Campaign Spending in Senate Elections: A Structural Analysis of a Two-Stage Election with Endogenous Candidate Entry*

Draft - please do not circulate or cite without permission of the author

Jun Zhao†

November 6, 2020

Abstract

This paper quantifies the effect of campaign spending on electoral outcomes in the U.S. Senate elections. To this end, I develop a two-stage contest model where in the first stage potential candidates decide whether to enter, and in the second stage those that decide to enter participate in an election consisting of within-party primaries and a general election where they make campaign spending decisions to influence the probability of winning. I also specify voters' decisions through a latent utility model. I estimate this game-theoretic model using data on the vote shares of Senate candidates in primary and general elections, data on candidates' campaign spending, and state-year level covariates for Senate elections between 1994 and 2018. Taking account of endogenous entry, I find that incumbents spend more and run more relative to potential challengers, and that campaign spending is more effective in increasing votes in the primary compared to the general election. Moreover, I show that the primary election in the incumbent's party is important in the sense that it can increase the electoral competition. I also show in a counterfactual analysis that moving to a top-two primary results in a slightly less competitive election and lower spending relative to the commonly used primary system.

Keywords: Campaign Spending, Empirical Contest, Endogenous Entry, the U.S. Senate Election

JEL Codes: C57, D72, L10

---

*I am deeply indebted to Tong Li for his guidance and support. I am also grateful to Joshua Clinton, Atsushi Inoue, Mattias Polborn, and Pedro Sant'Anna for their invaluable comments and feedback.

†Department of Economics, Vanderbilt University. Email: jun.zhao@vanderbilt.edu.

1 Introduction

The effect of money on electoral outcomes has attracted much attention across the U.S. society, among policy makers and journalists to the voters at large. In 2018, the spending of elections for Congress topped $5.7 billion, according to *CNN politics*, creating the most expensive midterm as the political battle for the control over the House and Senate. In line with a poll of representative U.S. adults conducted by the *Pew Research Center* in 2018, 77% supported limits on political campaign spending.[1] Given these concerns, it becomes increasingly critical to understand to what extent campaign spending can facilitate or impede electoral competition, especially in a representative democracy where competitive elections are of fundamental importance.

On the other hand, politicians are strategic when deciding to seek election (Copeland (1989)). This choice is based on a decision calculus, which incorporates the value of office, the probability of winning, the risks and/or costs involved in seeking it, and other political and/or policy factors (see, e.g., Jacobson and Kernell (1981), Abramson, Aldrich, and Rohde (1987), and Meserve, Pemstein, and Bernhard (2009)). As a result, ignoring politicians' decisions about whether to run for election can lead to biased estimates of the influence of campaign spending. As noted by Diermeier, Keane, and Merlo (2005), the endogenous entry of politicians is important because the selection bias induced by politicians' decisions whether to run for office is not negligible when studying the returns to congressional careers. However, the existing literature rarely takes into account or analyzes formally the role of politicians' endogenous entry when analyzing the effect of campaign spending on elections.

This paper fills the gap, through developing and estimating a two-stage game-theoretic model to characterize the strategic interactions among candidates and quantify the impact of campaign spending on the U.S. Senate elections, where candidates' participation and spending behavior is modeled as an equilibrium result from a contest with incomplete information and endogenous entry. Therefore, I explicitly model the entry stage by assuming that candidates strategically decide to participate in the race (whom I call "potential candidates"). This is the first and the most important contribution of this paper, which is to provide new evidence on how the strategic interactions among political candidates affect the composition and the level of competitiveness of elections, conditioning on candidates choosing their entry decisions endogenously.

I also model voters' decisions using a latent utility model that combines the within-party primaries

1 The Pew Research Center reported Americans said new laws would be effective in reducing role of money in politics.

and general elections together and that depends on candidates' campaign spending, demographic and economic covariates, and importantly, the election-level unobserved heterogeneity. The voter model describes how campaign spending translates into votes, thus determining the winning probabilities of candidates. The campaign spending, as an equilibrium outcome from the two-stage contest model, can be potentially endogenous in the voter model. Hence, the failure to account for the unobserved heterogeneity in the estimation of the voter model can lead to biased estimates.

Using data on the U.S. Senate elections from 1994 to 2018, I find that incumbents are more likely to run, and upon entry they will make strictly positive spending and spend more in campaigns, compared to challengers. Further, more campaign spending translates to larger latent utility for voters, then leads to more votes for candidates, especially in the primary elections. The findings indicate that if an election is more competitive with more candidates choosing to enter, these candidates that enter will make less campaign spending since their winning probabilities are smaller relative to a contest with fewer candidates. When deciding whether to run, more potential candidates are associated with a larger entry cost on average and more candidates choosing to run in the following election.

The second contribution of this paper is to demonstrate the important role of primaries in the U.S. Senate elections. The structural model allows me to characterize the role of candidate entry by accounting for candidates' strategic interactions in both participation and election stages. The results show that the effect of the number of potential candidates on candidate behavior and electoral outcomes depends on which party's number being changed. Increasing numbers of potential candidates in *both* the incumbent and opposition parties has a negative effect on campaign spending, the vote share of incumbents in the primary, and the probability of entry, regardless of the asymmetry in the size of the primary between two parties. However, its effect on the winning chance of incumbents in the general or the probability of having an open-seat election is not monotonic. If letting the incumbent be unopposed in the primary, increasing the number of potential candidates in the opposition party does not have notable effect on the behavior of incumbents. In contrast, increasing the number of potential candidates in the incumbent party can reduce the relative strength of incumbents, in terms of less campaign spending, smaller winning probability in both primary and general elections, and higher propensity of having an open-seat election. This implies that the primary election of the incumbent's party can be important in terms of improving electoral competition.

With the structural model, I also evaluate an alternative election institution, such as a top-two primary system being used in California and Washington.[2] As the third contribution of this paper, I emphasize that incorporating the endogenous entry significantly alters the assessment of the counterfactual format. If one ignores the participation behavior of potential candidates and treats entry as fixed, the simulated campaign spending is larger and the winning probability of incumbents is smaller in a top-two primary. However, once I account for strategic entry, I find that the alternative top-two primary mechanism creates a slightly less competitive election with the entry proportion dropping by 1% compared to the within-party primary system. Moreover, although incumbents will spend relatively more, challengers will spend much less, and total spending will fall in equilibrium by 10 percent point. In consequence, although the top-two primary system potentially reduces the spending by candidates, it impedes the competition level of elections to a certain degree. The failure to take into account the strategic entry decisions of potential candidates can produce misleading implications.

Lastly, the fourth contribution of this paper lies in the structural approach, which controls for the election-specific unobserved heterogeneity when estimating the entry and election models. The endogeneity of campaign spending (or advertising as a main part of the spending) has drawn much attention, when studying the effect of campaign spending on electoral outcomes, see Green and Krasno (1988), Erikson and Palfrey (1998), Gerber (1998), Erikson and Palfrey (2000), Lau and Pomper (2002), Gordon and Hartmann (2016), and Sovey and Green (2011) for a survey on the instrumental variables estimation in political science, among others. In this paper, the endogeneity of candidates' campaign spending in the voter model is accounted for via the election-specific unobserved heterogeneity. I introduce additional variables and possible unobserved ones that can explain electoral heterogeneity and candidates' participation/spending behavior, using identification strategies such as Matzkin (2003) and Guerre, Perrigne, and Vuong (2009). Specifically, I use *lagged campaign spending* by Senate incumbents and challengers, constructed by summing the spending by candidates in the *previous* Senate general election in the state.

This paper is related to several branches of the literature on empirical games and political economy. First of all, the model of candidates' endogenous entry builds on an extensive literature on the structural analysis of auction models with entry, see Li and Zheng (2009), Athey, Levin, and Seira (2011), Krasnokutskaya and Seim (2011), Roberts and Sweeting (2013), and Gentry and Li (2014),

2 In the top two primary system, all candidates who participate in the primary election are listed on the same ballot regardless of their party affiliations, and the top two winners of the primary proceed to the general election and compete for office.

among others. However, political competition contest models are fairly under-developed. Theoretically papers by Fu, Jiao, and Lu (2015) and Gu, Hehenkamp, and Leininger (2019) model players' entry in contest models with complete information. On the empirical side, this paper is the first one to construct a two-stage contest model with the first stage of entry and use it to analyze the campaign spending, to the best of my knowledge. Prior studies either treat the entry decision as an auxiliary component (Kawai and Sunada (2015)), or use the reduced-form analysis to study political entry (Avis, Ferraz, Finan, and Varjão (2017)).

This study also relates to a large literature on the effects of campaign spending on electoral outcomes (e.g., Jacobson (1978), Green and Krasno (1988), Levitt (1994), Gerber (1998), Erikson and Palfrey (2000), and Sovey and Green (2011)). Although the aforementioned findings are mixed, most work on estimating the effect of campaign spending on elections relies on reduced-form analyses. This paper improves our understanding of how campaign spending affects electoral competition by using a structural estimation to account for the impact of endogenous candidate entry in a two-stage election.

More broadly, this paper builds on an extensive literature studying contests and all-pay auctions in the context of political lobbying, campaigning, and advertising, including theoretical comparison of the lottery and all-pay auction models of lobbying (Fang (2002)) and empirical analyses regarding campaign spending in the U.S. House of Representatives election (He and Huang (2018)), firms' lobbying expenditures and policy enactment (Kang (2016)), as well as advertising decisions in the U.S. presidential elections (Gordon and Hartmann (2016)).[3] Relative to the previous literature, this paper fits in the framework of incomplete information contest model and incorporates both the endogenous entry and primary election, in order to obtain a more complete picture of the political contest.

This paper also adds to the research using structural models to study political phenomena. For instance, Diermeier, Eraslan, and Merlo (2003) examine the impacts of institutional features on the government formation process and the stability of the government in West European parliamentary democracies. Deltas, Herrera, and Polborn (2016) analyze the trade-off between learning and co-ordination among voters, through estimating a sequential election model in the U.S. presidential primary system. There is a growing literature studying a class of dynamic games to evaluate electoral competition, e.g., Sieg and Yoon (2017) who investigate term limits in the U.S. gubernatorial elections and Acharya, Grillo, Sugaya, and Turkel (2019) who estimate the popularity process of politicians.

3 For nonparametric identification of Bayesian games, see Li, Zhang, and Zhao (2020).

This paper assumes away the dynamic evolution of campaign financing and builds a static model. By doing so, I can focus my attention on campaign spending in the presence of candidates' endogenous entry and both primary and general elections, which are relatively under-studied components but of great importance.

The rest of the paper is organized as follows. In Section 2, I introduce the background and the data, along with the main variables I use for the later empirical analysis. Section 3 characterizes the two-stage game-theoretic model and establishes existence of the model equilibrium. Section 4 is devoted to the structural analysis of the data. Section 5 provides a set of counterfactual analyses. Section 6 concludes. Mathematical proofs, the detailed estimation approach, and various sensitivity analyses are gathered in the Appendix.

2 Background and Data

2.1 Data

The U.S. Senate elections are held for one class of the Senate every two years, when approximately one third of the 100 Senators face election or reelection.[4] A Senator is elected based on the plurality rule in a general election that includes candidates from all qualified parties. Most states also hold primary elections to decide which candidates can be on the ballot in the general election for each party. In these states, winning office requires winning both the primary and general elections.[5] Candidates' campaign spending plays an important role in this election process.[6]

To determine the effects of campaign spending and candidate entry on elections, two main sources of data are matched to each other. First, I collect the electoral results of Senate elections in the U.S. states between 1994 and 2018, documented by the Federal Election Commission (FEC).[7] This data contains information on candidates' vote shares in the primary and general elections, the state and the year, as well as the incumbent status of the candidates. If there is no incumbent in a Senate election, it is an open-seat election. I call individuals who decide to run for the Senate election the set of *actual candidates*. Second, I obtain the data on campaign spending by candidates for the same

4 Article I, section 3 of the Constitution requires that the Senate to be divided into three classes for purposes of elections.

5 I assume that candidates are office-motivated (see Adams and Merrill (2008)).

6 Here I refer to the aggregate campaign spending by candidates, thus do not differentiate between the amounts spent in the primary and general elections. This is due to the data limitation that I only observe a total value of campaign spending.

7 The FEC website: www.fec.gov.

period, from a different source collected by FEC.[8] This data provides candidates' campaign finance information including money raised, money spent, cash on hand and debt, together with the election years, incumbent status, and party affiliations for every individual who raises more than $5,000 for a race.[9] Note that not every individual who appears in this data actually runs for office, because it is required that every ongoing committee and new campaigns that raise or spend $5,000 or more must file quarterly reports and this applies even if the candidate plans to retire, withdraws from the race prior to the primary election, or drops out of the race prior to the general election. I call individuals who intend to run for the election and raise at least $5,000 the set of *potential candidates*. As a result, the second spending data nests the first vote data, since potential candidates may decide not to enter into the election.

I match the vote data with the spending data by the candidates' full names, the state and the year, as well as the party affiliation, in order to generate the vote-spending data for the structural analysis. Since this study mainly focuses on the Senate election in which the Democratic and the Republican parties have within-party primaries to decide the general election's nominations, I exclude the states where a top-two primary system is adopted including California, Louisiana, and Washington.[10] When a state has both full-term and special elections in the same year, the corresponding spending data only has a combined group of candidates who may attend either one of the elections, thus I cannot differentiate between the potential groups of candidates for the full-term and special elections. In this case, I delete both elections from the data. Some states rely on party conventions to nominate the candidates for the general election (e.g., Connecticut), which are also removed from the data.

It is also possible that some actual candidates lack the financial data, I replace the missing spending data by the same candidates' spending data from two years ago or two years later. This is valid because in the spending data it appears that the consecutive election years of each candidate have at least a four-year gap. This implies that in general one candidate will not prepare for both Senator offices in one state, which in turn indicates that the financial data two years before or later can serve the election

8 For details, see https://www.fec.gov/data/browse-data/?tab=candidates.

9 FEC defines the spending (or expenditure) as a purchase, payment, distribution, loan, advance, deposit or gift of money or anything of value to influence a federal election. However in the finance data, I only observe the disbursement reported by campaigns, which is a broader term that covers both the spending and other kinds of payments (those not made to influence a federal election). Since I lack the data on the spending itself, I use the disbursement as a proxy.

10 Further, I exclude the elections with run-offs after the primary or the general elections (e.g., Alabama in some years), the elections with electoral fusion where two or more parties on a ballot list the same candidate and such candidate's votes are pooled together (e.g., New York), and the elections with candidates who drop out from the race prior to the general election that induces a replacement of one candidate in the general election (e.g., Vermont in some years).

in this year, given that this candidate actually runs for this election. Lastly, since I concentrate on the two major parties, I remove candidates from third parties (e.g., the Libertarian Party, the Green Party, the Constitution Party) and write-in candidates.

In the end, this data set consists of 2283 potential candidates, 1679 actual candidates, and 306 elections. I complement the main vote-spending data with a collection of state-year level covariates drawn from a variety of sources. I include: (1) local political preferences; (2) demographic and economic variables; (3) variables that affect voters' turnout decisions.

First, the Cook Partisan Voting Index (PVI) is a measure of how strongly a state leans toward the Democratic or Republican party relative to the country as a whole. PVIs are calculated by comparing a state's average Democratic or Republican Party share of the two-party presidential vote in the past two presidential elections to the national average share for these elections. I also include an indicator for whether the incumbent governor's party is the same as the President's.

Second, I add a set of demographic and economic variables for each election year. I use the state-level percentage of the population in three age bins (i.e., 25-44, 45-64, 65 and up) from the U.S. Census Bureau, the state-level unemployment rate from the U.S. Bureau of Labor Statistics, and the state-level median household income from the U.S. Census Bureau.

Third, for the sensitivity analysis where voters' turnout is taken into account, I include variables that should solely affect voters' decisions to turnout in the general election. I use the data on the state-level voting eligible population (VEP) to calculate the voter turnout, from the U.S. Elections Project conducted by Michael P. McDonald at University of Florida. I also include the state-level estimates of rain and snowfall on the general election dates, by taking averages of the precipitation and snowfall from all stations within one state on the general election dates from the National Centers for Environmental Information. Furthermore, I use a dummy variable to indicate if a presidential election occurs in the same year.

Lastly, I correct for the endogeneity of campaign spending when estimating the voter model by controlling for the election-specific unobserved heterogeneity, because the failure to account for the unobserved heterogeneity will lead to biased estimates.[11] In order to identify the unobserved heterogeneity following the identification strategies in Matzkin (2003) and Guerre, Perrigne, and

11 One can think of some reasons that spending levels are influenced by the unobserved electoral conditions including: the corruption level of the local government might influence the underlying fundraising environment, which in turn affects candidates' spending levels; or if the state voters hold expectations of the victory, these expectations will drive spending decisions, see Erikson and Palfrey (2000).

Vuong (2009), I introduce additional variable or instrumental variable (IV): *lagged campaign spending*, with the exact variable being the total spending by candidates in the *previous* Senate general election in the same state.[12] Due to the staggered nature of the Senate elections, Senators in one state are elected to six-year terms. Thus, the previous Senate election in one state for the other Senate seat rarely involves the same incumbent or challenger as the current race. Therefore, the *lagged campaign spending* affects the candidates' entry and spending behavior only through the correlation between this lagged variable and the latent fundraising environment or the underlying political climate in one state.

Table 1 presents the summary statistics of the data. Around 69% candidates have strictly positive campaign spending.[13] The average amount of campaign spending is about $4 million, which includes all disbursements by a representative candidate during the election. The entry proportion, calculated as the ratio of the number of actual candidates to the number of potential candidates for each election, is about 0.78 on average, meaning that on average more than three quarters of the potential candidates will decide to run for the elections. Among all 306 elections in the data, only 18% are open-seat elections, implying that incumbents are more likely to participate in the elections. The variable PVI is assigned positive (negative) value if the state under investigation leans toward the Republican (Democratic) party. The lagged spending variable has a larger sample mean than the spending variable, which makes sense because the lagged spending is defined as the sum of campaign spending by the general election candidates in the same state's previous Senate election, while the spending is defined for all candidates.

2.2 Stylized Facts

The campaign spending of actual candidates and the entry decisions of potential candidates are of primary interest of this paper. Therefore I provide some stylized facts to indicate the patterns regarding the behavior of candidates.

Figure 1 presents the (unconditional) spending distributions of incumbents and challengers in incumbent-challenger elections and that of candidates in open-seat elections. In incumbent-challenger elections, the spending density of incumbents is shifted rightward from that of challengers, indicating a higher level of campaign spending by incumbents relative to challengers on average. In open-seat

12 This is similar to the variable used in Gerber (1998) being the *lagged campaign spending* by Senate incumbents and challengers in the *previous* election of a state. However, this is different from the variable used by Green and Krasno (1988), which is a candidate's own lagged spending.

13 Candidates with zero campaign spending still receive strictly positive (but small) vote shares. When modeling candidates' spending behavior in Section 3, I assume that candidates can choose not to spend if they find it unprofitable, even though they still have a strictly positive winning chance.

Table 1: Summary statistics of candidate- and election-specific variables

	Observations	Mean	SD
Non-zero spending	1679	0.6945	0.461
Spending	1166	4.2521	7.261
No. of potential candidates	306	7.4608	4.766
No. of actual candidates	306	5.4869	3.659
Entry proportion	306	0.7785	0.211
Open-seat election	306	0.1830	0.387
PVI	306	2.0163	7.913
Governor partisanship	306	0.4379	0.497
% Aged 25-44	306	0.3071	0.030
% Aged 45-64	306	0.2782	0.061
% Aged 65 and up	306	0.1563	0.039
Unemployment rate	306	5.2964	1.870
Median household income	306	0.0589	0.009
VEP	306	3.6125	3.224
Rain (in.)	306	0.1147	0.198
Snowfall (in.)	306	0.0562	0.248
Same day as presidential election	306	0.4706	0.500
Lagged spending	306	11.8170	11.246

Notes: PVI stands for the state-year-level Cook Partisan Voting index with negative (positive) value meaning Democratic (Republican) leaning. VEP is the state-year-level voting eligible population. The means and SDs of the variables spending, median household income, VEP, and lagged spending are scaled down by 10^6.

elections, the spending density of candidates is slightly shifted rightward from that of challengers and shifted leftward from that of incumbents in incumbent-challenger elections. This implies that candidates in open-seat elections tend to make larger campaign spending than challengers in incumbent-challenger elections on average. However, unlike incumbents, candidates in open-seat elections and challengers in incumbent-challenger elections have positive probabilities of making zero spending.

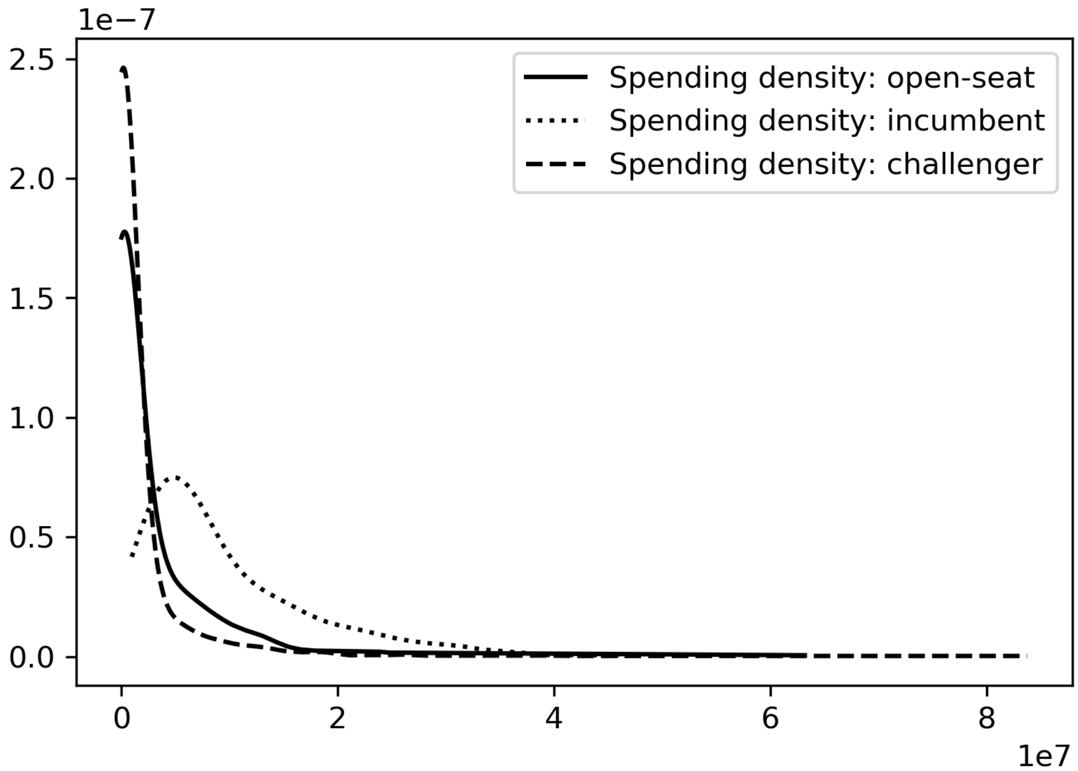


Figure 1: Spending distributions of incumbents and challengers in incumbent-challenger elections, as well as candidates in open-seat elections

Turning to the entry behavior of potential candidates, Table 2 shows the means and standard deviations of entry decisions for incumbents and challengers, where the entry decision is a dummy that equals 1 if a potential candidate decides to run for the race and 0 otherwise. Clearly, reelection probabilities in the Senate are indeed high, even unconditionally. However, there is still substantial selection among challengers in terms of who run for office.

Although these observations regarding the entry and spending behavior of candidates are important, they call for a further structural analysis if one wants to unfold the private information of actual and potential candidates, such as their private values of holding the office and entry costs of running for office, and examine the impacts of covariates. By estimating a structural model, I can not only verify

Table 2: Summary statistics of entry dummies for incumbents and challengers

	Observations	Mean	SD
Incumbent	306	0.8170	0.387
Challenger	1977	0.7228	0.448

Notes: The entry decision of a potential candidates is a dummy variable that equals 1 if this candidates decides to run for office and 0 otherwise.

the patterns, but implement counterfactual analyses as well.

3 The Model

I propose a two-stage game-theoretic model for the Senate election that incorporates both the entry stage of potential candidates and the election stage of actual candidates. Hence this model describes the strategic interactions among candidates. I also model decisions of voters in the sense that voters treat candidates like products in a differentiated goods market. Assume there are N Democratic potential candidates and M Republican potential candidates, corresponding to $1 \leq n \leq N$ Democratic actual candidates and $1 \leq m \leq M$ Republican actual candidates in a generic election.

Denote the election-level heterogeneity as (X, u), where X is the vector of observed covariates and u is the unobserved heterogeneity.[14] Throughout this section, the distributions of the entry cost and the private value are conditional on (X, u), but to economize on notations, I suppress the conditional set. In the entry stage, each potential candidate holds a private entry cost for office, denoted by c.[15] In the election stage, each actual candidate has a private value for office, denoted by v.[16] Suppose the distribution of the entry cost is $H(\cdot)$ over the support $[\underline{c}, \bar{c}] \subset \mathbb{R}^+$, from which potential candidates draw their entry costs independently at the beginning of the entry stage of one election. Suppose the distribution of the private value is $F(\cdot)$ over the support $[\underline{v}, \bar{v}] \subset \mathbb{R}^+$, from which actual candidates draw their private values independently prior to the election stage upon entry.[17] I assume that conditional on the election-level heterogeneity (X, u), each candidate's entry cost and private value are independent

14 The election I refer to here is defined as the state-year-level race including both the primary and general elections.

15 I define the entry cost for a potential candidate as a reservation value (or opportunity cost suggested by Lu (2009)), with which she compares the expected payoff of winning the election to decide whether to run for the race. Hall (2019) also states that the cost of running for office includes the amount of salary a candidate forgoes while running. This provides a simple way to conceptualize the endogenous entry decisions of potential candidates.

16 The private value is interpreted by Baron (1989) as the expected stream of benefits associated with winning office and any future election opportunities if successful, which include the monetary value of winning the office, the ability to implement preferred policies, and/or simply the "hunger" for office (see Gordon and Hartmann (2016)).

17 This implies that candidates pay the entry costs to learn their private values.

from each other.

3.1 Voters

I model the voters' behavior to derive the winning probability of an actual candidate, assuming that all voters vote sincerely. This winning probability is called the Contest Success Function (CSF) in the contest literature. Let D_i, $i = 1, \cdots, n$ denote the candidates in Democratic primary, and R_k, $k = 1, \cdots, m$ denote the candidates in Republican primary. A representative voter receives the latent utility, expressed as $M_D(D_i) + \iota_{D_i}$, from a Democratic candidate D_i in the Democratic primary, and the latent utility, expressed as $M_R(R_k) + \iota_{R_k}$, from a Republican candidate R_k in the Republican primary. Turning to the general election, an arbitrary voter receives the latent utility $M_G(D_i) + \iota_{D_i}$ from a Democratic candidate D_i given that she wins the Democratic primary, and the latent utility $M_G(R_k) + \iota_{R_k}$ from a Republican candidate R_k given that she wins the Republican primary. A voter will choose the most preferred candidate by ranking all candidates' latent utilities.

The unmeasured components, ι_{D_i} and ι_{R_k}, capture the candidates' ideological characteristics that determine the policy positions during the election, which the candidates do not know at the time they choose their campaign spending. I assume that a candidate reveals her ideological preference during the primary, and thus this unmeasured component is identical for the primary and the general election.

The ιs are assumed to be i.i.d. and follow type-1 extreme-value distribution.[18] Without loss of generality, I focus on a representative actual candidate in the Democratic primary D_i, where $i \in \{1, \cdots, n\}$, and use $P(D_i R_k)$ to represent the probability that this candidate D_i wins the Democratic primary and the general election with the candidate R_k as the general election opponent who wins the Republican primary, where $k \in \{1, \cdots, m\}$. This probability has a closed-form expression (see Adams and Merrill (2008)), which can be written as:

$$P(D_i R_k) = \frac{1}{\sum_{j \neq i}^{n} exp(-W_{D_j}) + 1 + exp(-W_G)\left[\sum_{l \neq k}^{m} exp(-W_{R_l}) + 1\right]} \cdot \frac{1}{\sum_{l \neq k}^{m} exp(-W_{R_l}) + 1}, \quad (3.1)$$

where I define the differences as

$$W_{D_j} \equiv M_D(D_i) - M_D(D_j),\ for\ j = 1, \cdots, i-1, i+1, \cdots, n;$$
$$W_{R_l} \equiv M_R(R_k) - M_R(R_l),\ for\ l = 1, \cdots, k-1, k+1, \cdots, m;$$

18 This assumption is widely used in both empirical studies of voting behavior (Whitten and Palmer (1996); Gordon and Hartmann (2016)) and theoretical models of elections (Schofield and Sened (2005); Schofield and Sened (2006); and Adams and Merrill (2008)).

$$W_G \equiv M_G(D_i) - M_G(R_k),$$

and the detailed derivation can be found in Appendix A1.[19]

The probability of a Democratic candidate D_i winning both the Democratic primary and the general election with the Republican candidate R_l as the opponent who wins the Republican primary, denoted as $P(D_iR_l)$, has similar expression, for $l \neq k$ and $l \in \{1, \cdots, m\}$. Therefore the probability of winning the final office for the Democratic candidate D_i is $P(D_i) \equiv \sum_{k=1}^{m} P(D_iR_k)$, given that all these events are disjoint. As manifested in the particular forms of $M_D(D_i)$ and $M_G(D_i)$ ($M_R(R_k)$ and $M_G(R_k)$), in Section 4 I specify these functions to depend on the campaign spending of the candidate e_{D_i} (e_{R_k}), coupled with the election-level heterogeneity (X, u). Further, suppose that each voter's latent utility from electing a candidate is increasing with diminishing returns in the candidate's campaign spending.[20] Finally, the probability of the candidate D_i winning the office is denoted as $CSF(e_{D_i}; e_{-D_i}) = P(D_i)$, meaning this is a CSF that depends on this candidate's campaign spending e_{D_i} and the spending profile of the rest actual candidates in the Democratic and Republican parties e_{-D_i}.

3.2 Candidates: Election Stage

Because the entry decision of a potential candidate is based on the pre-entry expected profit, I begin by characterizing the spending strategy in the election stage. I apply a contest model with incomplete information to the election stage. Without loss of generality, as in the previous subsection, I consider a representative actual candidate in the Democratic party D_i with the realization of the private value v_{D_i}, where $i \in \{1, \cdots, n\}$. The Bayesian Nash equilibrium (BNE) notion is adopted. This candidate chooses how much she wants to spend in order to alter the electoral result, which is the equilibrium strategy given her private value v_i. Recall that the CSF derived in the previous subsection is $CSF(e_{D_i}; e_{-D_i})$, which determines the probability of this Democratic candidate winning the office. I define a cost function $g(\cdot)$ that describes the effective cost of the campaign spending with $g(0) = 0$, $g'(\cdot) > 0$, and $g''(\cdot) \geq 0$. Thus her maximized expected payoff when she spends e_i can be written as:

$$\pi_{D_i}(v_{D_i}|a_{-D_i}) \equiv \max_{e_{D_i}} \quad v_{D_i} \cdot \mathbb{E}_{e_{-D_i}}[CSF(e_{D_i}; e_{-D_i})|v_{D_i}; a_{-D_i}] - g(e_{D_i}), \tag{3.2}$$

19 I also derive the probability where the turnout of voters in the general election is taken into account (see Appendix A1).

20 Thus campaign spending can "impress" the voters directly, see Grossman and Helpman (1996), Pastine and Pastine (2002), and Gordon and Hartmann (2016).

where, as before, e_{D_i} denotes the campaign spending of this Democratic candidate, the expectation is taken over the campaign spending profile of the rest of actual candidates in the Democratic and Republican parties e_{-D_i}, and

$$a_{-D_i} \in A_{-D_i} \equiv \{(\alpha_{D_j})_{j \neq i, j \in N}, (\alpha_{R_k})_{k \in M}\}$$

being one possibility of the composition of the entry behavior of $N-1$ potential Democratic candidates and M potential Republican candidates, where $\alpha_{D_j} = 1$ ($\alpha_{R_k} = 1$) if a potential candidate in the Democratic (Republican) party decides to run for office.[21]

I consider a continuously differentiable and pure-strategy BNE, $e_{D_i} = s_{D_i}(v_{D_i})$. This Democratic candidate has the following first-order condition (FOC) for the maximization problem:

$$v_{D_i} \cdot \mathbb{E}_{e_{-D_i}} \left[\frac{\partial CSF(s_{D_i}(v_{D_i}); e_{-D_i})}{\partial e_{D_i}} \middle| v_{D_i}; a_{-D_i} \right] \leq g'(s_{D_i}(v_{D_i})), \tag{3.3}$$

where the derivative of CSF is with respect to the first argument and both interior and corner solutions are allowed. When the inequality holds, the candidate will spend zero thus $e_{D_i} = 0$. One can think of the private value v_{D_i} as a structural parameter translating the probability of winning into monetary units, which places both sides of the condition (3.3) in equivalent terms.[22] Intuitively, the FOC balances the (monetary) value of increase in the winning probability relative to the marginal cost of campaign spending.

3.3 Candidates: Entry Stage

In the beginning of the entry stage, each potential candidate knows her own entry cost, the distribution of the entry cost $H(\cdot)$, and the distribution of the private value $F(\cdot)$. Without loss of generality, I again focus on the potential candidate from the Democratic party D_i with the entry cost c_{D_i}, where $i \in \{1, \cdots, N\}$. Define the pre-entry expected payoff of this candidate as Π_{D_i}, which is written as:

$$\Pi_{D_i} \equiv \sum_{a_{-D_i} \in A_{-D_i}} \mathbb{E}_{v_{D_i}} [\pi_{D_i}(v_{D_i} | a_{-D_i})] \cdot Pr(a_{-D_i} | \alpha_{D_i} = 1),$$

where I integrate out the private value v_{D_i} because this candidate only draws the realization upon entry. $Pr(a_{-D_i} | \alpha_{D_i} = 1)$ is the probability of the composition a_{-D_i} of entry behavior of the potential

21 In practice, it is rare that a general election only has one candidate who is unopposed. Thus when I estimate this game-theoretic model, I assume that there is at least one actual candidate in both parties' primary elections.

22 Alternatively, the private value v_{D_i} can be viewed as the candidate's financial strength, which is a policy-invariant parameter being independent of campaign fundraising (see Gordon and Hartmann (2016)).

candidates other than the candidate D_i, conditional on this candidate D_i entering into the election stage. Note that this probability depends on the entry probabilities of all potential candidates. Therefore, in equilibrium, the entry probability of this candidate, denoted by p_{D_i} is given by $p_{D_i} = Pr(C_{D_i} < \Pi_{D_i})$.

3.4 Equilibrium

The equilibrium of the model consists of two parts: entry equilibrium and spending equilibrium. The following proposition establishes existence of the model equilibrium.[23]

Proposition 1 *In the two-stage game-theoretic model with incomplete information and endogenous entry, there exists a pure-strategy Bayesian Nash equilibrium in continuous and strictly increasing strategy for the election stage. There also exists a Bayesian Nash equilibrium for the entry stage, where each potential candidate enters into the election following a threshold rule.*

Proof. See Appendix A2. ■

To prove the existence results in Proposition 1, I follow Wasser (2013) and Ewerhart (2014) who apply Athey (2001) for the election stage, by showing that the expected payoff of the actual candidate exhibits decreasing differences in the private value and the campaign spending, together with the continuity of the expected payoff. Furthermore, the existence of the entry equilibrium is equivalent to the existence of the entry probabilities, which is shown through applying Brouwer's fixed point theorem (see Li and Zhang (2015)).

4 Structural Analysis

I estimate the game-theoretic model proposed in the previous section using data on the vote shares of actual candidates and the campaign spending of potential candidates in the U.S. Senate elections during 1994 and 2018. The goals are to recover the entry cost distribution $H(\cdot)$ and the private value distribution $F(\cdot)$ in the candidate model, and to estimate the voter model. I adopt a fully parametric approach, and assume that I have an i.i.d. sample of elections, indexed by l for $l \in \{1, \cdots, L\}$.

4.1 Specifications

I use the type-2 tobit model to specify the log of campaign spending by actual candidates, taking into consideration that the campaign spending of some candidates can be zero. The campaign spending

23 The uniqueness of the model equilibrium is hard to prove. As in Krasnokutskaya and Seim (2011) and Gordon and Hartmann (2016), I verify the uniqueness of the equilibrium within the estimation routine by trying different initial values to see whether I obtain similar results.

is modeled as following:

$$I_{i_l,l} = X_l'\alpha + u_l + \varepsilon_{1,i_l,l}, \quad i_l = 1, \cdots, n_l,$$
$$\begin{cases} \log(e_{i_l,l}) = Z_{i_l,l}'\beta + u_l + \varepsilon_{2,i_l,l} \quad if\ I_{i_l,l} > 0, \\ e_{i_l,l} = 0 \quad if\ I_{i_l,l} \leq 0 \end{cases}$$

where I pool the candidates in the election l together regardless of their party affiliations, with the total number of actual candidates being n_l. u_l is the election-level unobserved heterogeneity. Suppose that $(\varepsilon_{1,i_l,l}, \varepsilon_{2,i_l,l})$ follows a bivariate normal distribution with zero mean, variances being σ_1^2 and σ_2^2, and σ_{12} as the covariance, where σ_1 is normalized to be 1. In the selection equation $I_{i_l,l} = X_l'\alpha + u_l + \varepsilon_{1,i_l,l}$, X_l contains the election-level covariates: a dummy variable to indicate whether the election is open-seat (OPEN), two variables to measure the local political preference including the Cook Partisan Voting index (PVI) and a dummy for wether the state governor is in the same party as the President (GOV), three variables to denote the population percentage falling into three age bins (25-44: YOUNG-PER; 45-64: MID-PER; 65 and up: OLD-PER), the state-year level unemployment rate (UNEMP), the log of state-year level median household income (LOGINC), the total number of potential candidates (PNCAN), and the total number of actual candidates (ANCAN). In the log(spending) equation $\log(e_{i_l,l}) = Z_{i_l,l}'\beta + u_l + \varepsilon_{2,i_l,l}$, other than the election-level covariates in the selection equation, I include an additional dummy equal to 1 for incumbents and 0 for others (INCUM) into $Z_{i_l,l}$.[24] Therefore, conditional on being strictly positive, the campaign spending of a candidate is specified as a log-normal distribution truncated from below at zero.[25]

The cost function of the campaign spending $g(\cdot)$ in the expected payoff expression (3.2) is assumed to be a linear function such that $g(e) = e$, as in He and Huang (2018). This leads to a cost derivative function being a constant 1 appearing in the FOC (3.3), which is used to back out the private value distribution of actual candidates after obtaining the estimates of the campaign spending distribution.[26]

The entry cost distribution of potential candidates is parameterized as a log-normal distribution

24 It turns out that in the data, all incumbents who run for office have non-zero campaign spending. Thus I do not include INCUM for the selection equation.

25 In the benchmark model, I consider the case where there is no upper boundary for the campaign spending. Later in the sensitivity analysis, I consider an extension of the benchmark model, where the campaign spending is supposed to be bounded from above by a finite number, which is nonparametrically estimated following Guerre, Perrigne, and Vuong (2000), see Appendix A5.2.

26 In Appendix A5.4, I alternatively specify a quadratic cost function as a sensitivity analysis. The results imply that a linear cost function can characterize the structural model more accurately, because the estimated quadratic term is essentially zero.

with mean $Z_{i_l,l}^{C}{}'\delta + u_l$ for $i_l = 1, \cdots, N_l$ with N_l potential candidates, and a constant group-specific variance depending on whether the potential candidate is an incumbent or not, denoted as λ_I^2 and λ_C^2, respectively.[27] $Z_{i_l,l}^{C}$ includes the same covariates as $Z_{i_l,l}$ but excluding OPEN and ANCAN, since at the entry stage, whether the election will be open-seat or not and how many candidates will finally decide to run have not been determined yet.

As for the voter model, I specify the latent utility for an arbitrary voter as:

$$u_{i_l,l}^{P} = \gamma \log(1 + e_{i_l,l}) + X_l^{P\prime}\gamma_X + u_l + \iota_{i_l,l}, \quad i_l = 1, \cdots, n_l,$$

for the primary election, and

$$u_{i_l,l}^{G} = \omega \log(1 + e_{i_l,l}) + X_l^{G\prime}\omega_X + u_l + \iota_{i_l,l}, \quad i_l = 1, \cdots, n_l,$$

given that this candidate i_l wins the primary election and proceeds to the general election. Note that adding the first three parts in the latent utility together yields $M_D(D_i)$ in Section 3.1, taking the Democratic candidate D_i in the primary election as one example. To reduce the burden on notations, here I pool the candidates from both parties together. In the voter model, X_l^P may be different from X_l^G. An econometric issue arises when analyzing the voter model: the coefficients γ_X and ω_X cannot be identified. This is because when a voter compares two candidates who attend the same election (either primary or general), the parts associated with election-level covariates are cancelled out and only the difference of the latent utility is crucial.[28] Therefore the terms $X_l^{P\prime}\gamma_X$, $X_l^{G\prime}\omega_X$, and u_l being the same for all candidates in an election are cancelled out, while I still control for the election-level heterogeneity (both observed and unobserved).

One way to solve this issue is to introduce voters' turnout decision in both primary and general

27 In the benchmark model, I do not distinguish between the two parties. However in the sensitivity analysis, I consider a case with the difference between Democratic and Republican parties (see Appendix A5.3).

28 Recall that in Equation (3.1), I model the probability $P(D_iR_k)$ through the differences W_{D_j}, W_{R_l}, and W_G.

elections into the voter model, as in Gordon and Hartmann (2016), in order to identify γ_X and ω_X.[29] However, this is not feasible in my case due to the data limitation that the turnout information is not available, at least for the primary election.[30] As a result, I can only identify γ and ω, which measure how campaign spending of candidates affects their probabilities of winning in the primary and general elections.

Last but not least, I need to control for the election-level unobserved heterogeneity u_l that affects the private value and the entry cost of candidates. This unobserved heterogeneity will in turn influence candidates' participation and spending behavior, and consequently the campaign spending will be endogenous in the voter model without controlling for the unobserved heterogeneity. I use *log of lagged spending* made by the candidates in the *previous* Senate general election in the same state as the IV.

Denote the private value of the i_l-th actual candidate as $v_{i_l,l}$ and the entry cost of the i_l-th potential candidate as $c_{i_l,l}$ in the election indexed by l. The IV being $log(lagged\ spending)_l$ satisfies the exclusion restrictions: (i) $v_{i_l,l} \perp\!\!\!\perp log(lagged\ spending)_l|(Z_{i_l,l}, u_l)$ for $i_l = 1,\cdots,n_l$; and (ii) $c_{i_l,l} \perp\!\!\!\perp log(lagged\ spending)_l|(Z^C_{i_l,l}, u_l)$ for $i_l = 1,\cdots,N_l$. Since u_l is unobserved, I follow Matzkin (2003) and Guerre, Perrigne, and Vuong (2009) to assume that $log(lagged\ spending)_l$ is modeled as $X_l^{IV'}\eta + u_l$, which is strictly increasing in u_l, and satisfies $X_l^{IV} \perp\!\!\!\perp u_l$. I additionally include two weather variables of rain and snowfall on the general election days (PRCP and SNOW) and a dummy variable indicating whether a presidential election occurs in the same year (PRE) in X_l^{IV}.[31]

29 The data on the voter turnout in the primary elections is limited, due to three practical reasons. First, as of 2019, some states do not ask voters to indicate their party preferences when registering, for example, Alabama, North Dakota, and Tennessee. Thus, it is impossible to approximate the by-party voter registration figures for these states. Second, different states adopt different types of primary elections. Some states (e.g., Arizona) use open primaries for the Senate elections, where a voter either does not have to formally affiliate with a political party to vote in the primary, or a voter previously affiliated with a different party can declare the party affiliation at the polls on the day of primary. Some states (e.g., Delaware) use closed primaries where voters have to formally affiliate with a political party before the election date to participate in the party's primary. Other states also use other kinds of primaries, such as semi-closed primaries and the non-partisan top-two primaries. Thus, due to the large variations of the primary election types, even with the total numbers of by-party voter registrations in the primaries, there is no clear rule about how to calculate the voting eligible population for the primary elections. Third, It is also invalid to use the information on eligible voters in general elections, since usually there are much more voters who turnout in the general elections compared to the primaries.

30 In the sensitivity analysis (see Appendix A5.1), I consider an extension of the benchmark model, where with the availability of the turnout information of the voters for each state's general election, I generate the probability of a candidate winning both the primary and general elections, in which the parameter ω_X can be identified and estimated.

31 As a reminder, a constant term is included in all specifications (CONST).

4.2 Estimation Method

In summary, I have a multi-stage structural model, which contains three parts: the voter model, the election model, and the entry model. To estimate the election model, I account for the left-censored campaign spending, which produces two parts: the selection part and the spending level part.

I adopt a multi-stage estimation method for the structural model, given the specifications.[32] In the first step, I regress the $log(lagged\ spending)_l$ on X_l^{IV} to estimate η via the ordinary linear squares regression (OLS), and obtain the residuals as estimates for u_l, the election-level unobserved heterogeneity. In the second step, together with the estimated unobserved heterogeneity for each election, I use the data on campaign spending of all candidates who actually run for the Senate elections to estimate the parameters in the spending model: α, β, σ_2, and σ_{12}, through the maximum likelihood estimation (MLE). In the third step, I use MLE to estimate the parameters in the voter model, i.e., γ and ω. In order to recover the private value distribution, I follow Jofre-Bonet and Pesendorfer (2003) and Athey, Levin, and Seira (2011) to estimate the value density after obtaining the pseudo private value from the FOC (3.3) and the simulated campaign spending. With the estimates of the spending and voter models, together with the estimated unobserved heterogeneity, I simulate the expected payoff for each potential candidate at the entry stage, and follow Li and Zhang (2015) to estimate the distribution of the entry cost, in order to get the estimates for δ and two λs. Therefore, the parameter vector being estimated is defined as $\theta = (\eta; \alpha, \beta, \sigma_2, \sigma_{12}; \gamma, \omega; \delta, \lambda_I, \lambda_C)$.

Only the subset of challengers is used to estimate the selection part of the campaign spending model, because incumbents surely spend non-zero in campaigns given that they enter into the elections. For each possible composition of actual candidates when estimating the entry model, I use 500 repetitions to simulate the expected payoffs, where the expectation in the first-order condition is calculated with 200 simulations.

4.3 Estimation Results

Table 3 summarizes the results of the structural analysis.[33] The first two columns show the estimates for the selection part of the campaign spending model together with the covariance matrix of the two errors, thus $(\hat{\alpha}, \hat{\sigma}_2, \hat{\sigma}_{12})$. The third and fourth columns correspond to the log(spending) equation of the campaign spending model, thus $\hat{\beta}$. The estimated standard deviation of the error in the log(spending)

32 The details of the estimation method can be found in Appendix A3.

33 The estimates of the first stage (i.e., $\hat{\eta}$) are reported in Appendix A4, which are of second-order importance.

equation is about 2.47, and the estimated covariance of the two errors in the selection equation and in the log(spending) equation is about 1.42, which yields an estimated correlation coefficient being around 0.58. A positive correlation coefficient means that the decision of making non-zero spending and the amount of spending of a candidate are positively correlated.

I now turn to the effects of explanatory variables. Incumbents will make non-zero campaign spending once they decide to run for the race, and the amount of spending is larger compared to challengers in Senate elections. If one election is open-seat, candidates tend to have a higher probability to make non-zero spending, and the amount of such spending tends to be larger, because candidates may think that without the incumbency advantage from the established reputation, they can have a larger chance of winning the office. I then focus on the estimated coefficients on the number of potential candidates and the number of actual candidates. Less actual candidates induce both a higher probability of making non-zero spending and a larger amount of spending, because actual candidates want to increase their winning chance in a less competitive environment where the marginal benefit of campaign spending can be larger than the marginal cost. The coefficient on the number of potential candidates is positive, being insignificant in the selection equation and significant in the log(spending) equation. However, since changing the number of potential candidates will accordingly influence the number of actual candidates, the impact of the number of potential candidates on spending remains unclear and has to be analyzed by incorporating the entry model.

For most of the covariates I include in the spending model specification, the signs in the selection equation and those in the log(spending) equation are the same. The only exceptions are the percent of population aged 25-44 and the unemployment rate. However, the estimated coefficients of these two variables are not significant. The log of median household income has a significant and positive coefficient in the log(spending) equation, indicating that in a relatively richer economy candidates are capable of making larger amounts of campaign spending.

The fifth and sixth columns present the estimated results for the voter model, i.e., $(\hat{\gamma}, \hat{\omega})$. These two parameters measure the marginal effects of campaign spending on the winning probabilities in the primary and general elections, respectively. Both estimates being positive implies that campaign spending has a positive effect on the candidates' winning probabilities, which means that voters appreciate the money candidates spend and may view it as something positively related to candidates' ability or valence, which is unobserved in practice. Note that $\hat{\omega} > \hat{\gamma}$,meaning that in the general election, voters appreciate campaign spending more, compared to the primary election, given the same amount of spending. However, since the specification of the latent utility for voters entails the log of one plus campaign spending, one needs to compare the effect of campaign spending on votes in the primaries to that in the general elections by taking account of the different amounts of campaign spending made by primary and general candidates. Therefore I calculate the average marginal effects of campaign spending on votes in the primary and general elections, which are 0.0511 and 0.0047 respectively.[34] This may indicate in the primaries candidates are more substitutive relative to the general elections, thus the campaign spending by candidates in the primary election is a bit more effective on average.

The last two columns give the estimates in the entry model, thus $(\hat{\delta}, \hat{\lambda}_I, \hat{\lambda}_C)$. First of all, incumbents have larger expected entry costs, compared to challengers. In footnote 15, I model the entry stage of potential candidates by referring entry cost to reservation value. This interpretation helps explain the higher entry probabilities by incumbents, which are consistent with the observed data. Specifically, incumbents may hold higher expectation regarding the value of holding the office, which turns to larger reservation value meaning that only when the expected payoff of winning the office exceeds the reservation value, they will decide to run.[35] Therefore, in Section 2.2, the reason that incumbents have higher entry probabilities is not because they have lower expected entry costs, but rather they can have higher expected payoffs if they win the office. Further, the standard deviation of entry costs for incumbents is marginally smaller than that of challengers in the entry cost distribution, which suggests that incumbents have a more concentrated entry cost distribution. Thus although incumbents have larger entry costs compared to challengers, among incumbents the difference of entry costs is not as large as that among challengers. Most of the estimated coefficients of the covariates have the expected signs, but the effects are insignificant. Specifically, the effect of the number of potential candidates on expected entry costs is positive (though insignificant). However since changing this number will affect the number of actual candidates that further has a significant effect on campaign spending, which in turn influences the entry behavior of potential candidates, it is still uncertain how the number of potential candidates affects outcomes in the entry stage, which will be studied in Section 5.1.

An important component in the structural analysis is the equilibrium strategy function mapping

34 Evaluated at the means of campaign spending by primary and general candidates, which are 2.8 million and 7.4 million respectively, the marginal effects of campaign spending on latent utilities of voters are around 5.74 and 2.98 (both scaled up by 10^8) in the primary and general elections respectively.

35 This is also consistent with the outcomes in Diermeier, Keane, and Merlo (2005), stating that the returns to outside opportunities depend on congressional experience. Hence incumbents hold higher opportunity costs.

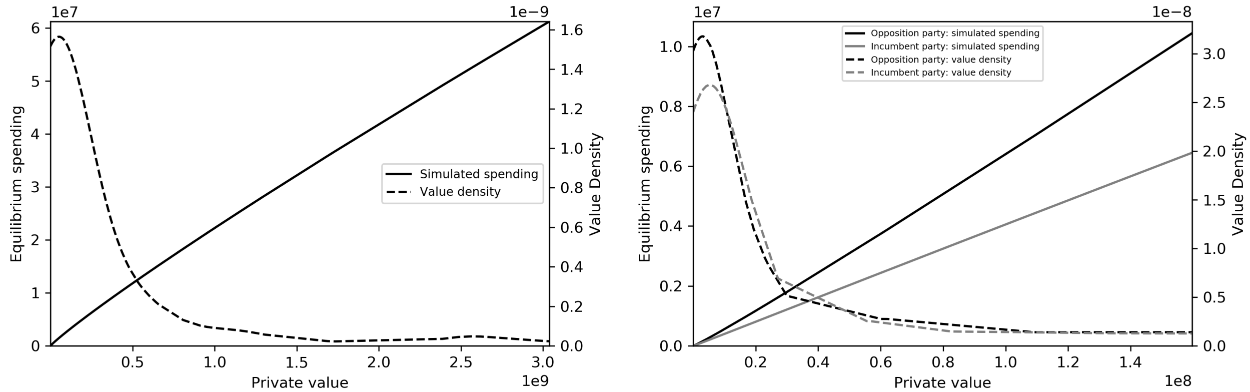
Table 3: **Estimation results from structural analysis**

	Spending: selection equation		Spending: log(spending) equation		Voter		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_2	2.4708	0.073						
σ_{12}	1.4211	0.224						
γ					0.1631	0.005		
ω					0.2203	0.014		
λ_I							0.4391	0.023
λ_C							0.5278	0.013
INCUM			4.0523	0.173			0.3627	0.112
OPEN	0.3867	0.151	1.2969	0.192				
PVI	−0.0082	0.009	−0.0097	0.012			0.0009	0.009
GOV	−0.1739	0.133	−0.0971	0.153			0.0603	0.143
YOUNG-PER	0.8538	2.511	−1.6093	2.897			−4.0242	2.696
MID-PER	−0.9982	1.675	−0.0710	2.151			2.4293	2.043
OLD-PER	0.1881	2.595	4.6085	3.364			−5.6572	3.192
UNEMP	0.0001	0.046	−0.0304	0.054			−0.0372	0.058
LOGINC	0.4749	0.555	1.8120	0.661			0.1646	0.521
PNCAN	0.0322	0.030	0.1346	0.036			0.0031	0.014
ANCAN	−0.0779	0.033	−0.2865	0.040				
CONST	−4.4034	5.998	−7.6213	7.172			0.4723	5.770

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. The estimated results are obtained through the estimation method described in Section 4.2. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

from the private value to the campaign spending in a given Senate election. This equilibrium strategy function can be different depending on the status of a candidate, whose private value following a type-specific distribution. I simulate the equilibrium strategy functions and the private value densities for the incumbent, the challenger from the opposition party, and the challenger from the incumbent party, using 200 draws, for the election indexed 1 in the data as a representation (this is not an open-seat election), shown in Figure 2. I also simulate the equilibrium strategy functions and the private value densities for the candidates from two parties in an open-seat election (indexed 19) as shown in Figure 3.

Figure 2 presents the equilibrium strategy functions and the private value densities in an incumbent-challenger election. Apparently, the equilibrium strategy function is strictly increasing for all candidates. Comparing the curves for the incumbent and for the two types of challengers, it is clear that incumbents value the office more than the challengers, and thus are willing to spend more to raise their winning probabilities. As the sub-figure 2b indicates, the distribution of the opposition party challenger's value is slightly shifted leftward from that of the incumbent party challenger's value, which indicates that on average the challengers in the incumbent party have slightly larger values relative to the challengers



(a) Incumbent **(b)** Challenger

Figure 2: Simulated strategy functions and value distributions for incumbent and challengers in a representative incumbent-challenger election

in the opposition party probably due to a more appealing office with the existence of the incumbent. However, the scales of private values for the two types of challengers are similar, but the challengers from the opposition party seem to spend more. Remember that this representative election is not an open-seat election. Therefore challengers from the incumbent party may think they have a smaller chance of winning the primary election facing an incumbent, thus reducing their spending.

I now turn to the open-seat election. Figure 3 gives the equilibrium strategy functions and the private value distributions for candidates from the incumbent and opposition parties. Although there is no incumbent participating in the election, I still define the parties as above for the ease of illustration and comparison. Now the distribution of the opposition party candidate's value is slightly shifted rightward from that of the incumbent party challenger's value, but the two distributions are still comparable. This is probably because without the incumbent, the challengers from the opposition party now have larger values in a battleground. The scales of private value and campaign spending of candidates in an open-seat election are comparable to those of incumbents, and much bigger than those of challengers in an incumbent-challenger election. Although there are differences between candidates from both parties, the differences are not large proportionally when compared to the case with the existence of an incumbent. In this representative open-seat election, it appears that candidates in the opposition party spend relatively more than those in the incumbent party. This can be potentially explained by the impression of candidates from the opposition party that it is possible to flip the state if the incumbent decides not to run, because the state may be at the margin.

In Appendix A5, I conduct a series of extensive sensitivity analyses, including considering the

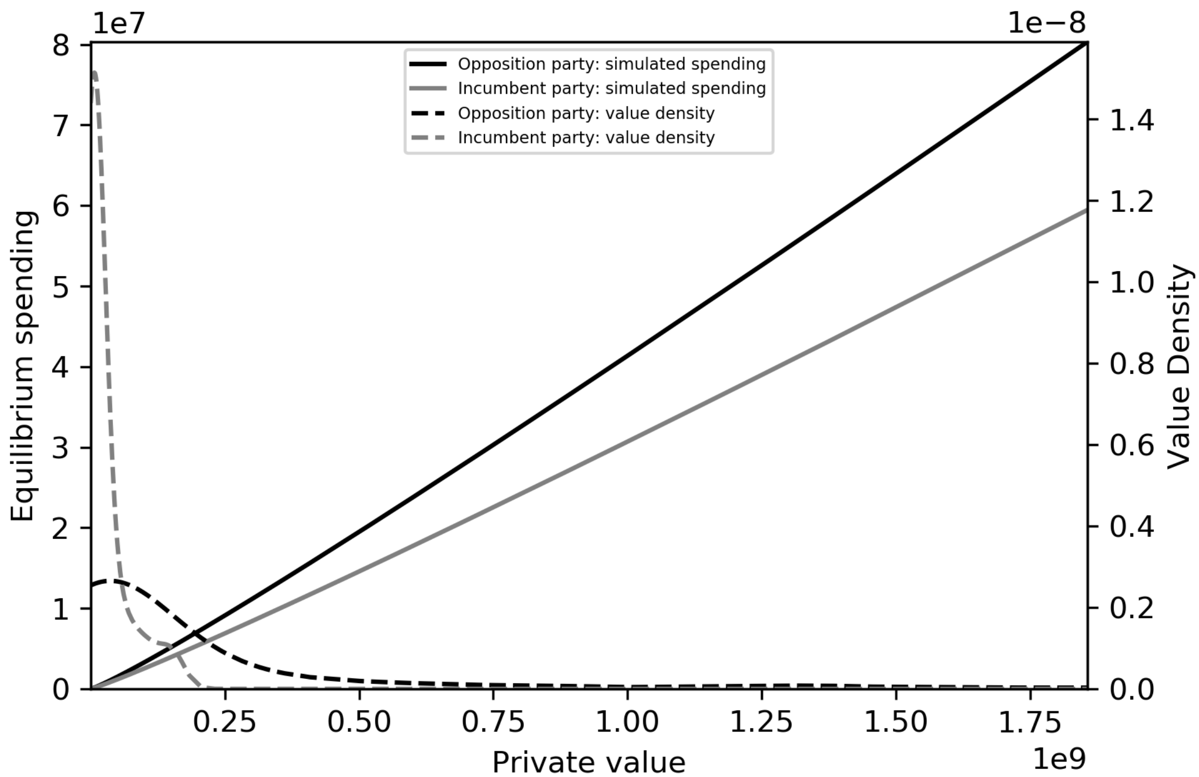


Figure 3: Simulated strategy functions and value distributions for candidates in a representative open-seat election

turnout of voters in the general elections (Appendix A5.1), a truncated spending distribution from above (Appendix A5.2), taking into account the difference between two political parties (Appendix A5.3), a quadratic cost function (Appendix A5.4), an alternative voter model specification (Appendix A5.5), including time-specific constants in all specifications (Appendix A5.6), as well as adding a dummy variable that is assigned one for the period of post-2008 and zero for that of pre-2008 to the specifications in the model (Appendix A5.7). I find the results are robust to these sensitivity analyses.

4.4 Model Fit

To assess the fit of the model, I employ the in-sample prediction, where I use the estimates from the structural analysis to simulate the data and calculate some key moments displayed in Table 4. The predicted first moments of some key features of the model are reported in order to compare with the observed counterparts through 500 simulations, together with the simulated standard errors. The overall fit of the predicted data to the actual data is good in both the level and the trend.

Table 4 shows the actual and predicted moments regarding the spending model, the voter model, and the entry model. In particular, the campaign spending distribution simulated from the estimated structural parameters fits the observed data well. The spending made by incumbents is a bit larger in the simulation, but in both predicted and actual data the spending of incumbents is always larger than that of challengers. For the voter model, I simulate the moments of vote shares of incumbents in the primary and general elections, to echo the specification where I attach importance to how incumbents behave differently from challengers in Senate elections. The vote shares of incumbents are on average larger in the within-party primaries, revealed in both predicted and actual data. This can also be reflected via the ratio of incumbent-challenger elections (either primary or general) incumbents win. I report the predicted means of the number of actual candidates, the entry proportion, and the proportion of open-seat elections for the entry model. Typically, I predict the entry patterns well, with slightly larger entry probabilities that lead to both bigger entry proportion and larger number of actual candidates. The percentage of open-seat elections is predicted to be slightly smaller, because I predict the campaign spending mean to be larger for incumbents that is associated with more expected payoff and higher entry probability of incumbents.

Table 4: Model fit results

	Spending		Voter		Entry	
	Predicted	Observed	Predicted	Observed	Predicted	Observed
Non-zero spending	0.6901 (0.009)	0.6945				
Spending	4.6132 (0.326)	4.2521				
Spending of incumbents	12.168 (1.239)	9.4280				
Spending of challengers	2.8793 (0.271)	2.8395				
Vote shares of incumbents in primary elections			0.7925 (0.008)	0.8840		
Primary elections incumbents win			0.9144 (0.017)	0.9820		
Vote shares of incumbents in general elections			0.7633 (0.011)	0.6016		
General elections incumbents win			0.8966 (0.019)	0.8440		
No. of actual candidates					5.8939 (0.060)	5.4869
Entry proportion					0.8227 (0.008)	0.7785
Open-seat election					0.1038 (0.017)	0.1830

Notes: The simulated standard errors are reported in parenthesis. The predicted and observed figures for spending, spending of incumbents, and spending of challengers are scaled down by 10^6.

5 Counterfactual Analyses

I use the estimation results to conduct two sets of counterfactual analyses to demonstrate the importance of accounting for the endogenous entry of candidates and to compare the effects of proposed institutional reforms.

5.1 The Role of Entry

In the structural analysis, the number of potential candidates is exogenous as I define potential candidates as those who file reports about campaign financing and thus appear in the spending data. In this counterfactual analysis, I investigate the impacts of varying the number potential candidates on the spending, vote and entry results. To compare behavior in alternative environments, I do so for representative elections. I first select two elections with different degrees of asymmetries between two parties: Election one has 9 challengers in both incumbent and opposition parties (indexed 21 in the data); Election two has 19 challengers in the incumbent party and 10 challengers in the opposition party (indexed 107). I assess the effects of decreasing (and increasing) the numbers of potential candidates in *both* incumbent and opposition parties (denoted as N and M) in the left panels of Figure 4 on spending outcomes, Figure 5 on vote outcomes, and Figure 6 on entry outcomes via simulations.

Despite different degrees of asymmetries between two parties in Election one and Election two, the profiles present similar patterns regarding the spending, voting, and entry outcomes. When the numbers of potential candidates in *both* incumbent and opposition parties increase, the non-zero spending ratio and the spending level decrease on average, regardless of the candidate status, as shown in the left panel of Figure 4. The reason that the non-zero spending ratio decreases is that the candidate with any level of private type will decrease the spending, corresponding to a homogenous effect. Moreover, increasing the numbers of potential candidates in *both* incumbent and opposition parties has a negative effect on the entry proportion and a positive effect on the number of actual candidates, because more potential candidates lead to larger entry cost and smaller entry probability on average, which in turn results in larger number of actual candidates (thus a more competitive environment) when combined with more potential candidates (see the left panel of Figure 6). In contrast, the impact on the probability of being an open-seat election is not monotonic, reflecting the combined effects of smaller spending by both incumbents and challengers. Turning to the vote outcomes in the left panel of Figure 5 where I focus on the incumbent, increasing the numbers of potential candidates in *both* incumbent and opposition parties only has a decreasing effect on the vote share of the incumbent in the primary, due to more actual

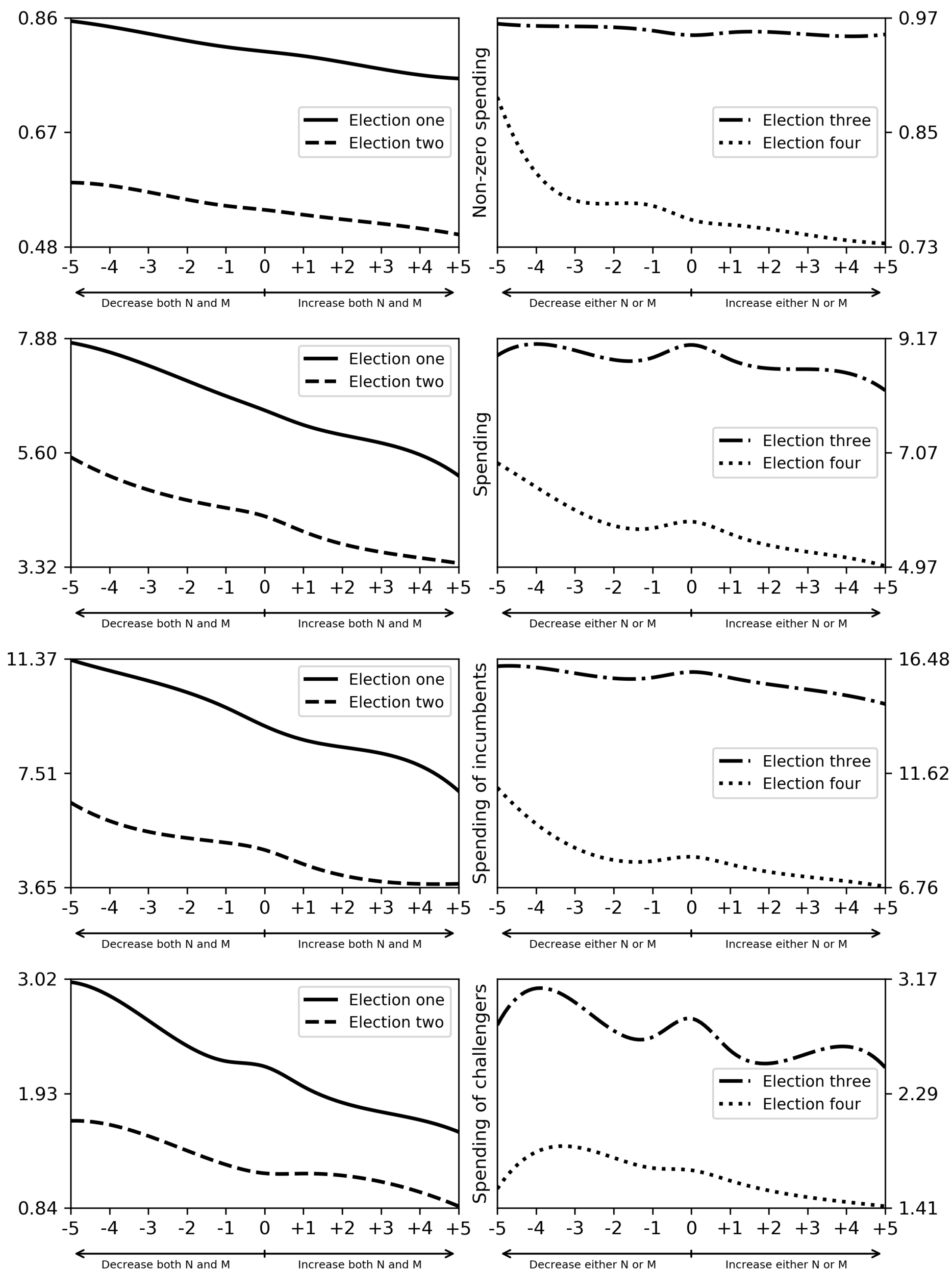


Figure 4: The effects of varying the number of potential candidates on spending outcomes (scaled down by 10^6)

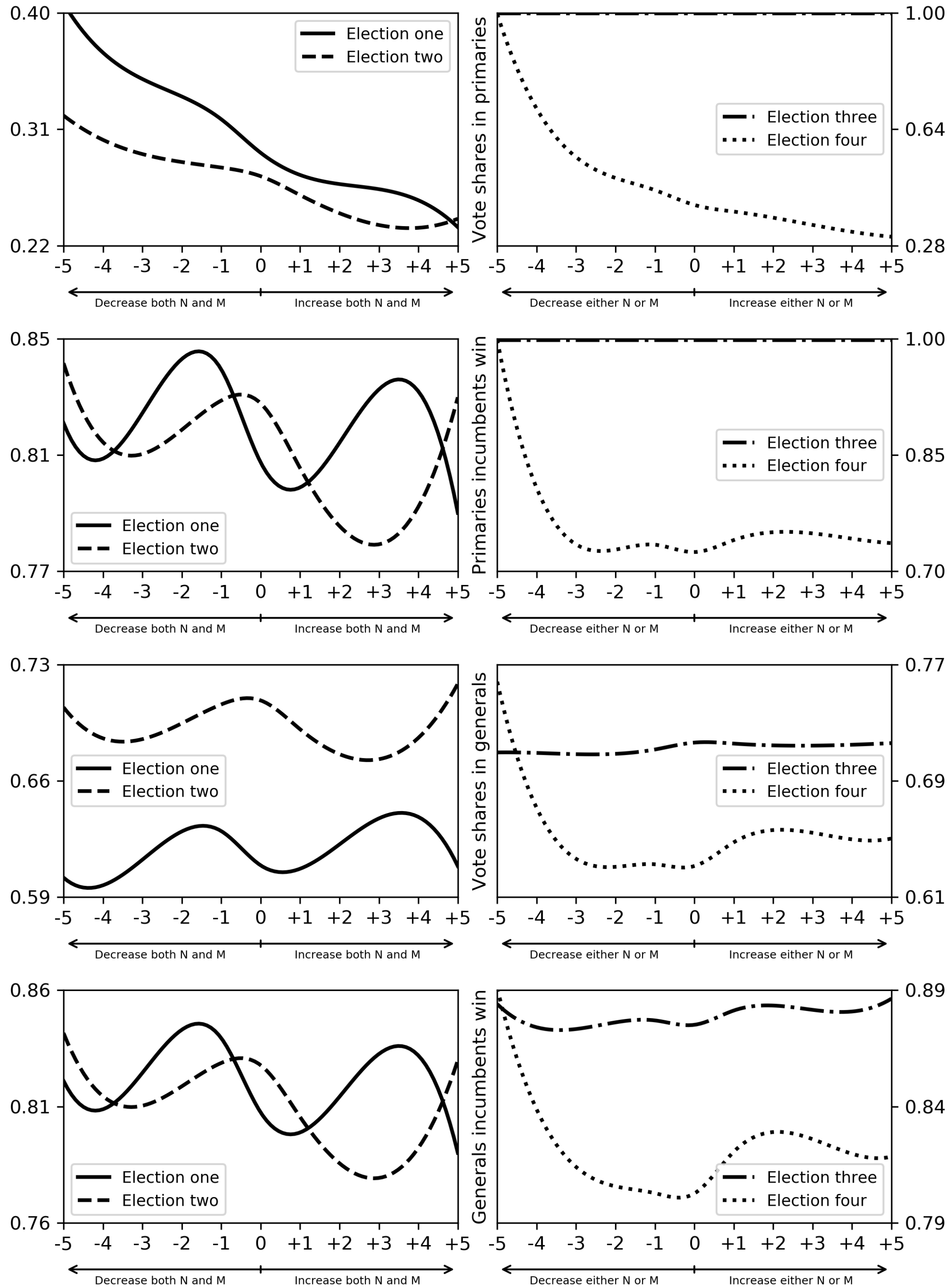


Figure 5: The effects of varying the number of potential candidates on vote outcomes for incumbents

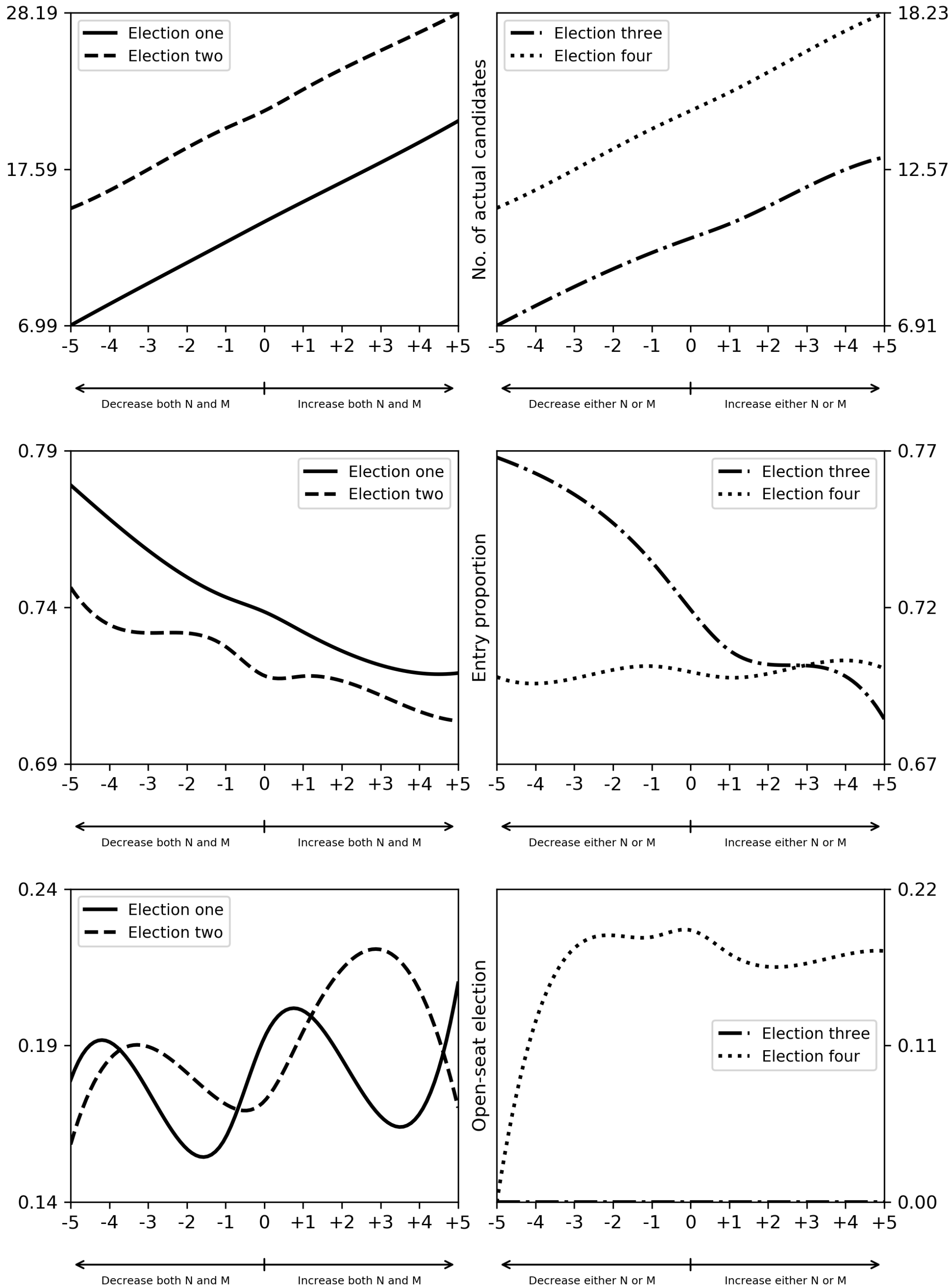


Figure 6: The effects of varying the number of potential candidates on entry outcomes

candidates in the election. However, the effects on the primary winning probability, the vote share in the general, and the general winning probability are volatile, consistent with the non-monotonic pattern on the probability of having an open-seat election.

Next, I investigate two new elections: Election three has no challengers in the incumbent party and 13 challengers in the opposition party (indexed 29); Election four has 5 challengers in the incumbent party and 15 challengers in the opposition party (indexed 40), and redo the above analysis. This time, I only decrease (and increase) the number of potential candidates in the opposition party for Election three, thus keeping the incumbent being unopposed in the primary. For Election four, I decrease (and increase) the number of potential candidates in the incumbent party. The simulated outcomes are exhibited in the right panels of Figure 4, Figure 5, and Figure 6.

The results indicate different effects of varying the number of potential candidates in the opposition or the incumbent party. If letting the incumbent be unopposed in the primary (as in Election three), increasing the number of potential candidates in the opposition party does not have notable or monotonic effects on the non-zero spending ratio and the average spending level. Because there is no challenger in the incumbent party, the primary vote share and the probability of winning primary for the incumbent remain 1 while the probability of having an open-seat election remains 0. The general vote outcomes for the incumbent appear steady. Putting together, this implies stable entry behavior of the incumbent. In contrast, challengers experience non-monotonically decreasing spending and monotonically decreasing entry probability on average, resulting in decreasing entry proportion and increasing number of actual candidates on average. In summary, varying the number of potential candidates in the opposition party does not change the behavior of the incumbent much.

On the other hand, if holding the opposition party unchanged (as in Election four), increasing the number of potential candidates in the incumbent party can reduce the non-zero spending ratio and the average spending level. In terms of the scale, the incumbent's spending decreases more compared to challengers. This is consistent with the vote outcomes showing that the incumbent's vote shares and winning chances of the primary and general elections are decreasing. The entry outcomes also corroborate the relatively stronger challengers and weaker incumbent as the number of potential candidates in the incumbent party increases, since the entry proportion is almost unchanged but the probability of having an open-seat election is increasing. Furthermore, more potential candidates induce more actual candidates. In summary, a more effective way to bring in more competitiveness to the election from more challengers who confront the incumbent is to introduce more potential candidates in the incumbent party's primary, but not the opposition party's primary.[36]

5.2 Top-Two Primary

In this subsection, I implement a top-two primary system counterfactual in which all candidates who participate in the primary election are listed on the same ballot, and the top two winners of the primary proceed to the general election and compete for office.[37] This counterfactual admits a different form of CSF that determines the winner of the election. Under the design of this paper, the candidate who has the highest winning probability in the primary will also possess the highest winning probability in the general election. Thus the CSF in the top-two primary can be presented by the usual multinomial logit form.

As in Section 4.3, I present the equilibrium strategy functions for different types of candidates in a representative incumbent-challenger election and a representative open-seat election. Using the same incumbent-challenger election as in Figure 2, I draw the equilibrium strategy functions under the counterfactual setting in Figure 7.[38]

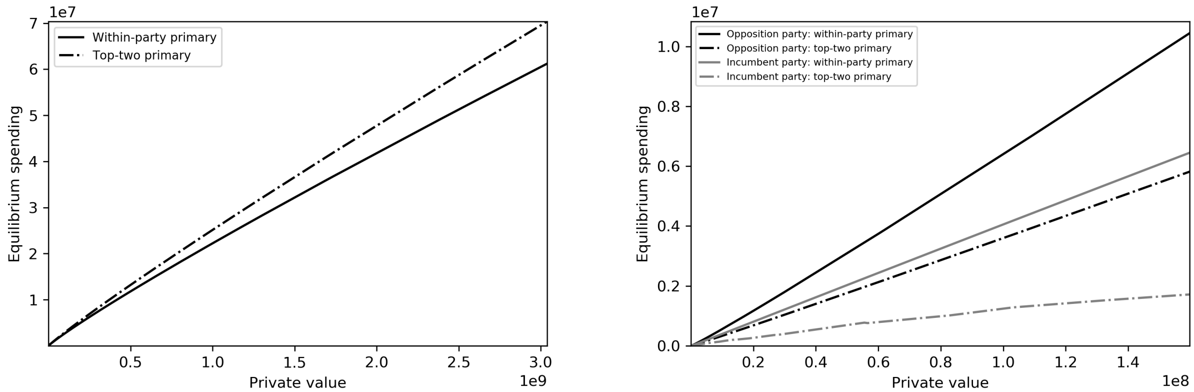
Figure 7 compares the simulated strategy functions in the counterfactual framework (lines of dashes and dots) and in the benchmark model (solid lines). It appears that the change of the spending behavior in the counterfactual is different depending on the candidate status. For incumbents, they spend more under the top-two primary system; while for challengers of both types, they spend less in the counterfactual equilibrium. The scale of change is also worth discussing. Although incumbents tend to spend more, the increasing magnitude is not large, compared to the decreasing magnitude of challengers, where they tend to spend much less. This is because under the top-two primary system,

---

36 This indication complements what is found in Hirano and Snyder (2014) who emphasize the importance of the primary election in safe constituencies where one party's candidate will have a large advantage in the general election. I further suggest that increasing the number of potential candidates in the incumbent party that is always the advantaged party can bring in more competitiveness to the election.

37 Recently, several states have incorporated alternative primary system into their governing constitutions: the top-two primaries. For example, Washington passed a new primary system in 2004 and implemented the top-two primary, which applies to federal, state, and local elections, from 2008. Moreover, California started using the top-two primary system (excluding presidential elections) in 2012. Louisiana also enacted the top-two primary system from as early as 1975, which differs from those of Washington and California because if one candidate receives more than 50% of the total vote in the primary, that candidate is declared the winner and no runoff election is held.

38 It is valid to compare the equilibrium strategies for the same election, though the entry behavior of potential candidate may change due to different expected payoffs under different primary systems, which leads to different composition of actual candidates under different systems. This is because the entry equilibrium in the current model setup is characterized by entry probabilities of potential candidates. Therefore, the observed composition in the actual data is assigned with different probabilities under different primary systems. And this will not affect how I compare the equilibrium strategy functions in the benchmark model and in the counterfactual analysis, given the composition of actual candidates.



(a) Incumbent

(b) Challenger

Figure 7: Simulated strategy functions in the top-two primary system for incumbent and challengers in a representative incumbent-challenger election

voters face more choices in the primary election, thus incumbency advantage may be amplified in the early stage of the election. Challengers from both parties think their chance of advancing to the general election is lower with the existence of an incumbent and more candidates, and consequently reduce their spending, because spending is less effective in the primary and they find it less beneficial to spend too much. On the other hand, the incumbent spends a bit more, in order to lock the win. However, since the marginal effect of spending on winning is diminishing and the incumbent's spending level is already high, she will not spend too much more.

Similarly, using the same open-seat election as in Figure 3, I plot the equilibrium strategy functions under the counterfactual setting in Figure 8.

Figure 8 depicts the simulated equilibrium strategy functions under the top-two primary system that are shown in lines of dashes and dots, where the equilibriums in the benchmark model are included via solid lines. In this case, candidates in open-seat election will spend less if the primary election takes the form of the top-two system. However, the magnitude of change is not large relative to the scale in the open-seat election.

I use the numerical routine to simulate the first moments of the same key features in the spending, voter, and/or entry models as in Section 4.4 under two different scenarios: (i) I hold entry fixed at the benchmark level that is observed from the data; and (ii) I allow entry to adjust with a different primary system where potential candidates decide whether they will run for office endogenously. Table 5 and Table 6 report the simulated results under these two scenarios.

First of all, if I study the campaign spending by incumbents and challengers separately, I find that

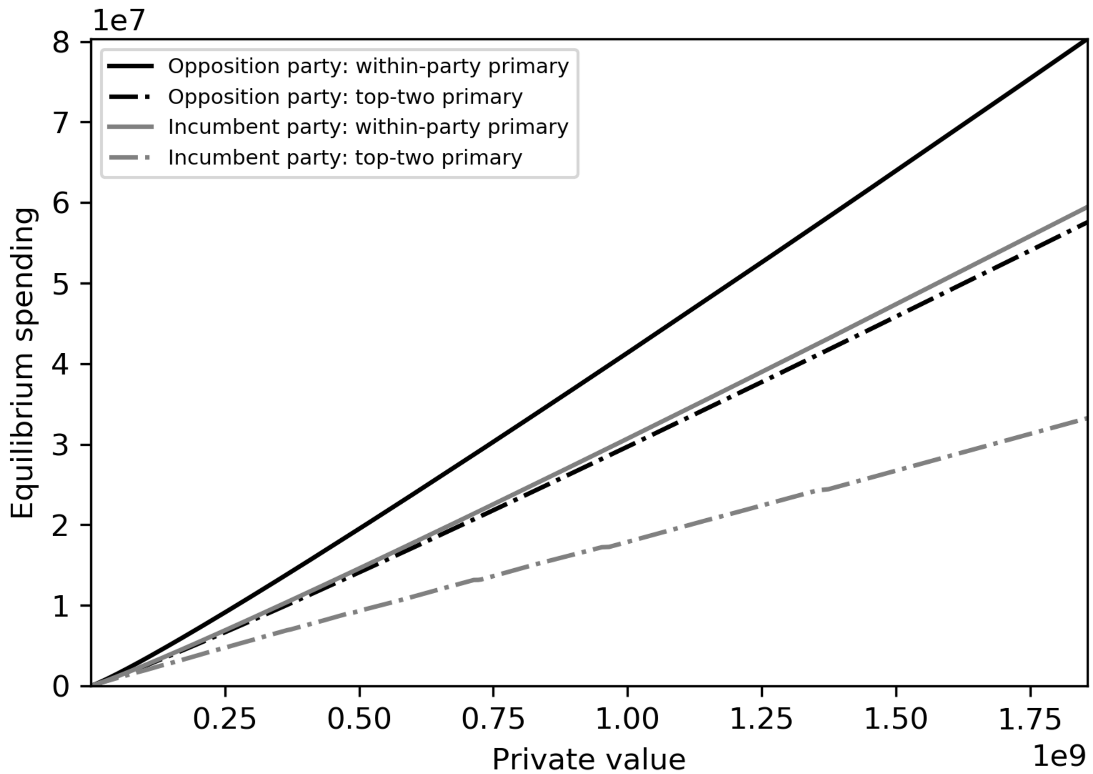


Figure 8: Simulated strategy functions in the top-two primary system for candidates in a representative open-seat election

Table 5: Counterfactual: top-two primary with fixed entry

	Spending	Voter
Non-zero spending	0.6901 (0.031)	
Spending	5.7620 (0.898)	
Spending of incumbents	15.265 (2.589)	
Spending of challengers	1.5780 (0.428)	
Vote shares of incumbents in elections		0.4107 (0.014)
Elections incumbents win		0.8170 (0.027)

Notes: The simulated standard errors are reported in parenthesis. The figures for spending, spending of incumbents, and spending of challengers are scaled down by 10^6.

Table 6: Counterfactual: top-two primary with endogenous entry

	Spending	Voter	Entry
Non-zero spending	0.6900 (0.029)		
Spending	4.1648 (1.865)		
Spending of incumbents	13.117 (0.421)		
Spending of challengers	1.2405 (0.620)		
Vote shares of incumbents in elections		0.4329 (0.062)	
Elections incumbents win		0.8947 (0.096)	
No. of actual candidates			5.8237 (0.064)
Entry proportion			0.8154 (0.008)
Open-seat election			0.1053 (0.017)

Notes: The simulated standard errors are reported in parenthesis. The figures for spending, spending of incumbents, and spending of challengers are scaled down by 10^6.

on average incumbents spend more and challengers spend less in the counterfactual model regardless of fixed or endogenous entry. This is consistent with Figure 7 and Figure 8 above. However, the average campaign spending appears differently conditional on fixed or endogenous entry. With fixed entry, the average campaign spending increases under the alternative top-two primary system; while with endogenous entry, the average spending is smaller.[39] With fixed entry, the campaign spending of candidates regardless of the candidate status is on average larger than that when entry is endogenous. The reason is probably that with endogenous entry, candidates can choose to not run for office if they find it unworthy. But if entry is fixed, they can only increase their spending level more to offset some cost because they cannot secede.

As for the voter model, note that in the counterfactual model the incumbent who wins the top-two primary election will have the highest winning probability and receive the highest vote shares in the general election, thus win the office eventually. As a result, I only report the vote shares of incumbents in primary elections and the ratio of incumbent-challenger primary elections incumbents win, which are more informative. Compared with the benchmark model, in the counterfactual the average vote share of incumbents drops drastically, regardless of fixed or endogenous entry. This is because under the new system, excess candidates reduce the votes the incumbent can obtain.[40] However, the ratio of primary elections incumbents win remains stable with endogenous entry. When entry is fixed, this ratio decreases marginally. When entry is endogenous, since incumbents who decide to run for the race on average have larger spending, their winning chance is still high compared to challengers, thus resulting in similar probability that an election is won by the incumbent. When entry is fixed, those incumbents who find entry unworthy cannot drop out and thus have lower chance of winning the primaries on average.

I next turn to the entry model in Table 6. Compared with the benchmark model, the entry proportion and the number of actual candidates are slightly smaller, while the ratio of open-seat elections is higher,

39 Sparks (2018) analyzes the state legislative general elections in California and Washington, which include both two-party and one-party elections under the top-two primary setting, and finds that increasing challenger campaign spending will generate greater vote share per dollar spent in one-party than two-party contests, as the expenditure made by challengers facing same-party opponents is more effective due to the absence of differentiating party labels in one-party elections. However, in the counterfactual analysis, I do not change how voters appreciate the campaign spending by candidates (i.e., $\hat{\gamma}$ and $\hat{\omega}$).

40 In Beck and Henrickson (2013), they also find that the switch to the top-two primary reduces the likelihood of having multiple Democratic candidates in a race, but does not have a significant result on that of Republican candidates in Washington State Legislative Primaries in 2004, 2006, 2008, and 2010. This can be explained by that since the votes in the top-two primary are split within one party with excess candidates, that party may end up with no candidates in the general election.

all due to the smaller entry probabilities of candidates. The differences are rather modest compared to those in the spending model. This can be explained by the equilibrium behavior of candidates. In the counterfactual model, the candidate who seeks to maximize her expected payoff will adjust her spending level in order to balance between the expected chance of winning and the cost of the spending, upon entry. Therefore, although the equilibrium spending made by the candidate changes significantly regardless of the candidate status, the resulting change in the entry equilibrium is attenuated because the adjusted expected payoff's change is attenuated.

In conclusion, the top-two primary system may give voters more choices in the primary, but at the cost of reducing their choices in the general election, where usually more voters participate. From this perspective, the top-two primary system may not lead to more competition in the election. With fixed entry, I find that campaign spending increases on average despite that challengers reduce their spending levels, and that average winning probability of incumbents is marginally smaller compared to the benchmark model. However, when endogenous entry is accounted for, the results show that on average the top-two primary system leads to relatively less campaign spending of candidates, and similar winning chance of incumbents. More importantly, with endogenous entry, this primary system seems to have a negative effect on the entry probabilities of candidates. Therefore, this implies that it is important to take into account the endogenous entry of potential candidates.[41]

6 Conclusion

In this paper, I develop and estimate a two-stage game-theoretic contest model to quantify the effect of campaign spending on electoral outcomes given endogenous candidate entry, using data on the U.S. Senate elections from 1994 to 2018. The model consists of two stages, with the first entry stage where potential candidates decide whether to run for the race, and the second election stage where actual candidates attend a Senate election with both within-party primaries and a general election and choose the spending in campaigns strategically. I also specify voters' decisions via a latent utility model that depends on candidates' campaign spending, demographic and economic covariates, and the election-specific unobserved heterogeneity. Taking the structural approach, I obtain estimates of

41 Although beyond the scope of this paper, the top-two primary system also affects the ideological position of the selected winner. For example, Amorós, Puy, and Martínez (2016) conduct a theoretical analysis with candidate entry and both primary and general elections and show that top-two primaries contribute to political moderation. Bullock and Clinton (2011) find that the blanket primary appears to produce more moderate representatives only in less partisan districts in the U.S. House of Representatives and California Assembly.

parameters in the spending model for actual candidates, which can be used to simulate the equilibrium strategy function in the election stage. I also get the estimates of parameters in the voter model, which indicate how the campaign spending translates into votes. The estimated entry cost distribution is helpful to characterize the entry behavior of potential candidates.

I find that incumbents tend to have larger private values of the office, spend more in campaigns, and enter more frequently into the election, compared to challengers. When incumbent does not participate in the election, actual candidates have larger amount of campaign spending. I also find that a more competitive election can induce the candidates to spend less, because they have a smaller chance to win the office upon entry. For voters, candidates with more campaign spending are appreciated more by voters, especially in the primary elections, thus enhancing their winning probabilities of the race. The impact of the number of potential candidates is examined through simulations. The main insight is that in the entry stage prior to the election, more potential candidates can generate larger entry cost on average and more actual candidates in the following election.

Moreover, how the number of potential candidates influences campaign spending and vote outcomes depends on which party's number being changed. Increasing numbers of potential candidates in both incumbent and opposition parties has negative effects on campaign spending, the vote share of incumbents in the primary, and entry proportion on average, regardless of the degree of asymmetry in the size of the primary between two parties. However, its effect on the winning chance of incumbents in the general or the probability of having an open-seat election is not monotonic. If letting the incumbent being unopposed in the primary, increasing the number of potential candidates in the opposition party does not have notable effect on the behavior of incumbents. In contrast, increasing the number of potential candidates in the incumbent party can reduce the relative strength of incumbents, in terms of less campaign spending, smaller winning probability in both primary and general, and lower propensity of incumbent entry. This implies that the incumbent party's primary election can be important in terms of improving electoral competition via introducing more challengers to confront the incumbent.

I then use the structural estimates to conduct the counterfactual analysis and study the top-two primary system. The results show that although the top-two primary system gives voters more choices in the primary, in the general election where more voters participate the choice set is more restricted. Because candidates have smaller entry probabilities, this alternative primary system induces less competition in the Senate election. Further, in this counterfactual, while incumbents will spend slightly more, challengers will spend much less in campaigns, which on average can generate a smaller amount of campaign spending. Importantly, when the participation behavior of potential candidates is ignored and entry is assumed to be fixed, the campaign spending appears to increase and the winning probability of incumbents is smaller in the alternative top-two format, both compared to the benchmark results, thus indicating opposite patterns compared to the situation with endogenous entry. This reveals the key role of endogenous entry in the analysis.

Appendix

A1 The Derivation of CSF in the Voter Model

Under the assumptions in Section 3.1, I show how to derive the winning probability, or the CSF for a representative actual candidate D_i from Democratic party, with R_k being the opponent in the general election from the Republican party, where $i \in \{1, \cdots, n\}$ and $k \in \{1, \cdots, m\}$ (see Web Supplement of Adams and Merrill (2008)). I use the notations appearing in the main text, and first consider the benchmark model where the turnout of voters in the general election is ruled out.

Fix i and k. The winning probability as a joint probability is expressed as following:

$$\begin{aligned}
&P(D_iR_k)\\
=&Pr\begin{pmatrix} M_D(D_i)+\iota_{D_i} > M_D(D_j)+\iota_{D_j},\ j=1,\cdots,i-1,i+1,\cdots,n;\\ M_R(R_k)+\iota_{R_k} > M_R(R_l)+\iota_{R_l},\ l=1,\cdots,k-1,k+1,\cdots,m;\\ M_G(D_i)+\iota_{D_i} > M_G(R_k)+\iota_{R_k} \end{pmatrix}\\
=&Pr\begin{pmatrix} \iota_{D_j} < \iota_{D_i}+M_D(D_i)-M_D(D_j),\ j=1,\cdots,i-1,i+1,\cdots,n;\\ \iota_{R_l} < \iota_{R_k}+M_R(R_k)-M_R(R_l),\ l=1,\cdots,k-1,k+1,\cdots,m;\\ \iota_{R_k} < \iota_{D_i}+M_G(D_i)-M_G(R_k) \end{pmatrix}\\
=&Pr\begin{pmatrix} \iota_{D_j} < \iota_{D_i}+W_{D_j},\ j=1,\cdots,i-1,i+1,\cdots,n;\\ \iota_{R_l} < \iota_{R_k}+W_{R_l},\ l=1,\cdots,k-1,k+1,\cdots,m;\\ \iota_{R_k} < \iota_{D_i}+W_G \end{pmatrix}.
\end{aligned}$$

Given ι_{D_i} and ι_{R_k}, this joint probability can be written as:

$$P(D_iR_k|\iota_{D_i},\iota_{R_k}) = \begin{cases} \prod_{j\neq i}^{n} F_{D_j}(\iota_{D_i}+W_{D_j}) \times \prod_{l\neq k}^{m} F_{R_l}(\iota_{R_k}+W_{R_l}), & \text{if } \iota_{R_k} < \iota_{D_i}+W_G\\ 0, & \text{otherwise} \end{cases}.$$

Integrating out ι_{D_i} and ι_{R_k}, I get the joint probability written as the following:

$$P(D_iR_k) = \int_{-\infty}^{\infty} \prod_{j\neq i}^{n} F_{D_j}(\iota_{D_i}+W_{D_j}) \times f_{D_i}(\iota_{D_i}) \times \int_{-\infty}^{\iota_{D_i}+W_G} \prod_{l\neq k}^{m} F_{R_l}(\iota_{R_k}+W_{R_l}) \times f_{R_k}(\iota_{R_k})\, d\iota_{R_k} d\iota_{D_i}.$$

Let $s=\iota_{D_i}$ and $t=\iota_{R_k}$, I have

$$P(D_iR_k) = \int_{-\infty}^{\infty} exp[-e^{-s}\sum_{j\neq i}^{n} e^{-W_{D_j}}]\cdot exp[-e^{-s}]e^{-s}\cdot \int_{-\infty}^{s+W_G} exp[-e^{-t}\sum_{l\neq k}^{m} e^{-W_{R_l}}]\cdot exp[-e^{-t}]e^{-t}\, dtds.$$

I first derive the inner part of the double integral:

$$\int_{-\infty}^{s+W_G} \prod_{l\neq k}^{m} exp[-e^{-t}\sum_{l\neq k}^{m} e^{-W_{R_l}}] \cdot exp[-e^{-t}]e^{-t}\, dt$$

$$\stackrel{e^{-t}=u}{=} \int_{e^{-(s+W_G)}}^{\infty} exp[-u(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1)]\, du$$

$$= \frac{exp[-e^{-(s+W_G)}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1)]}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1}.$$

Therefore,

$$\begin{aligned}
&P(D_iR_k)\\
=&\int_{-\infty}^{\infty} exp[-e^{-s}(\sum_{j\neq i}^{n} e^{-W_{D_j}}+1)]\cdot exp[-e^{-s}e^{-W_G}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1)]e^{-s}ds\cdot\frac{1}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1}\\
=&\int_{-\infty}^{\infty} exp[-e^{-s}(\sum_{j\neq i}^{n} e^{-W_{D_j}}+1+e^{-W_G}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1))]e^{-s}ds\cdot\frac{1}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1}\\
\stackrel{e^{-s}=v}{=}&\int_{0}^{\infty} exp[-v(\sum_{j\neq i}^{n} e^{-W_{D_j}}+1+e^{-W_G}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1))]dv\cdot\frac{1}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1}\\
=&\frac{1}{\sum_{j\neq i}^{n} exp(-W_{D_j})+1+exp(-W_G)\left[\sum_{l\neq k}^{m} exp(-W_{R_l})+1\right]}\cdot\frac{1}{\sum_{l\neq k}^{m} exp(-W_{R_l})+1}.
\end{aligned}$$

Next, I consider the extension where the turnout of voters in the general election is included. If a voter does not turnout for the general election, I assume that she receives a utility of ι_{G_0} from the outside good. This ι_{G_0} follows type-1 extreme-value distribution, independent from all other ιs.

In this case, the winning probability of the representative Democratic party candidate D_i is then:

$$\begin{aligned}
&P(D_iR_k)\\
=&Pr\begin{pmatrix}
M_D(D_i)+\iota_{D_i} > M_D(D_j)+\iota_{D_j},\ j=1,\cdots,i-1,i+1,\cdots,n\\
M_R(R_k)+\iota_{R_k} > M_R(R_l)+\iota_{R_l},\ l=1,\cdots,k-1,k+1,\cdots,m;\\
M_G(D_i)+\iota_{D_i} > M_G(R_k)+\iota_{R_k};\\
M_G(D_i)+\iota_{D_i} > \iota_{G_0}
\end{pmatrix}
\end{aligned}$$

$$
=Pr\begin{pmatrix} \iota_{D_j} < \iota_{D_i} + M_D(D_i) - M_D(D_j),\ j = 1,\cdots,i-1,i+1,\cdots,n; \\ \iota_{R_l} < \iota_{R_k} + M_R(R_k) - M_R(R_l),\ l = 1,\cdots,k-1,k+1,\cdots,m; \\ \iota_{R_k} < \iota_{D_i} + M_G(D_i) - M_G(R_k); \\ \iota_0 < \iota D_i + M_G(D_i) \end{pmatrix}
$$

$$
=Pr\begin{pmatrix} \iota_{D_j} < \iota_{D_i} + W_{D_j},\ j = 1,\cdots,i-1,i+1,\cdots,n; \\ \iota_{R_l} < \iota_{R_k} + W_{R_l},\ l = 1,\cdots,k-1,k+1,\cdots,m; \\ \iota_{R_k} < \iota_{D_i} + W_G; \\ \iota_{G_0} < \iota_{D_i} + M_G(D_i) \end{pmatrix}.
$$

Given ι_{D_i} and ι_{R_k}, this joint probability can be written as:

$$
P(D_iR_k|\iota_{D_i},\iota_{R_k}) = \begin{cases} \prod_{j\neq i}^{n} F_{D_j}(\iota_{D_i}+W_{D_j}) \times \prod_{l\neq k}^{m} F_{R_l}(\iota_{R_k}+W_{R_l}) \times \times F_{G_0}(\iota_{D_i}+M_G(D_i)), & \text{if } \iota_{R_k} < \iota_{D_i} + W_G \\ 0, & \text{otherwise} \end{cases}.
$$

Integrating out ι_{D_i} and ι_{R_k}, I get the joint probability expressed as the following:

$$
\begin{aligned}
&P(D_iR_k) \\
=&\int_{-\infty}^{\infty} \prod_{j\neq i}^{n} F_{D_j}(\iota_{D_i}+W_{D_j}) \times F_{G_0}(\iota_{D_i}+M_G(D_i)) \times f_{D_i}(\iota_{D_i}) \times \int_{-\infty}^{\iota_{D_i}+W_G} \prod_{l\neq k}^{m} F_{R_l}(\iota_{R_k}+W_{R_l}) \times f_{R_k}(\iota_{R_k})\, d\iota_{R_k} d\iota_{D_i}.
\end{aligned}
$$

Let $s = \iota_{D_i}$ and $t = \iota_{R_k}$, I have

$$
\begin{aligned}
&P(D_iR_k) \\
=&\int_{-\infty}^{\infty} exp[-e^{-s}\sum_{j\neq i}^{n} e^{-W_{D_j}}] \cdot exp[-e^{-s}e^{-M_G(D_i)}] \cdot exp[-e^{-s}]e^{-s} \cdot \int_{-\infty}^{s+W_G} exp[-e^{-t}\sum_{l\neq k}^{m} e^{-W_{R_l}}] \cdot exp[-e^{-t}]e^{-t}\ dtds.
\end{aligned}
$$

The inner part of the double integral remains the same as before:

$$
\int_{-\infty}^{s+W_G} \prod_{l\neq k}^{m} exp[-e^{-t}\sum_{l\neq k}^{m} e^{-W_{R_l}}] \cdot exp[-e^{-t}]e^{-t}\ dt = \frac{exp[-e^{-(s+W_G)}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1)]}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1}.
$$

Therefore,

$$
\begin{aligned}
&P(D_iR_k) \\
=&\int_{-\infty}^{\infty} exp[-e^{-s}(\sum_{j\neq i}^{n} e^{-W_{D_j}} + e^{-M_G(D_i)} + 1)] \cdot exp[-e^{-s}e^{-W_G}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1)]e^{-s} ds \cdot \frac{1}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1} \\
=&\int_{-\infty}^{\infty} exp[-e^{-s}(\sum_{j\neq i}^{n} e^{-W_{D_j}} + e^{-M_G(D_i)} + 1 + e^{-W_G}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1))]e^{-s} ds \cdot \frac{1}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1}
\end{aligned}
$$

$$\overset{e^{-s}=v}{=}\int_0^\infty exp[-v(\sum_{j\neq i}^{n} e^{-W_{D_j}}+e^{-M_G(D_i)}+1+e^{-W_G}(\sum_{l\neq k}^{m} e^{-W_{R_l}}+1))]dv\cdot\frac{1}{\sum_{l\neq k}^{m} e^{-W_{R_l}}+1}$$
$$=\frac{1}{\sum_{j\neq i}^{n} exp(-W_{D_j})+e^{-M_G(D_i)}+1+exp(-W_G)\left[\sum_{l\neq k}^{m} exp(-W_{R_l})+1\right]}\cdot\frac{1}{\sum_{l\neq k}^{m} exp(-W_{R_l})+1}.$$

A2 Proof of Proposition 1

In order to show the existence of the election stage equilibrium, I apply results by Athey (2001), which are also used in Wasser (2013) and Ewerhart (2014) for the private information contest models with different forms of CSFs. Without loss of generality, as in the main text, I focus on a representative actual candidate from the Democratic party D_i, for $i\in\{1,\cdots,n\}$. Her expected payoff when she spends e_{D_i} is

$$\pi_{D_i}(v_{D_i}|a_{-D_i})\equiv\max_{e_{D_i}}\quad v_{D_i}\cdot\mathbb{E}_{e_{-D_i}}[CSF(e_{D_i};e_{-D_i})|v_{D_i};a_{-D_i}]-g(e_{D_i}),$$

whose second order derivative $\partial^2\pi_{D_i}/\partial v_{D_i}\partial e_{D_i}$ is $\partial\mathbb{E}_{e_{-D_i}}\left[CSF(e_{D_i};e_{-D_i})|v_{D_i};a_{-D_i}\right]/\partial e_{D_i}$. With the interchangeability of the integration and the differentiation, I next show that $\partial CSF(e_{D_i};e_{-D_i})/\partial e_{D_i}\geq 0$, in order to show that $\partial^2\pi_{D_i}/\partial v_{D_i}\partial e_{D_i}\geq 0$.

Note that $CSF(e_{D_i};e_{-D_i})=P(D_i)=\sum_{k=1}^{m}P(D_iR_k)$. Thus it suffices to show that $\partial P(D_iR_k)/\partial e_{D_i}\geq 0$ for $k\in\{1,\cdots,m\}$, where recall that

$$P(D_iR_k)=\frac{1}{\sum_{j\neq i}^{n} exp(-W_{D_j})+1+exp(-W_G)\left[\sum_{l\neq k}^{m} exp(-W_{R_l})+1\right]}\cdot\frac{1}{\sum_{l\neq k}^{m} exp(-W_{R_l})+1},$$

and recall that the differences are defined as

$$W_{D_j}\equiv M_D(D_i)-M_D(D_j),\ for\ j=1,\cdots,i-1,i+1,\cdots,n;$$
$$W_{R_l}\equiv M_R(R_k)-M_R(R_l),\ for\ l=1,\cdots,k-1,k+1,\cdots,m;$$
$$W_G\equiv M_G(D_i)-M_G(R_k).$$

I can rewrite the probability $P(D_iR_k)$ as follows:

$$P(D_iR_k)=\frac{1}{\sum_{j\neq i}^{n} exp(-W_{D_j})+1+exp(-W_G)\left[\sum_{l\neq k}^{m} exp(-W_{R_l})+1\right]}\cdot\frac{1}{\sum_{l\neq k}^{m} exp(-W_{R_l})+1}$$
$$=\frac{1}{\sum_{j\neq i}^{n} exp(-W_{D_j})+1}\cdot\frac{1}{\sum_{l\neq k}^{m} exp(-W_{R_l})+1}\cdot\frac{\sum_{j\neq i}^{n} exp(-W_{D_j})+1}{\sum_{j\neq i}^{n} exp(-W_{D_j})+1+exp(-W_G)\left[\sum_{l\neq k}^{m} exp(-W_{R_l})+1\right]}$$
$$=\frac{exp(M_D(D_i))}{\sum_{j=1}^{n} exp(M_D(D_j))}\cdot\frac{exp(M_R(R_k))}{\sum_{l=1}^{m} exp(M_R(R_l))}\cdot\frac{exp(M_G(D_i))}{exp(M_G(D_i))+exp(M_G(R_k))\frac{\sum_{l\neq k}^{m} exp(-W_{R_l})+1}{\sum_{j\neq i}^{n} exp(-W_{D_j})+1}}$$

$$=P_{pri}(D_i)\cdot P_{pri}(R_k)\cdot\frac{exp(M_G(D_i))}{exp(M_G(D_i))+exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}},$$

where $P_{pri}(D_i)\equiv exp(M_D(D_i))/\sum_{j=1}^{n}exp(M_D(D_j))$ denoting the probability of the candidate D_i winning the Democratic primary election, and $P_{pri}(R_k)\equiv exp(M_R(R_k))/\sum_{l=1}^{m}exp(M_R(R_l))$ denoting the probability of the candidate R_k winning the Republican primary election.

Therefore, the derivative $\partial P(D_iR_k)/\partial e_{D_i}$ can be written as follows via the total derivative rule:

$$\frac{\partial P(D_iR_k)}{\partial e_{D_i}}=\frac{\partial P(D_iR_k)}{\partial P_{pri}(D_i)}\frac{\partial P_{pri}(D_i)}{\partial exp(M_D(D_i))}\frac{\partial exp(M_D(D_i))}{\partial e_{D_i}}+\frac{\partial P(D_iR_k)}{\partial exp(M_G(D_i))}\frac{\partial exp(M_G(D_i))}{\partial e_{D_i}}.$$

Note that

$$\frac{\partial P(D_iR_k)}{\partial P_{pri}(D_i)}=P_{pri}(R_k)\left(\frac{exp(M_G(D_i))}{exp(M_G(D_i))+exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}}\right)^2,$$

$$\frac{\partial P_{pri}(D_i)}{\partial exp(M_D(D_i))}=\frac{\sum_{j\neq i}exp(M_D(D_j))}{\left(\sum_{j=1}^{n}exp(M_D(D_j))\right)^2},$$

$$\frac{\partial exp(M_D(D_i))}{\partial e_{D_i}}=exp(M_D(D_i))\frac{\partial M_D(D_i)}{\partial e_{D_i}},$$

$$\frac{\partial P(D_iR_k)}{\partial exp(M_G(D_i))}=P_{pri}(D_i)P_{pri}(R_k)\frac{exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}}{\left(exp(M_G(D_i))+exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}\right)^2},$$

$$\frac{\partial exp(M_G(D_i))}{\partial e_{D_i}}=exp(M_G(D_i))\frac{\partial M_G(D_i)}{\partial e_{D_i}},$$

and each of the terms above is positive, by noting that the voter's latent utility from electing a candidate is increasing with diminishing returns in the candidate's campaign spending thus $\partial M_D(D_i)/\partial e_{D_i}$ and $\partial M_G(D_i)/\partial e_{D_i}$ are both positive. Hence, the *Single crossing condition for games of incomplete information* in Athey (2001) is satisfied. And existence of an equilibrium in nondecreasing strategy is established where $e_{D_i}=s_{D_i}(v_{D_i})$.

Next I turn to the entry stage, recall that for the representative Democratic potential candidate D_i, $i\in\{1,\cdots,N\}$, the entry probability is given by $p_{D_i}=Pr(C_{D_i}<\Pi_{D_i})$, where Π_{D_i} depends on the entry probabilities of all potential candidates. Therefore, if I stack the entry probability decision rules of all potential candidates together, it forms a mapping from $[0,1]^{N+M}\to[0,1]^{N+M}$, which is continuous in the vector of all entry probabilities. A fixed point of the vector of all potential candidates follows Brouwer's fixed point theorem.

A3 Omitted Details of the Estimation Method

I derive the log-likelihood function of the non-negative campaign spending distribution for the representative election indexed by l, where the notations follow those in Section 4.1:

$$\log(L_l) = \sum_{n_{l,0}} \log\left(1-\Phi(X_l'\alpha+u_l)\right)$$
$$+\sum_{n_{l,1}} \log\Phi\left\{\frac{1}{\sqrt{1-\frac{\sigma_{12}^2}{\sigma_2^2}}}\left[X_l'\alpha+u_l+\sigma_{12}\sigma_2^{-2}\left(\log(e_{i_l,l})-Z_{i_l,l}'\beta-u_l\right)\right]\right\}$$
$$+\sum_{n_{l,1}} \phi\left(\frac{\log(e_{i_l,l})-Z_{i_l,l}'\beta-u_l}{\sigma_2}\right)-n_{l,1}\log(\sigma_2)-\sum_{n_{l,1}}\log(e_{i_l,l}),$$

where $n_{l,0}$ denotes the zero campaign spending, and $n_{l,1}$ denotes the non-zero campaign spending in election l; and $\Phi(\cdot)$ and $\phi(\cdot)$ are the CDF and PDF of the standard normal distribution. Remember that the incumbent always spends strictly positive amount in the data.

I next show how to estimate the voter model through MLE via a two-step procedure where the primary voter model and the general voter model are estimated separately. First of all, for a given election where the Democratic candidate D_i and the Republican candidate R_k compete in the general election, with D_i being the winner, the probability is given by:

$$P(D_iR_k) = P_{pri}(D_i)\cdot P_{pri}(R_k)\cdot \frac{exp(M_G(D_i))}{exp(M_G(D_i))+exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}},$$

and it is already shown in Appendix A2 that $\partial P(D_iR_k)/\partial P_{pri}(D_i)$ is positive. Further,

$$\frac{\partial P(D_iR_k)}{\partial P_{pri}(R_k)} = P_{pri}(D_i)\left\{\frac{exp(M_G(D_i))}{exp(M_G(D_i))+exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}}+\frac{exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}}{\left(exp(M_G(D_i))+exp(M_G(R_k))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}\right)^2}\right\},$$

which is also positive. Therefore, to maximize the log-likelihood function, I first pool all the primary elections together and maximize the corresponding log-likelihood function; I then use the estimated winning probabilities of the primary elections and maximize the log-likelihood function for the general elections. The log-likelihoods of the two-party primaries for the representative election are as follows:

$$\log(L_{pri,D}) = \sum_{j=1}^{n} vot_{D_j}\left(\gamma\log(1+e_{D_j})\right)-\log\left(\sum_{j=1}^{n} exp\left(\gamma\log(1+e_{D_j})\right)\right),$$
$$\log(L_{pri,R}) = \sum_{l=1}^{m} vot_{R_l}\left(\gamma\log(1+e_{R_l})\right)-\log\left(\sum_{l=1}^{m} exp\left(\gamma\log(1+e_{R_l})\right)\right),$$

and the log-likelihood of the general election is as follows:

$$
\begin{aligned}
\log(L_{gen}) =& vot_{D_i}\left[\omega\log(1+e_{D_i}) - \log\left(exp(\omega\log(1+e_{D_i})) + exp(\omega\log(1+e_{R_k}))\frac{P_{pri}(D_i)}{P_{pri}(R_k)}\right)\right] \\
&+ vot_{R_k}\left[\omega\log(1+e_{R_k}) - \log\left(exp(\omega\log(1+e_{R_k})) + exp(\omega\log(1+e_{D_i}))\frac{P_{pri}(R_k)}{P_{pri}(D_i)}\right)\right],
\end{aligned}
$$

where for this representative election, the Democratic primary log-likelihood function is $\log(L_{pri,D})$ with the winner being D_i and for each candidate D_j the vote share is vot_{D_j} and the campaign spending is e_{D_j}; and the Republican primary log-likelihood function is $\log(L_{pri,R})$ with the winner being R_k and for each candidate R_l the vote share is vot_{R_l} and the campaign spending is e_{R_l}. For the general election, $\log(L_{gen})$ denotes the log-likelihood, where vot_{D_i} and vot_{R_k} represent the vote shares of the general election candidates. Note that in the general election's log-likelihood function, $P_{pri}(R_k)$ and $P_{pri}(D_i)$ can be estimated through MLE of the primary elections.

As for the entry stage, I need to compute the equilibrium entry probabilities that are fixed points determined by $p_{D_i} = Pr(C_{D_i} < \Pi_{D_i})$ for a representative potential candidate D_i from the Democratic party, for example. The estimation of the entry cost distribution thus includes two loops: the inner loop that uses fixed point finder to solve for the equilibrium entry probabilities, and the outer loop that uses the nonlinear least squares regression (NLS) to estimate the entry model parameters, both relying on the entry equilibrium decision rule. Since this estimation method is computationally intensive, I follow Li and Zhang (2015) and change the loop order. For the initial values of the equilibrium entry probabilities, I adopt a reduced-form probit model; I then estimate the parameters given the entry probabilities, and update the entry probabilities via fixed point finder. With the new entry probabilities, I estimate the parameters again. I repeat the above procedure until the estimates of parameters and the equilibrium entry probabilities converge.

Lastly, for the inference, I adopt a clustered bootstrap method at the election-level following Marmer and Shneyerov (2012), in order to correct for the multi-step estimation procedure.

A4 First Stage Results

In this Appendix, I report the estimated results from the first stage regression of the IV, i.e., *log(lagged spending)* from the *previous* Senate election in the same state, on explanatory variables, which is modeled as a linear function $X_l^{IV\prime}\eta + u_l$. The main purpose of this estimation stage is to recover the pseudo unobserved heterogeneity at the election level, and the estimated coefficients $\hat{\eta}$ are presented in Table A.1 to demonstrate how the election-specific variables affect the log of lagged spending from the previous Senate election.

Table A.1: First stage IV regression results

	Estimate	Sd. error
OPEN	−0.0955	0.117
PVI	−0.0080	0.007
GOV	0.0053	0.094
PRCP	0.2221	0.199
SNOW	−0.5549	0.215
PRE	0.1320	0.094
YOUNG-PER	1.3612	1.899
MID-PER	2.0341	1.243
OLD-PER	3.6003	2.024
UNEMP	0.0301	0.033
LOGINC	0.3982	0.408
PNCAN	0.0433	0.011
CONST	9.4789	4.371

Notes: Estimated coefficients in the corresponding specification are reported. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

Overall, there are some explanatory variables that exhibit significant impacts on the IV. Particularly, the variable SNOW, constructed from averaging all reported snowfall from all stations within one state on the general election day, has a significantly positive effect on the lagged campaign spending. This is probably because the snowfall can reflect some geographic component related to political reason that causes the campaign spending to be consistently lower. For instance, states that generally snow in the beginning of November are Alaska, Colorado, Montana, Nebraska, and Wyoming without metropolitans such as Houston, Chicago, Seattle, Boston, and Phoenix around which larger donors are concentrated. Therefore, the lack of high contributions induces smaller campaign spending in these states. The second reason may be that snowfall has annual dependence, meaning that the recurrence of snow on the same day is positively correlated from one year to another. And bad weather can result in a smaller votes cast in the general election, thus reducing the campaign spending by candidates.

Another significant coefficient is on the percentage of the population aged 65 and up (OLD-PER) that is positive. The reason can be that in the U.S., the oldest citizens are the most likely to cast their ballots. Thus this group of voters have political clout beyond their numbers alone. The number of potential candidates (PNCAN) also has a significantly positive effect on the lagged spending. This can be due to the same underlying political climate that makes both the previous and the current races to have a larger pool of potential candidates, which causes the campaign spending in the previous race to be lower. Lastly, the constant term is also significant.[42]

A5 Sensitivity Analyses

In this Appendix section, I conduct a series of sensitivity analyses in order to justify the assumptions implicitly made in the model and the structural estimation, and to show the robustness of the structural results across different hypotheses.

A5.1 Turnout in General Elections

In the general election, a representative voter can choose to vote or not, depending on the latent utility ι_{G_0} obtained through the outside good, distributed as a type-1 extreme-value distribution (see Appendix A1). As in Section 3.1, I focus on the first actual candidate in the Democratic primary D_i, and still use $P(D_iR_k)$ to represent the probability that this candidate D_i wins the Democratic primary, as well as the general election with the candidate R_k as the general election opponent. By the derivation in Appendix A1, $P(D_iR_k)$ has the following form:

$$P(D_iR_k) = \frac{1}{\sum_{j\neq i}^{n} exp(-W_{D_j}) + e^{-M_G(D_i)} + 1 + exp(-W_G)\left[\sum_{l\neq k}^{m} exp(-W_{R_l}) + 1\right]} \cdot \frac{1}{\sum_{l\neq k}^{m} exp(-W_{R_l}) + 1},$$

where for ease of illustration, I still use the notations in Section 3.1. Therefore, due to the existence of the term $exp(-M_G(D_i))$ in the denominator of the first ratio above, I can identify the coefficient vector ω_X, along with ω, in the following specification of Section 4.1

$$u_{i_l,l}^{G} = \omega \log(1 + e_{i_l,l}) + X_l^{G'} \omega_X + u_l + \iota_{i_l,l}, \quad i_l = 1, \cdots, n_l,$$

42 Since I use linear regression in this estimation stage, one may wonder the model fit result. However, when Matzkin (2003) provides the nonparametric identification of the distribution of the unobserved variable, which is u_l in the model, she does not impose parametric assumption on this distribution. Hence, I follow Matzkin (2003) and do not parameterize the distribution of u_l. If, in any sense it is interesting, I assume that u_l follows a normal distribution with zero mean, the usual F-test can be applied, with a estimated F-statistic being 9.30, implying a relatively good model fit (see Staiger and Stock (1997)).

for a generic election l with n_l actual candidates. In the data, the voter turnout data is calculated using the ratio of the total votes cast in the general Senate election to the VEP of a certain state in a certain year. The results are contained in Table A.2 as follows. Since the selection equation and the log of spending equation are the same as in the benchmark model, I only report the results of this sensitivity analysis for the voter and entry models. The estimation results for the entry model are affected by considering the voter turnout in general elections, because this influences the simulated expected payoffs of all elections.

The result of the primary part in the voter model is the same as in Table 3. Once the voter turnout is taken into account, though the effect of campaign spending on votes in the general election is still positive, it is not as prominent as in the benchmark model. As for the estimated coefficients in the general latent utility, most of them have expected signs, but are insignificant. Being an open-seat election reduces the latent utility for a representative voter in the general election, which can be justified by the incumbency advantage, because incumbents already build the reputation among voters after holding the office for a period. The number of actual candidates has a negative effect on the latent utility of general election voters, related to the structural result in the benchmark model that a more competitive election reduces the campaign spending made by candidates. This can indicate that a more competitive election can reduce the latent utility of voters, due to the reason that voters may have a hard time figuring out who they should select. Therefore, candidates in a more competitive election receive less expected payoffs, and spend less in campaigns consequently. The only significant coefficient is on the dummy variable PRE which equals 1 if a presidential election occurs on the same Senate general election day. This dummy variable has a significantly positive effect on the latent utility of voters, thus a significantly negative effect on the voter turnout. If a presidential election occurs in a year, it can spur additional turnout for the Senate general election in the same year.

Although the expected payoffs of candidates are now simulated through the voter model with the voter turnout in the general elections, the estimation results of the entry model remain stable.

A5.2 Truncated Spending Distribution

In this sensitivity analysis, I assume that the campaign spending is distributed as a log-normal truncated from above when the spending is strictly positive. Since the model is conditional on the election-level heterogeneity (X,u), the resulting distribution of the spending will also condition on (X,u), so is the upper boundary. Therefore, I need to estimate the upper boundary of non-zero spending

Table A.2: Sensitivity analysis: turnout in general elections

	Voter		Entry	
	Estimate	Sd. error	Estimate	Sd. error
γ	0.1631	0.005		
ω	0.0200	0.038		
λ_I			0.4391	0.021
λ_C			0.5278	0.013
INCUM			0.3627	0.112
OPEN	−0.1421	0.149		
PVI	−0.0031	0.020	0.0008	0.009
GOV	0.0869	0.133	0.0602	0.132
YOUNG-PER	−1.4270	2.275	−4.0241	2.658
MID-PER	0.3320	1.753	2.4294	1.982
OLD-PER	0.7019	2.574	−5.6570	3.073
UNEMP	−0.0384	0.057	−0.0372	0.057
LOGINC	0.1962	0.520	0.1645	0.528
PNCAN			0.0027	0.015
ANCAN	−0.0477	0.056		
PRCP	0.0759	0.279		
SNOW	0.1217	0.217		
PRE	0.6063	0.298		
CONST	−3.2340	5.609	0.4723	5.803

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

varying with the election-level heterogeneity.

The upper boundary of non-zero spending is estimated nonparametrically following Guerre, Perrigne, and Vuong (2000). Consider the vector of election-level covariates $W \subset X$ and $W \in \mathbb{R}^d$ with the support $\mathscr{W} = [\underline{w}, \bar{w}]$ (which is assumed to be known or can be readily estimated), I partition $\mathscr{W}$ to k_d bins $\{\mathscr{W}_k : k = 1, \cdots, k_d\}$ of equal length $\Delta_d \propto (logL/L)^{1/(d+1)}$. The estimate of the upper boundary of non-zero spending is the maximum of all non-zero spending whose corresponding realization of w belongs to $\mathscr{W}_k$. In the data, I specify W to include two continuous and unbounded covariates: the normalized unemployment rate and the log of median household income. Since the voter model is the same as in the benchmark model, I only report the results of estimating the truncated spending distribution and the entry model, which are contained in Table A.3.

Table A.3: Sensitivity analysis: truncated spending distribution

	Spending: selection equation		Spending: log(spending) equation		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_2	2.8189	0.073				
σ_{12}	1.5073	0.225				
λ_I					0.4391	0.023
λ_C					0.5278	0.013
INCUM			7.7142	0.174	0.3628	0.115
OPEN	0.4092	0.149	1.5132	0.197		
PVI	−0.0102	0.008	−0.0226	0.012	0.0010	0.010
GOV	−0.1775	0.128	−0.0820	0.144	0.0604	0.137
YOUNG-PER	0.8286	2.575	−1.5245	2.817	−4.0240	2.906
MID-PER	−0.8661	1.832	1.4636	2.409	2.4294	2.135
OLD-PER	0.1442	2.737	3.4832	3.453	−5.6571	3.214
UNEMP	−0.0017	0.048	−0.0409	0.059	−0.0370	0.060
LOGINC	0.4745	0.559	1.8320	0.688	0.1647	0.546
PNCAN	0.0437	0.031	0.2330	0.035	0.0031	0.015
ANCAN	−0.0955	0.034	−0.4371	0.041		
CONST	−4.3754	5.976	−7.7340	7.389	0.4725	5.908

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

Most of the results remain consistent with those of the benchmark model. The estimated standard deviation of the error in the log of spending equation is now slightly larger: 2.82, and the estimated covariance of the two errors in the spending model is also slightly larger: 1.51, resulting in a correlation being about 0.54, which is close to the estimate of 0.58 in the benchmark model. The estimated coefficients in the selection equation have the same signs and similar magnitudes as in the benchmark model shown in Table 3, in most cases. The unemployment rate now has a negative effect on the decision of making non-zero spending, which is consistent with the positive effect of the log of median household income, although this effect is still insignificant. With some small discrepancies, the estimates of the entry model remain stable in this sensitivity analysis, compared with the benchmark model.

A5.3 Difference Between Political Parties

I do not take into account the political party difference in the benchmark model, which I aim to capture in this subsection by including the interaction terms with the party affiliation dummy that equals 1 for the Democratic party, and 0 for the Republican party.

Due to the relatively large number of total parameters needed to be estimated, I focus on the difference between political parties among challenger candidates. Therefore, when analyzing the spending model and the primary voter model, I categorize the data into four different groups: incumbents (I), the Democratic challengers (DEM), the Republican challengers (REP), and candidates in open-seat elections (OPEN).[43] For the general voter model, I consider three groups: incumbents (I), challengers (C), and open-seat candidates (OPEN). Lastly when estimating the entry model, since it is not revealed yet whether the election is open-seat, I differentiate among three groups: incumbents (I), the Democratic challengers (DEM), and the Republican challengers (REP). For the corresponding linear specification parts in the model, I add the interaction terms of these group indexes with the intercept by adding an indicator being 1 for the Democratic challenger (DEM-IDX), and I also contain the interaction terms of this indicator with the continuous covariates (YOUNG-PER, MID-PER, OLD-PER, UNEMP, and LOGINC).

Table A.4 presents the results of this sensitivity analysis of taking into consideration the difference between political parties. First of all, the Democratic party challenger index and the interaction terms of it with the continuous covariates yield insignificant coefficients in the selection equation, the log of campaign spending equation, and the mean equation of entry cost distribution. Therefore, I do not find strong evidence of the difference between political parties regarding the effects of the election-level

43 I do not consider the group-specific standard deviation for the spending model, due to the computative infeasibility that the optimization algorithm cannot converge.

Table A.4: Sensitivity analysis: difference between political parties

	Spending: selection equation		Spending: log(spending) equation		Voter		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_2	2.4719	0.072						
σ_{12}	1.4641	0.220						
γ_I					0.1733	0.009		
γ_{DEM}					0.1180	0.013		
γ_{REP}					0.1372	0.010		
γ_{OPEN}					0.2261	0.017		
ω_I					0.1562	0.060		
ω_C					0.1151	0.063		
ω_{OPEN}					0.3800	0.066		
λ_I							0.4387	0.023
λ_{DEM}							0.5160	0.013
λ_{REP}							0.5149	0.015
INCUM			4.0656	0.173			0.3799	0.126
OPEN	0.3692	0.170	1.2750	0.213				
DEM-IDX	10.2573	8.710	−5.1792	12.164			12.8429	7.561
PVI	−0.0083	0.009	−0.0093	0.011			0.0013	0.009
GOV	−0.1718	0.135	−0.1179	0.152			0.0679	0.143
YOUNG-PER	0.9169	3.011	0.7058	3.444			−6.8047	3.114
MID-PER	−1.9044	1.860	−0.5647	2.315			2.3815	2.318
OLD-PER	2.0601	2.909	5.4751	3.770			−5.3437	3.612
UNEMP	0.0215	0.051	0.0059	0.057			−0.0045	0.066
LOGINC	0.7082	0.669	1.5593	0.694			0.6999	0.602
YOUNG-PER × DEM-IDX	0.2618	4.373	−8.7650	6.653			7.5699	4.230
MID-PER × DEM-IDX	3.9263	3.437	1.7336	6.337			0.3565	2.296
OLD-PER × DEM-IDX	−8.1985	4.668	−2.7353	8.406			−1.3540	4.258
UNEMP × DEM-IDX	−0.0978	0.072	−0.2090	0.126			−0.0958	0.053
LOGINC × DEM-IDX	−0.8864	0.802	0.8088	1.126			−1.3239	0.711
PNCAN	0.0331	0.031	0.1279	0.037			0.0038	0.015
ANCAN	−0.0788	0.033	−0.2786	0.041				
CONST	−7.1056	7.212	−5.7461	7.519			−4.7797	6.643

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

covariates on the candidates' decisions of making non-zero spending and how much they prefer to make through campaigns, as well as the candidates' entry cost expectations. Secondly, the estimated coefficients of the election-level covariates in the equations have the same signs and similar magnitudes as in the benchmark model shown in Table 3.

The difference between political parties are reflected in the estimates of the voter model. In the primary voter model, the incumbents have a relatively higher return of campaign spending than challengers in the incumbent-challenger elections, in terms that their spending can translate to more votes in the primary election. The challengers have different returns of campaign spending: for the Democratic challenger, the estimated coefficient on the spending term in the vote equation is around 0.12, while that coefficient is about 0.14 for the Republican challenger. However, this difference between parties is not very big. Overall, the campaign spending can translate to more votes in open-seat elections, compared to incumbent-challenger elections, which may be due to the reason that in open-seat elections, voters cannot evaluate the candidates' valence depending on their performance as the incumbent, thus they tend to rely more on the spending the candidates made through campaigns. This discrepancy between open-seat and incumbent-challenger elections is more prominent in the general election. The estimates coefficient on the spending term is about 0.38 for open-seat general elections, while I estimate the coefficients to be 0.16 and 0.12 for incumbents and challengers respectively for incumbent-challenger elections. The latter estimates are close to those in the primary voter model, indicating that the different return of campaign spending between primary and general elections found in the benchmark model is largely driven by that difference in open-seat elections, in the sense that the campaign spending can translate to more votes in the general election than in the primary election, given the same amount of spending.

A5.4 Quadratic Cost Function

In the benchmark model, I consider a linear cost function of the campaign spending $g(\cdot)$ such that $g(e) = e$. In this sensitivity analysis, I alternatively specify a quadratic cost function that is parameterized as $g(e) = \kappa e^2 + \zeta e$, where the constant term is excluded since I use the first-order condition 3.3 which only contains the derivative of the cost function in the estimation.

Following Campo, Guerre, Perrigne, and Vuong (2011), I use the equality of Equation 3.3, through assuming that one quantile of the private value distribution is constant and unknown. Since the strictly positive spending is unbounded from above, without loss of generality, I consider a constant and unknown median of the private value distribution. This median is constant, meaning that it is unrelated to the election-level heterogeneity (X,u), i.e., $Med(v|X,u) = Med(v)$; and this median is unknown, because I treat it as a parameter which can be identified and estimated along with the parameters in the cost function (κ, ζ).

In the data, I observe that the incumbents will always spend non-zero through campaigns once they decide to run for office. Therefore in equilibrium, the first-order condition always holds as an equality for the incumbents. I rely on the subset of incumbents' campaign spending to estimate $(\kappa, \zeta, Med(v))$ through NLS. For incumbents, the expectation in Equation 3.3 is simulated by 500 repetitions regarding the spending profile of their opponents. Since the spending model and the voter model are the same as the benchmark model, I report the results for the cost function and the entry model, as follows.

Table A.5: Sensitivity analysis: quadratic cost function

	Voter		Entry	
	Estimate	Sd. error	Estimate	Sd. error
κ	0.0000	0.000		
ζ	0.9545	0.327		
$Med(v)$	1819442	626664		
λ_I			0.4391	0.021
λ_C			0.5278	0.013
INCUM			0.3626	0.109
PVI			0.0006	0.010
GOV			0.0603	0.131
YOUNG-PER			−4.0233	2.733
MID-PER			2.4308	2.023
OLD-PER			−5.6554	3.000
UNEMP			−0.0375	0.058
LOGINC			0.1639	0.551
PNCAN			0.0000	0.017
CONST			0.4723	6.026

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

Table A.5 presents the estimates of the sensitivity analysis. The estimated constant $Med(v)$ is around 18 million, which is significant with an estimated standard error being about 6 million. This figure is larger than the mean of non-zero campaign spending shown in Table 1. Interestingly, although the cost function is assumed to be quadratic, the results show that the estimated cost function is linear, with $\hat{\zeta} = 0.95$, thus leading to a linear function very similar to the one considered in the benchmark model. Because the estimated cost function does not change much, the resulting estimates for the entry model are similar to those in the benchmark model. The exception is the effect of the number of potential candidates, which is estimated to be a tiny positive number that is essentially zero.

A5.5 Alternative Voter Model Specification

In this sensitivity analysis, I consider an alternative voter model specification where the campaign spending affects voters' latent utility linearly, for a generic election index by l, and a representative candidate indexed by i_l, as follows:

$$u^P_{i_l,l} = \gamma e_{i_l,l} + X_l^{P'} \gamma_X + u_l + \iota_{i_l,l}, \quad i_l = 1, \cdots, n_l,$$

for the primary election, and

$$u^G_{i_l,l} = \omega e_{i_l,l} + X_l^{G'} \omega_X + u_l + \iota_{i_l,l}, \quad i_l = 1, \cdots, n_l,$$

given that this candidate i_l wins the primary election and enters into the general election.

Since the spending model is unchanged, I only report the results from estimating the alternative voter model and the entry model in the following table.

Table A.6 gives the results in this sensitivity analysis. For the voter model, in both the primary and general elections, the campaign spending made by the candidates affects the latent utility of the voters positively and significantly. Unlike the result in the benchmark model, when the campaign spending enters voters' latent utility linearly, the effect of spending is larger in the primary election, almost doubling that in the general election. This is consistent with the benchmark implication and can be due to the reason that the campaign spending made by the candidates who participate in the general elections usually spend much more than the rest of the candidates. Thus when the latent utility is assumed to be a linear function where the spending dose not exhibit diminishing marginal effect, the marginal effect of the spending on the latent utility of voters is averaged out. Recall that in the benchmark model where the effect of campaign spending in the voter model is assumed to take the log form, the more spending candidates make, the smaller the marginal effect of spending entails. It

Table A.6: Sensitivity analysis: alternative voter model specification

	Voter		Entry	
	Estimate	Sd. error	Estimate	Sd. error
γ	0.1393	0.012		
ω	0.0726	0.011		
λ_I			0.4834	0.029
λ_C			0.5716	0.016
INCUM			0.3392	0.112
PVI			−0.0202	0.013
GOV			0.0301	0.134
YOUNG-PER			−4.0509	2.727
MID-PER			2.4033	2.030
OLD-PER			−5.6832	3.210
UNEMP			−0.0513	0.058
LOGINC			0.1572	0.510
PNCAN			0.0008	0.016
CONST			0.4403	5.615

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 500 nonparametric bootstrap at the election level. The estimates and Sd. errors of γ and ω are scaled up by 10^6.

is important to compare the roles of the spending in the primary and general voter models given the same level of the campaign spending. The results from estimating the entry model are consistent with those of the benchmark model shown in Table 3. In this sensitivity analysis, the estimated standard deviations of the entry cost distributions for the incumbents and challengers are slightly larger than those in the benchmark model. The estimated coefficients on the explanatory variables have the same signs and similar magnitude as in the benchmark model.

A5.6 Dynamic Change

In the structural analysis of the benchmark model, I maintain an implicit assumption that there is no time trend regarding the private value of actual candidates and the entry cost of potential candidates. However, since the data covers a relatively long time span from 1994 to 2018, it is more natural and realistic to take into consideration the possible dynamic change of the private information possessed by candidates. As a result, I introduce time trend into the benchmark analysis in this exercise. Specifically, I allow for time-specific effects in the specifications of the campaign spending distribution and the entry cost distribution for candidates, which control for time fixed effects in the estimation. The voter model remains the same as in the benchmark model, because I want to focus on the dynamic discussion on candidates' private information.[44]

I present the estimated results from the model with time fixed effects in Table A.7. The estimates of the time-specific constants are not reported and will be discussed later in Figure 9. Note that since the voter model is unchanged, the results only contain the campaign spending model with the selection equation and the level equation, and the entry cost model.

As shown in Table A.7, most of the estimates are consistent with those in Table 3. The estimated standard deviation of the error in the log of spending equation, σ_2, and the estimated covariance of the two errors in the selection equation and in the level equation of campaign spending, σ_{12}, are similar to those derived from estimating the benchmark model. Furthermore, in the specifications of campaign spending and entry cost distributions, the estimated coefficients on the covariates have expected signs and magnitudes. Compared with challengers, incumbents tend to spend non-zero and larger amount through campaigns in elections. Candidates in an open-seat election are more likely to

44 As a sensitivity analysis in Appendix A5.7, instead of using time-specific constants in the specifications, I also estimate an alternative model where a dummy variable being one for the period of post-2008 and zero for that of pre-2008 in order to capture the effect of time. The voter model estimates show no significant difference on how voters appreciate the campaign spending made by candidates before and after 2008, for both primary and general elections.

Table A.7: Estimation results with time fixed effects

	Spending: selection equation		Spending: log(spending) equation		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_2	2.4377	0.074				
σ_{12}	1.3562	0.237				
λ_I					0.4406	0.023
λ_C					0.5232	0.013
INCUM			3.9989	0.174	0.3484	0.114
OPEN	0.3174	0.172	1.2327	0.218		
PVI	−0.0021	0.010	−0.0003	0.013	0.0131	0.011
GOV	−0.2216	0.137	−0.0540	0.162	0.1302	0.158
YOUNG-PER	2.9662	5.847	3.1465	6.285	−5.1482	6.215
MID-PER	0.2007	5.820	6.0058	6.047	5.8030	6.420
OLD-PER	4.9271	4.817	8.7122	5.321	−2.1961	4.992
UNEMP	−0.0759	0.072	−0.0088	0.083	−0.0190	0.087
LOGINC	0.7587	0.614	1.9367	0.703	0.7624	0.522
PNCAN	0.0259	0.033	0.1263	0.038	0.0034	0.015
ANCAN	−0.0661	0.036	−0.2752	0.040		
CONST	−8.4307	7.698	−12.3870	8.537	−6.9235	6.446

Notes: Time fixed effects are included in the selection equation of campaign spending, the level equation of campaign spending, and the entry cost distribution. For explanatory variables, estimated coefficients in corresponding specifications are reported. The estimated results are obtained through the estimation method described in Section 4.2. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

have larger probability of making non-zero spending and higher level of such spending by candidates. The numbers of potential candidates and actual candidates have similar effects on the choice and level of campaign spending by candidates, when comparing with the benchmark model results. The number of potential candidates has a positive effect and the number of actual candidates has a negative effect, where in the selection model of campaign spending the effects are not very significant while in the log(spending) model the effects are significant. Turning to the entry model, incumbents have relatively higher entry costs, when comparing to challengers. The number of potential candidates has a positive effect on the entry cost on average, which is again consistent with the implication of the benchmark model.

In the following figure I report the estimated time-specific constants in the three models: the selection model of campaign spending, the log of non-zero campaign spending, and the entry cost model. The year of 1994 serves as a benchmark and thus is excluded.

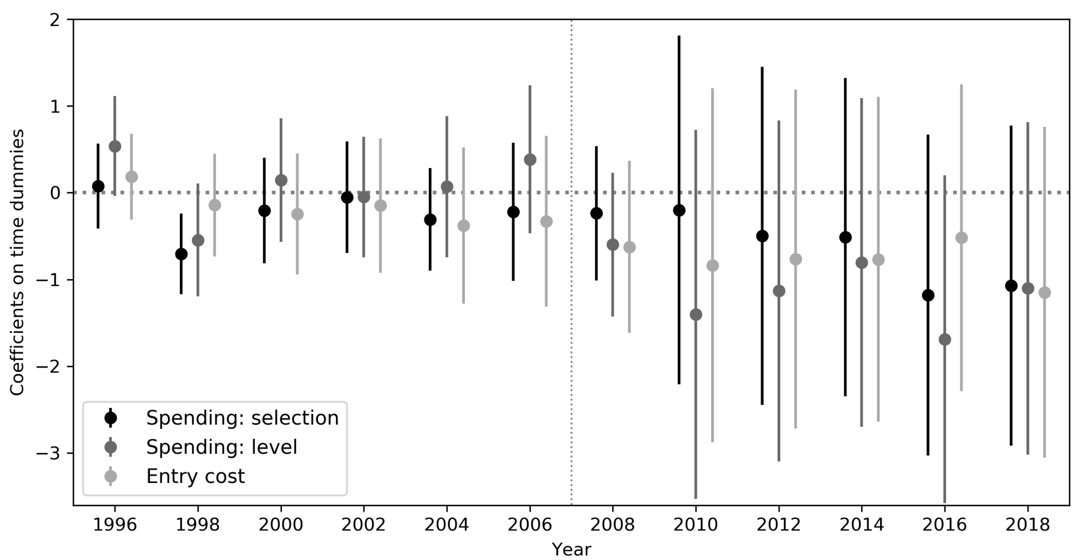


Figure 9: Coefficients on time dummies in multiple regressions with 90% confidence intervals

As reflected in Figure 9, most of the estimated time-specific effects are negative. If I divide the sample into two periods: pre-2008 and post-2008, the results before 2008 are mixed while those after 2008 are consistently negative with wider confidence intervals. Although all but one of the estimates are insignificant at the 90% significance level, the trend itself implies the devaluation of candidates for the Senate office across time, if any, because campaign spending is a strictly increasing function of the private value of holding the office for actual candidates. The entry costs of potential candidates also fall over the same period. Since the entry costs of running for office and the valuations of holding the office both decrease over time, it is hard to determine whether the net benefits of the Senate office decrease as well. The estimates correspond to an average entry probability of 0.8050 before 2008, and an average entry probability of 0.7838 after 2008. Therefore, with the falling entry costs of running for office and the falling valuations of holding office, the net benefits of office have a minuscule drop over the years, leading to slightly decreasing entry probabilities of potential candidates.

One reason why it may be worth discussing the time trend is that the decrease of the net benefits of office for a period of time can form possible driving force of political polarization in terms of candidates' ideology, according to Hall (2019).[45] McCarty, Poole, and Rosenthal (2006) and Voorheis, McCarty, and Shor (2015) provide evidence for the sharp growth in legislative polarization in the U.S. since the 1970s and link this polarization to the rising income inequality simultaneously, which explains the polarization phenomenon from the demand-side: the changing ideological preferences of voters. On the other side, if I assume that candidates cannot change their political positions easily throughout the whole election process, this rigidity of candidate positions may provide new perspective on why the polarization happens from the angle of supply-side, under the assumption that candidates also care about ideological benefits of office.[46] Hall (2019) proposes a theory showing that when the non-ideological net benefits of office decrease, the set of potential candidates willing to run for office becomes more ideologically extreme. The intuition is that keeping the non-ideological net benefits of office equal for both moderate and extreme candidates, the ideological cost of not running for extreme candidates will be larger than that for moderate candidates because of the Median Voter Theorem. Thus, when the non-ideological net benefits of office fall, the marginal extreme candidates will choose to run for office while the marginal moderate candidates will choose not to run, which drives the pool of actual candidates more extreme. The time trend estimated from the model in this part shows the potential decrease in the non-ideological net benefits of Senate office from 1994 to 2018, if any. Therefore, the results in this part can be viewed as weak evidence for the ideological polarization of political candidates over this period of time (see Hirano, Snyder, Ansolabehere, and Hansen (2010)

45 In this paper I do not consider candidates' choices of ideological position as equilibrium results. This is partially because due to the Median Voter Theorem, candidates will eventually move to the middle in response to electoral competition, under the assumption that candidates can flip-flop their political policies or opinions after they proceed into the general election stage and that there is a lack of commitment regarding candidates' ideological positions.

46 The reasons behind the rigidity of candidate positions can be due to the strong personal views of candidates, see Aldrich and Rohde (2001) and Cox and McCubbins (2007), and/or due to the fact that voters may punish candidates for changing positions, see Tomz and Van Houweling (2012) and Debacker (2015).

and Theriault and Rohde (2011)).

A5.7 The Great-Recession Effect

In Section A5.6, I introduce the full set of time dummies into the benchmark model so that the dynamic change of the results can be taken into account. In this sensitivity analysis, I conduct an alternative exercise adding a dummy variable that is assigned one for the period of post-2008 and zero for that of pre-2008 to the specifications in the model. Figure 9 in Section A5.6 shows a persistent decrease of the propensity to spend through campaigns, the amount of positive campaign spending, and the entry cost of candidates after 2008, which may be due to the Great-Recession in 2008. Therefore, by introducing a dummy variable to indicate the period of post-2008, it may form additional evidence for the falling of the benefit of office by comparing the difference before and after 2008. Further, in this exercise, I allow the effects of campaign spending on voters to differ before and after 2008 in primary and general elections.

Table A.8: Sensitivity analysis: difference over time

	Spending: selection equation		Spending: log(spending) equation		Voter		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_2	2.4774	0.074						
σ_{12}	1.4431	0.226						
$\gamma_{pre-2008}$					0.1576	0.010		
$\gamma_{post-2008}$					0.1694	0.010		
$\omega_{pre-2008}$					0.1969	0.028		
$\omega_{post-2008}$					0.2454	0.038		
λ_I							0.4396	0.023
λ_C							0.5274	0.013
INCUM			4.0648	0.174			0.3630	0.112
OPEN	0.3942	0.151	1.3058	0.192				
PVI	−0.0069	0.009	−0.0070	0.012			0.0024	0.009
GOV	−0.1761	0.133	−0.0880	0.154			0.0582	0.142
YOUNG-PER	1.4943	5.305	2.0650	5.841			−3.7232	5.554
MID-PER	−0.5449	4.138	3.2381	4.571			2.3272	4.140
OLD-PER	0.7460	4.442	7.4776	5.319			−5.3445	4.438
UNEMP	0.0024	0.048	−0.0291	0.055			−0.0327	0.059
LOGINC	0.47099	0.555	1.7827	0.663			0.1875	0.521
PNCAN	0.0321	0.030	0.1371	0.036			0.0035	0.014
ANCAN	−0.0780	0.034	−0.2903	0.040				
CONST	−4.7339	6.767	−9.5297	7.990			0.0891	6.302
POST-2008	−0.1320	0.905	−0.7574	1.007			−0.0335	0.847

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. The variable "POST-2008" equals one for post-2008 period, and zero for pre-2008 period. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

I present the results of this sensitivity analysis in Table A.8. Most of the estimates have the same signs and similar magnitudes as those in the benchmark model shown in Table 3. Since I care about the effect of the Great-Recession in 2008, I focus on the coefficients on the variable "POST-2008" in the corresponding specifications for the spending and entry models, as well as the estimated parameters in the voter model. As in Section A5.6, the dummy variable indicating the period of post-2008 has a negative coefficient in both the selection equation and the level equation of the spending model, implying that after 2008 actual candidates are less likely to spend non-zero in campaigns, and make smaller amount of campaign spending. This dummy variable also has a negative coefficient in the entry cost specification, showing a possible decline of entry costs after 2008. However, these effects are not significant, as shown in Figure 9. For voter model, I observe a difference regarding the estimated effects of campaign spending on voters' utility before and after 2008, and this difference is more pronounced in general elections (but the difference is still not big). Voters tend to value how much candidates spend in the general election more after 2008. This can explain why I find null effect of time dummies in Section A5.6 and here. The devaluing of office may exist over time among candidates that should result in the decrease of the campaign spending in equilibrium, but because voters value the spending made by candidates through campaigns more over time, therefore in the end two conflicting effects on the campaign spending cancel out with each other and lead to no significant change of the campaign spending across time.

References

Abramson, P. R., Aldrich, J. H., and Rohde, D. W. (1987). Progressive Ambition Among United States Senators: 1972–1988. *Journal of Politics*, *49*(1), 3–35.

Acharya, A., Grillo, E., Sugaya, T., and Turkel, E. (2019). Dynamic Campaign Spending. *Working Paper*, Stanford University, Collegio Carlo Alberto, Stanford University, and Stanford University.

Adams, J., and Merrill, S. (2008). Candidate and Party Strategies in Two-Stage Elections Beginning with a Primary. *American Journal of Political Science*, *52*(2), 344–359.

Aldrich, J. H., and Rohde, D. W. (2001). The Logic of Conditional Party Government: Revisiting the Electoral Connection. In *Congress Reconsidered*, (pp. 269–92). CQ Press.

Amorós, P., Puy, M. S., and Martínez, R. (2016). Closed Primaries Versus Top-Two Primaries. *Public Choice*, *167*, 21–35.

Athey, S. (2001). Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information. *Econometrica*, *69*(4), 861–889.

Athey, S., Levin, J., and Seira, E. (2011). Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions. *Quarterly Journal of Economics*, *126*(1), 207–257.

Avis, E., Ferraz, C., Finan, F., and Varjão, C. (2017). Money and Politics: The Effects of Campaign Spending Limits on Political Competition and Incumbency Advantage. *Working Paper*, HEC Montréal, University of British Columbia and PUC-Rio, UC Berkeley, and Stanford University.

Baron, D. P. (1989). Service-Induced Campaign Contributions and the Electoral Equilibrium. *Quarterly Journal of Economics*, *104*(1), 45–72.

Beck, J. H., and Henrickson, K. E. (2013). The Effect of the Top Two Primary on the Number of Primary Candidates. *Social Science Quarterly*, *94*(3), 1–18.

Bullock, W., and Clinton, J. D. (2011). More a Molehill than a Mountain: The Effects of the Blanket Primary on Elected Officials' Behavior from California. *Journal of Politics*, *73*(3), 1–16.

Campo, S., Guerre, E., Perrigne, I., and Vuong, Q. (2011). Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders. *Review of Economic Studies*, *78*(1), 112–147.

Copeland, G. W. (1989). Choosing to Run: Why House Members Seek Election to the Senate. *Legislative Studies Quarterly*, *14*(4), 549–565.

Cox, G. W., and McCubbins, M. D. (2007). *Legislative Leviathan: Party Government in the House*. Cambridge University Press.

Debacker, J. M. (2015). Flip-Flopping: Ideological Adjustment Costs in the United States Senate. *Economic Inquiry*, *53*(1), 108–128.

Deltas, G., Herrera, H., and Polborn, M. K. (2016). Learning and Coordination in the Presidential Primary System. *Review of Economic Studies*, *83*(4), 1544–1578.

Diermeier, D., Eraslan, H., and Merlo, A. (2003). A Structural Model of Government Formation. *Econometrica*, *71*(1), 27–70.

Diermeier, D., Keane, M., and Merlo, A. (2005). A Political Economy Model of Congressional Careers. *American Economic Review*, *95*(1), 347–373.

Erikson, R. S., and Palfrey, T. R. (1998). Campaign Spending and Incumbency: An Alternative Simultaneous Equations Approach. *Journal of Politics*, *60*(2), 355–73.

Erikson, R. S., and Palfrey, T. R. (2000). Equilibria in Campaign Spending Games: Theory and Data. *American Political Science Review*, *94*(3), 595–609.

Ewerhart, C. (2014). Unique Equilibrium in Rent-Seeking Contests with a Continuum of Types. *Economics Letters*, *125*(1), 115–118.

Fang, H. (2002). Lottery Versus All-Pay Auction Models of Lobbying. *Public Choice*, *112*, 351–371.

Fu, Q., Jiao, Q., and Lu, J. (2015). Contests with Endogenous Entry. *International Journal of Game Theory*, *44*(2), 387–424.

Gentry, M. L., and Li, T. (2014). Identification in Auctions with Selective Entry. *Econometrica*, *82*(1), 315–344.

Gerber, A. (1998). Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables. *American Political Science Review*, *92*(2), 401–411.

Gordon, B. R., and Hartmann, W. R. (2016). Advertising Competition in Presidential Elections. *Quantitative Marketing and Economics*, *14*(1), 1–40.

Green, D. P., and Krasno, J. S. (1988). Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections. *American Journal of Political Science*, *32*(4), 884–907.

Grossman, G. M., and Helpman, E. (1996). Electoral Competition and Special Interest Politics. *Review of Economic Studies*, *63*(2), 265–286.

Gu, Y., Hehenkamp, B., and Leininger, W. (2019). Evolutionary Equilibrium in Contests with Stochastic Participation: Entry, Effort and Overdissipation. *Journal of Economic Behavior and Organization*, *164*, 469–485.

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal Nonparametric Estimation of First-Price Auctions. *Econometrica*, *68*(3), 525–574.

Guerre, E., Perrigne, I., and Vuong, Q. (2009). Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions. *Econometrica*, *77*(4), 1193–1227.

Hall, A. B. (2019). *Who Wants to Run?: How the Devaluing of Political Office Drives Polarization*. University of Chicago Press.

He, M., and Huang, Y. (2018). Structural Analysis of Tullock Contests with an Application to U.S. House of Representatives Elections. *Working Paper*, University of Technology Sydney and Hong

Kong University of Science and Technolog.

Hirano, S., and Snyder, J. M. (2014). Primary Elections and the Quality of Elected Officials. *Quarterly Journal of Political Science*, *9*(4), 473–500.

Hirano, S., Snyder, J. M., Ansolabehere, S., and Hansen, J. M. (2010). Primary Elections and Partisan Polarization in the U.S. Congress. *Quarterly Journal of Political Science*, *5*(2), 169–191.

Jacobson, G. C. (1978). The Effects of Campaign Spending in Congressional Elections. *American Political Science Review*, *72*(2), 469–491.

Jacobson, G. C., and Kernell, S. (1981). *Strategy and Choice in Congressional Elections*. Yale University Press.

Jofre-Bonet, M., and Pesendorfer, M. (2003). Estimation of a Dynamic Auction Game. *Econometrica*, *71*(5), 1443–1489.

Kang, K. (2016). Policy Influence and Private Returns from Lobbying in the Energy Sector. *Review of Economic Studies*, *83*(1), 269–305.

Kawai, K., and Sunada, T. (2015). Campaign Finance in U.S. House Elections. *Working Paper*, University of California, Berkeley and University of Pennsylvania.

Krasnokutskaya, E., and Seim, K. (2011). Bid Preference Programs and Participation in Highway Procurement Auctions. *American Economic Review*, *101*(6), 2653–2686.

Lau, R. R., and Pomper, G. M. (2002). Effectiveness of Negative Campaigning in U.S. Senate Elections. *American Journal of Political Science*, *46*(1), 47–66.

Levitt, S. D. (1994). Using Repeat Challengers to Estimate the Effect of Campaign Spending on Election Outcomes in the U.S. House. *Journal of Political Economy*, *102*(4), 777–798.

Li, T., and Zhang, B. (2015). Affiliation and Entry in First-Price Auctions with Heterogeneous Bidders: An Analysis of Merger Effects. *American Economic Journal: Microeconomics*, *7*(2), 188–214.

Li, T., Zhang, J., and Zhao, J. (2020). Nonparametric Identification of Bayesian Games Under Exclusion Restrictions. *Working Paper*, Vanderbilt University, University of Technology Sydney, and Vanderbilt University.

Li, T., and Zheng, X. (2009). Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions. *Review of Economic Studies*, *76*(4), 1397–1429.

Lu, J. (2009). Auction Design with Opportunity Cost. *Economic Theory*, *38*, 73–103.

Marmer, V., and Shneyerov, A. (2012). Quantile-Based Nonparametric Inference for First-Price Auctions. *Journal of Econometrics*, *167*(2), 345–357.

Matzkin, R. L. (2003). Nonparametric Estimation of Nonadditive Random Functions. *Econometrica*, *71*(5), 1339–1375.

McCarty, N. M., Poole, K. T., and Rosenthal, H. (2006). *Polarized America: The Dance of Ideology* *and Unequal Riches*. MIT Press Cambridge.

Meserve, S. A., Pemstein, D., and Bernhard, W. T. (2009). Political Ambition and Legislative Behavior in the European Parliament. *Journal of Politics*, *71*(3), 1015–1032.

Pastine, I., and Pastine, T. (2002). Consumption Externalities, Coordination, and Advertising. *International Economic Review*, *43*(3), 919–944.

Roberts, J. W., and Sweeting, A. (2013). When Should Sellers Use Auctions? *American Economic Review*, *103*(5), 1830–61.

Schofield, N., and Sened, I. (2005). Multiparty Competition in Israel, 1988-1996. *British Journal of Political Science*, *35*(4), 1988–1996.

Schofield, N., and Sened, I. (2006). *Multiparty Democracy: Elections and Legislative Politics*. Cambridge University Press.

Sieg, H., and Yoon, C. (2017). Estimating Dynamic Games of Electoral Competition to Evaluate Term Limits in US Gubernatorial Elections. *American Economic Review*, *107*(7), 1824–1857.

Sovey, A. J., and Green, D. P. (2011). Instrumental Variables Estimation in Political Science: A Readers' Guide. *American Journal of Political Science*, *55*(1), 188–200.

Sparks, S. (2018). Campaign Spending and the Top-Two Primary: How Challengers Earn More Votes per Dollar in One-Party Contests. *Electoral Studies*, *54*, 56–65.

Staiger, D., and Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica*, *65*(3), 557–586.

Theriault, S. M., and Rohde, D. W. (2011). The Gingrich Senators and Party Polarization in the U.S. Senate. *Journal of Politics*, *73*(4), 1011–1024.

Tomz, M., and Van Houweling, R. P. (2012). Candidate Repositioning. *Working Paper*, Stanford University and University of California, Berkeley.

Voorheis, J., McCarty, N. M., and Shor, B. (2015). Unequal Incomes, Ideology and Gridlock: How Rising Inequality Increases Political Polarization. *Working Paper*, University of Oregon, Princeton University, and Georgetown University.

Wasser, C. (2013). Incomplete Information in Rent-Seeking Contests. *Economic Theory*, *53*(1), 239–268.

Whitten, G. D., and Palmer, H. D. (1996). Heightening Comparativists' Concern for Model Choice: Voting Behavior in Great Britain and the Netherlands. *American Journal of Political Science*, *40*(1), 231–260.