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$, we can control for selection bias by including a nonparametric function of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ in equation (2), as follows:

$$\begin{aligned} vote_I &= \beta_I \ln d_I + \beta_C \ln d_C + \beta_{ten} ten_I + D_I \beta_X X + q_I - q_C + \varepsilon \\ &= \beta_I \ln d_I + \beta_C \ln d_C + \beta_{ten} ten_I + D_I \beta_X X + q_I - g(P_e(\mathbf{s}), E[M|\mathbf{s}]) + \tilde{\varepsilon}, \end{aligned}$$

where $\tilde{\varepsilon} = g(P_e(\mathbf{s}), E[M|\mathbf{s}]) - q_C + \varepsilon$. Note that $\tilde{\varepsilon}$ is uncorrelated with s by construction. Intuitively, control variables such as ten_I and X have a direct effect on the vote share, as well as an indirect effect through the change in $E[q_C|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}]$. However, by conditioning on P_e and $E[M]$, we can control for the indirect effect and identify just the direct effect.

Note that, thus far, our discussion closely follows the identification strategy used in OP. Unlike in their paper, however, the sufficient statistics in our setting depend, in part, on q_I, which is unobserved; i.e., $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are functions of q_I. Hence, we still need an extra step to make this approach work. To do so, we now consider how to invert out q_I from the policy function of uncontested incumbents.

Recall the problem of the incumbent when she is uncontested:

$$v_I^{2(c)}(\mathbf{s}) = \max_{w'_I \geq 0, d_I \geq 0} B - \tilde{C}_I(w'_I + d_I - w_I, q_I) + H(d_I) + \delta E_{\mathbf{s}'}[V_I(\mathbf{s}')].$$

The policy functions associated with this problem are how much to save, $w'_I = w'_I(\mathbf{s})$, and how much to spend, $d_I = d_I(\mathbf{s})$. Note that the policy functions can be viewed as mappings from q_I to (w'_I, d_I), holding the other state variables fixed. If the mapping $q_I \mapsto (w'_I, d_I)$ is one-to-one (given w_I, ten_I, X and D_I), then we can uniquely solve for q_I using these policy functions as $q_I = q_I(w'_I, d_I; w_I, ten_I, X, D_I) = q_I(\mathbf{s}_{NC})$, where $\mathbf{s}_{NC}$ denotes the vector of state variables and actions in the uncontested period. In the Appendix, we give a simple proof that this mapping is one-to-one under the functional form for $\tilde{C}_I$ and H that we adopt in our estimation. The proof shows that the first-order condition associated with d_I can be solved in terms of q_I.[21] In particular, we prove that q_I can be expressed as a function of a single variable – $(\ln d_I)^{-1/2} d_I^{-1} fr_I (\ln fr_I)^{-1}$, where $fr_I = w'_I + d_I - w_I$ – rather than as a function of the whole vector, $\mathbf{s}_{NC}$, which greatly reduces the estimation burden.

The inversion of the policy function enables us to express $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$, as functions of $(\mathbf{s}_{NC}, w_I, ten_I, X)$ – all of which are observed – instead of $(q_I, w_I, ten_I, X, D_I)$. In particular, this allows us to estimate $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ as functions of observables, $(\mathbf{s}_{NC}, w_I, ten_I, X)$ using a subset of observations for which (1) we observe the incumbent in more than two periods, and (2) there is a period in which the incumbent was uncontested.[22] Once we estimate $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$, we can use them to control for selection bias. For the purpose of identification, can take $P_e(\mathbf{s})$ and

[21] The first-order condition for d_I is as follows,

$$\frac{\partial}{\partial d_I} H_I(d_I) - \frac{\partial}{\partial d_I} \tilde{C}_I(w'_I + d_I - w_I, q_I) = 0.$$

As long as the marginal cost ($\frac{\partial}{\partial x} \tilde{C}_I(x, \cdot)$) is decreasing in the second argument, q_I can be expressed in terms of d_I and $w'_I + d_I - w_I$.

[22] In practice, we identify the first period in which each incumbent is uncontested, and use subsequent reelection attempts to estimate P_e and $E[M]$.

$E[M|\mathbf{s}]$ as known from now on.

3.1.3 Endogeneity of Spending

We now discuss how we deal with the endogeneity of spending. Recall that there are two sources of endogeneity in identifying the vote share function: the correlation between q_I and the spending variables, and the correlation between q_C and the spending variables. The inversion of the policy function eliminates the first endogeneity, because we can explicitly control for q_I by replacing out q_I as functions of $\mathbf{s}_{NC}$, as $q_I(\mathbf{s}_{NC})$. Hence, we discuss how we deal with the remaining correlation between q_C and (d_I, d_C).

Our idea for dealing with the correlation between q_C and (d_I, d_C) is to exploit variation in s that is orthogonal to q_C. We can do so by considering a projection of the vote share equation on a particular set of variables. Formally, consider projecting the vote shares on $\Omega \equiv \{\mathbf{s}_{NC}, w_I, ten_I, X, D_I, \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}\}$:

$$\begin{aligned} vote_I &= E[vote_I|\Omega] + (votes_I - E[votes_I|\Omega]) \\ &= E[\beta_I \ln d_I + \beta_C \ln d_C + \beta_{ten} ten_I + D_I \beta_X X + q_I - q_C + \varepsilon|\Omega] + \epsilon \\ &= \beta_I E[\ln d_I|\Omega] + \beta_C E[\ln d_C|\Omega] + \beta_{ten} ten_I + D_I \beta_X X + q_I(\mathbf{s}_{NC}) - g(P_e, E[M]) + \epsilon \end{aligned} \tag{10}$$

where $\epsilon \equiv (votes_I - E[votes_I|\Omega])$ and $E[\epsilon|\Omega] = 0$ by construction. We use the fact that $q_I = q_I(\mathbf{s}_{NC})$, $E[\varepsilon|\Omega] = 0$, and $E[q_C|\Omega] = g(P_e, E[M])$ to go from the second line of the expression to the third.

The conditioning set Ω in expression (10) corresponds to the information available to the incumbent before q_C is realized. Projecting the vote share and the regressors on Ω is attractive for two reasons: First, $E[\ln d_I|\Omega]$ and $E[\ln d_C|\Omega]$ are uncorrelated with ϵ by construction.[23] Second, all of the variables that constitute Ω are observable to the researcher. This means that we can estimate $E[\ln d_I|\Omega]$ and $E[\ln d_C|\Omega]$. In fact, all of the regressors in equation (10) – $E[\ln d_I|\Omega]$, $E[\ln d_C|\Omega]$, ten_I, X, $\mathbf{s}_{NC}$, P_e, $E[M]$ – are either observed or can be estimated. Given that the error term ϵ in equation (10) is mean independent of all the regressors, (10) is a

[23] $E[\ln d_I|\Omega]$ is measureable with respect to Ω and $E[\epsilon|\Omega] = 0$. Hence, they are uncorrelated.

partially linear regression with generated regressors that can be estimated.

Note that unlike in OP (or Levinsohn and Petrin, 2003), q_I is expressed as a function of actions and state variables of some past period. This means that we do not have collinearity issues with respect to $\mathbf{s}_{NC}$. Also, Ω includes w_I that is excluded from (10), which ensures that there are no collinearity issues with respect to P_e and $E[M]$.[24] Equation (10) serves as both an identification equation and an estimation equation. All of the coefficients of the vote share equation, as well as $q_I(\cdot)$ and $g(\cdot, \cdot)$, are identified in equation (10). In particular, incumbent quality, $q_I = q_I(\mathbf{s}_{NC})$, is identified.

In order to give a more intuitive explanation of identification, consider two incumbents who are of similar quality. Given that q_I is a function of $\mathbf{s}_{NC}$, choosing incumbents of similar quality can be accomplished by choosing incumbents with similar $\mathbf{s}_{NC}$. Now consider periods in which these incumbents are contested and let $\{\mathbf{s}_1, \mathbf{s}_2\}$ be the state variables. By considering states $\{\mathbf{s}_1, \mathbf{s}_2\}$ so that the ex-ante probability of entry, P_e, and the number of potential entrants, $E[M]$, are the same – i.e., $P_e(\mathbf{s}_1) = P_e(\mathbf{s}_2)$ and $E[M|\mathbf{s}_1] = E[M|\mathbf{s}_2]$ – the expected quality of the challenger will also be the same in these states. Note that the state variables themselves need not be identical in order for P_e and $E[M]$ to be the same. For example, one incumbent may have a higher war chest than the other incumbent, but the other state variables, (ten_{I1}, X_1) may be less favorable, so that $P_e(\mathbf{s}_1) = P_e(\mathbf{s}_2)$ and $E[M|\mathbf{s}_1] = E[M|\mathbf{s}_2]$.

Given state variables $\{\mathbf{s}_1, \mathbf{s}_2\}$ such that $P_e(\mathbf{s}_1) = P_e(\mathbf{s}_2)$ and $E[M|\mathbf{s}_1] = E[M|\mathbf{s}_2]$, consider the average difference in the vote shares in these two states:

$$
\begin{aligned}
&E[votes_1|\mathbf{s}_1, \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}] - E[votes_2|\mathbf{s}_2, \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}] \\
&\quad = \beta_I(E[\ln d_I|\Omega_1] - E[\ln d_I|\Omega_2]) + \beta_C(E[\ln d_C|\Omega_1] - E[\ln d_C|\Omega_2]) \\
&\qquad + \beta_{ten}(ten_{I1} - ten_{I2}) + \beta_X(D_I X_1 - D_I X_2),
\end{aligned}
$$

where we use the fact that the conditional expectation of challenger quality is the same under $\mathbf{s}_1$ and $\mathbf{s}_2$. This expression offers intuition as to where identifi-

[24] See Ackerberg, Caves and Frazer (2006) for a discussion of collinearity issues in production function estimation.

cation comes from: β_I and β_C are identified from comparing the vote share of periods that have the same expected challenger quality ($E[q_C|\Omega_1] = E[q_C|\Omega_2]$), but have different expected spending ($E[\ln d_I|\Omega_1] \neq E[\ln d_I|\Omega_2]$, $E[\ln d_C|\Omega_1] \neq E[\ln d_C|\Omega_2]$). That is, we can identify the coefficients of (2) by correlating the differences in the spending of the incumbents ($E[\ln d_I|\Omega_1] - E[\ln d_I|\Omega_2]$) and challengers ($E[\ln d_C|\Omega_1] - E[\ln d_C|\Omega_2]$) etc., to the differences in the vote shares. Basically, our approach exploits variation in $\{w_I, ten_I, X\}$, which moves d_I and d_C, but keeps P_e and $E[M]$ constant.

Time varying q_I Finally, we discuss how our approach can be extended to accommodate time varying q_I. To do so, we first specify the evolution of $q_{I,t}$ and also the information set of the players at each point in time. One possible specification is that (1) $q_{I,t}$ evolves as a random walk as $q_{I,t} = q_{I,t-1} + \xi_t$ and that (2) ξ_t is revealed before the candidates decide how much to spend, save, and raise, but after the challenger makes her entry decision (that is, after 1 and before 2 in the timeline). This would be the case if the challenger makes an entry decision based on what she knows from the previous election and learns the innovation only as she starts to compete for the seat. If the timing is such, then P_e and $E[M]$ are functions of $q_{I,t-1}$, $w_{I,t}$, $ten_{I,t}$, X_t, D_I.

Under this specification, we cannot use $\mathbf{s}_{NC}$ from more than two periods before to substitute out $q_{I,t}$ because $q_{I,t}$ is not fixed. We need to use $\mathbf{s}_{NC}$ from one period before. The estimation of the vote share equation can then be conducted on the subset of the samples in which (1) the incumbent was uncontested and (2) the incumbent was contested in the following period. The vote share equation is

$$\begin{aligned} vote_{I,t} \;&=\; \beta_I \ln d_{I,t} + \beta_C \ln d_{C,t} + \beta_{ten} ten_{I,t} + D_I \beta_X X_t + q_I(\mathbf{s}_{NC,t-1}) - g(P_e, E[M]) \\ &\quad + \xi_t + (q_{C,t} - g(P_e, E[M])) + \varepsilon_t. \end{aligned}$$

The econometric error term is $\xi_t + (q_{C,t} - g(P_e, E[M])) + \varepsilon_t$, where $\xi_t = (q_{I,t} - q_I(\mathbf{s}_{NC,t-1}))$. Given that the expectation of the error term conditional on $\Omega_t \equiv$

$\{\mathbf{s}_{NC,t-1}, w_{I,t}, ten_{I,t}, X_t, \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}\}$ is 0, we can proceed as before.[25][26] In our estimation, we opted not to work with time varying q_I, however, because of data limitations.

3.2 Second Step

3.2.1 Overview

We now consider identification and estimation of challenger quality, (q_C) the cost functions, (C, $\tilde{C}$), the function that accounts for the consumption value of spending (H) and the standard error in the vote share equation (σ_ε). In this section, we take q_I and the parameters of the vote share equation as given. We also take the discount factor, δ, as known and proceed with normalization $B = 1$, where B is the benefit of winning.[27] The key condition we exploit in our identification and estimation is the candidates' first-order conditions. That is, we use the restriction that at the true parameter and at the true realization of q_C, the actions chosen by the candidates satisfy the first-order conditions.

Our estimation borrows from Bajari, Benkard and Levin (2007). They propose a method of simulating the value function by first estimating the policy function and then using the policy function to generate sample paths of outcomes and actions, which can then be averaged to compute the continuation value. Because we do not observe the full set of state variables in contested periods (we do not observe q_C in contested periods), we estimate the distribution of the actions and outcomes conditional on observed state variables instead of the actual policy function. This turns out to be sufficient for the purpose of forward-simulating the value function. We then feed the value function into the first-order conditions that are associated with the problem of choosing the optimal level of spending and savings.

[25] Note that $\ln d_{I,t}$ and ξ_t are correlated, but $E[\ln d_{I,t}|\Omega_t]$ and ξ_t are not. The latter is all we need to apply our method.

[26] If the challenger observes ξ_{t+1} when deciding whether to enter or not, our empirical method will not extend in a straightforward way.

[27] Identifying the discount factor in dynamic games is known to be very difficult. We follow the literature in taking δ as given. See Magnac and Thesmar (2002) for a detailed discussion. Regarding B, we need to normalize it to some number because eveything can be scaled up or down by a constant factor. The normalization of B does not affect the results of our conterfactuals.

3.2.2 First-Order Conditions

We start by discussing the key restrictions on which our identification and estimation rely. Associated with the incumbent's spending and saving decisions, there are two first-order conditions:

$$\underbrace{\frac{\partial C_I}{\partial d_I}(w_I^{/} + d_I - w_I, q_I)}_{\text{MC of fund-raising}} = \underbrace{\frac{\beta_I}{\sigma_\varepsilon d_I}\phi\left(K\right)\cdot\left(B + \delta E_{\mathbf{s}'}[V_I]\right) + \frac{\partial H_I}{\partial d_I}(d_I)}_{\text{MB of spending}}, \text{ and} \tag{11}$$

$$\underbrace{\frac{\partial C_I}{\partial w_I'}(w_I^{/} - w_I + d_I, q_I)}_{\text{MC of fund-raising}} = \underbrace{\delta\Phi\left(K\right)\frac{\partial}{\partial w_I'}E_{\mathbf{s}'}[V_I]}_{\text{MB of saving}}, \tag{12}$$

$$\text{where } K = \frac{1}{\sigma_\varepsilon}(\beta_I \ln d_I + \beta_C \ln d_C + q_I - q_C + \beta_{ten} ten_I + D_I \beta_x X).$$

The first expression equates the marginal cost of raising money to the marginal benefit of spending, and the second expression equates the marginal cost of raising money to the marginal benefit of saving. $\Phi(\cdot)$ and $\phi(\cdot)$ are the c.d.f and the p.d.f of the standard normal. The expressions can be obtained by substituting $\Pr(vote_I > 0.5)$ with $\Phi(K)$ in equation (3) and taking the derivative. We obtain analogous expressions corresponding to the first-order conditions of incumbents in uncontested periods and those corresponding to the challengers.

Note that q_I appears in expressions (11) and (12), but we take the value of q_I as given, as we discussed above. To the extent that q_I can be recovered without any error, the first-order condition must hold with equality for each state variable with d_I, $w_I' > 0$. Of course, in practice, q_I is estimated nonparametrically, and the first-order conditions will only hold with noise.[28]

The two first-order conditions can be re-expressed as follows:

[28] With a finite sample size, we cannot recover q_I perfectly, and this means that the expression will not hold exactly when evaluated at the estimated quality, $\hat{q}_I$. In a finite sample, the model is not rejected because $\hat{q}_I$ is estimated for each incumbent with some noise.

$$
\begin{aligned}
\frac{\partial C_I}{\partial d_I}(w_I' + d_I - w_I, q_I) &= \frac{\beta_I}{\sigma_\varepsilon d_I}\phi\left(\Phi^{-1}\left(\frac{\partial C_I}{\partial w_I'}\Big/\delta\frac{\partial E_{\mathbf{s}'}[V_I]}{\partial w_I'}\right)\right)(B + \delta E_{\mathbf{s}'}[V_I]) + \frac{\partial H_I}{\partial d_I} \\
K &= \Phi^{-1}\left(\frac{\partial C_I}{\partial w_I'}\Big/\delta\frac{\partial E_{\mathbf{s}'}[V_I]}{\partial w_I'}\right) \\
K &= \frac{1}{\sigma_\varepsilon}(\beta_I \ln d_I + \beta_C \ln d_C + q_I - q_C + \beta_{ten} ten_I + D_I\beta_x X),
\end{aligned}
\tag{13}
$$

These expressions are obtained by solving for K in equation (12) and substituting out K in equation (11). Note that K no longer appears in the first line of the expression. Given that the continuation value $E_{\mathbf{s}'}[V_I]$ can be expressed as a function of C, $\tilde{C}$, H and σ_ε, as we discuss below, the first line of the expression can be considered as an equation in C, $\tilde{C}$, H and σ_ε.[29] Hence, the first line of the expression identifies these parameters. Once C, $\tilde{C}$, H and σ_ε are identified, the second line of the expression identifies K, and the third line can be used to identify q_C.

3.2.3 Evaluating the Continuation Value by Simulation.

When we exploit the first-order conditions for estimating the parameters of the model, the key step involves expressing the continuation value as a function of the model's parameters. We discuss this issue in this section. Our approach is to use simulation methods first proposed by Hotz and Miller (1993) and Hotz, Miller, Sanders and Smith (1994) in the context of single-decision problems and then extended to games by Bajari, Benkard and Levin (2007). The simulation methods proposed by these papers involve (1) estimating the transition of the state variables and the equilibrium policy functions nonparametrically; and (2) forward-simulating the value function as a function of the parameters, given the estimated transition probabilities and the policy function. Importantly, it does *not* involve solving for the equilibrium for each parameter value. We adapt their approach in our context.

The first step of the simulation method is to estimate the transition of the state variables and the policy functions. When the state variables and the actions are all observed, this step is data-intensive, but relatively straightforward. The only

[29]Note that equation (13) has to hold for each value of $\mathbf{s}$.

requirement is that the policy functions be estimated as flexibly as possible (preferably nonparametrically) because they are equilibrium objects. In our application, however, we do not observe q_C, which is one of the arguments of the policy functions (note that we take q_I as given here). Therefore, we modify the standard procedure by estimating the distribution of actions and outcomes conditional on observable states. In particular, we estimate the distribution of spending ($F_d(\cdot|\tilde{\Omega})$), savings ($F_{w'}(\cdot|\tilde{\Omega})$), fund-raising ($F_{fr}(\cdot|\tilde{\Omega})$), and the probability of incumbent victory conditional on $\tilde{\Omega}$, where $\tilde{\Omega} \equiv \{q_I, w_I, ten_I, X, \boldsymbol{\chi}' * \mathbf{1}\}$. All of the variables in $\tilde{\Omega}$ are observable (in particular, $\tilde{\Omega}$ does not include q_C) and, hence, these objects are estimable.

For the purpose of forward-simulating the value function, we need only the distribution of the actions conditional on $\tilde{\Omega}$ and not the policy function itself, which would depend on q_C. This is the case because the utility of the incumbent does not directly depend on q_C: It depends on q_C only indirectly through the change in the amount of money that needs to be raised and the change in the probability that the incumbent loses office. As a concrete example, consider evaluating the conditional expectation of C_I next period:

$$\begin{aligned} E_{\mathbf{s}'}[C_I(fr_I)|\mathbf{s}] &= E_{\mathbf{s}'|\mathbf{s}}\left[\int C_I(fr_I(\mathbf{s}', q_C))dF_{q_C}(q_C|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1, \mathbf{s}'\})\right] \\ &= E_{\mathbf{s}'|\mathbf{s}}\left[\int C_I(fr_I)dF_{fr_I}(fr_I|\tilde{\Omega})\right]. \end{aligned}$$

Note that integrating over $C_I(fr_I(\mathbf{s}', \cdot))$ with respect to q_C (first line) is equivalent to integrating over $C_I(\cdot)$ with respect to fr_I. As long as we can estimate the transition of s and the conditional distribution of fr_I, we can simulate the continuation value as functions of C_I, $\tilde{C}_I$, H_I, and σ_ε. The Online Appendix contains a more detailed explanation.

Once we can express the continuation value in equation (13) as functions of C_I, $\tilde{C}_I$, H_I, and σ_ε, we can use (13) as moment conditions to estimate these parameters.[30] The moment condition has to hold state by state. Once these model

[30]In practice, we also form additional moment conditions using the incumbent's first-order conditions in uncontested periods.

parameters are identified, the second and third lines of expression (13) identify the realization of q_C. Lastly, C_C is identified by the first-order conditions of the challengers.

4 Data

We now explain the data. The campaign finance data come from the Federal Election Commission. These data contain information on the amount of campaign contributions, candidate spending, and savings of all U.S. House candidates from 1984 to 2008. We augment these data with data on electoral returns, candidate characteristics and demographic characteristics of Congressional Districts. We collected the data on electoral returns and candidate characteristics from the database of the CQ Press and the demographic characteristics from the Census and the Bureau of Labor Statistics.

Before we present the summary statistics, recall that our empirical procedure requires using the state variables and the actions of the incumbent in uncontested periods ($\mathbf{s}_{NC}$). This means that we use only a subset of the elections for our estimation. In particular, for each incumbent i, we identify the first period $\overline{t_i}$ in which the incumbent is uncontested. We then keep only the set of elections that involve incumbent i that take place at or after $\overline{t_i}$. If an incumbent is contested on every occasion, we drop all elections involving that candidate.[31] The reason that we do not use elections prior to $\overline{t_i}$ is to avoid selection bias. By using actions and state variables in period $\overline{t_i}$ as control variables, we are implicitly conditioning on the event that incumbent i survived until period $\overline{t_i}$. If we were to include elections before $\overline{t_i}$, we would be selecting on unobservables; e.g., selection on the realizations of ε in the vote share equation.[32]

[31] In addition to incumbents who are never contested, we also drop Louisiana elections, elections in Texas that were affected by the Supreme Court ruling and elections involving a scandal. See the Online Appendix for more information on data construction.

[32] If an incumbent is uncontested in period $\overline{t_i}$, it implies that the incumbent won in all elections prior to $\overline{t_i}$. Hence, observing an incumbent in period $\overline{t_i}$ introduces a positive selection on ε for elections before $\overline{t_i}$ – the incumbent could not have experienced very low realizations of ε before $\overline{t_i}$. Otherwise, the incumbent would have lost, and we would not have observed her in period $\overline{t_i}$.

	(1)	(2)	(3)	(4)		(5)	(6)
Incumbent					Challenger		
Spending	460.4	299.4	470.3	277.7	Spending	98.2	153.5
	(358.1)	(214.0)	(357.9)	(204.8)		(232.7)	(287.6)
Fund-raising	475.5	367.4	501.4	347.0	Fund-raising	99.8	155.4
	(351.8)	(257.1)	(351.1)	(237.2)		(232.7)	(290.1)
War Chest	167.3	190.3	108.0	135.4	War Chest	0.1	0.5
	(189.1)	(228.3)	(161.3)	(188.8)		(0.6)	(4.1)
Savings	182.6	272.2	139.5	212.0	Savings	1.7	2.5
	(246.4)	(409.5)	(199.7)	(311.7)		(6.9)	(8.6)
Tenure	7.5	7.1	5.2	5.8	Tenure	0	0
	(4.2)	(3.9)	(4.0)	(4.2)		(0.0)	(0.0)
Vote Share	0.684		0.661	1.0	Vote Share	0.316	0.339
	(0.098)		(0.095)	(0.0)		(0.098)	(0.095)
Sample Size	963	387	4177	806		963	4177

Table 1: Discriptive Statistics of Incumbents and Challengers. Dollar values reported in units of $\$1,000$.

Table 1 presents the summary statistics of the key variables, where dollar values are normalized to 1984 dollars and reported in units of $1,000. Column (1) corresponds to sample statistics for the incumbents in contested elections (after $\overline{t_i}$), and column (5) corresponds to the challengers. In contested elections, incumbents start out with an average war chest of about $\$167$,000 and raise about $\$476$,000. The average spending is about $\$460$,000 and average saving is about $\$183$,000. The challengers, on the other hand, typically start out with zero war chest and raise about $\$100$,000, almost all of which is spent. Average incumbent vote share is 68.4%.

The second column corresponds to the sample statistics for the incumbents in uncontested elections (at or after $\overline{t_i}$). Uncontested incumbents start with an average war chest of $\$190$,000, which is slightly higher than the average war chest of contested incumbents. The average amount of money raised in uncontested periods is about $\$367$,000 and the average amount spent is about $\$299$,000. Incumbents save substantially more ($\$272$,000) in uncontested than in contested races ($\$183$,000).

In columns (3) and (4), we report the sample statistics of incumbents for all contested and uncontested elections, not just elections taking place after $\overline{t_i}$. The corresponding sample statistics for the challengers are reported in column (6). There are some interesting differences between our sample and the full sample, the biggest

	(1)		(2)		(3)		(4)	
	Rep	Dem	Rep	Dem	Rep	Dem	Rep	Dem
% Unemployed	0.053	0.064	0.050	0.063	0.054	0.061	0.052	0.064
	(0.017)	(0.021)	(0.017)	(0.026)	(0.019)	(0.021)	(0.018)	(0.025)
% White	0.860	0.789	0.855	0.766	0.869	0.800	0.855	0.765
	(0.086)	(0.135)	(0.096)	(0.146)	(0.085)	(0.137)	(0.097)	(0.145)
Sample Size	397	566	154	233	1951	2226	314	492

Table 2: Demographic Characteristics of Congressional Districts.

differences being the challenger's spending and fund-raising variables. For the sample we constructed, the mean challenger spending and amount raised are considerably less ($\$98{,}000$ and $\$100{,}000$) than the mean challenger spending and amount raised in the full sample ($\$153{,}000$ and $\$155{,}000$). The vote share for the incumbent is also slightly different, with 68.4% and 66.1%, respectively. This suggests that, on average, the challengers in the sample we constructed are weaker than in the full sample. Given our sample selection procedure, this seems natural. After all, we are selecting incumbents who were uncontested in at least one election – presumably, these incumbents are less vulnerable than incumbents who are always contested. Note that the similarity between the constructed sample and the full sample is desirable for generalizability, but it is not necessary for the consistency of our estimates. That is, our empirical approach controls for sample selection and gives consistent estimates of the primitives even if the sample that is used for estimation is not representative.

Table 2 reports the summary statistics of the demographic characteristics of each Congressional District by sample and by party of the incumbent. Columns (1) and (2) correspond to the summary statistics of contested and uncontested elections in our sample. Columns (3) and (4) correspond to the same summary statistics of the full sample. The first row reports the average unemployment, and the second row reports the average fraction of Whites in each Congressional District by party of the incumbent. We include these variables in our control variable, X. Note that unemployment tends to be high and the fraction of Whites low in districts held by a Democratic incumbent.

5 Specification and Estimation Results

5.1 Specification

We specify our model as follows. First, the benefit of office, B, is normalized to 1; the discount factor δ is set to 0.9; and the interest rate r is set to 10%.[33] Because Congressional elections take place every two years, $\delta = 0.9$ and $r = 10\%$ correspond to an annual discount of roughly 0.95 and an interest rate of about 5%. We specify the cost function of the incumbent in uncontested periods as quadratic in the log amount raised, $\tilde{C}_I(fr_I; q_I) = c(q_I)(\ln fr_I)^2$, where fr_I denotes the amount raised and $c(q_I)$ is a decreasing function of q_I. In practice, we estimate $c(\cdot)$ using B-splines of $z_I \equiv (\ln d_{NC})^{-1/2} d_{NC}^{-1} fr_{NC} (\ln fr_{NC})^{-1}$, where d_{NC} and fr_{NC} are spending and amount raised by the incumbent in the uncontested period (recall that q_I is a function of $\mathbf{s}_{NC}$, which, in turn, becomes a function of z_I under our specification. See, also the Online Appendix). We specified the cost function of the incumbent in contested periods, $C_I(\cdot; q_I)$, and the cost function of the challengers, $C_C(\cdot; q_C)$, by scaling $\tilde{C}_I$ as $C_I(\cdot; q_I) = \theta_I \tilde{C}_I(\cdot; q_I)$ and $C_C(\cdot; q_C) = \theta_C \tilde{C}_I(\cdot; q_I)$. We specified H as $H(d_I) = \gamma\sqrt{\ln d_I}$ and we specified the retirement probability, $\lambda(\mathbf{s})$, as a function of ten_I.

5.2 Parameter Estimates

First-Stage Estimates Table 3 presents the parameter estimates of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$, which are the probability of challenger entry and the average number of entrants. $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are estimated as functions of $\mathbf{s} = (q_I, w_I, ten_I, D_I, X)$, which are, in turn, functions of (z_I, w_I, ten_I, D_I, X). We estimate P_e using a Probit in w_I, ten_I, $D_I X$ and B-spline bases of z_I. $E[M|\mathbf{s}]$ was estimated as a linear function of w_I, ten_I, $D_I X$ and B-spline bases of z_I.[34] The Online Appendix contains a detailed explanation of the estimation.

In the first column, we find that the coefficient on $\ln w_I$ is -0.056 and $\ln ten_I$ is

[33] We assume that the savings of the candidates collect 10% interest by the next election.

[34] Because P_e and $E[M]$ are both equilibrium objects, they should be estimated nonparametrically. Given the moderate sample size and the number of state variables, however, there is a limit to how flexibly we can estimate these objects in practice.

$P_e(s)$		$E[M\|s]$	
const.	1.448 (0.475)	const.	2.353 (0.373)
$\ln w_I$	-0.056 (0.029)	$\ln w_I$	-0.065 (0.034)
$\ln ten_I$	0.165 (0.089)	$\ln ten_I$	0.238 (0.076)
$D_I \times$Unemply	-2.558 (1.982)	$D_I \times$Unemply	-3.333 (1.663)
$D_I \times$% White	0.184 (0.049)	$D_I \times$% White	0.240 (0.130)
B-spline of z_I	Yes	B-spline of z_I	Yes

Table 3: Estimates of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$. Bootstrapped standard errors reported in parentheses.

0.165, which suggest that a one standard deviation increase in incumbent war chest translates to a decrease in the probability of challenger entry by about 3.50%, and a one standard deviation increase in ten_I leads to an increase in the entry probability by about 6.29% at the mean. The estimated coefficients on $D_I\beta_X$ imply that a higher unemployment rate decreases entry, while a higher fraction of Whites increases entry when the incumbent is a Democrat (Recall that D_I is a dummy variable that takes the value 1 if the incumbent is a Democrat and -1 if the incumbent is a Republican.). The opposite is true when the incumbent is a Republican. In the second column, we find that $\ln w_I$, $\ln ten_I$ and X affect $E[M|\mathbf{s}]$ in the same way they affect $P_e(\mathbf{s})$. Note that our results suggest that sample selection is potentially important. To the extent that variables such as $\ln ten_I$ and X affect $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ (and hence $E[q_C|\mathbf{s}]$), these variables cannot be included in the vote share regression as exogenous variables.

Next, we report the estimates of the vote share function, which was specified as follows:

$$vote_I = \beta_I \ln d_I + \beta_C \ln d_C + \beta_{ten} ten_I + D_I \beta_X X + q_I - q_C + \varepsilon. \qquad (2)$$

As discussed in Section 3.1, we estimate this equation by including additional re-

(Control Function)		(OLS)	
β_I	5.56×10^{-3} (1.83×10^{-2})	β_I	-1.82×10^{-2} (3.78×10^{-3})
β_C	-2.39×10^{-2} (1.14×10^{-2})	β_C	-2.92×10^{-2} (1.67×10^{-3})
β_{ten}	2.47×10^{-2} (3.41×10^{-2})	β_{ten}	-1.55×10^{-2} (5.30×10^{-3})
β_{Un}	-3.58×10^{-1} (5.37×10^{-1})	β_{Un}	1.23×10^{-1} (1.21×10^{-1})
β_{Wh}	-8.65×10^{-2} (5.50×10^{-2})	β_{Wh}	-5.49×10^{-3} (8.87×10^{-3})

Table 4: Parameter Estimates of the Vote Share Equation. First column corresponds to the estimates obtained by using the procedure proposed in the paper. Second column corresponds to OLS estimates. Standard errors reported in parentheses.

gressors ($\widehat{P_e}$, $\widehat{E[M]}$) and projecting the variables on Ω.[35] The parameters are estimated using sieve minimum distance estimator (see, e.g., Ai and Chen (2005)).[36] The first column of Table 4 reports our parameter estimates. Our point estimates of β_I and β_C are about 5.6×10^{-3} and -2.4×10^{-2}, respectively, which implies that a one standard deviation increase in the spending of the incumbent leads to an increase in the incumbent vote share of about 0.4%, while a one standard deviation increase in the challengers' spending leads to a decrease in the incumbent's vote share of about 3.9%. Regarding β_{ten} and $\beta_X = (\beta_{Un}, \beta_{Wh})$, we estimate β_{ten} to be around 2.47×10^{-2}, β_{Un} to be -0.36 and β_{Wh} to be -8.65×10^{-2}.

In the second column of Table 4, we report the OLS estimates of equation (2) for comparison. In the OLS estimation, $(q_I - q_C + \varepsilon)$ is treated as the error term in equation (2). Note that the OLS estimate of β_I is *negative*, at -1.82×10^{-2}. If we were to interpret this estimate as a causal effect, it means that a one standard deviation increase in incumbent spending *decreases* incumbent vote share by about 1.3%. We also find that OLS overestimates the magnitude of β_C relative to the estimate in the first column.

[35]The Online Appendix contains a more detailed discussion of our estimation.
[36]Standard errors are bootstrapped.

In Figure **??**, we plot the histogram of the estimated incumbent quality, q_I. Recall that our measure of quality is in units of vote share: if candidate 1 and candidate 2 have quality q_1 and q_2, respectively, candidate 1 obtains $(q_1 - q_2)$ higher vote share than candidate 2, on average, everything else equal. While the level of these measures is not informative (In our setting, q_I is estimated only relative to q_C), quality differences between candidates are informative about the differential vote-getting ability of the candidates. For example, the interquartile range for q_I is about 0.021 in Figure **??**, which implies that an incumbent at the 75% quantile of the quality distribution obtains about a 2.1% higher vote share than an incumbent at the 25% quantile, ceteris paribus.

Second-Stage Estimates In Table 5, we report the results of the second-stage estimation. This estimation requires a preliminary estimate of the distribution of the policy function, but we focus here on the estimates of the structural parameters.[37] We specify the cost functions $C_I(\cdot)$ and $C_C(\cdot)$ by rescaling $\tilde{C}_I(\cdot)$ as $C_I = \theta_I \tilde{C}_I$ and $C_C = \theta_C \tilde{C}_I$. Hence, θ_I measures the relative fund-raising cost that the incumbent incurs in contested versus uncontested periods. Likewise, θ_C measures the relative cost that the challenger incurs compared to incumbents in contested periods. Note that H is specified as $H(y) = \gamma\sqrt{y}$. We estimate θ_I to be 5.02 and θ_C to be 7.70, suggesting that it is about 50% more ($7.70/5.02$) costly for the challenger to raise the same amount of money as the incumbents in contested elections. Our estimate of σ_ε, the standard deviation of the error term in expression (2), is about 0.04.

Next, in the top panel of Figure **??**, we plot the histogram of q_C. In the bottom panel, we superimpose the distribution of q_I (See Figure **??**) on q_C. We find that the distribution of q_C has a large dispersion, where the top 1% are comparable to the incumbents, while the median value of q_C is around -0.32, which is more than 0.3 points lower than the median of q_I. This implies that the median challenger receives a 30% smaller vote share, on average, than the median incumbent, ceteris paribus.

[37]The distribution of the actions are estimated using a (first-order) Hermite series approximation following Gallant and Nychka (1987). See Online Appendix for details.

θ_I	5.020 (12.051)
θ_C	7.698 (26.486)
γ	3.249×10^{-2} (4.935×10^{-2})
σ_ε	4.158×10^{-2} (2.564×10^{-3})

Table 5: Second-Stage Parameter Estimates. Standard Errors given in parentheses.

6 Counterfactual Analysis

In our counterfactual analysis, we consider the effect of two campaign finance policies. In particular, we consider the effect of government subsidies to challengers in the form of a public fund that matches either one dollar or two dollars for every dollar that a challenger raises. Often, public financing of campaigns comes with a spending limit in exchange for public funds (e.g., public financing of Presidential campaigns), with the result that candidates who can easily raise money opt out of these programs. The counterfactual policies that we consider approximate this by targeting the challengers only.

In order to compare the counterfactual results with the baseline case of no subsidies, we first recompute the equilibrium with the estimated parameters.[38] In column (0) of Table 6, we report the mean log incumbent spending ($\ln d_I$), log challenger spending ($\ln d_C$), incumbent vote share ($vote_I$) and probability of challenger victory, for the baseline case of no subsidies. These numbers are obtained by solving for the equilibrium at the estimated parameter values and evaluating the policy functions at the observed state variables and candidate quality. We find that mean log incumbent spending is about 12.71 and mean log challenger spending is about 9.29. Given that the mean log incumbent spending and log challenger spending are 12.86 and 10.49 in our sample, the model does a reasonable job of fitting incumbent

[38]Because we do not estimate the model of the Primary, computing the equilibrium requires one to compute the equilibrium P_e and F_{q_C} without solving the model of the Primary. We discuss how this can be done in the Online Appendix.

spending, while it slightly underpredicts challenger spending.[39] This is because the model predicts that low quality challengers with little chance of winning to spend very little, while the data show that even low quality challengers seem to engage in some spending. If we focus on the very high quality challengers, however, the model actually slightly over predicts challenger spending. In the last two rows of column (0), we report the mean incumbent vote share and mean incumbent winning probability. We find the mean incumbent vote share to be 70.25% and the mean winning probability to be 8.55%.

In the next three columns under (1-a), we report the equilibrium effects of the dollar-for-dollar subsidy. The column labeled Δmean reports the mean change, the column labeled $\Delta 5\%$ reports the 5% quantile of the change, and the column labeled $\Delta 95\%$ reports the 95% quantile of the change. We find that the equilibrium mean log spending of the incumbents increases slightly by 0.01 compared to column (0). We find that the 5% and the 95% quantiles of the change in $\ln d_I$ are -0.17 and 0.21, respectively. This means that 5% of the incumbents would change log spending by -0.17, or less; and 95% of the incumbents would change log spending by 0.21, or less. The incumbent spending increases in states when the incumbent is vulnerable in order to offset increased challenger spending, while it decreases when the incumbent is in a formidable state.

In the next row, we report the change in equilibrium mean log spending of the challengers. We find that the log spending of the challengers changes by 0.06 at the mean, and by -0.00 and 0.30 at the 5% and the 95% quantiles, respectively. We find that there is substantial crowding out of challenger spending, leaving the challenger in some states with almost the same amount of spending as the baseline case. We also find that the correlation coefficient between the change in the incumbents' spending and the change in the challengers' spending is about 0.4, suggesting that there are strong strategic complementarities.

In the next two rows, we report the changes in mean incumbent vote share and the mean challenger winning probability. We find that the mean incumbent vote

[39] Note that if we take the log of mean incumbent and challenger spending in Table 1, we get 13.04 and 11.49. The reason these numbers are different than in the main text is because average log is different to log average.

share decreases by about 0.13%, on average, while decreasing by about 0.63% at the 5% quantile and increasing by about 0.03% at the 95% quantile. The change in the probability of challenger victory is about 0.09% at the mean.

Overall, the dollar-for-dollar subsidy induces only modest changes in the equilibrium outcomes. This is in contrast to the results that we present in column (1-b) where we ignore the equilibrium responses of the agents. In this column, we fix incumbent spending at the original level while doubling the challengers' spending level. Unlike in column (1-a), we now find that the effect of the subsidy is quite strong, reducing incumbent vote share by 1.65%, on average, and increasing the challengers' winning probability by 1.46%, on average. The increase in the winning probability represents about a 17% increase from the baseline probability of 8.55%. The difference between columns (1-a) and (1-b) underscores the importance of factoring in equilibrium responses of the players in evaluating counterfactual policies.

The three columns under (2-a) report the effects of a subsidy that matches every dollar raised by a challenger with two dollars of public money. In the first row, we find that incumbent spending changes very little at the mean, while the changes are somewhat more amplified than in (1-a) for the 5% and the 95% quantiles. In the second row, we find that challenger spending increases only by about 0.08 in equilibrium, suggesting that there is a lot of crowding out as before. Overall, this policy also has modest equilibrium effects, decreasing mean incumbent vote share by 0.20% and increasing the winning probability of the challengers by 0.21%.

In the last column, we report the results from an experiment where we triple the challengers' spending relative to the baseline level and keep everything else constant. Now, we find a very large effect, with mean incumbent vote share decreasing by 2.62% and increasing the winning probability of the challenger by 2.39%.

7 Conclusion

In this paper, we estimate a dynamic model of campaign financing with unobserved candidate heterogeneity. We develop a control function approach that allows us to consistently estimate the vote share equation in the presence of endogeneity and selection. Our approach also allows us to recover candidate quality, which can be

	(0)	(1-a)			(1-b)	(2-a)			(2-b)
		Δmean	$\Delta 5\%$	$\Delta 95\%$	Δmean	Δmean	$\Delta 5\%$	$\Delta 95\%$	Δmean
$\ln d_I$	12.71	+0.01	-0.17	+0.21	0	+0.00	-0.23	+0.29	0
$\ln d_C$	9.29	+0.06	-0.00	+0.30	+0.69	+0.08	-0.00	+0.38	+1.10
$vote_I$	70.25%	-0.13%	-0.63%	+0.03%	-1.65%	-0.20%	-0.70%	+0.02%	-2.62%
win_C	8.55%	+0.09%	-0.00%	+0.68%	+1.46%	+0.21%	-0.00%	+1.28%	+2.39%

Table 6: Counterfactual Results. Column (0) corresponds to the baseline with no subsidies. The three columns under (1-a) correspond to the dollar-for-dollar subsidy and columns under (2-a) correspond to the two dollars-for-dollar subsidy. Columns (1-b) and (2-b) report the effect of the policies ignoring any equilibrium responses.

useful for studying other topics in political economy.

Our estimates of the vote share function suggest that a one standard deviation increase in the incumbent's spending yields about a 0.4% increase in her vote share, while a standard deviation increase in the challenger's spending yields about a 3.9% increase in her vote share. While these estimates are in line with the results reported in the previous literature, our counterfactual policy analysis underscores the importance of accounting for the equilibrium responses of the candidates.

References

[1] Ackerberg, Daniel, Kevin Caves, and Garth Frazer (2006), "Structural Identification of Production Functions," Unpublished Manuscript, UCLA Economics Department.

[2] Ai, Chunrong and Xiaohong Chen (2003), "Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions," *Econometrica*, Vol. 71, No 6, pp. 1795–1843.

[3] Ansolabehere, Stephen and James Snyder (2004), "Using Term Limits to Estimate Incumbency Advantages When Officeholders Retire Strategically," *Legislative Studies Quarterly,* Vol. 29, No 4, pp. 487–515.

[4] Bajari, Patrick, C. Lanier Benkard and Jonathan Levin (2007), “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, Vol.75, No 5, pp. 1331–1370.

[5] Cox, Gary W., and Michael F. Thies (2000) “How Much Does Money Matter?: ”Buying” Votes in Japan, 1967-1990,” *Comparative Political Studies* Vol. 33, No. 1, pp. 37–57.

[6] da Silveira, Bernardo, and Joaõ de Mello (2011) “Campaign Advertising and Election Outcomes: Quasi-natural Experiment Evidence from Gubernatorial Elections in Brazil,” *Review of Economic Studies*, Vol. 78 (2), 590-612.

[7] Diermeier, Daniel, Michael Keane and Antonio Merlo (2005), “Political Economy Model of Congressional Careers,” *American Economic Review* , Vol. 95, No. 1 (Mar., 2005), pp. 347-373

[8] Ericson, Richard, and Ariel Pakes (2001), “Markov-Perfect Industry Dynamics: A Framework for Empirical Work,” *Review of Economic Studies*, Vol. 62, No. 1, pp. 53-82

[9] Erikson, Robert, and Thomas Palfrey (2000) “Equilibria in Campaign Spending Games: Theory and Data,” *American Political Science Review*, Vol. 94, No. 3, pp. 595-609

[10] Gallant, A. Ronald, and Douglas W. Nychka, “Semi-Nonparametric Maximum Likelihood Estimation,” *Econometrica*, Vol. 55(2), 1987, pp. 363-390.

[11] Gerber, Alan (1998)“Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables,” *American Political Science Review*, Vol. 92, No. 2, pp. 401-411

[12] Green, Donald, and Jonathan Krasno, (1988) “Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections,” *American Journal of Political Science*, Vol. 32, No. 4, pp. 884-907

[13] Hotz, V. Joseph and Robert Miller (1993), "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies,* Vol. 60, No. 3 , pp. 497-529

[14] Hotz, V. Joseph, Robert Miller, Seth Sanders and Jeffery Smith (1994), "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies,* Vol. 61, No. 2 , pp. 265-289

[15] Jacobson, Gary C. (1978) "The Effects of Campaign Spending in Congressional Elections," *American Political Science Review*, Vol. 72, No. 2, pp. 469-491

[16] Levinsohn, James and Amil Petrin (2003) "Estimating Production Function Using Inputs to Control for Unobservables," *Review of Economic Studies*, Vol. 70, No. 2, pp 317-341.

[17] Levitt, Steven (1994) "Using Repear Challengers to Estimate the Effect of Campaign Spending on Election Outcomes in the U.S. House," *Journal of Political Economy*, Vol. 102, No. 4, pp. 777-798.

[18] Magnac, Thierry and David Thesmar (2002), "Identifying Dynamic Discrete Decision Processes," *Econometrica*, Vol. 70, No. 2, pp. 801-816.

[19] Maskin, Eric and Jean Tirole (1988). "A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs," *Econometrica, Vol.* 56, pp. 549-69.

[20] Olley, Steven and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, Vol. 64, No. 6 , pp. 1263-1297

[21] Stratmann, Thomas (2005), "Some talk: Money in politics. A (partial) review of the literature," *Public Choice,* 124,135–156

[22] United States General Accounting Office, "Campaign Finance Reform: Early Experiences of Two States that Offer Full Public Funding for Political Candidates," May 2003.

[23] Whitt, Ward (1980), "Representation and Approximation of Noncooperative Sequential Games," *SIAM Journal of Control and Optimization*, Vol. 18 (1980), pp. 33–48

8 Online Appendix [For Online Publication]

In the Online Appendix, we provide the proofs that we ommitted from the main text as well as details regarding estimation, data construction and computation of the counterfactuals. In Section 1, we prove the sufficient statistics property; and in Section 2, we prove that q_I can be inverted as a function of the state variables and the actions of the incumbent in uncontested periods. In Section 3, we discuss how we forward-simulate the continuation value. Section 4 explains data construction, Section 5 discusses details on estimation, and Section 6 explains how we compute the counterfactuals.

8.1 Sufficient Statistics

Recall that our identification and estimation both make use of the fact that the distribution of challenger quality, $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, depends on $\mathbf{s}$ only through sufficient statistics. The set of sufficient statistics, however, depends on how the distribution of N (the number of potential entrants) is specified. For example, when the distribution of N is a constant that does not depend on $\mathbf{s}$, it is easy to see that $P_e(\mathbf{s})$ is a sufficient statistic. When the distribution of N is specified as a negative binomial function as $NB(r(\mathbf{s}), p)$ – which is the specification we use in the estimation – we claim that $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are sufficient statistics. In this Appendix, we first prove this result and then discuss how the result extends to the case when we relax the distributional assumption on N.

Lemma 1 *Suppose that $F_N(\cdot|\mathbf{s})$ is given by the negative binomial distribution, $F_N(\cdot|\mathbf{s}) = NB(r(\mathbf{s}), p)$, where $r(\cdot)$ can be an arbitrary function of $\mathbf{s}$. Then, $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are sufficient statistics for $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$ and $p(q_C, \mathbf{s})$.*

Proof. Recall from expression (7) and (8) that $p(q_C, \mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$ have the following expressions:

$$\begin{aligned} &F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}) \\ &\quad = \mathbf{E}_N \left[\sum_{M=1}^{N} Bin(N, M; 1 - F_{q_C}(\bar{q}_C)) \frac{\int_{\bar{q}_C}^{t} \int_{\bar{q}_C}^{+\infty} \cdots \int_{\bar{q}_C}^{+\infty} M\pi(q_{C,n}, \mathbf{q}_{C,-n})(dF_{q_C})^M}{(1 - F_{q_C}(\bar{q}_C(\mathbf{s})))^M} \middle| \mathbf{s} \right], \end{aligned}$$

and

$$p(q_C, \mathbf{s}) = \mathbf{E}_N \left[\sum_{M=1}^{N} Bin(N, M; 1 - F_{q_C}(\bar{q}_C)) \frac{\int_{\bar{q}_C}^{+\infty} \cdots \int_{\bar{q}_C}^{+\infty} \pi(q_{C,n}, \mathbf{q}_{C,-n})(dF_{q_C})^{M-1}}{(1 - F_{q_C}(\bar{q}_C))^{M-1}} \middle| \mathbf{s} \right].$$

Note that $\bar{q}_C(\mathbf{s})$ and $r(\mathbf{s})$ are sufficient statistics, by inspection. That is, as long as $\bar{q}_C(\mathbf{s}) = \bar{q}_C(\mathbf{s}')$ and $r(\mathbf{s}) = r(\mathbf{s}')$, we have $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}) = F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}')$ (also $p(q_C, \mathbf{s}) = p(q_C, \mathbf{s}')$). In order to show that $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are also sufficient statistics, it then suffices to show that whenever $P_e(\mathbf{s}) = P_e(\mathbf{s}')$ and $E[M|\mathbf{s}] = E[M|\mathbf{s}']$, we have $\bar{q}_C(\mathbf{s}) = \bar{q}_C(\mathbf{s}')$ and $r(\mathbf{s}) = r(\mathbf{s}')$.

Recall from expression (9) that

$$E[M|\mathbf{s}] = \mathbf{E}_N \left[\sum_{M=1}^{N} M \cdot Bin(N, M; 1 - F_{q_C}(\bar{q}_C(\mathbf{s}))) \middle| \mathbf{s} \right] = (1 - F_{q_C}(\bar{q}_C(\mathbf{s}))) \mathbf{E}_N [N|\mathbf{s}].$$

When $F_N(\cdot|\mathbf{s})$ is the negative binomial distribution, the expression for $E[M|\mathbf{s}]$ becomes

$$E[M|\mathbf{s}] = (1 - F_{q_C}(\bar{q}_C(\mathbf{s}))) \frac{pr(\mathbf{s})}{1 - p}, \tag{14}$$

where we have used the fact that the mean of the negative binomial distribution $NB(r, p)$ is just $\frac{pr}{1-p}$.

On the other hand, $P_e(\mathbf{s})$ has the following form (see equation (6)):

$$\begin{aligned} P_e(\mathbf{s}) &= \mathbf{E}_N\left[1 - F_{q_C}(\bar{q}_C(\mathbf{s}))^N \middle| \mathbf{s}\right] \\ &= 1 - \sum_{N=0}^{+\infty} \binom{N + r(\mathbf{s}) - 1}{N} p^N (1-p)^{r(\mathbf{s})} F_{q_C}(\bar{q}_C(\mathbf{s}))^N \\ &= 1 - \sum_{N=0}^{+\infty} \binom{N + r(\mathbf{s}) - 1}{N} (F_{q_C}(\bar{q}_C(\mathbf{s}))p)^N (1 - F_{q_C}(\bar{q}_C(\mathbf{s}))p)^{r(\mathbf{s})} \times \frac{(1-p)^{r(\mathbf{s})}}{(1 - F_{q_C}(\bar{q}_C(\mathbf{s}))p)^{r(\mathbf{s})}} \\ &= 1 - \left(\frac{1-p}{1 - F_{q_C}(\bar{q}_C(\mathbf{s}))p}\right)^{r(\mathbf{s})}, \end{aligned} \tag{15}$$

where the second line follows from the definition of the probability mass function of the negative binomial distribution, and the fourth line follows from the fact that the probability mass function sums up to one. In order to show that $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ are sufficient statistics, it suffices to show that we can uniquely solve for $r(\mathbf{s})$ and $\bar{q}_C(\mathbf{s})$ in equations (14) and (15) as functions of $E[M|\mathbf{s}]$ and $P_e(\mathbf{s})$.[40]

With that in mind, we first take expression (14) and solve for $F_{q_C}(\bar{q}_C(\mathbf{s}))$:

$$F_{q_C}(\bar{q}_C(\mathbf{s})) = 1 - \frac{E[M|\mathbf{s}](1-p)}{pr(\mathbf{s})}. \tag{16}$$

We then substitute out $F_{q_C}(\bar{q}_C(\mathbf{s}))$ from expression (15):

$$\begin{aligned} P_e(\mathbf{s}) &= 1 - \left(\frac{1-p}{1 - \left(1 - \frac{E[M|\mathbf{s}](1-p)}{pr(\mathbf{s})}\right)p}\right)^{r(\mathbf{s})} \\ &= 1 - \left(\frac{1-p}{1 - p + \frac{E[M|\mathbf{s}](1-p)}{r(\mathbf{s})}}\right)^{r(\mathbf{s})} = 1 - \left(\frac{1}{1 + \frac{E[M|\mathbf{s}]}{r(\mathbf{s})}}\right)^{r(\mathbf{s})}. \end{aligned} \tag{17}$$

If we can show that the right-hand side of expression (17) is monotone in $r(\mathbf{s})$, this implies that we can express $r(\mathbf{s})$ uniquely as a function of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$. The

[40] Suppose that we can uniquely solve for $r(\mathbf{s})$ and $\bar{q}_C(\mathbf{s})$ as functions of $E[M|\mathbf{s}]$ and $P_e(\mathbf{s})$. Then, if we have $E[M|\mathbf{s}] = E[M|\mathbf{s}']$ and $P_e(\mathbf{s}) = P_e(\mathbf{s}')$, we would have $r(\mathbf{s}) = r(\mathbf{s}')$ and $\bar{q}_C(\mathbf{s}) = \bar{q}_C(\mathbf{s}')$. Given that $r(\mathbf{s})$ and $\bar{q}_C(\mathbf{s})$ are sufficient statistics, this means that $E[M|\mathbf{s}]$ and $P_e(\mathbf{s})$ are also sufficient statistics.

proof would then be done because, together with equation (16), the monotonicity would ensure that both $r(\mathbf{s})$ and $\bar{q}_C(\mathbf{s})$ are expressed uniquely as a function of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$.

In order to show that the right-hand side of expression (17) is monotone in $r(\mathbf{s})$, consider a function $f(x)$ defined as follows:

$$f(x) = \frac{1}{x}\ln(1+x),\ (x>0)$$

It is easy to see from Figure 2 that $f(x)$ is monotone decreasing in x ($x > 0$).

Given that $f(x)$ is monotone decreasing, $\exp(-\alpha f(\alpha/x))$ is monotone decreasing for $x > 0$ for any constant $\alpha > 0$, where

$$\exp(-\alpha f(\alpha/x)) = \left(\frac{1}{1+\frac{\alpha}{x}}\right)^x.$$

By inspection, it is now easy to see that the right-hand side of expression (17) is monotone increasing in $r(\mathbf{s})$, and we are done. ■

Note that even if we do not impose any assumption on $F_N(\cdot|\mathbf{s})$), $\bar{q}_C(\mathbf{s})$ and $F_N(\cdot|\mathbf{s})$ are still sufficient statistics for $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$. Hence, if we relax our distributional assumption on $F_N(\cdot|\mathbf{s})$ so that $F_N(\cdot|\mathbf{s})$ now depends on $r_1(\mathbf{s})$,...,$r_L(\mathbf{s})$, – i.e., $F_N(\cdot|\mathbf{s}) = F_N(\cdot|r_1(\mathbf{s}),...,r_L(\mathbf{s}))$ – then $\bar{q}_C(\mathbf{s})$, $r_1(\mathbf{s})$,...,$r_L(\mathbf{s})$ would be sufficient statistics. When $L > 1$, we would need more functionals of $F_M(\cdot|\mathbf{s})$ for the sufficient statistic result to hold, unlike the case above where we needed just two ($P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$). However, it is intuitive to see that we can obtain a similar sufficient statistic result by increasing the set of variables (e.g., quantiles of M) to condition on in addition to $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$.

8.2 Inversion of q_I from uncontested periods

Recall that we use a control function approach to express the unobserved incumbent quality, q_I, as a function of states and incumbent's actions in uncontested periods. As explained in the main text, this requires that we invert the policy function of the incumbent when she is uncontested. Recall the problem of the incumbent when she is uncontested:

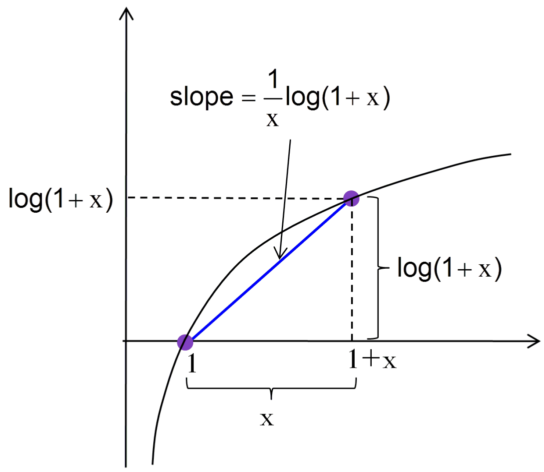

Figure 2: A graphical explanation of why $f(x)$ is monotone. Note that $\frac{1}{x}\log(1+x)$ corresponds to the slope of $\log(t)$ between $t = 1$ and $t = 1 + x$. Because $log(t)$ is concave, $f(x)$ is decreasing.

$$v_I^{2(c)}(\mathbf{s}) = \max_{w_I' \geq 0, d_I \geq 0} B - \tilde{C}_I(w_I' + d_I - w_I, q_I) + H_I(d_I) + \delta E_{r,x'}[V_I(\mathbf{s}')].$$

The policy functions associated with this problem are how much to save, $w_I' = w_I'(\mathbf{s})$, and how much to spend, $d_I = d_I(\mathbf{s})$. Note that the policy functions can be viewed as mappings from q_I to (w_I', d_I), holding fixed the other state variables. Below, we show that the mapping $q_I \mapsto (w_I', d_I)$ is one-to-one for the functional form we use in our estimation.[41]

Lemma 2 *Suppose that $\tilde{C}_I(y; q_I) = c(q_I)(\ln y)^2$, where $c(\cdot)$ is any decreasing function, and $H(y) = \gamma\sqrt{\ln y}$ as specified in our estimation. Then, the mapping from q_I to (w_I', d_I) is one-to-one given other state variables.*

[41] More precisely, the one-to-one property holds only when $d_I > 0$. This condition is always satisfied in our sample.

Proof. Note that the first-order condition for d_I implies

$$\underbrace{\frac{\partial}{\partial d_I} H_I(d_I)}_{\text{MB of spending}} - \underbrace{\frac{\partial}{\partial d_I} \tilde{C}_I(w_I' + d_I - w_I, q_I)}_{\text{MC of fund-raising}} = 0.$$

Substituting $\tilde{C}_I(y; q_I) = c(q_I)(\ln y)^2$ and $H(y) = \gamma\sqrt{\ln y}$ into the previous expression, we obtain

$$\begin{aligned} & \frac{1}{2}\gamma(\ln d_I)^{-1/2}(d_I)^{-1} - 2c(q_I)(\ln fr_I)(fr_I)^{-1} = 0 \qquad (18) \\ \iff \quad & c(q_I) = \frac{1}{4}\gamma(\ln d_I)^{-1/2}(d_I)^{-1}(\ln fr_I)^{-1} fr_I \\ \iff \quad & q_I = c^{-1}\left(\frac{1}{4}\gamma\frac{fr_I}{(\ln d_I)^{1/2} d_I \ln fr_I}\right), \end{aligned}$$

where fr_I denotes the amount raised ($fr_I = w_I' + d_I - w_I$). We have used the fact that $c(\cdot)$ is monotone to obtain the last line of the above expression. Equation (18) implies that if spending (d_I) and amount raised are the same (given other state variables), q_I has to be the same. This concludes the proof. ■

Lemma 3 *Suppose, alternatively, that $\tilde{C}_I(y; q_I) = c(q_I)y^\alpha$, where $c(\cdot)$ is any decreasing function, and $H(y) = \gamma y^\beta$. Then, the mapping from q_I to (w_I', d_I) is one-to-one given other state variables.*

Proof. Note that the first-order condition for d_I implies

$$\underbrace{\frac{\partial}{\partial d_I} H_I(d_I)}_{\text{MB of spending}} - \underbrace{\frac{\partial}{\partial d_I} \tilde{C}_I(w_I' + d_I - w_I, q_I)}_{\text{MC of fund-raising}} = 0.$$

Substituting $\tilde{C}_I(y; q_I) = c(q_I)y^\alpha$ and $H(y) = \gamma y^\beta$ into the previous expression, we obtain

$$\begin{aligned} & \frac{1}{2}\gamma(\ln d_I)^{-1/2}(d_I)^{-1} - 2c(q_I)(\ln fr_I)(fr_I)^{-1} = 0 \qquad (19) \\ \iff \quad & c(q_I) = \frac{1}{4}\gamma(\ln d_I)^{-1/2}(d_I)^{-1}(\ln fr_I)^{-1} fr_I \\ \iff \quad & q_I = c^{-1}\left(\frac{1}{4}\gamma \frac{fr_I}{(\ln d_I)^{1/2} d_I \ln fr_I}\right), \end{aligned}$$

where fr_I denotes the amount raised ($fr_I = w_I' + d_I - w_I$). We have used the fact that $c(\cdot)$ is monotone to obtain the last line of the above expression. Equation (18) implies that if spending (d_I) and amount raised are the same (given other state variables), q_I has to be the same. This concludes the proof. ■

Note that given our functional form assumptions, q_I can be expressed as a function of $\frac{fr_I}{(\ln d_I)^{1/2} d_I \ln fr_I}$ ($= z_I$) only. The fact that we can control for q_I just by conditioning on a one-dimensional object, z_I, simplifies our estimation immensely. It would be practically impossible to implement our procedure if we had to condition on the full vector of actions and state variables, $\mathbf{s}_{NC}$.

8.3 Forward-Simulation of the Continuation Value

In this section, we discuss how we forward-simulate the continuation value. As we discussed in Section 3.2, our idea is based on Hotz and Miller (1993), Hotz, Miller, Sanders, and Smith (1994) and Bajari, Benkard and Levin (2007). These papers propose a method of simulating the value function by first estimating the policy function and then using the policy function to generate sample paths of outcomes and actions, which can then be averaged to compute the continuation value. Because we do not observe the full set of state variables in contested periods (we do not observe q_C in contested periods), we estimate the distribution of the actions and outcomes conditional on observed state variables instead of the actual policy function (which depends on q_C). Below, we describe the details of our procedure.

8.3.1 Estimation of the transition probability of the states

First, we estimate the transition probability of X:

Transition of X Estimate the transition probability of demographic characteristics from X to X'. We assume an exogenous AR(1) process for each element of X as $X^k_{t+1} = \alpha^k_0 + \alpha^k_1 X^k_t + \xi^k_{t+1}$, and $\xi^k_{t+1} \sim N(0, \sigma^k_\xi)$, where X^k_{t+1} is the k-th component of X. We let the first component of X be the fraction of Whites in the Congressional District and the second component be the fraction of unemployed people in the Congressional District.[42]

8.3.2 Estimation of the distribution of actions conditional on observed state variables

The second set of objects that we estimate are the projections of the policy functions on observed state variables. The relevant objects we estimate are as follows:

Distribution of d_I conditional on $\tilde{\Omega}$ in contested periods Recall that the equilibrium spending of the incumbent in contested periods maps $(\tilde{\Omega}, q_C)$ to a non-negative number, where $\tilde{\Omega} \equiv \{q_I, w_I, ten_I, X, D_I, \{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}\}$. The projection of the policy function on $\tilde{\Omega}$ is just the conditional distribution of d_I in contested periods given observable states, denoted as $F_{d_I}(\cdot|\tilde{\Omega})$. We estimate the conditional distribution using nonparametric maximum likelihood (Gallant and Nychka 1987).

Distribution of fr_I conditional on $\tilde{\Omega}$ in contested periods We estimate the conditional distribution of the amount raised by the incumbent in contested periods given $\tilde{\Omega}$, denoted as $F_{fr_I}(\cdot|\tilde{\Omega})$. We let fr_I ($= w'_I + d_I - w_I$) denote the amount raised by the incumbent.

Distribution of w'_I conditional on $\tilde{\Omega}$ in contested periods We estimate the conditional distribution of incumbent savings in contested periods, given incumbent

[42] First component of X is the fraction of Whites and the second component is the fraction of unemployed. We estimated $\alpha^1_0 = 0.006$, $\alpha^1_1 = 0.988$, $\sigma^1_\xi = 0.023$ and $\alpha^2_0 = 0.017$, $\alpha^2_1 = 0.697$, $\sigma^2_\xi = 0.014$

victory and $\tilde{\Omega}$, denoted as $F_{w'_I}(\cdot|\tilde{\Omega}, \{v_I > 0.5\})$. For simulating the value function, we need the distribution of savings conditional on winning.

Policy functions in uncontested periods We estimate the mapping from observable states to spending, savings and amount raised in uncontested periods. In uncontested periods, we can estimate the policy function itself because the state variables are all observed.

8.3.3 Estimation of the distribution of outcomes conditional on observed state variables

Lastly, we estimate the retirement probability, $\lambda(\mathbf{s})$, and the probability that the incumbent wins, $P_{win}(\tilde{\Omega})$.

Retirement probability, $\lambda(\mathbf{s})$ We estimate the probability that the incumbent retires as a function of $\mathbf{s}$. As we explained in the main text, we specify $\lambda(\mathbf{s})$ to be only a function of ten_I, because none of the other state variables, such as w_I and q_I, has a statistically significant effect on retirement.

Probability of winning, $P_{win}(\tilde{\Omega})$, in contested periods We also estimate the probability that the incumbent wins in contested periods given $\tilde{\Omega}$, denoted as $P_{win}(\tilde{\Omega})$. $P_{win}(\tilde{\Omega})$ is estimated as a Probit in w_I, w_I^2, ten_I, $D_I X$, q_I, and q_I^2.

8.3.4 Computation of the continuation value

Once we obtain estimates of the distribution of actions and outcomes conditional on observed states, it is possible to simulate the continuation value for each profile of parameters. The idea is to generate sample paths of outcomes and actions using the estimated distributions, which can then be averaged to compute the continuation value. The key to our approach is that the incumbent's utility does not depend directly on q_C, which is unobservable, but only indirectly through actions and outcomes such as d_I, fr_I, etc. We compute the continuation value, $E_x[V_I(\mathbf{s}')]$, as follows:

1. Randomly draw X', given X using the estimated transition matrix, which gives us a new state vector, $\mathbf{s}' = \{q_I, w_I, ten_I, D_I, X'\}$. Note that we take q_I as given here. Draw a random variable U_{RET} from a uniform distribution. If U_{RET} is less than $\lambda(\mathbf{s}')$, then the incumbent retires and we terminate the process.

2. Draw a random variable U_{ENT} from a uniform distribution. If U_{ENT} is less than the probability of entry, i.e., $U_{ENT} \leq P_e(\mathbf{s}')$, then there is entry (Recall that P_e is estimated in Sec 3.1.2). If $U_{ENT} > P_e(\mathbf{s}')$, then there is no entry.

3. Depending on whether or not there is challenger entry in the previous step, draw d_I and fr_I using the conditional distributions ($F_{d_I}(\cdot|\tilde{\Omega})$ and $F_{fr_I}(\cdot|\tilde{\Omega})$) estimated above or compute them using the estimated policy function in un-ontested periods. In case of entry, further draw a random variable U_{WIN} from a uniform distribution.

4. The period utility function is computed as $u_I = B - \tilde{C}_I(fr_I, q_I) + H_I(d_I)$ in the case of no entry and as either $u_I = B - C_I(fr_I, q_I) + H_I(d_I)$ or $u_I = -C_I(fr_I, q_I) + H_I(d_I)$, depending on whether U_{WIN} is smaller or bigger than $P_{win}(\tilde{\Omega})$. A draw of U_{win} with a smaller value than $P_{win}(\tilde{\Omega})$ is interpreted as a victory for the incumbent, while a draw with a larger value is interpreted as a victory for the entrant.

5. Terminate the process if the incumbent loses to the entrant. Otherwise, draw w'_I from $F_{w'_I}(\cdot|\tilde{\Omega}, \{v_I > 0.5\})$. This determines the amount of savings.

6. The state variables become $\{q_I, w'_I, ten_I + 1, D_I, X'\}$. Go back to step 1 and repeat until termination. Take the discounted sum of u_I.

7. Repeat steps 1 through 6 and take the average.

Note that in computing the continuation value, we do not need to know the joint distribution of the actions, d_I, w'_I, and fr_I. Knowledge of the marginal distributions alone is enough to compute the continuation value. This follows from the additive separability of u_I and it greatly simplifies the computation.

8.4 Data Construction

We created the sample we use for our estimation as follows: We first dropped all House elections in Louisiana.[43] We also dropped elections in Texas in years 1996 and 2006 which were deemed unconstitutional in the Supreme Court.[44] We also dropped special elections held outside of the regular election cycle, elections that occur right after special elections, instances in which two incumbents run against each other, and elections in which a scandal broke out.[45] Some samples were also dropped because of missing data.[46] We are left with a base sample of $4,177$ contested elections and 806 uncontested elections.

8.5 Details on the Estimation

We now discuss the details on the estimation that we omitted from the main text.

Estimation of $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ We estimate $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ by least squares. We specify $P_e(\mathbf{s})$ as a Probit in $\ln w_I$, $\ln ten_I$, $D_I \times X$ and B-spline bases of $z_I = \frac{fr_I}{(\ln d_I)^{1/2} d_I \ln fr_I}$. We take 7 knots, corresponding to $(1/8,...,7/8)$ quantiles of z_I. We specify $E[M|\mathbf{s}]$ as a linear function in $\ln w_I$, $\ln ten_I$, $D_I \times X$ and B-spline bases of z_I.

Estimation of the vote share function We approximate $g(P_e(\mathbf{s}),\ E[M|\mathbf{s}])$ with linear terms in $P_e(\mathbf{s})$ and $E[M|\mathbf{s}]$ and their interaction term, $P_e(\mathbf{s}) \times E[M|\mathbf{s}]$.

[43]Louisiana has a run-off election unlike any other U.S. state.

[44]Elections in Louisiana use a run-off system unlike any other State in the United States. The Congressional Elections that were affected by the Supreme Court ruling are TX03, TX05, TX06, TX07, TX08, TX09, TX18, TX22, TX24, TX25, TX26, TX29 and TX30 in 1996 and TX15, TX21, TX23, TX25, and TX28 in 2006.

[45]The list of elections that were dropped because of a scandal are CA17 (1990), MA04 (1990), MN06 (1992), NY15 (1992), NY15 (2000) FL24 (2006), PA10 (2006), VA05 (2006), VA07 (2006), WV01 (2006), WA04 (2008). These events were identified by going through the biography of candidates in the CQ press Congressional Collection.

[46]Some of the entries in the FEC data set is clearly incorrect. Some candidates are listed as having run in a different State, for example. Most of these missing data is easily identifiable because the vote shares do not add up to one or there are multiple candidates from the same party. Where the accuracy of the data is suspect, Open Secrets (http://www.opensecrets.org/) was used as a cross-check in order to correct the mistakes. The full list of changes that were made is available upon request.

We then project the residual of the vote share equation on a set of basis functions consisting of exogenous variables, i.e., a constant, $\ln ten_I$, $(\ln ten_I)^2$, $D_I \times \ln ten_I$, $D_I \times X$, $\ln w_I$, $(\ln w_I)^2$, $D_I \times \ln w_I$, $\ln ten_I \times \ln w_I$, and B-spline bases of z_I. We then minimize the squared sum (See Ai and Chen, 2003).

Estimate of the distribution of actions in contested periods conditional on $\tilde{\Omega}$. We use a (first-order) Hermite series approximation to estimate the distribution of outcomes and actions conditional on $\tilde{\Omega}$, by (nonparametric) maximum likelihood. Because ten_I is a discrete variable, we estimate separate distributions for $ten_I \in [1, 3]$, $ten_I \in [4, 7]$, and $ten_I \in [7, \infty]$. The distribution of incumbent's savings is estimated conditional on incumbent winning the election.

Estimate of the distribution of actions in uncontested periods conditional on $\tilde{\Omega}$. We estimate the amount of spending and saving in uncontested periods by least squares. The function is specified as a linear function of a constant, $\ln w_I$, $(\ln w_I)^2$, $D_I \times X$, $D_I \times X^2$, $\ln ten_I$, $(\ln ten_I)^2$, q_I, and q_I^2.

Estimate the distribution of outcomes conditional on $\tilde{\Omega}$. The probability that the incumbent wins, $P_{win}(\tilde{\Omega})$, is estimated as a Probit in $\ln w_I$, $(\ln w_I)^2$, ten_I, $D_I X$, q_I, and q_I^2.

Estimate of σ_ε^2, θ_I, γ and q_C. We use the incumbents' first-order condition to estimate σ_ε^2, θ_I and γ. In particular, for each value of σ_ε^2, θ_I and γ, we simulate the continuation value ($E_{\mathbf{s}'}[V_I]$) and evaluate the first-order condition (13). At the true parameter value, equation (13) should hold with equality. Our estimates of σ_ε^2, θ_I and γ are obtained by choosing the parameter values that make equation (13) come as close to holding with equality as possible. Once we obtain estimates of σ_ε^2, θ_I and γ, we can identify the right hand side of the second line of equation (13), which in turn, identifies K and q_C. In practice, when we estimate σ_ε^2, θ_I and γ, we augment equation (13) with the incumbents' first-order condition in uncontested periods as well as exclusion restrictions regarding ε in the vote share equation.[47]

[47] For each σ_ε^2, θ_I and γ, we can compute the implied challenger quality, q_C, using expressions (13). This, in turn, allows us to compute the implied value of the residual (ε) in the vote share

In evaluating the right hand side of expression (13), we need to compute the derivative of the value function with respect to w_I. To do so, we approximated the value function with polynomials of the state variables and used the coefficient on w_I.[48] For some of the samples, we encountered trouble computing the inverse of Φ (see equation 13), because the argument inside Φ^{-1} was not between $(0,1)$. Consequently, we replaced the value of $\Phi^{-1}(\cdot)$ with 10^{-5} or $1-10^{-5}$ depending on whether or not the argument is below 0 or above 1. We also excluded 5% of the sample when computing the criterion function.[49]

Estimate of θ_C We used the challengers' first-order condition to estimate θ_C, the cost function parameter of the challenger. We use the polynomial approximation of the incumbents' value function in order to approximate the continuation value of the challenger. This is possible because the challenger becomes the incumbent conditional on winning. Our estimate of θ_C minimizes the difference between the marginal benefit of spending and the marginal cost of fund-raising of the challengers.

8.6 Computation of Counterfactual Policy Experiment

Computing P_e, and F_{q_C} without Solving the Model of the Primary This section explains how we compute our counterfactual policy experiment. The computation is somewhat complicated by the fact that we do not fully estimate the model of the Primary, i.e., we do not estimate $\pi(q_C, \mathbf{q}_{C,-n})$, κ, and R. This means that we cannot simply solve the model of the Primary in computing our counterfactual.

In order to compute the equilibrium, we begin by fixing the probability of entry ($P_e^{cf} = \Pr(\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$) and the distribution of challenger quality ($F_{q_C}^{cf} =$

eqution, (2). Note that the residual (ε) in the vote share equation (2) is independent with respect to the state variables and the choices of the players. We use this condition here.

[48]The alternative approach is to use numerical differentiation, but we found the numerical derivative to be less stable, depending quite a lot on the step size. This may be because we are not allowing the distribution of actions to be sufficiently flexible in our estimation. We excluded 5% of the samples with very small values of w_I in deriving the value function approximation.

[49]We acknowledge that this is not ideal. However, note that even at the true parameter values of σ_ε^2, θ_I and γ, the argument of Φ^{-1} can fall outside $(0,1)$ when the other parameters (such as the distribution of outcomes and actions) are estimated with noise.

$F_{q_C}(\cdot|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$).[50] Note that once we fix P_e^{cf} and $F_{q_C}^{cf}$, we are able to compute the profile of strategies, $\boldsymbol{\sigma}(P_e^{cf}, F_{q_C}^{cf})$, which are generated given P_e^{cf} and $F_{q_C}^{cf}$. $\boldsymbol{\sigma}(P_e^{cf}, F_{q_C}^{cf})$ is just the profile of strategies for the incumbent and the challenger that determine how much to spend, how much to save, etc., such that the incumbent's strategy is a best response to the challenger's strategy and vice versa. Since P_e^{cf} and $F_{q_C}^{cf}$ are equilibrium objects, however, we need to find P_e^{cf} and $F_{q_C}^{cf}$ that are consistent with the strategy profile they generate. That is, if we let $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ be the entry probability and challenger distribution generated by $\boldsymbol{\sigma}$, we have to find the fixed point of $(P_e^{cf}(\boldsymbol{\sigma}(\cdot)), F_{q_C}^{cf}(\boldsymbol{\sigma}(\cdot)))$.[51]

The problem with not estimating the model of the Primary is that we cannot solve the model of the Primary to find $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ – if we knew $\pi(q_C, \mathbf{q}_{C,-n})$, κ, and R, we would be able to compute $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ for each $\boldsymbol{\sigma}$ by solving the model. Given that we do not estimate the model of the Primary, however, we need another way to find $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ for each $\boldsymbol{\sigma}$. Our strategy for finding $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ relies on identifying $P_e(\mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$ in the actual data that correspond to $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$. As long as we can show that $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ correspond to some $P_e(\mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$, we can find $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ without solving the model of the Primary.

In order to identify $P_e(\mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s})$ in the actual data that correspond to $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$, it is useful to note that $P_e^{cf}(\boldsymbol{\sigma})$ and $F_{q_C}^{cf}(\boldsymbol{\sigma})$ both depend only on $r(\cdot)$ and $\bar{q}_C(\cdot)$, where $r(\mathbf{s})$ is the parameter of F_N and $\bar{q}_C(\mathbf{s})$ is the challenger type that is indifferent between entering and not entering. This follows from the sufficient statistics property that we showed earlier: $\bar{q}_C(\mathbf{s})$ and $r(\mathbf{s})$ are sufficient statistics for F_{q_C} and $P_e(\mathbf{s})$. Moreover, given that $r(\mathbf{s})$ is known, identifying the corresponding $P_e(\mathbf{s})$ and $F_{q_C}(t|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1\}, \mathbf{s}$ boils down to finding the appropriate $\bar{q}_C(\mathbf{s})$.

We begin by fixing some $\mathbf{s} = \mathbf{s}_0$. Suppose that $r(\mathbf{s}_0) = \overline{r}$. Consider varying $\mathbf{s}$ while maintaining $r(\mathbf{s}) = \overline{r}$ under the actual data-generating process. For each $\mathbf{s}$, we can identify $\bar{q}_C(\mathbf{s})$ and $v_C^{(2a)}(\mathbf{s}, \bar{q}_C(\mathbf{s}))$ for the actual data-generating process. By

[50]The superscript cf stands for counterfactual.

[51]Whether or not to enter is a choice variable for the challenger: Hence, P_e and F_{q_C} are endogenous objects. This means that we need to make sure that P_e^{cf}, $F_{q_C}^{cf}$ that we fixed at the outset is indeed consistent with the profile of incumbent and challenger strategies that they generate.

varying s, while fixing $r(\mathrm{s}) = \bar{r}$, we can map out the relationship between $\bar{q}_C(\mathrm{s})$ and $v_C^{2(a)}(\mathbf{s},\bar{q}_C(\mathbf{s}))$, which we denote as $c(\cdot)$ (See Figure 3).

Next, let us denote the challenger's value function (for fixed P_e^{cf} and $F_{q_C}^{cf}$) as $v_C^{2(a)}(\mathbf{s}_0,q_C;P_e^{cf},F_{q_C}^{cf})$. Note that $v_C^{2(a)}(\mathbf{s}_0,q_C;P_e^{cf},F_{q_C}^{cf})$ is just the value function of the challenger when the incumbent and the challenger play strategy $\boldsymbol{\sigma}(P_e^{cf},F_{q_C}^{cf})$. Now, define q^* as the intersection between $c(\cdot)$ and $v_C^{2(a)}(\mathbf{s}_0,\cdot;P_e^{cf},F_{q_C}^{cf})$. If we let s^* be the state such that $\bar{q}_C(\mathrm{s}^*) = q^*$, then the distribution of challenger quality and entry probability at $\mathrm{s} = \mathrm{s}_0$ in the counterfactual are the same as the observed quality distribution and entry probability at $\mathrm{s} = \mathrm{s}^*$ in the data. That is, $P_e^{cf}(\boldsymbol{\sigma}(\cdot|P_e^{cf},F_{q_C}^{cf})|\mathbf{s_0}) = P_e(\mathrm{s}^*)$ and $F_{q_C}^{cf}(\boldsymbol{\sigma}(\cdot|P_e^{cf},F_{q_C}^{cf})|\mathbf{s_0}) = F_{q_C}(\cdot|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1, \mathrm{s}^*\})$.

To see why this holds, it suffices to show that q^* satisfies the entry condition with equality, i.e., $p(q^*,\mathrm{s}_0)v_C^{2(a)}(\mathrm{s}_0,q^*;P_e^{cf},F_{q_C}^{cf}) - \kappa = R$, where $p(q^*,\mathrm{s}_0)$ is generated by $\bar{r}$ and $\bar{q}_C = q^*$.[52] Note first, that $p(q^*,\mathrm{s}^*)v_C^{2(a)}(\mathrm{s}^*,q^*) - \kappa = R$, where $v_C^{2(a)}$ is the value of the challenger in the actual data-generating process. This is by virtue of the fact that q^* lies on $c(\cdot)$. Moreover, we have $v_C^{2(a)}(\mathrm{s}^*,q^*) = v_C^{2(a)}(\mathrm{s}_0,q^*;P_e^{cf},F_{q_C}^{cf})$ because q^* is the intersection between $c(\cdot)$ and $v_C^{2(a)}(\mathrm{s}_0,\cdot;P_e^{cf},F_{q_C}^{cf})$. Finally, we have $p(q^*,\mathrm{s}^*) = p(q^*,\mathrm{s}_0)$, because they are both generated by the same sufficient statistics, $\bar{r}$ and $\bar{q}_C = q^*$. This implies that $p(q^*,\mathrm{s}_0)v_C^{2(a)}(\mathrm{s}_0,q^*;P_e^{cf},F_{q_C}^{cf}) - \kappa = R$, which means that $P_e^{cf}(\boldsymbol{\sigma}(\cdot|P_e^{cf},F_{q_C}^{cf})|\mathbf{s_0}) = P_e(\mathrm{s}^*)$ and $F_{q_C}^{cf}(\boldsymbol{\sigma}(\cdot|P_e^{cf},F_{q_C}^{cf})|\mathbf{s_0}) = F_{q_C}(\cdot|\{\boldsymbol{\chi}' * \mathbf{1} \geq 1, \mathrm{s}^*\})$.

Algorithm for Computing the Counterfactual We briefly summarize our algorithm for computing our counterfactual.

1. Start with an initial pair P_e^{cf} and $F_{q_C}^{cf}$.

2. Start with an initial profile of strategies $\boldsymbol{\sigma}^0 = (\sigma_I^0, \sigma_C^0)$ for incumbents and challengers.

3. Given $\boldsymbol{\sigma}^k = (\sigma_I^k, \sigma_C^k)$ ($k \in \{0, ..\}$), compute $\sigma_I^{k+1} = BR(\sigma_C^k)$ and $\sigma_C^{k+1} = BR(\sigma_I^k)$, the best response for incumbents and challengers given $\boldsymbol{\sigma}_k$.

4. Repeat step 3 until $\left\|\boldsymbol{\sigma}^{k+1} - \boldsymbol{\sigma}^k\right\| < \epsilon_1$.

[52] Recall that $\bar{r}$ and $\bar{q}_C$ are sufficient statistics for $p(q^*,\mathrm{s}_0)$.

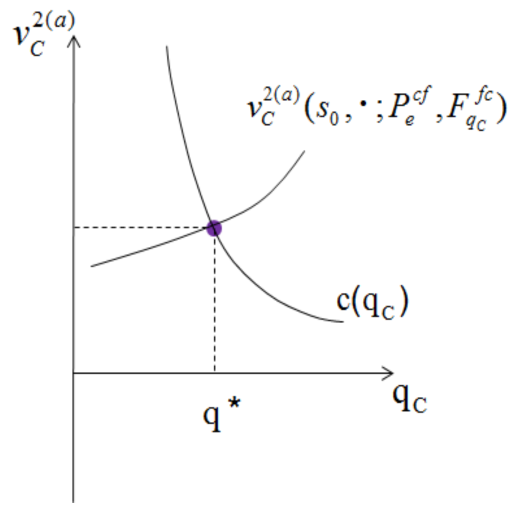

Figure 3: Finding P_e^{cf} and $F_{q_C}^{cf}$ that is Generated under a Profile of Strategies σ. Because we do not estimate the model of the Primary, we need a way to find P_e^{cf} and $F_{q_C}^{cf}$ without solving the model of the Primary. Given that the counferfactual entry probability and challenger distribution correspond to P_e and F_{q_C} in the actual data, the problem boils down to figuring out the corresponding P_e and F_{q_C}.

5. Find $P_e^{cf}(\boldsymbol{\sigma}^{k+1}(\cdot|P_e^{cf}, F_{q_C}^{cf}))$ and $F_{q_C}^{cf}(\boldsymbol{\sigma}^{k+1}(\cdot|P_e^{cf}, F_{q_C}^{cf}))$.
6. Replace initial P_e^{cf} and $F_{q_C}^{cf}$ with new ones and repeat steps 1-5 until $\| (P_e^{cf}, F_{q_C}^{cf}) - (P_e^{cf}(\boldsymbol{\sigma}^{k+1}), F_{q_C}^{cf}(\boldsymbol{\sigma}^{k+1})) \| < \epsilon_2$.

In practice, when we compute the distance between the strategies in step 3, we consider only the strategy for spending $(d_I(\mathbf{s}, q_C), d_C(\mathbf{s}, q_C))$ in contested periods. This is because spending is the only strategic variable that affects the opponent. If the distance between the new spending strategy and the old spending strategy is zero ($\left\|d_I^{k+1} - d_I^k\right\| + \left\|d_C^{k+1} - d_C^k\right\| = 0$), then we know that $\left\|\boldsymbol{\sigma}^{k+1} - \boldsymbol{\sigma}^k\right\| = 0$.[53]

The parameters of the cost function, C_C, are identified from the second equation because all the terms on the right-hand side are already identified. Once C_C is identified, H_C is also identified from the first equation.

[53]We discretize the spending amount into 100 bins, quality into 20 bins, tenure into 9 bins, war chest into 10 bins, and X into 2 bins.