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†

October 27, 2021

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 asymmetric 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 candidate-specific characteristics, as well as state-year level covariates for Senate elections between 1994 and 2018. Taking account of endogenous entry, I find that candidates with past legislative experience and 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 conduct a counterfactual analysis and show that a public funding policy targeting at challengers can reduce the amount of campaign spending by candidates, and increase the probability of challenger victory modestly.

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, Finance, and Legal Studied, The University of Alabama. Email: jzhao29@cba.ua.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 investigating 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 asymmetric political candidates and to 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 an asymmetric contest model with incomplete information and endogenous entry. Toward this end, I explicitly model the entry stage by assuming that candidates (whom I call "potential candidates") strategically decide to participate in the race. 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, together with other candidate-specific traits, 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 candidates with past legislative experience and incumbents are more likely to run, and upon entry they will make strictly positive spending. The structural estimation results show that female candidates, candidates with past legislative experience, and incumbents spend more in campaigns, compared to the rest of the challengers. Further, more campaign spending translates to larger latent utility for voters, then leads to more votes for candidates, especially in the primary elections. I also find that the incumbency advantage exists across different elections in the sense that incumbents attract more votes in both primary and general 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 is monotonic in most cases, regardless of which party's number being changed. Increasing the number of potential candidate has a negative effect on campaign spending, the winning chance of incumbents in both primary and general elections, and the entry probability. However, changing the number of potential candidates in the incumbent party induces relatively larger effects, compared to changing this number in the opposition party. This implies that the primary election of the incumbent's party can be important in terms of improving electoral competition.

I consider the effect of a government campaign finance policy in the counterfactual experiment. The counterfactual policy approximates the type of policy proposal in the form of a public matching fund for political challengers. As the third contribution of this paper, I emphasize that incorporating the endogenous entry significantly alters the assessment of the counterfactual policy. If one ignores the participation behavior of potential candidates and treats entry as fixed, the counterfactual campaign spending is larger relative to the benchmark model, and the winning probability of incumbents reduces to a great extent: by around 10% in primary elections and 12% in general elections. However, once the strategic entry is accounted for, I find that the public funding policy only creates a slightly more competitive election with the entry proportion rising by just 0.2% compared to the benchmark model. Moreover, the effect of the policy is rather modest with the equilibrium entry adjustment of potential candidates. The counterfactual campaign spending with endogenous entry decreases by 16%, and the probability of incumbent victory is down only by about 2% in both elections, corresponding to a much weaker effect of the counterfactual policy. This is because less powerful incumbents choose to opt out the race strategically when facing financially stronger challengers under the policy. Therefore The failure to take into account the strategic entry decisions of potential candidates can produce misleading policy 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), Gilens, Patterson, and Haines (2021), 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 use the simulated maximum likelihood estimation to estimate the unknown distribution of the unobserved heterogeneity, together with the spending distribution as the first step in the multi-stage estimation procedure. In the rest steps, I estimate other primitives of the model, such as the voter model and the distribution of the entry cost.

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), among others. However, political competition contest models are fairly under-developed. Theoretical papers by Fu, Jiao, and Lu (2015), Gu, Hehenkamp, and Leininger (2019), and Jia and Sun (2021) model players' entry in contest models with complete information; while more recent papers, like Liu and Lu (2019) and Boosey, Brookins, and Ryvkin (2020), work on theoretical contest models with incomplete information and endogenous entry. On the empirical side, this paper is the first one to construct a two-stage asymmetric 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), Sovey and Green (2011)), and Gilens, Patterson, and Haines (2021). Although the aforementioned findings are mixed, most work on estimating the influence of campaign spending on elections relies on reduced-form analyses. This paper improves the understanding of how campaign spending affects electoral competition via a structural analysis 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 (2021) and Shakhgildyan (2021)), firms' lobbying expenditures and policy enactment (Kang (2016)), as well as advertising decisions in the U.S. presidential elections (Gordon and Hartmann (2016)).[2] 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 coordination among voters, through estimating a sequential election model in the U.S. presidential

2 For nonparametric identification of Bayesian games that cover the Tullock contest model with incomplete information, see Li, Zhang, and Zhao (2020).

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. 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.[3] 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.[4] Candidates' campaign spending plays an important role in this election process.[5]

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).[6] This data contains information on candidates' vote shares in the primary and general elections, the state and

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

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

5 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 in the candidate section.

6 The FEC website: www.fec.gov.

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 period, from a different source collected by FEC.[7] 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.[8] 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.[9] 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

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

8 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.

9 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 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 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 dataset consists of 2282 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: (i) local political preferences; (ii) demographic and economic variables; and (iii) 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 days 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, in order to control for the candidate-level characteristics, I construct two dummy variables to indicate the gender and the past legislative experience of a potential candidate. If the candidate is a female, the gender dummy equals one. As for the past legislative experience, the experience dummy is encoded as one if the candidate ever held a legislative position in the past, such as House of representative, Senator, Councilwoman or Councilman, County Commissioner, and so on.[10]

Table 1 presents the summary statistics of the data. Around 69% candidates have strictly positive campaign spending.[11] The average amount of campaign spending is about $4 million, which includes all disbursements by a representative candidate during the election. Only 13% of the potential candidates are female, while 37% of them have some past legislative experience. 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 races. 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.

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 those of candidates in open-seat elections for different groups of candidates. The candidate-specific characteristics contain two dummies: the gender and the past legislative experience of each candidate, thus dividing the candidates into four different groups: female and experienced, male and experienced, female and inexperienced, as well as male and inexperienced candidates. Note that since all incumbents must have the past experience from winning the previous election cycle, Subfigures 1c and 1d show the spending densities only for challengers and open-seat election candidates for the latter two groups. As Subfigures 1a and 1b show, 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. As shown in all the subfigures, in open-seat elections, the spending density of candidates is slightly shifted rightward from

10 The profiles of political candidates who ever register and file reports can be found on websites as Our Campaigns and Ballotpedia.

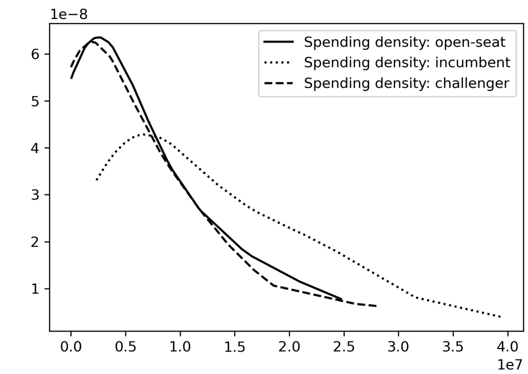
11 Candidates with zero campaign spending still receive strictly positive (but small) vote shares. When modeling candidates' spending behavior in Section 3, I assume that candidates can choose not to spend if they find it unprofitable, even though they still have a strictly positive winning chance.

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

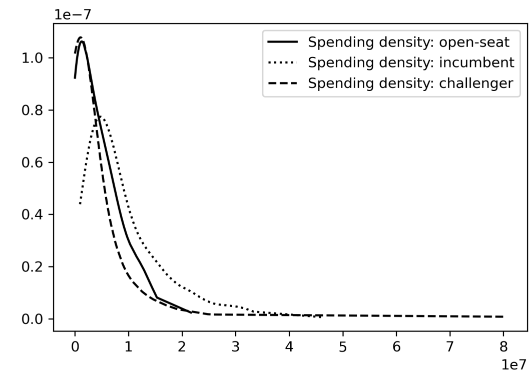
	Observations	Mean	SD
Non-zero spending	1679	0.6945	0.461
Spending	1166	4.2521	7.261
Female	2282	0.1301	0.337
Past legislative experience	2282	0.3723	0.469
No. of potential candidates	306	7.4575	4.762
No. of actual candidates	306	5.4869	3.659
Entry proportion	306	0.7787	0.210
Open-seat election	306	0.1830	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

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, and VEP are scaled down by 10^6.

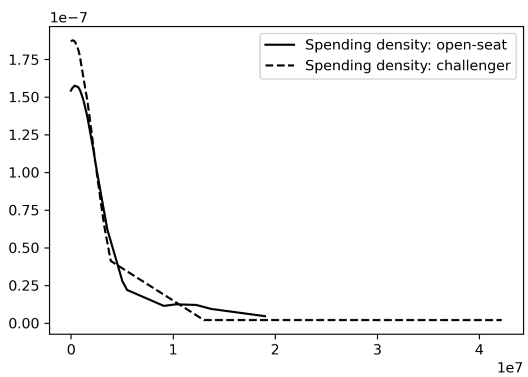
that of challengers and shifted leftward from that of incumbents in incumbent-challenger elections. This implies that candidates in open-seat elections tend to make larger campaign spending than challengers in incumbent-challenger elections on average. However, unlike incumbents, candidates in open-seat elections and challengers in incumbent-challenger elections have positive probabilities of making zero spending. If comparing across different subfigures vertically (thus different groups of candidates), experienced candidates tend to spend more than their inexperienced counterparts, which holds true for both genders. However, comparing the subfigures horizontally leads to less clearer results: for female candidates, the spending distributions spread over a smaller support (thus smaller standard deviations), but might have larger expected values because the male spending distributions are more right-skewed.



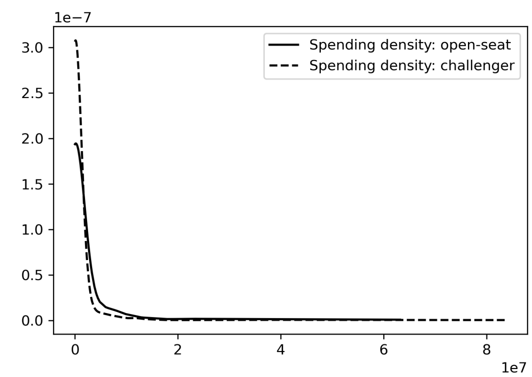
(a) Female and experienced candidates



(b) Male and experienced candidates



(c) Female and inexperienced candidates



(d) Male and inexperienced candidates

Figure 1: Spending distributions of incumbents and challengers in incumbent-challenger elections and 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 different groups of candidates in the same spirit of Figure 1, 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; and female incumbents enter relatively more compared to male ones. However, there is still substantial selection among challengers in terms of who run for office; and the entry probability shows nonessential difference across different groups of candidates among challengers.

Table 2: Summary statistics of entry dummies for different candidates

	Observations	Mean	SD
Female and experienced incumbents	29	0.9310	0.258
Female and experienced challengers	74	0.7568	0.432
Male and experienced incumbents	277	0.8051	0.397
Male and experienced challengers	367	0.7411	0.439
Female and inexperienced challengers	194	0.7268	0.447
Male and inexperienced challengers	1341	0.7159	0.451

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.

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 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.[12] Let the candidate-level heterogeneity be Z, which contains the dummy variables indicating the potential candidate's gender and past legislative experience status.

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

Throughout this section, the distributions of the entry cost and the private value are conditional on (Z, X, u), but to economize on notations, I suppress the conditional set. Note that due to the existence of Z, the model assumes asymmetric potential candidates in the entry stage and asymmetric actual candidates in the election stage.

In the entry stage, each potential candidate holds a private entry cost for office, denoted by c_{D_i}, $i = 1, \cdots, N$ for Democratic potential candidate D_i, and c_{R_k}, $k = 1, \cdots, M$, for Republican potential candidate R_k.[13] In the election stage, each actual candidate has a private value for office, denoted by v_{D_i}, $i = 1, \cdots, n$, for Democratic actual candidate i, and v_{R_k}, $k = 1, \cdots, m$, for Republican actual candidate k.[14] Suppose the distribution of Democratic potential candidate D_i's entry cost is $H_{D_i}(\cdot)$ ($H_{R_k}(\cdot)$ for Republican potential candidate R_k) 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 Democratic actual candidate D_i's private value is $F_{D_i}(\cdot)$ ($F_{R_k}(\cdot)$ for Republican actual candidate R_k) 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.[15] I assume that conditional on (Z, X, u), each candidate's entry cost and private value are independent 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

13 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.

14 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)).

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

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.[16] 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 $P(D_iR_k)$ 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:

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

where I define the differences as

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

and the detailed derivation can be found in Appendix A1.[17]

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 candidate-specific heterogeneity Z and the election-level heterogeneity (X, u). Further, suppose that each voter's latent utility from electing a candidate is increasing with diminishing returns

16 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)).

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

in the candidate's campaign spending.[18] 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_{D_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:

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

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.[19]

I consider a continuously differentiable and pure-strategy BNE, $e_{D_i} = s_{D_i}(v_{D_i})$. This Democratic

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

19 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.

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.[20] 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:

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

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 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.[21]

Proposition 1 *In the two-stage game-theoretic model with incomplete information and endogenous*

---

20 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)).

21 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.

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 $H_{D_i}(\cdot)$ ($H_{R_k}(\cdot)$) and the private value distribution $F_{D_i}(\cdot)$ ($F_{R_k}(\cdot)$) in the candidate model, and to estimate the voter model and the distribution of the election-level unobserved heterogeneity. 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 is modeled as following:

$$I_{i_l,l} = {Z^E_{i_l,l}}' \alpha_Z + X_l' \alpha_X + u_l + \varepsilon_{1,i_l,l}, \quad i_l = 1, \cdots, n_l,$$
$$\begin{cases} \log(e_{i_l,l}) = Z_{i_l,l}' \beta_Z + X_l' \beta_X + u_l + \varepsilon_{2,i_l,l} \quad if \ I_{i_l,l} > 0, \\ e_{i_l,l} = 0 \quad if \ I_{i_l,l} \leq 0 \end{cases}$$

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

$I_{i_l,l} = Z^{E\,\prime}_{i_l,l}\alpha_Z + X_l'\alpha_X + u_l + \varepsilon_{1,i_l,l}$, $Z^E_{i_l,l}$ includes the candidate-specific dummies indicating the gender (FEMALE) and the past legislative experience (EXPC). 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 $\log(e_{i_l,l}) = Z_{i_l,l}\beta_Z + X_l'\beta_X + u_l + \varepsilon_{2,i_l,l}$, other than FEMALE and EXPC in the selection equation, I include an additional dummy equal to 1 for incumbents and 0 for others (INCUM) into $Z_{i_l,l}$.[22] Therefore, conditional on being strictly positive, the campaign spending of a candidate is specified as a log-normal distribution truncated from below at zero.[23]

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 (2021). 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.[24]

The entry cost distribution of potential candidates is parameterized as a log-normal distribution with mean $Z_{i_l,l}'\delta_Z + X_l^{C\,\prime}\delta_X + u_l$ 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.[25] X_l^C includes the same election-level covariates as X_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.

22 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.

23 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 A4.2.

24 In Appendix A4.4, 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.

25 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 A4.3).

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

$$u^P_{i_l,l} = \gamma \log(1+e_{i_l,l}) + Z'_{i_l,l}\gamma_Z + X^{P'}_l \gamma_X + u_l + \iota_{i_l,l}, \quad i_l = 1,\cdots,n_l,$$

for the primary election, and

$$u^G_{i_l,l} = \omega \log(1+e_{i_l,l}) + Z'_{i_l,l}\omega_Z + X^{G'}_l \omega_X + u_l + \iota_{i_l,l}, \quad i_l = 1,\cdots,n_l,$$

given that this candidate i_l wins the primary election and proceeds to the general election. Note that adding the first four 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^P_l may be different from X^G_l. 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.[26] Therefore the terms $X^{P'}_l\gamma_X$, $X^{G'}_l\omega_X$, and u_l being the same for all candidates in an election are canceled 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 elections into the voter model, as in Gordon and Hartmann (2016), in order to identify γ_X and ω_X.[27] 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.[28] 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

26 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.

27 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.

28 In the sensitivity analysis (see Appendix A4.1), 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.

elections.[29]

Lastly, I assume that u_l has a normal distribution with zero mean and unknown variance σ_u^2 to be estimated. Further u_l is assumed to be independent from $Z_{i_l,l}$ and X_l.

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.[30] In the first step, 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 and the distribution of the election-level unobserved heterogeneity: $\alpha \equiv (\alpha_Z, \alpha_X)$, $\beta \equiv (\beta_Z, \beta_X)$, σ_2, σ_{12}, and σ_u, through the simulated maximum likelihood estimation (SMLE). In the second step, I use MLE to estimate the parameters in the voter model, i.e., γ, γ_Z and ω, ω_Z. 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 simulated 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 $\delta \equiv (\delta_Z, \delta_X)$ and two λs. Therefore, the parameter vector being estimated is defined as $\theta \equiv (\sigma_u; \alpha, \beta, \sigma_2, \sigma_{12}; \gamma, \gamma_Z, \omega, \omega_Z; \delta, \lambda_I, \lambda_C)$.

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. The standard errors of estimators are calculated through nonparametric bootstrapping at the game level. For all the simulation methods, I use 100 repetitions.

4.3 Estimation Results

Table 3 summarizes the results of the structural analysis. The first two columns show the estimates for the distribution of election-specific unobserved heterogeneity, together with the selection part of the campaign spending model together with the covariance matrix of the two errors, thus $(\hat{\sigma}_u, \hat{\alpha}, \hat{\sigma}_2, \hat{\sigma}_{12})$. The third and fourth columns correspond to the log(spending) equation of the campaign spending model,

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

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

thus $\hat{\beta}$. For the distribution of election-level unobserved heterogeneity, the estimated variance is around 0.84 being significant, which indicates the existence of the election-level unobserved heterogeneity modeled as a random variable. The estimated standard deviation of the error in the log(spending) equation is about 2.25, and the estimated covariance of the two errors in the selection equation and in the log(spending) equation is about 1.34, which yields an estimated correlation coefficient being around 0.60. 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. Compared to male candidates, female candidates are equally likely to spend non-zero amount of money in elections, because the effect of the gender dummy is insignificant in the spending selection equation. However, female candidates significantly spend more compared to male candidates, which can be explained by two different reasons: one is that this indicates a possibly inferior position of female in the political society that induces female candidates to spend more to attract voters; or female candidates are more attractive than their male opponents to voters, thus on one hand they have more financial strength compared to male candidates thus value the office more, on the other hand they can raise more money among contributors. One can decide which factor is more prominent by looking at the results of voter model. Candidates with past legislative experience are more likely to make non-zero spending and spend more money in elections, with both effects being significant. 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. These are probably due to the reason that candidate with past experience and incumbent candidates have relatively stronger financial strength, thus either value the office more or have more ability to raise contributions. 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 (insignificantly) and a larger amount of spending (significantly), 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.[31] 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 estimated coefficients in the selection equation and those in the log(spending) equation are insignificant. The only exception is the effect of median household income on the amount of spending. 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 in primary elections, i.e., $(\hat{\gamma}, \hat{\gamma}_Z)$. The seventh and eighth columns show the results for the voter model in general elections: $(\hat{\omega}, \hat{\omega}_Z)$. The two parameters γ and ω 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 0.0392 and 0.0052 respectively.[32] 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.

31 Another reason why actual candidates increase the spending in a less competitive election can be that candidates in one election tend to seek contribution supports in the same resource pool, and less actual candidates can decrease their costs to raise the funds, thus increase the money they can raise and the corresponding spending level.

32 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 4.4 and 3.3 (both scaled up by 10^8) in the primary and general elections respectively.

The estimated γ_Z and ω_X show the effects of corresponding candidate-specific covariates on voters' latent utilities in primary and general elections, respectively. Overall, being a female candidate, a candidate with past legislative experience, or an incumbent can increase the utility of a voter in both primary and general elections. However, the effects in primary elections are all significant, while in general elections there is only one significant effect being the incumbency status. This could be due to the difference between primary and general elections mentioned above: when the party label is missed in the primary election, voters need to rely more on some candidate-specific characteristics to decide who they like more and to vote, such as the gender and the past legislative experience even when candidates have the same amount of spending. Once party label is accounted for in the general election, voters are less concerned about candidate-specific characteristics. On the other hand, the incumbency advantage exists across different elections, and the effects are similar. As a last note, that the estimated coefficient of the gender dummy being positive in voter model implies that the reason why female candidates spend more money could be that female candidates are more attractive than their male opponents to voters, but not that they have a relatively inferior position compared to male ravels.

The last two columns give the estimates in the entry model, thus $(\hat{\delta}, \hat{\lambda}_I, \hat{\lambda}_C)$. First of all, candidates with past legislative experience and incumbents have larger expected entry costs, compared to challengers. In footnote 13, I model the entry stage of potential candidates by referring entry cost to reservation value. This interpretation helps explain the higher entry probabilities by experienced candidates and incumbents, which are consistent with the observed data. Specifically, experienced candidates and 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.[33] Another possible explanation why incumbents hold larger expected entry costs is due to the choice of the electorate, because the electorate eventually wants a change of the person in charge of the final political office. 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

33 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.

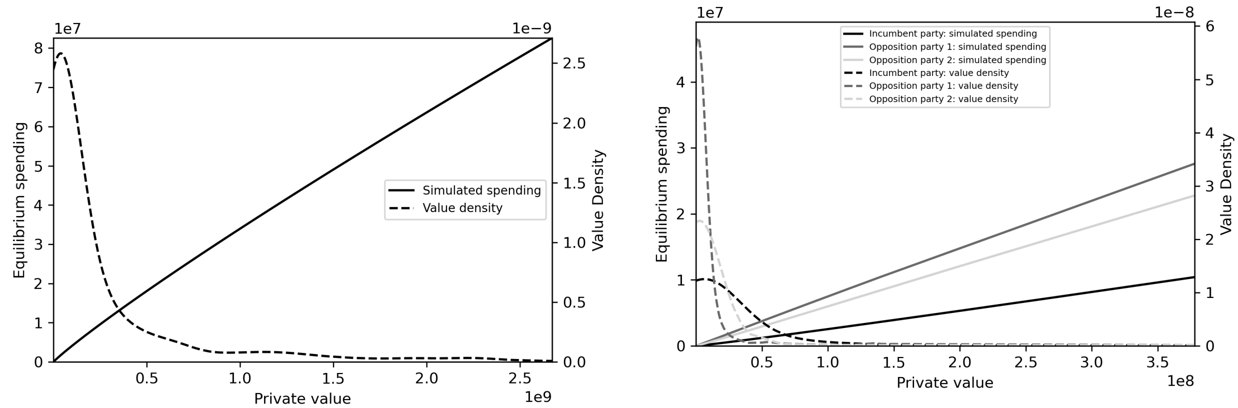
although incumbents have larger entry costs compared to challengers on average, among incumbents the difference of entry costs is not as large as that among challengers. Most of the estimated coefficients of the covariates are insignificant, with two exceptions including the effect of the percentage of the population in the oldest age bin (65 and up) and the effect of the median household income, which is possibly due to the reason that (i) the total votes cast reduces with more older citizens in the society, which makes the office less attractive and deters candidates from running for the office; and (ii) a richer economy entails more valuable outside options rather than running for the office, thus leads to higher entry costs. 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.

Table 3: Estimation results from structural analysis

	Spending: selection equation		Spending: log(spending) equation		Voter: primary election		Voter: general election		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_u	0.9175	0.009								
σ_2	2.2478	0.118								
σ_{12}	1.3404	0.197								
γ					0.1250	0.008				
ω							0.2444	0.046		
λ_I									4.1243	0.301
λ_C									5.4718	0.424
FEMALE	0.4598	0.362	0.3623	0.183	0.2257	0.051	0.1362	0.169	0.1364	0.195
EXPC	1.5790	0.204	2.6999	0.215	0.4665	0.063	0.1004	0.119	2.0924	0.242
INCUM			2.1639	0.168	0.3529	0.090	0.3319	0.078	1.8237	0.275
OPEN	0.4402	0.135	0.9837	0.155						
PVI	0.0016	0.009	−0.0040	0.011					−0.0573	0.086
GOV	−0.0907	0.150	0.0360	0.170					−0.1174	0.295
YOUNG-PER	1.7911	2.569	−0.1684	3.059					−4.4392	3.937
MID-PER	−1.3249	1.990	−0.5188	2.287					−0.5306	2.525
OLD-PER	−0.3069	3.177	4.1681	3.546					6.7749	3.716
UNEMP	−0.0020	0.046	−0.0186	0.046					0.1465	0.261
LOGINC	0.4607	0.500	1.5676	0.640					1.6898	0.774
PNCAN	0.0183	0.045	0.1013	0.037					0.0197	0.155
ANCAN	−0.0574	0.057	−0.2115	0.043						
CONST	−4.6626	5.296	−6.2732	6.834					−6.7966	8.131

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 100 nonparametric bootstrap at the election level.

An important component in the structural analysis is the equilibrium strategy function mapping 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 candidate-specific distribution. I simulate the equilibrium strategy functions and the private value densities for different candidates 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 different candidates in an open-seat election (indexed 19) as shown in Figure 3.

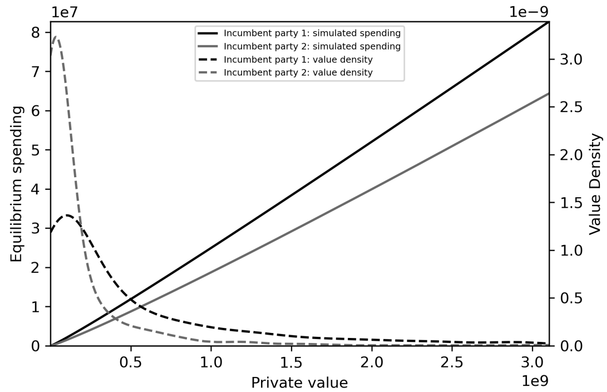


(a) Incumbent: male

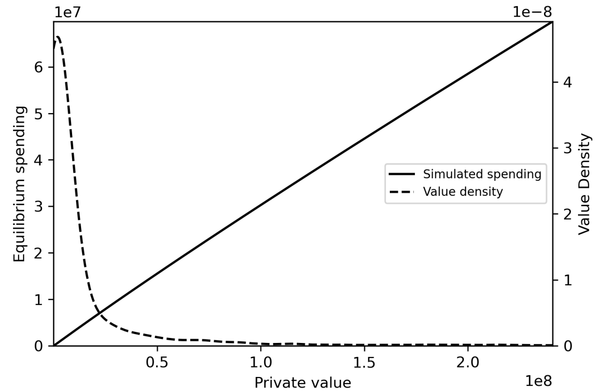
(b) Challengers: all are inexperienced; op-party 1 is female and op-party 2 is male

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

Figure 2 presents the equilibrium strategy functions and the private value densities in an incumbent-challenger election. Election 1 has four actual candidates: a male incumbent, one challenger from the incumbent party (male and inexperienced), and two inexperienced challengers from the opposition party (one is female and the other is male). Apparently, the equilibrium strategy function is strictly increasing for all candidates. Comparing the curves for the incumbent and for challengers, it is clear that incumbents value the office more than challengers, and thus are willing to spend more to raise their victory probabilities. In Subfigure 2b, the distributions of the opposition party challengers' values are 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 in the opposition party probably due to a more appealing office with the existence of the incumbent. However, although the scales of private values for different challengers are similar, the challengers from the opposition party seem to spend more, where female challenger spends the most, consistent with the estimation results of the spending and the voter models. 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.



(a) Incumbent party candidates: both are male; in-party 1 is experienced and in-party 2 is inexperienced



(b) Opposition party candidates: male and inexperienced

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

I now turn to the open-seat election. Figure 3 gives the equilibrium strategy functions and the private value distributions for different candidates. Although there is no incumbent participating in the election, I still define the parties as above for the ease of illustration and comparison. Election 19 has seven actual candidates belonging to three groups (all are male): for the incumbent party, both experienced and inexperienced candidates exist; for the opposition party, candidate is inexperienced. Comparing Subfigures 3a and 3b shows that the incumbent party candidates value the office more than the opposition party candidates, and the value of the experienced candidate from the incumbent party is the largest. This is probably because the incumbent party has some advantage in terms of the political resources or policies even though it is an open-seat election. However, taking a closer look at the equilibrium spending of these three groups of candidates shows that the experienced candidate from the incumbent party spends the most, the spending of the opposition party candidate ranks the next, and the inexperienced candidate from the incumbent party spends the least, which is possibly due to the existence of a strong rival in the same party.

In Appendix A4, I conduct a series of extensive sensitivity analyses, including considering the turnout of voters in the general elections (Appendix A4.1), a truncated spending distribution from above (Appendix A4.2), taking into account the difference between two political parties (Appendix A4.3), a quadratic cost function (Appendix A4.4), an alternative voter model specification (Appendix A4.5), as well as adding a dummy variable that is assigned one for the period of post-2010 and zero for that of pre-2010 to the specifications in the model (Appendix A4.6). 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 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: Model fit results

	Spending		Voter		Entry	
	Predicted	Observed	Predicted	Observed	Predicted	Observed
Non-zero spending	0.6901 (0.008)	0.6945				
Spending	4.0606 (0.961)	4.2521				
Spending of incumbents	12.421 (0.616)	9.4280				
Spending of challengers	2.1017 (0.683)	2.8395				
Vote shares of incumbents in primary elections			0.7483 (0.078)	0.8840		
Primary elections incumbents win			0.9020 (0.082)	0.9820		
Vote shares of incumbents in general elections			0.7359 (0.073)	0.6016		
General elections incumbents win			0.8922 (0.084)	0.8440		
No. of actual candidates					5.4944 (0.992)	5.4869
Entry proportion					0.7745 (0.138)	0.7785
Open-seat election					0.0929 (0.022)	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.

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 for the full set of candidates simulated from the estimated structural parameters fits the observed data well. The spending made by incumbents is a bit larger in the simulation and the spending made by challengers is relatively smaller 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 relatively perfectly, with slightly larger entry probabilities of incumbents (which is also consistent with the predicted higher winning chance of incumbents in general elections) that lead to a slightly smaller percentage of open-seat elections.

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 the proposed government policy.

5.1 The Role of Entry

In the structural estimation, 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, voter and entry results. To compare behavior in alternative environments and to avoid computational complication, I simulate the outcomes for a subsample of elections observed in the data: specifically, the subsample included in this exercise has at least two potential challengers in the incumbent party and at least three potential challengers in the opposition party. I asses the effects of decreasing (increasing) the numbers of potential candidates in (i) the incumbent party; (ii) the opposition party; and (iii) both the incumbent and opposition parties on the spending, vote and entry results, given the estimated structural model. The results are shown in Tables 5 to 7.

Table 5: Effects of changing the numbers of potential candidates in the incumbent party

		Decrease by 2	Decrease by 1	No change	Increase by 1	Increase by 2
Spending	Non-zero spending	0.6813 (0.014)	0.6762 (0.016)	0.6719 (0.017)	0.6703 (0.016)	0.6682 (0.017)
	Spending	3.8016 (1.046)	3.3642 (1.073)	3.0710 (1.022)	2.9680 (1.003)	2.8468 (0.986)
	Spending of incumbents	11.846 (3.304)	11.133 (3.400)	10.728 (3.491)	10.418 (3.548)	10.088 (3.573)
	Spending of challengers	2.1226 (0.736)	1.9795 (0.730)	1.9245 (0.715)	1.8600 (0.693)	1.8077 (0.697)
Voter	Vote shares of incumbents in primary elections	0.7077 (0.089)	0.6119 (0.113)	0.5397 (0.113)	0.4887 (0.107)	0.4514 (0.102)
	Primary elections incumbents win	0.8907 (0.094)	0.8567 (0.130)	0.8393 (0.136)	0.8294 (0.143)	0.8192 (0.151)
	Vote shares of incumbents in general elections	0.7285 (0.081)	0.7022 (0.102)	0.6904 (0.113)	0.6828 (0.119)	0.6761 (0.126)
	General elections incumbents win	0.8850 (0.096)	0.8534 (0.124)	0.8384 (0.136)	0.8288 (0.144)	0.8188 (0.152)
Entry	No. of actual candidates	7.2833 (1.352)	7.9472 (1.436)	8.6231 (1.512)	9.3034 (1.591)	10.007 (1.678)
	Entry proportion	0.7241 (0.141)	0.7221 (0.135)	0.7175 (0.130)	0.7143 (0.125)	0.7102 (0.122)
	Open-seat election	0.1032 (0.025)	0.1355 (0.033)	0.1506 (0.035)	0.1603 (0.036)	0.1702 (0.037)

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: Effects of changing the numbers of potential candidates in the opposition party

		Decrease by 2	Decrease by 1	No change	Increase by 1	Increase by 2
Spending	Non-zero spending	0.6771 (0.014)	0.6762 (0.015)	0.6719 (0.017)	0.6706 (0.017)	0.6686 (0.017)
	Spending	3.2918 (1.007)	3.2192 (1.028)	3.0710 (1.022)	3.0219 (1.038)	2.9369 (1.033)
	Spending of incumbents	11.297 (3.354))	11.011 (3.401)	10.728 (3.491)	10.493 (3.557)	10.250 (3.573)
	Spending of challengers	1.8870 (0.678)	1.8621 (0.682)	1.9245 (0.715)	1.8964 (0.706)	1.8642 (0.708)
Voter	Vote shares of incumbents in primary elections	0.5397 (0.107)	0.5413 (0.110)	0.5397 (0.113)	0.5388 (0.115)	0.5366 (0.116)
	Primary elections incumbents win	0.8487 (0.138)	0.8457 (0.123)	0.8393 (0.136)	0.8348 (0.140)	0.8307 (0.143)
	Vote shares of incumbents in general elections	0.7033 (0.106)	0.6993 (0.108)	0.6904 (0.113)	0.6860 (0.115)	0.6846 (0.118)
	General elections incumbents win	0.8500 (0.128)	0.8468 (0.131)	0.8384 (0.136)	0.8330 (0.140)	0.8292 (0.143)
Entry	No. of actual candidates	7.3726 (1.336)	7.9715 (1.428)	8.6231 (1.512)	9.2789 (1.595)	9.9729 (1.674)
	Entry proportion	0.7374 (0.138)	0.7253 (0.135)	0.7175 (0.130)	0.7122 (0.126)	0.7078 (0.122)
	Open-seat election	0.1408 (0.034)	0.1440 (0.034)	0.1506 (0.035)	0.1553 (0.035)	0.1592 (0.358)

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 7: Effects of changing the numbers of potential candidates in both parties

		Decrease by 2	Decrease by 1	No change	Increase by 1	Increase by 2
Spending	Non-zero spending	0.6871	0.6795	0.6719	0.6690	0.6657
		(0.012)	(0.015)	(0.017)	(0.017)	(0.017)
	Spending	4.1030	3.436	3.0710	2.8806	2.7590
		(0.987)	(1.073)	(1.022)	(1.012)	(1.014)
	Spending of incumbents	12.300	11.451	10.728	10.144	9.6807
		(2.890)	(3.321)	(3.401)	(3.584)	(3.622)
	Spending of challengers	1.9684	1.8933	1.9245	1.8443	1.7496
		(0.669)	(0.699)	(0.715)	(0.695)	(0.689)
Voter	Vote shares of incumbents in primary elections	0.7077	0.6138	0.5397	0.4875	0.4517
		(0.086)	(0.110)	(0.113)	(0.109)	(0.106)
	Primary elections incumbents win	0.8958	0.8629	0.8393	0.8238	0.8128
		(0.089)	(0.118)	(0.136)	(0.148)	(0.157)
	Vote shares of incumbents in general elections	0.7381	0.7109	0.6904	0.6679	0.6709
		(0.076)	(0.098)	(0.113)	(0.122)	(0.130)
	General elections incumbents win	0.8935	0.8615	0.8384	0.8226	0.8117
		(0.090)	(0.118)	(0.136)	(0.148)	(0.157)
Entry	No. of actual candidates	5.9773	7.2872	8.6231	9.9560	11.309
		(1.114)	(1.336)	(1.512)	(1.672)	(1.825)
	Entry proportion	0.7579	0.7309	0.7175	0.7096	0.7023
		(0.137)	(0.139)	(0.130)	(0.122)	(0.115)
	Open-seat election	0.0973	0.1292	0.1506	0.1659	0.1769
		(0.248)	(0.033)	(0.035)	(0.036)	(0.375)

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 5 corresponds to the effects of changing the number of potential candidates in the incumbent party on the spending, voter, and entry outcomes; while Table 6 shows the results if I change the number of potential candidates in the opposition party. On the other hand, the results presented in Table 7 are consistent with the change of the numbers of potential candidates in both incumbent and opposition parties. Despite the difference across the three Tables regarding whose party I change the number of potential candidates, the results show similar patterns. In general, increasing the number of potential candidates can decrease the non-zero spending ratio and the average spending level of all candidates. When the number of potential candidates increase, the incumbents will win less primary elections; they will also win less general elections with smaller vote shares. Turning to the entry results, increasing the number of potential candidates has positive effects on the number of actual candidates and the probability of being an open-seat election, as well as a negative effect on the entry proportion. The above effects remain the same across Tables 5, 6, and 7.

When increasing the number of potential candidates in the incumbent party, the spending level decreases on average, regardless of the candidate status. In contrast, only the incumbents' spending decreases as the number of potential candidates in the opposition party increases or the numbers of potential candidates in both parties increase. But because the changing magnitudes of the incumbents' spending is relatively larger, it still leads to the reduction of winning change for the incumbents which is reflected by the voter model outcomes. Increasing the number(s) of potential candidates in the opposition party (both parties) doesn't show a monotonic effect on the vote shares of incumbents in primary (general) elections, but the incumbents still wins less primary (general) elections in the end due to relatively stable spending made by the challengers. The entry behavior is impeded by the rise of the number of potential candidates universally, 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.

With a further investigation in the magnitudes of the changes regarding various outcomes when increasing the number of potential candidates, I find that changing the number of potential candidates in the incumbent party or both parties incurs relatively larger effects shown in Tables 5 and 7, compared to those shown in Table 6 when changing the number of potential candidates in the opposition party. 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.[34]

5.2 Government Subsidies to Challengers

In this counterfactual analysis, I consider the effect of a campaign finance policy, i.e., government subsidies to political candidates in Senate elections. Specifically, this takes the form of a public funding program, which provides eligible candidates with federal government funds to pay for the qualified expenses of their campaigns in both primary and general elections. Incumbents usually opt out of this program, because it comes with a spending limit along with some other terms that the candidate needs to agree to.[35] Hence, the public fund considered in this subsection targets at political challengers in Senate elections only, and can match one dollar for every dollar the challenger raises, which reduces the effective cost of the challenger's campaign spending.

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 Sections 4.4 and 5.1 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 the public fund program targeting at challengers so that potential candidates decide whether they will run for office endogenously. Table 8 and Table 9 report the simulated results in the counterfactual analysis under these two scenarios.

When entry is fixed, the simulated results in the spending and voter models are presented in Table 8. If the equilibrium entry behavior is ignored, the effects of the dollar-for-dollar subsidies are relatively strong. With the subsidies, incumbents will largely decrease the spending while challengers will largely increase the spending, leading to a rise of the average spending level for the same set of candidates as in the observed data. More importantly, the subsidies reduce the strength of incumbents in both primary and general elections, reflected by diminishing vote shares and winning chance of incumbents in both elections on average. The changes are economically significant, implying that incumbents' winning probability drops by around 10.25% (11.79%) from the baseline probability of 90.20% (89.22%) in primary (general) elections.

34 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.

35 One example of such public funding of presidential elections is the primary matching funds.

Table 8: Counterfactual: government subsidies to challengers with fixed entry

	Spending	Voter
Non-zero spending	0.6901 (0.025)	
Spending	5.9648 (0.679)	
Spending of incumbents	8.3187 (3.392)	
Spending of challengers	5.5313 (0.679)	
Vote shares of incumbents in primary elections		0.7016 (0.171)
Primary elections incumbents win		0.8095 (0.156)
Vote shares of incumbents in general elections		0.6438 (0.118)
General elections incumbents win		0.7870 (0.144)

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 9: Counterfactual: government subsidies to challengers with endogenous entry

	Spending	Voter	Entry
Non-zero spending	0.6900 (0.008)		
Spending	3.4108 (0.398)		
Spending of incumbents	8.9075 (0.444)		
Spending of challengers	2.9246 (0.396)		
Vote shares of incumbents in primary elections		0.7391 (0.079)	
Primary elections incumbents win		0.8839 (0.083)	
Vote shares of incumbents in general elections		0.7122 (0.073)	
General elections incumbents win		0.8739 (0.088)	
No. of actual candidates			5.5010 (0.984)
Entry proportion			0.7761 (0.137)
Open-seat election			0.0944 (0.022)

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.

However, once the equilibrium entry responses of potential candidates are taken into account, the dollar-for-dollar subsidies induce relatively modest changes in the electoral outcomes as shown in Table 9. Incumbents decrease the spending while challengers increase the spending, but the magnitudes of such changes are much smaller compared to the fixed entry scenario, especially for challengers. This leads to a smaller spending level on average for all candidates compared to the benchmark model, in drastic contrast to the result under fixed entry scenario. The reason is probably that with endogenous entry, candidates can choose not to run for office if they find it unworthy. But if entry is fixed, these incumbents can only decrease their spending when facing financially stronger challengers while challengers can only increase their spending more to offset some costs, because they cannot secede.

The modest effects on the spending results further transmit to the modest effects on the voter results. The vote shares and winning chance of incumbents in primary and general elections only decrease by a smaller percentage: in primary elections, the winning probability of incumbents drops by about 2.01%; while in general elections, the winning probability of incumbents falls by just 2.05% at the mean. Turning to the entry results, Table 9 shows that the probability of being an open-seat election increases slightly, indicating a slightly smaller entry probability of incumbents on average, compared to the benchmark model. On the other hand, challengers with subsidies enter more frequently into the elections, pushing up the entry proportion and the number of actual candidates at the mean.

In summary, the government dollar-for-dollar subsidies to political challengers in Senate elections can make the challengers financially stronger, which induces them to spend more in the elections; however, since weaker incumbents can choose to drop out of the races when encountering stronger challengers, the remaining stronger incumbents make the probability of challenger victory increase only slightly. In addition, if one ignores the equilibrium entry behavior of potential candidates, the effects of the subsidies may be amplified in the sense that it will introduce potentially powerful challengers in Senate elections to weaken the incumbent victory and intensify competition, but at the cost of increasing the spending level on average. Nevertheless, after accounting for the endogenous entry of potential candidates, the subsidies seem to reduce the average campaign spending, and cause only modest rise of challenger victory.

6 Conclusion

In this paper, I develop and estimate a two-stage game-theoretic asymmetric 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 asymmetric potential candidates decide whether to run for the race, and the second election stage where asymmetric 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, together with other candidate-specific traits, demographic and economic covariates, and the election-specific unobserved heterogeneity. Taking the structural approach, I obtain estimates of 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 candidates with past legislative experience and 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. Further, incumbents can attract more votes in both primary and general elections. 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, the number of potential candidates influences campaign spending, voter, and entry outcomes monotonically, regardless of which party's number being changed. Increasing numbers of potential candidates has negative effects on campaign spending, the winning chance of incumbents in primary and general elections, and entry proportion on average. However, increasing the number of potential candidates in the incumbent party leads to larger effects that weaken the strength of the incumbent, relative to the opposition party. 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 a public matching fund policy. The results show that under the counterfactual policy, challengers will spend more and incumbents will spend less, which can generate a smaller amount of campaign spending on average. However, the probability of incumbent victory only decreases modestly, because weaker incumbents can choose to not enter into the race strategically. Absent the equilibrium entry response of potential candidates, nevertheless, the policy effects are different and more salient. Ignoring the endogenous entry, the counterfactual policy can increase the average campaign spending level among all candidates, but it can also improve the probability of challenger victory largely, thus the efficacy of this policy is seemingly overestimated.

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:

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

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:

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

Therefore,

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

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:

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

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

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

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 $\partial \mathbb{E}_{e_{-D_i}}[CSF(e_{D_i}; e_{-D_i})|v_{D_i}; a_{-D_i}]/\partial e_{D_i}$. 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 $CSF(e_{D_i}; e_{-D_i}) = P(D_i) = \sum_{k=1}^{m} P(D_i R_k)$. Thus it suffices to show that $\partial P(D_i R_k)/\partial e_{D_i} \geq 0$ for $k \in \{1, \cdots, m\}$, where recall that

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

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 $P(D_i R_k)$ as follows:

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

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

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 $[0,1]^{N+M}\rightarrow[0,1]^{N+M}$, 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:

$$\begin{aligned}\log(L_l) = &\sum_{n_{l,0}} \log\left(1-\Phi(Z^{E'}_{i_l,l}\alpha_Z + X_l'\alpha_X + u_l)\right)\\ &+\sum_{n_{l,1}} \log\Phi\left\{\frac{1}{\sqrt{1-\frac{\sigma_{12}^2}{\sigma_2^2}}}\left[Z^{E'}_{i_l,l}\alpha_Z + X_l'\alpha_X + u_l + \sigma_{12}\sigma_2^{-2}\left(\log(e_{i_l,l}) - Z'_{i_l,l}\beta_Z - X_l'\beta_X - u_l\right)\right]\right\}\\ &+\sum_{n_{l,1}} \phi\left(\frac{\log(e_{i_l,l}) - Z'_{i_l,l}\beta_Z - X_l'\beta_X - u_l}{\sigma_2}\right) - n_{l,1}\log(\sigma_2) - \sum_{n_{l,1}}\log(e_{i_l,l}),\end{aligned}$$

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:

$$\log(L_{pri,D}) = \sum_{j=1}^{n} vot_{D_j}\left(\gamma\log(1+e_{D_j}) + Z'_{D_j}\gamma_Z\right) - \log\left(\sum_{j=1}^{n} exp\left(\gamma\log(1+e_{D_j}) + Z'_{D_j}\gamma_Z\right)\right),$$

$$\log(L_{pri,R}) = \sum_{l=1}^{m} vot_{R_l}\left(\gamma\log(1+e_{R_l}) + Z'_{R_l}\gamma_Z\right) - \log\left(\sum_{l=1}^{m} exp\left(\gamma\log(1+e_{R_l}) + Z'_{R_l}\gamma_Z\right)\right),$$

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

$$\log(L_{gen}) = vot_{D_i}\left[\omega\log(1+e_{D_i})+Z'_{D_i}\omega_Z - \log\left(exp(\omega\log(1+e_{D_i})+Z'_{D_i}\omega_Z)+exp(\omega\log(1+e_{R_k})+Z'_{R_k}\omega_Z)\frac{P_{pri}(D_i)}{P_{pri}(R_k)}\right)\right]$$
$$+vot_{R_k}\left[\omega\log(1+e_{R_k})+Z'_{R_k}\omega_Z - \log\left(exp(\omega\log(1+e_{R_k})+Z'_{R_k}\omega_Z)+exp(\omega\log(1+e_{D_i})+Z'_{D_i}\omega_Z)\frac{P_{pri}(R_k)}{P_{pri}(D_i)}\right)\right],$$

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 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.

A4.1 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

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

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.1 as follows. Since the distribution of the unobserved heterogeneity, the selection equation, the log of spending equation, and the voter model in the primary elections are the same as in the benchmark model, I only report the results of this sensitivity analysis for the voter model in the general election and entry model. 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.

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. Candidates with past legislative experience and incumbents attract more votes in the general elections. As for the estimated coefficients in the general latent utility, most of them have expected signs. Being an open-seat election significantly improves the latent utility for a representative voter in the general election; this can also be explained by the "getting-out-the-vote" phenomenon in open-seat elections since voter turnout rate is higher. In the U.S., the middle-aged citizens are the most likely to cast their ballots, which is reflected by a positive and significant estimated coefficient on the percentage of the population in the middle age bin (i.e., 45-64). A richer economy can facilitate the voting behavior of

Table A.1: Sensitivity analysis: turnout in general elections

	Voter: general election		Entry	
	Estimate	Sd. error	Estimate	Sd. error
ω	0.0379	0.029		
λ_I			4.1271	0.149
λ_C			4.8610	0.338
FEMALE	0.0532	0.034	−0.0525	0.192
EXPC	0.3339	0.067	2.1194	0.183
INCUM	0.6475	0.037	1.5810	0.180
OPEN	0.1329	0.021		
PVI	−0.0124	0.015	−0.0228	0.041
GOV	0.0396	0.028	0.1038	0.237
YOUNG-PER	−3.5041	1.337	−4.8303	3.894
MID-PER	3.3732	0.527	0.0586	2.541
OLD-PER	−2.3893	0.430	6.8967	3.727
UNEMP	−0.0613	0.023	0.1556	0.144
LOGINC	0.7821	0.311	1.7542	0.799
PNCAN			0.0146	0.083
ANCAN	−0.0678	0.013		
PRCP	0.0866	0.041		
SNOW	−0.0957	0.036		
PRE	0.5585	0.126		
CONST	−9.8909	3.159	−6.6644	8.139

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 100 nonparametric bootstrap at the election level.

voters, thus the unemployment rate has a negative effect while the median household income has a positive effect. 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. Severe weather, such as the snow on the election day, has a negative effect on the voter turnout. The last 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.

A4.2 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. I assume that the upper boundary of the spending conditions on (X, u), I need to estimate the upper boundary of non-zero spending 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.2.

Most of the results remain consistent with those of the benchmark model. The estimated coefficients in the selection equation have the same signs and similar magnitudes as in the benchmark model

Table A.2: Sensitivity analysis: truncated spending distribution

	Spending: selection equation		Spending: log(spending) equation		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_u	1.0009	0.294				
σ_2	2.2462	0.292				
σ_{12}	1.3394	0.212				
λ_I					4.3513	0.218
λ_C					5.0055	0.890
FEMALE	0.2508	0.144	0.3623	0.179	0.1531	0.061
EXPC	1.9872	0.205	2.6999	0.201	2.0900	0.051
INCUM			2.1139	0.147	1.6353	0.513
OPEN	0.2678	0.134	0.9837	0.156		
PVI	−0.0057	0.009	−0.0040	0.011	−0.0181	0.077
GOV	−0.1594	0.142	0.0360	0.162	−0.0771	1.148
YOUNG-PER	1.9239	2.559	−0.1684	3.052	−4.7707	4.311
MID-PER	−1.4342	1.858	−0.5313	2.085	−0.1914	3.092
OLD-PER	−0.2054	2.945	4.1556	3.224	6.6988	4.312
UNEMP	−0.0020	0.046	−0.0186	0.046	0.0754	0.153
LOGINC	0.4607	0.500	1.5676	0.640	1.7532	0.826
PNCAN	0.0183	0.031	0.1013	0.035	0.0124	0.122
ANCAN	−0.0574	0.035	−0.2115	0.038		
CONST	−4.6617	5.295	−6.2732	6.831	−7.2760	8.673

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 100 nonparametric bootstrap at the election level.

shown in Table 3, in most cases. Female candidate has a significant effect in the selection equation of campaign spending now. With some small discrepancies, the estimates of the entry model remain stable in this sensitivity analysis, compared with the benchmark model.

A4.3 Difference Between Political Parties

I take into account the political party difference 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).[36] 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.3 presents the results of this sensitivity analysis of taking into consideration the difference between political parties. First of all, most of 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 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

36 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.3: Sensitivity analysis: difference between political parties

	Spending: selection equation		Spending: log(spending) equation		Voter: primary election		Voter: general election		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_u	0.9304	0.321								
σ_2	2.4388	0.372								
σ_{12}	1.3011	0.239								
γ_I					0.1048	0.002				
γ_{DEM}					0.1332	0.022				
γ_{REP}					0.2289	0.045				
γ_{OPEN}					0.2911	0.031				
ω_I							0.1256	0.001		
ω_C							0.1387	0.040		
ω_{OPEN}							0.5084	0.094		
λ_I									4.1478	0.928
λ_{DEM}									4.7779	0.719
λ_{REP}									4.5998	0.843
FEMALE	0.2989	0.150	0.3590	0.186	0.1461	0.073	0.0179	0.223	0.5044	0.110
EXPC	1.9836	0.194	2.6782	0.204	0.0858	0.117	0.1017	0.171	0.2658	0.121
INCUM			2.0700	0.172	2.2635	2.037	0.6497	1.345	0.3016	0.162
DEM-IDX	12.0234	9.079	1.2340	11.809	1.0972	0.199	−1.9953	7.561		
OPEN	0.2360	0.147	0.9277	0.178						
PVI	−0.0052	0.023	−0.0032	0.013					−0.0129	0.013
GOV	−0.1613	0.211	0.0242	0.162					0.1769	0.147
YOUNG-PER	1.8812	3.139	2.1392	3.243					7.5838	3.729
MID-PER	−3.0933	1.980	−1.3683	2.202					−4.8568	2.069
OLD-PER	3.0638	3.156	5.6576	3.417					9.4998	4.017
UNEMP	0.0220	0.163	0.0164	0.043					0.2681	0.055
LOGINC	0.7350	0.629	1.4159	0.643					0.9443	0.678
YOUNG-PER × DEM-IDX	0.5519	5.420	−8.6303	6.432					4.5446	4.348
MID-PER × DEM-IDX	6.8406	3.768	5.1946	5.722					−2.8543	2.147
OLD-PER × DEM-IDX	−11.7866	5.643	−8.6161	8.173					3.6750	3.988
UNEMP × DEM-IDX	−0.1085	0.073	−0.2230	0.107					0.0102	0.050
LOGINC × DEM-IDX	−1.0671	0.853	0.2147	1.090					−1.3462	0.615
PNCAN	0.0142	0.078	0.0941	0.050					0.0136	0.017
ANCAN	−0.0516	0.038	−0.2016	0.038						
CONST	−7.7923	6.539	−5.4341	6.836					−1.6778	7.378

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 100 nonparametric bootstrap at the election level.

primary voter model, 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 0.13, while that coefficient almost doubles and is about 0.23 for the Republican challenger. On the other hand, the candidate-specific covariates present economically significant effect of Democratic challengers. This probably implies that for Republican primaries, voters appreciate the money spent by candidates more; in contrast, for Democratic primaries, candidate own feature is more important for voters (such as their policy positions). Further in primary elections, 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 0.51 for open-seat general elections, while I estimate the coefficients to be 0.13 and 0.14 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. Lastly, the incumbency advantage exists across different elections as in the benchmark model, but the effects are insignificant in this sensitivity analysis.

A4.4 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 informational set (Z, X, u), i.e., $Med(v|Z, X, u) = Med(v)$; 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 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.4: Sensitivity analysis: quadratic cost function

	Voter		Entry	
	Estimate	Sd. error	Estimate	Sd. error
κ	0.0125	0.011		
ζ	0.9250	0.159		
$Med(v)$	40158810	2951477		
λ_I			4.1230	0.299
λ_C			5.4181	0.308
FEMALE			0.1609	0.995
EXPC			2.1357	1.679
INCUM			1.6066	6.401
PVI			−0.0110	0.406
GOV			−0.0157	1.594
YOUNG-PER			−4.4285	5.200
MID-PER			−0.5352	3.384
OLD-PER			6.8500	5.005
UNEMP			0.0290	0.577
LOGINC			1.7472	1.081
PNCAN			0.0122	0.781
CONST			−6.8290	7.978

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 100 nonparametric bootstrap at the election level.

Table A.4 presents the estimates of the sensitivity analysis. The estimated constant $Med(v)$ is around 40 million, which is significant with an estimated standard error being about 3 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 since the estimated quadratic term is insignificant, with $\hat{\zeta} = 0.93$, 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.

A4.5 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:

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

for the primary election, and

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

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.5 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 much larger in the primary 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 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 candidate-specific

Table A.5: Sensitivity analysis: alternative voter model specification

	Voter: primary election		Voter: general election		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
γ	0.0750	0.018				
ω	0.0022	0.000				
λ_I					5.0169	0.355
λ_C					5.3747	0.291
FEMALE	0.3122	0.064	0.2134	0.165	0.4621	0.027
EXPC	0.9163	0.086	0.6077	0.114	2.0261	0.035
INCUM	0.5586	0.142	0.5603	0.085	2.3122	0.155
PVI					0.0233	0.136
GOV					−0.5069	1.308
YOUNG-PER					−4.8928	4.575
MID-PER					−0.4316	2.877
OLD-PER					6.8828	4.307
UNEMP					0.3015	0.215
LOGINC					1.6818	0.771
PNCAN					0.0888	0.156
CONST					−6.9784	8.079

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. Sd. errors are obtained through 100 nonparametric bootstrap at the election level. The estimates and Sd. errors of γ and ω are scaled up by 10^6.

characteristics now have larger effects in the voter model relative to the benchmark specification. 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 coefficients on the explanatory variables have the same signs and similar magnitude as in the benchmark model.

A4.6 The Super PAC Influence

In this sensitivity analysis, I introduce an indicator to denote the year of 2010, when the Supreme Court ruled in *Citizens United v. Federal Election Commission*, a controversial decision that reversed campaign finance restrictions and enabled corporations and other outside groups to spend unlimited funds on elections. The creation of a new type of political action committee (PAC), called the Super PAC, which can receive unlimited contributions from one single individual.

The existence of Super PACs certainly affects Senate elections, such as voters, campaign spending, as well as entry behavior of potential candidates systematically. Hence, I conduct an alternative exercise adding a dummy variable that is assigned one for the period of post-2010 and zero for that of pre-2010 to the specifications in the model (denoted as "SPAC").

Table A.6: Sensitivity analysis: Super PAC Indicator

	Spending: selection equation		Spending: log(spending) equation		Voter: primary election		Voter: general election		Entry	
	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error	Estimate	Sd. error
σ_u	0.9213	0.299								
σ_2	2.2493	0.460								
σ_{12}	1.3465	0.214								
$\gamma_{pre-2010}$					0.1144	0.009				
$\gamma_{post-2010}$					0.0939	0.007				
$\omega_{pre-2010}$							0.0829	0.260		
$\omega_{post-2010}$							0.0610	0.095		
λ_I									4.1571	0.428
λ_C									5.2565	0.576
FEMALE	0.4852	0.225	0.3614	0.181	0.2577	0.048	0.2188	0.379	0.1653	0.201
EXPC	1.5904	0.204	2.7036	0.200	0.6091	0.064	0.4931	0.398	2.1078	0.239
INCUM			2.1142	0.155	0.4938	0.078	0.5209	0.475	1.5817	0.215
SPAC	−0.0281	0.982	−0.9722	0.931					1.0477	1.211
OPEN	0.2760	0.135	0.9910	0.157						
PVI	−0.0034	0.012	−0.0008	0.011					−0.0113	0.083
GOV	0.0229	0.141	0.0485	0.161					−0.0153	0.285
YOUNG-PER	2.0975	5.347	4.4603	5.273					0.4273	7.148
MID-PER	−1.4415	4.474	3.7068	4.432					4.5578	5.593
OLD-PER	−0.0790	4.832	7.9431	5.048					10.6776	5.671
UNEMP	−0.0011	0.046	−0.0151	0.046					0.1046	0.253
LOGINC	0.4470	0.505	1.5333	0.648					1.6289	0.851
PNCAN	0.0192	0.037	0.1017	0.036					0.0080	0.157
ANCAN	−0.0585	0.152	−0.2122	0.039						
CONST	−4.5798	6.571	−8.7455	7.535					−9.4804	8.211

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. The variable "SPAC" equals one for post-2010 period, and zero for pre-2010 period. Sd. errors are obtained through 100 nonparametric bootstrap at the election level.

I present the results of this sensitivity analysis in Table A.6. 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 Super PAC in 2010, I focus on the coefficients on the variable "SPAC" in the corresponding specifications for the spending and entry models, as well as the estimated parameters in the voter model. The dummy variable has a negative coefficient in both the selection equation and the level equation of the spending model, implying that after 2010 actual candidates are less likely to spend non-zero in campaigns, and make smaller amount of campaign spending, which might indicate a "crowd-out" effect of Super PACs' spending on candidate spending. However, the effects in the spending model are insignificant, indicating a noisy influence of Super PACs on candidate's own campaign spending. This dummy variable has an insignificant positive effect in the entry cost specification, showing a possible growth of entry costs after the introduction of Super PACs after 2010.[37] Probably the most salient effect of Super PACs is on the voter model, especially in the general elections. The candidate spending's influence on votes in primary and general elections is smaller after 2010, and voters care more about candidates' own traits when voting. In contrast with primaries, candidate's role in general elections is fairly noisy in the voter model. This is probably because Super PACs specialize in independent expenditures, so they are more effective than candidates or parties in general elections.

37 These imprecise effects of Super PACs are in line with the noisy results that are found in Cox (2020).

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.

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.

Boosey, L., Brookins, P., and Ryvkin, D. (2020). Entry in Group Contests. *Working Paper*, Florida State University, University of South Carolina, and Florida State University.

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, C. (2020). Campaign Finance in the Age of Super PACs. *Working Paper*, Michigan State University.

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.

Gilens, M., Patterson, S., and Haines, P. (2021). Campaign Finance Regulations and Public Policy. *American Political Science Review*, *115*(3), 1074–1081.

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.

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. (2021). Structural Analysis of Tullock Contests with an Application to U.S. House of Representatives Elections. *International Economic Review*, *62*(3), 1011–1054.

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

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.

Jia, H., and Sun, C. (2021). The Optimal Entry Fee-Prize Ratio in Tullock Contests. *Journal of Mathematical Economics*, *96*, 102515.

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.

Liu, B., and Lu, J. (2019). The Optimal Allocation of Prizes in Contests with Costly Entry. *International Journal of Industrial Organization*, *66*, 137–161.

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.

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.

Shakhgildyan, K. (2021). Nonparametric Identication and Estimation of Contests with Uncertainty. *Working Paper*, Bocconi University and IGIER.

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.

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.