Essays on the Econometrics and Empirics of Bayesian Games

By

Jun Zhao

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

June 30th, 2021

Nashville, Tennessee

Approved:

Tong Li, Ph.D.

Joshua D. Clinton, Ph.D.

Atsushi Inoue, Ph.D.

Mattias Polborn, Ph.D.

Pedro H.C. Sant'Anna, Ph.D.

To my loving family for their endless love, support, and encouragement.

Acknowledgments

I would like to extend sincere thanks to the many people, in many places, who generously contributed to the work presented in this dissertation. First, I wish to thank my caring and enthusiastic supervisor, Tong Li, from whom I received enormous help and learned so much. Without his invaluable encouragement and guidance, I would not have made it. I am also appreciative to Joshua D. Clinton, Atsushi Inoue, Mattias Polborn, and Pedro H.C. Sant'Anna, who served as my dissertation committee members and provided great insights and comments on my research. I am grateful for feedback received from Peter Bils, Eric Bond, Mario J. Crucini, Andrew Daughety, Gregory Huffman, Jennifer Reinganum, Joel Rodrigue, Andrea Moro, John Weymark, and Matthew Zaragoza-Watkins during the presentation of my research. I am also appreciative to Tong Li, Pedro H.C. Sant'Anna, Jun Zhang, Dakshina G. De Silva, and Jun Nakabayashi, who are my coauthors in multiple papers. I learned a lot of innovative ideas, state-of-the-art techniques, positive attitudes, and teamwork spirits from them. I also thank my M.A. thesis advisor, Ruilong Yang, who inspired me to pursue a Ph.D., as well as Yang Song and Sanxi Li at Renmin University of China who supported me a lot during the Ph.D. application. Many thanks go to my dear friends, Tam Bui, Elijah Coleman, Thu Tran, Ji Hye Heo, Trang Hoang, Philip Huang, Ebehi Iyoha, Terry Kim, Jiatong Li, Zizheng Luo, Yukun Ma, Kayleigh McCrary, Oscar O'Flaherty, Qi Xu in the Vanderbilt Economics Department. Finally, but by no means least, this dissertation would not have been possible if not for the push and encouragement from my family. To my husband, Carl, and our cat, Dill: your love and understanding helped me through the dark times, especially when working from home. My mom and dad, Huiye Ren and Gaochang Zhao, are the most wonderful parents and earned this degree along with me.

Table of Contents

List of Tables

Table **Page**

List of Figures

Figure **Page**

Chapter 1: Campaign Spending in Senate Elections: A Structural Analysis of a Two-Stage Election with Endogenous Candidate Entry

1.1 Introduction

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

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

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

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

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

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

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

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

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

This paper is related to several branches of the literature on empirical games and political economy. First of all, the model of candidates' endogenous entry builds on an extensive literature on the

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

structural analysis of auction models with entry, see Li and Zheng (2009), Athey, Levin, and Seira (2011), Krasnokutskaya (2011), Roberts and Sweeting (2013), and Gentry and Li (2014), among others. However, political competition contest models are fairly under-developed. Theoretically papers by Fu, Jiao, and Lu (2015) and Gu, Hehenkamp, and Leininger (2019) model players' entry in contest models with complete information. On the empirical side, this paper is the first one to construct a two-stage contest model with the first stage of entry and use it to analyze the campaign spending, to the best of my knowledge. Prior studies either treat the entry decision as an auxiliary component (Kawai and Sunada (2015)), or use the reduced-form analysis to study political entry (Avis, Ferraz, Finan, and Varjão (2017)).

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

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

This paper also adds to the research using structural models to study political phenomena. For instance, Diermeier, Eraslan, and Merlo (2003) examine the impacts of institutional features on the government formation process and the stability of the government in West European parliamentary democracies. Deltas, Herrera, and Polborn (2016) analyze the trade-off between learning and co-ordination among voters, through estimating a sequential election model in the U.S. presidential primary system. There is a growing literature studying a class of dynamic games to evaluate electoral

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

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 1.2, I introduce the background and the data, along with the main variables I use for the later empirical analysis. Section 1.3 characterizes the two-stage game-theoretic model and establishes existence of the model equilibrium. Section 1.4 is devoted to the structural analysis of the data. Section 1.5 provides a set of counterfactual analyses. Section 1.6 concludes. Mathematical proofs, the detailed estimation approach, and various sensitivity analyses are gathered in the Appendix.

1.2 Background and Data

1.2.1 Data

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

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

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

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

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

7 The FEC website: www.fec.gov.

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

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

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

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

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

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

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

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

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

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

Lastly, I correct for the endogeneity of campaign spending when estimating the voter model by controlling for the election-specific unobserved heterogeneity, because the failure to account for the unobserved heterogeneity will lead to biased estimates.[11] In order to identify the unobserved

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

heterogeneity following the identification strategies in Matzkin (2003) and Guerre, Perrigne, and Vuong (2009), I introduce additional variable or instrumental variable (IV): *lagged campaign spending*, with the exact variable being the total spending by candidates in the *previous* Senate general election in the same state.[12] Due to the staggered nature of the Senate elections, Senators in one state are elected to six-year terms. Thus, the previous Senate election in one state for the other Senate seat rarely involves the same incumbent or challenger as the current race. Therefore, the *lagged campaign spending* affects the candidates' entry and spending behavior only through the correlation between this lagged variable and the latent fundraising environment or the underlying political climate in one state.

Table 1.1 presents the summary statistics of the data. Around 69% candidates have strictly positive campaign spending.[13] The average amount of campaign spending is about $4 million, which includes all disbursements by a representative candidate during the election. The entry proportion, calculated as the ratio of the number of actual candidates to the number of potential candidates for each election, is about 0.78 on average, meaning that on average more than three quarters of the potential candidates will decide to run for the elections. Among all 306 elections in the data, only 18% are open-seat elections, implying that incumbents are more likely to participate in the elections. The variable PVI is assigned positive (negative) value if the state under investigation leans toward the Republican (Democratic) party. The lagged spending variable has a larger sample mean than the spending variable, which makes sense because the lagged spending is defined as the sum of campaign spending by the general election candidates in the same state's previous Senate election, while the spending is defined for all candidates.

1.2.2 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 1.1 presents the (unconditional) spending distributions of incumbents and challengers in incumbent-challenger elections and that of candidates in open-seat elections. In incumbent-challenger elections, the spending density of incumbents is shifted rightward from that of challengers, indicating

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

13 Candidates with zero campaign spending still receive strictly positive (but small) vote shares. When modeling candidates' spending behavior in Section 1.3, 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 1.1: Summary statistics of candidate- and election-specific variables

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

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

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

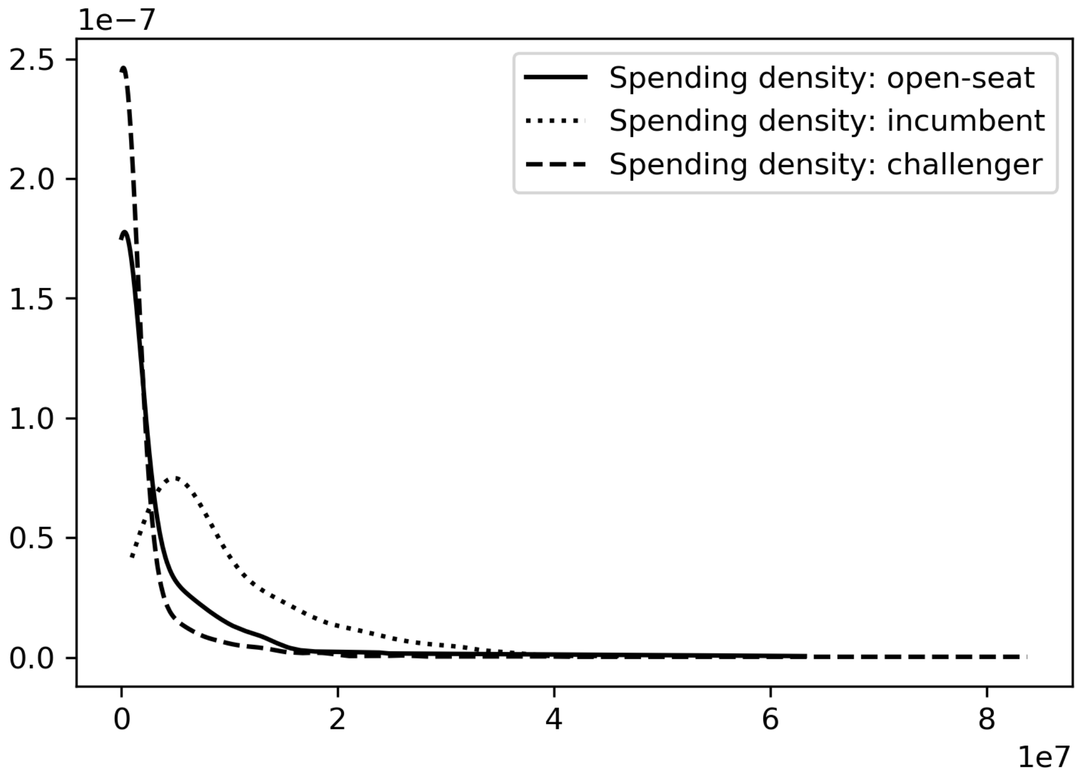


Figure 1.1: Spending distributions of incumbents and challengers in incumbent-challenger elections, as well as candidates in open-seat elections

Turning to the entry behavior of potential candidates, Table 1.2 shows the means and standard deviations of entry decisions for incumbents and challengers, where the entry decision is a dummy that equals 1 if a potential candidate decides to run for the race and 0 otherwise. Clearly, reelection probabilities in the Senate are indeed high, even unconditionally. However, there is still substantial selection among challengers in terms of who run for office.

Although these observations regarding the entry and spending behavior of candidates are important, they call for a further structural analysis if one wants to unfold the private information of actual and potential candidates, such as their private values of holding the office and entry costs of running for

Table 1.2: Summary statistics of entry dummies for incumbents and challengers

	Observations	Mean	SD
Incumbent	306	0.8170	0.387
Challenger	1977	0.7228	0.448

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

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.

1.3 The Model

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

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

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

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

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

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

the election-level heterogeneity (X, u), each candidate's entry cost and private value are independent from each other.

1.3.1 Voters

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

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

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

$$P(D_i R_k) = \frac{1}{\sum_{j \neq i}^{n} exp(-W_{D_j}) + 1 + exp(-W_G) \left[\sum_{l \neq k}^{m} exp(-W_{R_l}) + 1\right]} \cdot \frac{1}{\sum_{l \neq k}^{m} exp(-W_{R_l}) + 1}, \quad (1.3.1)$$

where I define the differences as

$$W_{D_j} \equiv M_D(D_i) - M_D(D_j), \; for \; j = 1, \cdots, i-1, i+1, \cdots, n;$$

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

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

and the detailed derivation can be found in Appendix 1.7.1.[19]

The probability of a Democratic candidate D_i winning both the Democratic primary and the general election with the Republican candidate R_l as the opponent who wins the Republican primary, denoted as $P(D_i R_l)$, 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 $P(D_i) \equiv \sum_{k=1}^{m} P(D_i R_k)$, 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 1.4 I specify these functions to depend on the campaign spending of the candidate e_{D_i} (e_{R_k}), coupled with the election-level heterogeneity (X, u). Further, suppose that each voter's latent utility from electing a candidate is increasing with diminishing returns in the candidate's campaign spending.[20] Finally, the probability of the candidate D_i winning the office is denoted as $CSF(e_{D_i}; e_{-D_i}) = P(D_i)$, meaning this is a CSF that depends on this candidate's campaign spending e_{D_i} and the spending profile of the rest actual candidates in the Democratic and Republican parties e_{-D_i}.

1.3.2 Candidates: Election Stage

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

19 I also derive the probability where the turnout of voters in the general election is taken into account (see Appendix 1.7.1).

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

$g''(\cdot) \geq 0$. Thus her maximized expected payoff when she spends e_i can be written as:

$$\pi_{D_i}(v_{D_i}|a_{-D_i}) \equiv \max_{e_{D_i}} \quad v_{D_i} \cdot \mathbb{E}_{e_{-D_i}}[CSF(e_{D_i}; e_{-D_i})|v_{D_i}; a_{-D_i}] - g(e_{D_i}), \tag{1.3.2}$$

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.[21]

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

$$v_{D_i} \cdot \mathbb{E}_{e_{-D_i}}\left[\frac{\partial CSF(s_{D_i}(v_{D_i}); e_{-D_i})}{\partial e_{D_i}}\middle| v_{D_i}; a_{-D_i}\right] \leq g'(s_{D_i}(v_{D_i})), \tag{1.3.3}$$

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 (1.3.3) in equivalent terms.[22] Intuitively, the FOC balances the (monetary) value of increase in the winning probability relative to the marginal cost of campaign spending.

1.3.3 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

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

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

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 $p_{D_i} = Pr(C_{D_i} < \Pi_{D_i})$.

1.3.4 Equilibrium

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

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

Proof. See Appendix 1.7.2. ■

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)).

1.4 Structural Analysis

I estimate the game-theoretic model proposed in the previous section using data on the vote shares of actual candidates and the campaign spending of potential candidates in the U.S. Senate elections during 1994 and 2018. The goals are to recover the entry cost distribution $H(\cdot)$ and the private value distribution $F(\cdot)$ in the candidate model, and to estimate the voter model. I adopt a fully parametric

23 The uniqueness of the model equilibrium is hard to prove. As in Krasnokutskaya (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.

approach, and assume that I have an i.i.d. sample of elections, indexed by l for $l \in \{1, \cdots, L\}$.

1.4.1 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:

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

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

The cost function of the campaign spending $g(\cdot)$ in the expected payoff expression (1.3.2) is assumed to be a linear function such that $g(e) = e$, as in He and Huang (2018). This leads to a cost

---

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

25 In the benchmark model, I consider the case where there is no upper boundary for the campaign spending. Later in the sensitivity analysis, I consider an extension of the benchmark model, where the campaign spending is supposed to be bounded from above by a finite number, which is nonparametrically estimated following Guerre, Perrigne, and Vuong (2000), see Appendix 1.7.5.

derivative function being a constant 1 appearing in the FOC (1.3.3), which is used to back out the private value distribution of actual candidates after obtaining the estimates of the campaign spending distribution.[26]

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

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

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

for the primary election, and

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

given that this candidate i_l wins the primary election and proceeds to the general election. Note that adding the first three parts in the latent utility together yields $M_D(D_i)$ in Section 1.3.1, taking the Democratic candidate D_i in the primary election as one example. To reduce the burden on notations, here I pool the candidates from both parties together. In the voter model, X_l^P may be different from X_l^G. An econometric issue arises when analyzing the voter model: the coefficients γ_X and ω_X cannot be identified. This is because when a voter compares two candidates who attend the same election (either primary or general), the parts associated with election-level covariates are cancelled out and only the difference of the latent utility is crucial.[28] Therefore the terms $X_l^{P\prime}\gamma_X$, $X_l^{G\prime}\omega_X$, and u_l being the same for all candidates in an election are cancelled out, while I still control for the election-level heterogeneity (both observed and unobserved).

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

26 In Appendix 1.7.5, 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.

27 In the benchmark model, I do not distinguish between the two parties. However in the sensitivity analysis, I consider a case with the difference between Democratic and Republican parties (see Appendix 1.7.5).

28 Recall that in Equation (1.3.1), I model the probability $P(D_iR_k)$ through the differences W_{D_j}, W_{R_l}, and W_G.

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

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

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

---

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

30 In the sensitivity analysis (see Appendix 1.7.5), I consider an extension of the benchmark model, where with the availability of the turnout information of the voters for each state's general election, I generate the probability of a candidate winning both the primary and general elections, in which the parameter ω_X can be identified and estimated.

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

1.4.2 Estimation Method

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

I adopt a multi-stage estimation method for the structural model, given the specifications.[32] In the first step, I regress the $log(lagged\ spending)_l$ on X_l^{IV} to estimate η via the ordinary linear squares regression (OLS), and obtain the residuals as estimates for u_l, the election-level unobserved heterogeneity. In the second step, together with the estimated unobserved heterogeneity for each election, I use the data on campaign spending of all candidates who actually run for the Senate elections to estimate the parameters in the spending model: α, β, σ_2, and σ_{12}, through the maximum likelihood estimation (MLE). In the third step, I use MLE to estimate the parameters in the voter model, i.e., γ and ω. In order to recover the private value distribution, I follow Jofre-Bonet and Pesendorfer (2003) and Athey, Levin, and Seira (2011) to estimate the value density after obtaining the pseudo private value from the FOC (1.3.3) and the simulated campaign spending. With the estimates of the spending and voter models, together with the estimated unobserved heterogeneity, I simulate the expected payoff for each potential candidate at the entry stage, and follow Li and Zhang (2015) to estimate the distribution of the entry cost, in order to get the estimates for δ and two λs. Therefore, the parameter vector being estimated is defined as $\theta = (\eta; \alpha, \beta, \sigma_2, \sigma_{12}; \gamma, \omega; \delta, \lambda_I, \lambda_C)$.

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

1.4.3 Estimation Results

Table 1.3 summarizes the results of the structural analysis.[33] The first two columns show the estimates for the selection part of the campaign spending model together with the covariance matrix of the two errors, thus $(\hat{\alpha}, \hat{\sigma}_2, \hat{\sigma}_{12})$. The third and fourth columns correspond to the log(spending) equation of the campaign spending model, thus $\hat{\beta}$. The estimated standard deviation of the error in the

32 The details of the estimation method can be found in Appendix 1.7.3.

33 The estimates of the first stage (i.e., $\hat{\eta}$) are reported in Appendix 1.7.4, which are of second-order importance.

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

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

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

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

The last two columns give the estimates in the entry model, thus $(\hat{\delta}, \hat{\lambda}_I, \hat{\lambda}_C)$. First of all, incumbents have larger expected entry costs, compared to challengers. In footnote 15, I model the entry stage of potential candidates by referring entry cost to reservation value. This interpretation helps explain the higher entry probabilities by incumbents, which are consistent with the observed data. Specifically, incumbents may hold higher expectation regarding the value of holding the office, which turns to larger reservation value meaning that only when the expected payoff of winning the office exceeds the reservation value, they will decide to run.[35] Therefore, in Section 1.2.2, the reason that incumbents have higher entry probabilities is not because they have lower expected entry costs, but rather they can have higher expected payoffs if they win the office. Further, the standard deviation of entry costs for incumbents is marginally smaller than that of challengers in the entry cost distribution, which suggests that incumbents have a more concentrated entry cost distribution. Thus although incumbents have larger entry costs compared to challengers, among incumbents the difference of entry costs is not as large as that among challengers. Most of the estimated coefficients of the covariates have the expected signs, but the effects are insignificant. Specifically, the effect of the number of potential candidates on expected entry costs is positive (though insignificant). However since changing this number will affect the number of actual candidates that further has a significant effect on campaign spending, which in turn influences the entry behavior of potential candidates, it is still uncertain how the number of potential candidates affects outcomes in the entry stage, which will be studied in Section 1.5.1.

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

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

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

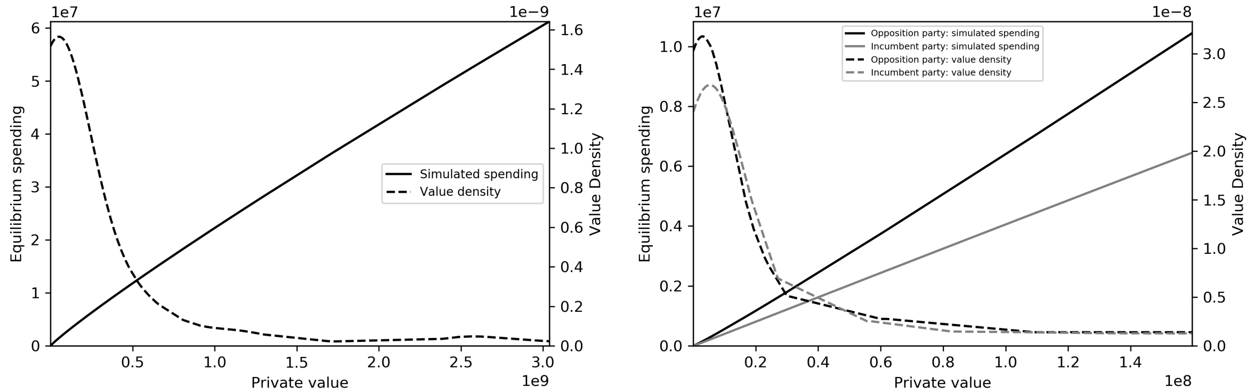
Table 1.3: Estimation results from structural analysis

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

Notes: For explanatory variables, estimated coefficients in corresponding specifications are reported. The estimated results are obtained through the estimation method described in Section 1.4.2. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

from the private value to the campaign spending in a given Senate election. This equilibrium strategy function can be different depending on the status of a candidate, whose private value following a type-specific distribution. I simulate the equilibrium strategy functions and the private value densities for the incumbent, the challenger from the opposition party, and the challenger from the incumbent party, using 200 draws, for the election indexed 1 in the data as a representation (this is not an open-seat election), shown in Figure 1.2. I also simulate the equilibrium strategy functions and the private value densities for the candidates from two parties in an open-seat election (indexed 19) as shown in Figure 1.3.

Figure 1.2 presents the equilibrium strategy functions and the private value densities in an incumbent-challenger election. Apparently, the equilibrium strategy function is strictly increasing for all candidates. Comparing the curves for the incumbent and for the two types of challengers, it is clear that incumbents value the office more than the challengers, and thus are willing to spend more to raise their winning probabilities. As the sub-figure 1.2b indicates, the distribution of the opposition party challenger's value is slightly shifted leftward from that of the incumbent party challenger's value, which indicates that on average the challengers in the incumbent party have slightly larger values



(a) Incumbent

(b) Challenger

Figure 1.2: Simulated strategy functions and value distributions for incumbent and challengers in a representative incumbent-challenger election

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

I now turn to the open-seat election. Figure 1.3 gives the equilibrium strategy functions and the private value distributions for candidates from the incumbent and opposition parties. Although there is no incumbent participating in the election, I still define the parties as above for the ease of illustration and comparison. Now the distribution of the opposition party candidate's value is slightly shifted rightward from that of the incumbent party challenger's value, but the two distributions are still comparable. This is probably because without the incumbent, the challengers from the opposition party now have larger values in a battleground. The scales of private value and campaign spending of candidates in an open-seat election are comparable to those of incumbents, and much bigger than those of challengers in an incumbent-challenger election. Although there are differences between candidates from both parties, the differences are not large proportionally when compared to the case with the existence of an incumbent. In this representative open-seat election, it appears that candidates in the opposition party spend relatively more than those in the incumbent party. This can be potentially explained by the impression of candidates from the opposition party that it is possible to flip the state if the incumbent decides not to run, because the state may be at the margin.

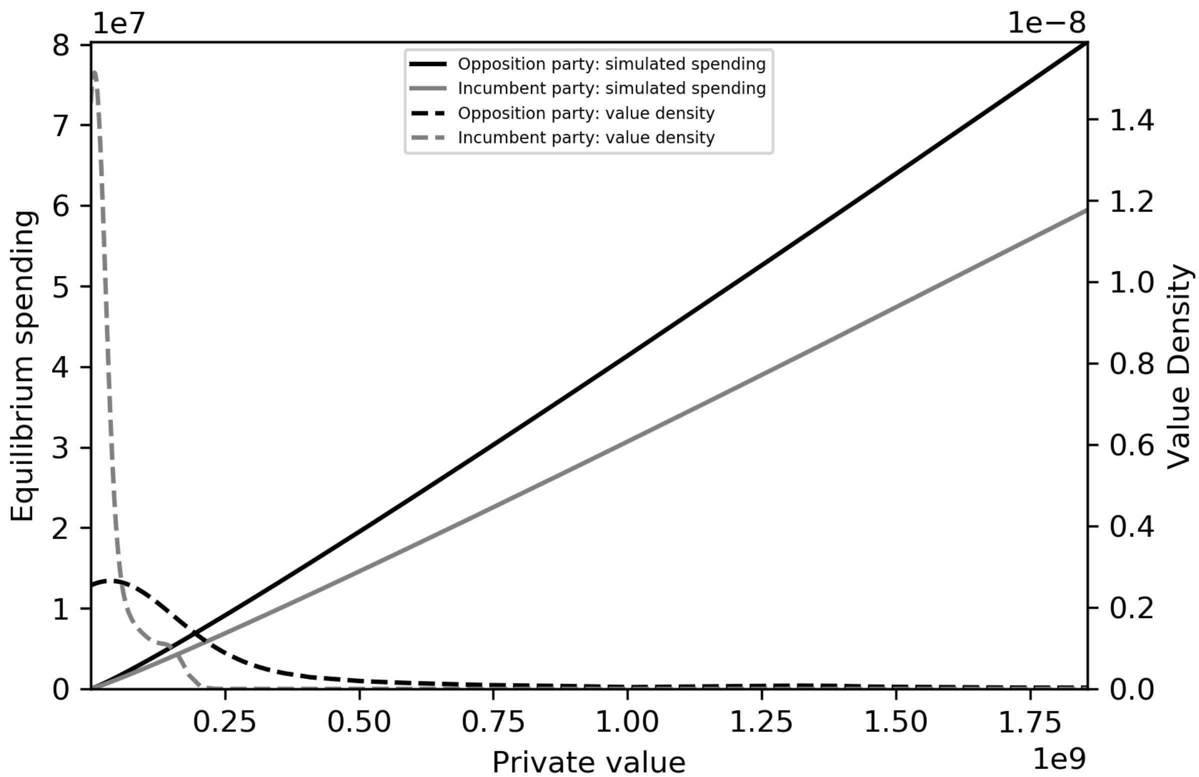


Figure 1.3: Simulated strategy functions and value distributions for candidates in a representative open-seat election

In Appendix 1.7.5, I conduct a series of extensive sensitivity analyses, including considering the turnout of voters in the general elections (Appendix 1.7.5), a truncated spending distribution from above (Appendix 1.7.5), taking into account the difference between two political parties (Appendix 1.7.5), a quadratic cost function (Appendix 1.7.5), an alternative voter model specification (Appendix 1.7.5), including time-specific constants in all specifications (Appendix 1.7.5), as well as adding a dummy variable that is assigned one for the period of post-2008 and zero for that of pre-2008 to the specifications in the model (Appendix 1.7.5). I find the results are robust to these sensitivity analyses.

1.4.4 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 1.4. The predicted first moments of some key features of the model are reported in order to compare with the observed counterparts through 500 simulations, together with the simulated standard errors. The overall fit of the predicted data to the actual data is good in both the level and the trend.

Table 1.4 shows the actual and predicted moments regarding the spending model, the voter model, and the entry model. In particular, the campaign spending distribution simulated from the estimated structural parameters fits the observed data well. The spending made by incumbents is a bit larger in the simulation, but in both predicted and actual data the spending of incumbents is always larger than that of challengers. For the voter model, I simulate the moments of vote shares of incumbents in the primary and general elections, to echo the specification where I attach importance to how incumbents behave differently from challengers in Senate elections. The vote shares of incumbents are on average larger in the within-party primaries, revealed in both predicted and actual data. This can also be reflected via the ratio of incumbent-challenger elections (either primary or general) incumbents win. I report the predicted means of the number of actual candidates, the entry proportion, and the proportion of open-seat elections for the entry model. Typically, I predict the entry patterns well, with slightly larger entry probabilities that lead to both bigger entry proportion and larger number of actual candidates. The percentage of open-seat elections is predicted to be slightly smaller, because I predict the campaign spending mean to be larger for incumbents that is associated with more expected payoff and higher entry probability of incumbents.

Table 1.4: Model fit results

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

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

1.5 Counterfactual Analyses

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

1.5.1 The Role of Entry

In the structural analysis, the number of potential candidates is exogenous as I define potential candidates as those who file reports about campaign financing and thus appear in the spending data. In this counterfactual analysis, I investigate the impacts of varying the number potential candidates on the spending, vote and entry results. To compare behavior in alternative environments, I do so for representative elections. I first select two elections with different degrees of asymmetries between two parties: Election one has 9 challengers in both incumbent and opposition parties (indexed 21 in the data); Election two has 19 challengers in the incumbent party and 10 challengers in the opposition party (indexed 107). I assess the effects of decreasing (and increasing) the numbers of potential candidates in *both* incumbent and opposition parties (denoted as N and M) in the left panels of Figure 1.4 on spending outcomes, Figure 1.5 on vote outcomes, and Figure 1.6 on entry outcomes via simulations.

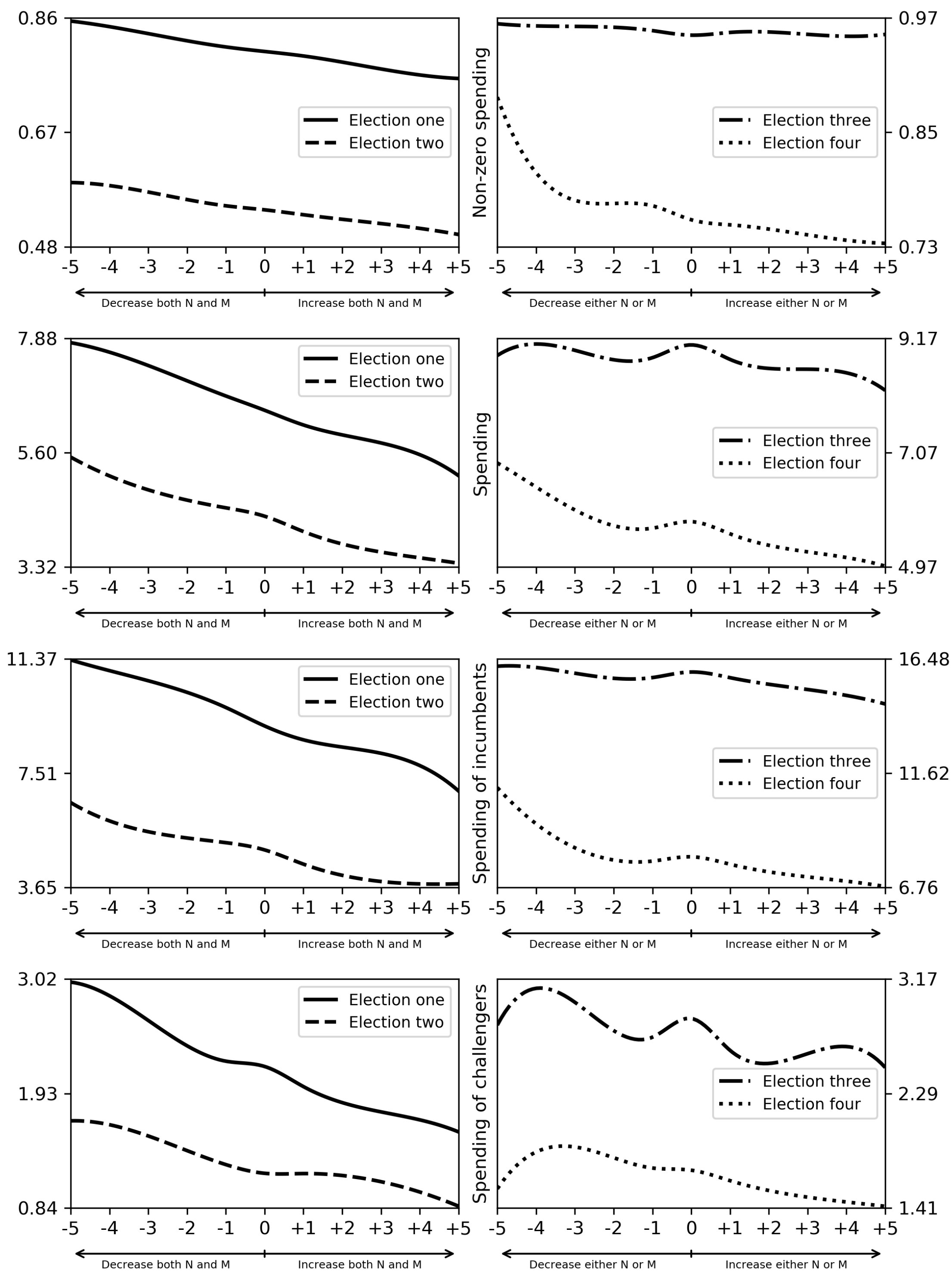


Figure 1.4: The effects of varying the number of potential candidates on spending outcomes (scaled down by 10^6)

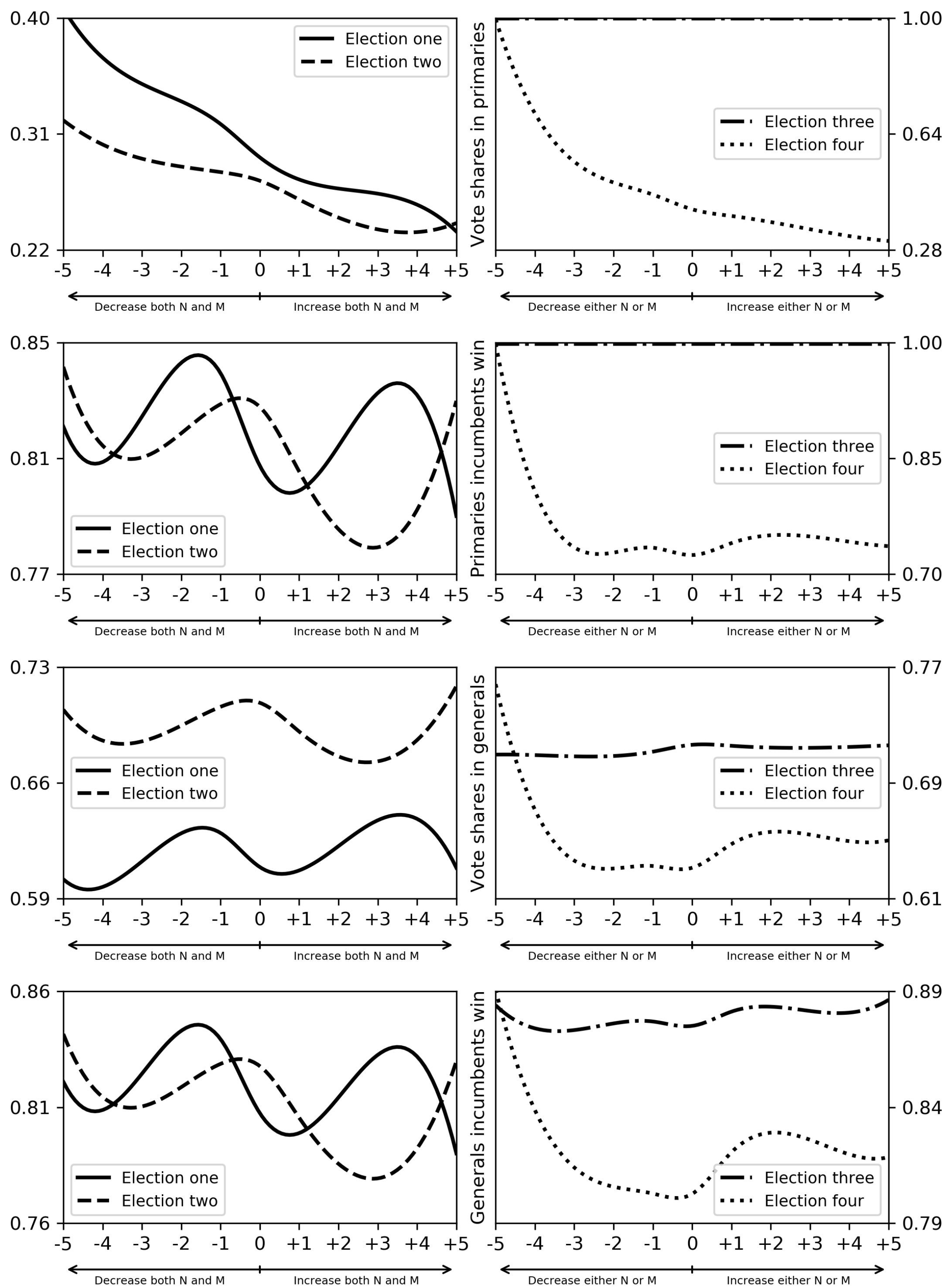


Figure 1.5: The effects of varying the number of potential candidates on vote outcomes for incumbents

Despite different degrees of asymmetries between two parties in Election one and Election two, the profiles present similar patterns regarding the spending, voting, and entry outcomes. When the numbers of potential candidates in *both* incumbent and opposition parties increase, the non-zero spending ratio and the spending level decrease on average, regardless of the candidate status, as shown in the left panel of Figure 1.4. The reason that the non-zero spending ratio decreases is that the candidate with any level of private type will decrease the spending, corresponding to a homogenous effect. Moreover, increasing the numbers of potential candidates in *both* incumbent and opposition parties has a negative effect on the entry proportion and a positive effect on the number of actual candidates, because more potential candidates lead to larger entry cost and smaller entry probability on average, which in turn results in larger number of actual candidates (thus a more competitive environment) when combined with more potential candidates (see the left panel of Figure 1.6). In contrast, the impact on the probability of being an open-seat election is not monotonic, reflecting the combined effects of smaller spending by both incumbents and challengers. Turning to the vote outcomes in the left panel of Figure 1.5 where I focus on the incumbent, increasing the numbers of potential candidates in *both* incumbent and opposition parties only has a decreasing effect on the vote share of the incumbent in the primary, due to more actual candidates in the election. However, the effects on the primary winning probability, the vote share in the general, and the general winning probability are volatile, consistent with the non-monotonic pattern on the probability of having an open-seat election.

Next, I investigate two new elections: Election three has no challengers in the incumbent party and 13 challengers in the opposition party (indexed 29); Election four has 5 challengers in the incumbent party and 15 challengers in the opposition party (indexed 40), and redo the above analysis. This time, I only decrease (and increase) the number of potential candidates in the opposition party for Election three, thus keeping the incumbent being unopposed in the primary. For Election four, I decrease (and increase) the number of potential candidates in the incumbent party. The simulated outcomes are exhibited in the right panels of Figure 1.4, Figure 1.5, and Figure 1.6.

The results indicate different effects of varying the number of potential candidates in the opposition or the incumbent party. If letting the incumbent be unopposed in the primary (as in Election three), increasing the number of potential candidates in the opposition party does not have notable or monotonic effects on the non-zero spending ratio and the average spending level. Because there is no challenger in the incumbent party, the primary vote share and the probability of winning primary for the incumbent remain 1 while the probability of having an open-seat election remains 0. The general

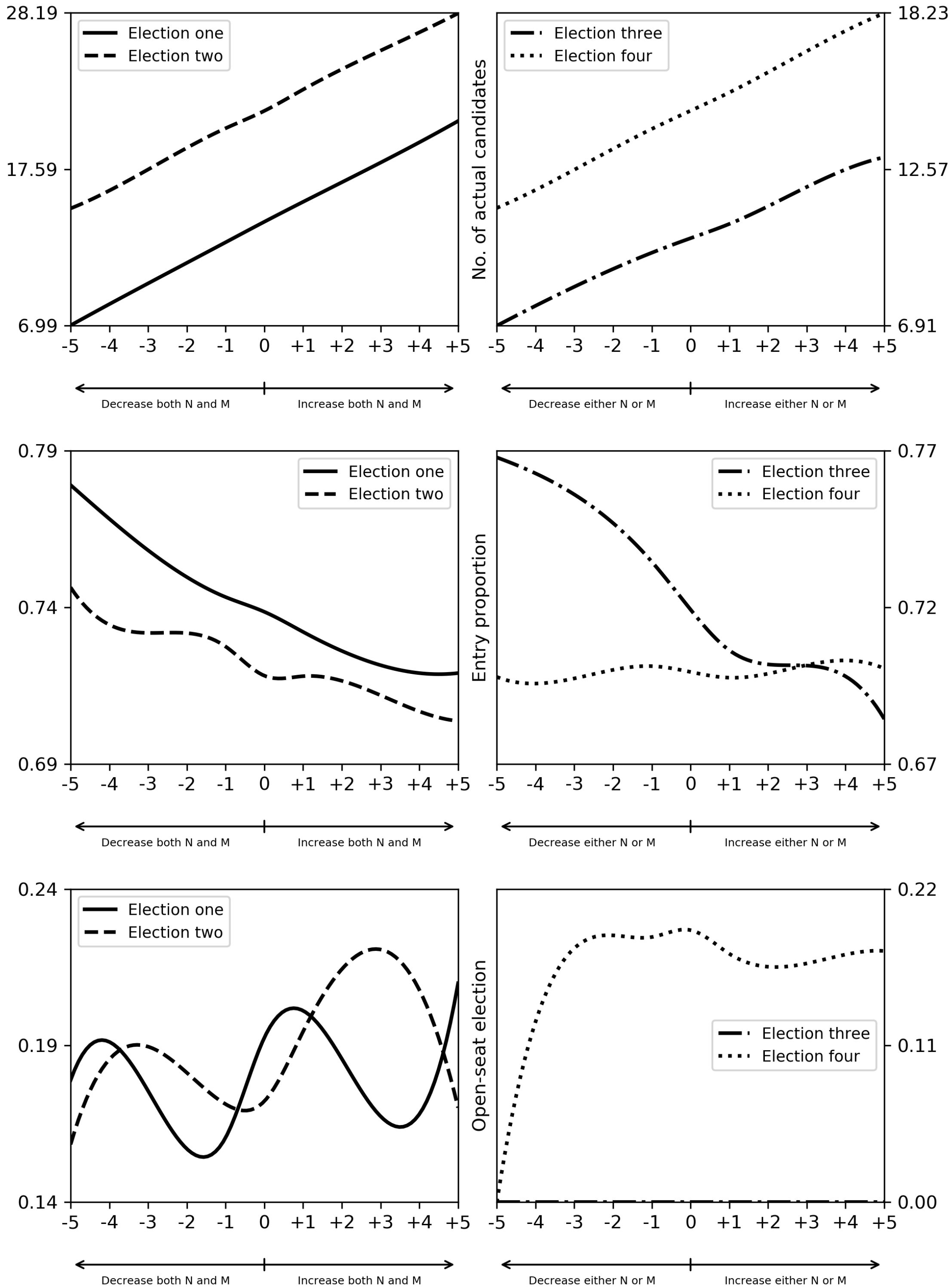


Figure 1.6: The effects of varying the number of potential candidates on entry outcomes

vote outcomes for the incumbent appear steady. Putting together, this implies stable entry behavior of the incumbent. In contrast, challengers experience non-monotonically decreasing spending and monotonically decreasing entry probability on average, resulting in decreasing entry proportion and increasing number of actual candidates on average. In summary, varying the number of potential candidates in the opposition party does not change the behavior of the incumbent much.

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

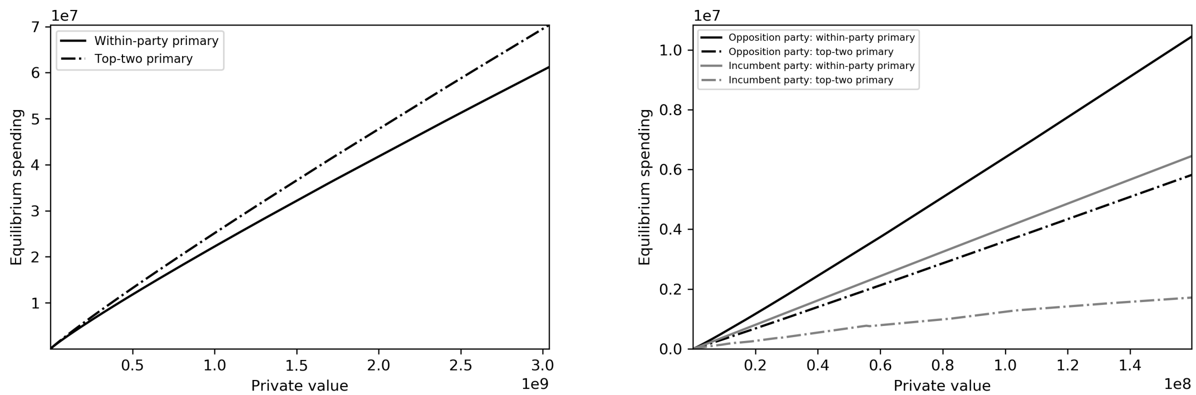
1.5.2 Top-Two Primary

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

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

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

As in Section 1.4.3, I present the equilibrium strategy functions for different types of candidates in a representative incumbent-challenger election and a representative open-seat election. Using the same incumbent-challenger election as in Figure 1.2, I draw the equilibrium strategy functions under the counterfactual setting in Figure 1.7.[38]



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

Figure 1.7: Simulated strategy functions in the top-two primary system for incumbent and challengers in a representative incumbent-challenger election

Figure 1.7 compares the simulated strategy functions in the counterfactual framework (lines of dashes and dots) and in the benchmark model (solid lines). It appears that the change of the spending behavior in the counterfactual is different depending on the candidate status. For incumbents, they spend more under the top-two primary system; while for challengers of both types, they spend less in the counterfactual equilibrium. The scale of change is also worth discussing. Although incumbents tend to spend more, the increasing magnitude is not large, compared to the decreasing magnitude of challengers, where they tend to spend much less. This is because under the top-two primary system, voters face more choices in the primary election, thus incumbency advantage may be amplified in the early stage of the election. Challengers from both parties think their chance of advancing to the general election is lower with the existence of an incumbent and more candidates, and consequently reduce their spending, because spending is less effective in the primary and they find it less beneficial to spend too much. On the other hand, the incumbent spends a bit more, in order to lock the win. However,

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

since the marginal effect of spending on winning is diminishing and the incumbent's spending level is already high, she will not spend too much more.

Similarly, using the same open-seat election as in Figure 1.3, I plot the equilibrium strategy functions under the counterfactual setting in Figure 1.8.

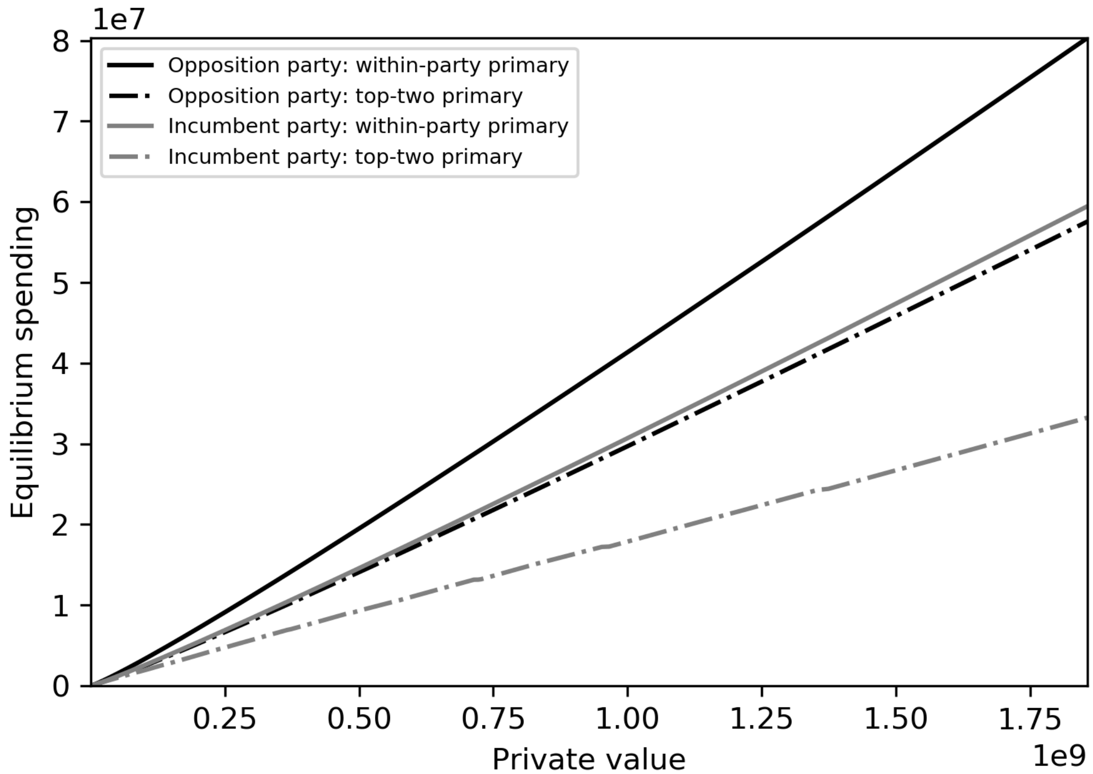


Figure 1.8: Simulated strategy functions in the top-two primary system for candidates in a representative open-seat election

Figure 1.8 depicts the simulated equilibrium strategy functions under the top-two primary system that are shown in lines of dashes and dots, where the equilibriums in the benchmark model are included via solid lines. In this case, candidates in open-seat election will spend less if the primary election takes the form of the top-two system. However, the magnitude of change is not large relative to the scale in the open-seat election.

I use the numerical routine to simulate the first moments of the same key features in the spending, voter, and/or entry models as in Section 1.4.4 under two different scenarios: (i) I hold entry fixed at the benchmark level that is observed from the data; and (ii) I allow entry to adjust with a different primary system where potential candidates decide whether they will run for office endogenously. Table 1.5 and Table 1.6 report the simulated results under these two scenarios.

Table 1.5: Counterfactual: top-two primary with fixed entry

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

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

First of all, if I study the campaign spending by incumbents and challengers separately, I find that on average incumbents spend more and challengers spend less in the counterfactual model regardless of fixed or endogenous entry. This is consistent with Figure 1.7 and Figure 1.8 above. However, the average campaign spending appears differently conditional on fixed or endogenous entry. With fixed entry, the average campaign spending increases under the alternative top-two primary system; while with endogenous entry, the average spending is smaller.[39] With fixed entry, the campaign spending of candidates regardless of the candidate status is on average larger than that when entry is endogenous. The reason is probably that with endogenous entry, candidates can choose to not run for office if they find it unworthy. But if entry is fixed, they can only increase their spending level more to offset some cost because they cannot secede.

As for the voter model, note that in the counterfactual model the incumbent who wins the top-two primary election will have the highest winning probability and receive the highest vote shares in the

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

Table 1.6: Counterfactual: top-two primary with endogenous entry

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

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

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

I next turn to the entry model in Table 1.6. Compared with the benchmark model, the entry proportion and the number of actual candidates are slightly smaller, while the ratio of open-seat elections is higher, all due to the smaller entry probabilities of candidates. The differences are rather modest compared to those in the spending model. This can be explained by the equilibrium behavior of candidates. In the counterfactual model, the candidate who seeks to maximize her expected payoff will adjust her spending level in order to balance between the expected chance of winning and the cost of the spending, upon entry. Therefore, although the equilibrium spending made by the candidate changes significantly regardless of the candidate status, the resulting change in the entry equilibrium is attenuated because the adjusted expected payoff's change is attenuated.

In conclusion, the top-two primary system may give voters more choices in the primary, but at the cost of reducing their choices in the general election, where usually more voters participate. From this perspective, the top-two primary system may not lead to more competition in the election. With fixed entry, I find that campaign spending increases on average despite that challengers reduce their spending levels, and that average winning probability of incumbents is marginally smaller compared to the benchmark model. However, when endogenous entry is accounted for, the results show that on average the top-two primary system leads to relatively less campaign spending of candidates, and

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

similar winning chance of incumbents. More importantly, with endogenous entry, this primary system seems to have a negative effect on the entry probabilities of candidates. Therefore, this implies that it is important to take into account the endogenous entry of potential candidates.[41]

1.6 Conclusion

In this paper, I develop and estimate a two-stage game-theoretic contest model to quantify the effect of campaign spending on electoral outcomes given endogenous candidate entry, using data on the U.S. Senate elections from 1994 to 2018. The model consists of two stages, with the first entry stage where potential candidates decide whether to run for the race, and the second election stage where actual candidates attend a Senate election with both within-party primaries and a general election and choose the spending in campaigns strategically. I also specify voters' decisions via a latent utility model that depends on candidates' campaign spending, demographic and economic covariates, and the election-specific unobserved heterogeneity. Taking the structural approach, I obtain estimates of parameters in the spending model for actual candidates, which can be used to simulate the equilibrium strategy function in the election stage. I also get the estimates of parameters in the voter model, which indicate how the campaign spending translates into votes. The estimated entry cost distribution is helpful to characterize the entry behavior of potential candidates.

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

Moreover, how the number of potential candidates influences campaign spending and vote outcomes depends on which party's number being changed. Increasing numbers of potential candidates in both

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

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

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

1.7 Appendix

1.7.1 The Derivation of CSF in the Voter Model

Under the assumptions in Section 1.3.1, 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:

$$P(D_i R_k)$$

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

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

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

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,

$$P(D_iR_k)$$

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

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:

$$
\begin{aligned}
&P(D_iR_k)\\
=&Pr\begin{pmatrix} M_D(D_i)+\iota_{D_i} > M_D(D_j)+\iota_{D_j},\ j=1,\cdots,i-1,i+1,\cdots,n\\ M_R(R_k)+\iota_{R_k} > M_R(R_l)+\iota_{R_l},\ l=1,\cdots,k-1,k+1,\cdots,m;\\ M_G(D_i)+\iota_{D_i} > M_G(R_k)+\iota_{R_k};\\ M_G(D_i)+\iota_{D_i} > \iota_{G_0} \end{pmatrix}\\
=&Pr\begin{pmatrix} \iota_{D_j} < \iota_{D_i}+M_D(D_i)-M_D(D_j),\ j=1,\cdots,i-1,i+1,\cdots,n;\\ \iota_{R_l} < \iota_{R_k}+M_R(R_k)-M_R(R_l),\ l=1,\cdots,k-1,k+1,\cdots,m;\\ \iota_{R_k} < \iota_{D_i}+M_G(D_i)-M_G(R_k);\\ \iota_0 < \iota D_i+M_G(D_i) \end{pmatrix}\\
=&Pr\begin{pmatrix} \iota_{D_j} < \iota_{D_i}+W_{D_j},\ j=1,\cdots,i-1,i+1,\cdots,n;\\ \iota_{R_l} < \iota_{R_k}+W_{R_l},\ l=1,\cdots,k-1,k+1,\cdots,m;\\ \iota_{R_k} < \iota_{D_i}+W_G;\\ \iota_{G_0} < \iota_{D_i}+M_G(D_i) \end{pmatrix}.
\end{aligned}
$$

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

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

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

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

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

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

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,

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

1.7.2 Proof of Proposition 1

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

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

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

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

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

and recall that the differences are defined as

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

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

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

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

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

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

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

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

1.7.3 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 1.4.1:

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

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 1.7.2 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:

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

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

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

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

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

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

1.7.4 First Stage Results

In this Appendix, I report the estimated results from the first stage regression of the IV, i.e., *log(lagged spending)* from the *previous* Senate election in the same state, on explanatory variables, which is modeled as a linear function $X_l^{IV'}\eta + u_l$. The main purpose of this estimation stage is to recover the pseudo unobserved heterogeneity at the election level, and the estimated coefficients $\hat{\eta}$ are presented in Table 1.7 to demonstrate how the election-specific variables affect the log of lagged spending from the previous Senate election.

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

Table 1.7: First stage IV regression results

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

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

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

1.7.5 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.

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 1.7.1). As in Section 1.3.1, 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 1.7.1, $P(D_iR_k)$ 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 1.3.1. 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

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

vector ω_X, along with ω, in the following specification of Section 1.4.1

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

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 1.8 as follows. Since the selection equation and the log of spending equation are the same as in the benchmark model, I only report the results of this sensitivity analysis for the voter and entry models. The estimation results for the entry model are affected by considering the voter turnout in general elections, because this influences the simulated expected payoffs of all elections.

The result of the primary part in the voter model is the same as in Table 1.3. Once the voter turnout is taken into account, though the effect of campaign spending on votes in the general election is still positive, it is not as prominent as in the benchmark model. As for the estimated coefficients in the general latent utility, most of them have expected signs, but are insignificant. Being an open-seat election reduces the latent utility for a representative voter in the general election, which can be justified by the incumbency advantage, because incumbents already build the reputation among voters after holding the office for a period. The number of actual candidates has a negative effect on the latent utility of general election voters, related to the structural result in the benchmark model that a more competitive election reduces the campaign spending made by candidates. This can indicate that a more competitive election can reduce the latent utility of voters, due to the reason that voters may have a hard time figuring out who they should select. Therefore, candidates in a more competitive election receive less expected payoffs, and spend less in campaigns consequently. The only significant coefficient is on the dummy variable PRE which equals 1 if a presidential election occurs on the same Senate general election day. This dummy variable has a significantly positive effect on the latent utility of voters, thus a significantly negative effect on the voter turnout. If a presidential election occurs in a year, it can spur additional turnout for the Senate general election in the same year.

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

Table 1.8: Sensitivity analysis: turnout in general elections

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

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

Truncated Spending Distribution

In this sensitivity analysis, I assume that the campaign spending is distributed as a log-normal truncated from above when the spending is strictly positive. Since the model is conditional on the election-level heterogeneity (X, u), the resulting distribution of the spending will also condition on (X, u), so is the upper boundary. Therefore, I need to estimate the upper boundary of non-zero spending 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 1.9.

Most of the results remain consistent with those of the benchmark model. The estimated standard deviation of the error in the log of spending equation is now slightly larger: 2.82, and the estimated covariance of the two errors in the spending model is also slightly larger: 1.51, resulting in a correlation being about 0.54, which is close to the estimate of 0.58 in the benchmark model. The estimated coefficients in the selection equation have the same signs and similar magnitudes as in the benchmark model shown in Table 1.3, in most cases. The unemployment rate now has a negative effect on the decision of making non-zero spending, which is consistent with the positive effect of the log of median household income, although this effect is still insignificant. With some small discrepancies, the estimates of the entry model remain stable in this sensitivity analysis, compared with the benchmark model.

Difference Between Political Parties

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

Table 1.9: Sensitivity analysis: truncated spending distribution

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

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

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

Table 1.10 presents the results of this sensitivity analysis of taking into consideration the difference between political parties. First of all, the Democratic party challenger index and the interaction terms of it with the continuous covariates yield insignificant coefficients in the selection equation, the log of campaign spending equation, and the mean equation of entry cost distribution. Therefore, I do not find strong evidence of the difference between political parties regarding the effects of the election-level 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 1.3.

The difference between political parties are reflected in the estimates of the voter model. In the primary voter model, the incumbents have a relatively higher return of campaign spending than challengers in the incumbent-challenger elections, in terms that their spending can translate to more votes in the primary election. The challengers have different returns of campaign spending: for the Democratic challenger, the estimated coefficient on the spending term in the vote equation is around 0.12, while that coefficient is about 0.14 for the Republican challenger. However, this difference between parties is not very big. Overall, the campaign spending can translate to more votes in open-seat elections, compared to incumbent-challenger elections, which may be due to the reason that in open-

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

Table 1.10: Sensitivity analysis: difference between political parties

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

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

seat elections, voters cannot evaluate the candidates' valence depending on their performance as the incumbent, thus they tend to rely more on the spending the candidates made through campaigns. This discrepancy between open-seat and incumbent-challenger elections is more prominent in the general election. The estimates coefficient on the spending term is about 0.38 for open-seat general elections, while I estimate the coefficients to be 0.16 and 0.12 for incumbents and challengers respectively for incumbent-challenger elections. The latter estimates are close to those in the primary voter model, indicating that the different return of campaign spending between primary and general elections found in the benchmark model is largely driven by that difference in open-seat elections, in the sense that the campaign spending can translate to more votes in the general election than in the primary election, given the same amount of spending.

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 1.3.3 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 1.3.3, through assuming that one quantile of the private value distribution is constant and unknown. Since the strictly positive spending is unbounded from above, without loss of generality, I consider a constant and unknown median of the private value distribution. This median is constant, meaning that it is unrelated to the election-level heterogeneity (X, u), i.e., $Med(v|X, u) = Med(v)$; and this median is unknown, because I treat it as a parameter which can be identified and estimated along with the parameters in the cost function (κ, ζ).

In the data, I observe that the incumbents will always spend non-zero through campaigns once they decide to run for office. Therefore in equilibrium, the first-order condition always holds as an equality for the incumbents. I rely on the subset of incumbents' campaign spending to estimate $(\kappa, \zeta, Med(v))$ through NLS. For incumbents, the expectation in Equation 1.3.3 is simulated by 500 repetitions regarding the spending profile of their opponents. Since the spending model and the voter model are the same as the benchmark model, I report the results for the cost function and the entry model, as follows.

Table 1.11 presents the estimates of the sensitivity analysis. The estimated constant $Med(v)$ is

Table 1.11: Sensitivity analysis: quadratic cost function

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

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

around 18 million, which is significant with an estimated standard error being about 6 million. This figure is larger than the mean of non-zero campaign spending shown in Table 1.1. Interestingly, although the cost function is assumed to be quadratic, the results show that the estimated cost function is linear, with $\hat{\zeta} = 0.95$, thus leading to a linear function very similar to the one considered in the benchmark model. Because the estimated cost function does not change much, the resulting estimates for the entry model are similar to those in the benchmark model. The exception is the effect of the number of potential candidates, which is estimated to be a tiny positive number that is essentially zero.

Alternative Voter Model Specification

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

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

for the primary election, and

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

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

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

Table 1.12 gives the results in this sensitivity analysis. For the voter model, in both the primary and general elections, the campaign spending made by the candidates affects the latent utility of the voters positively and significantly. Unlike the result in the benchmark model, when the campaign spending enters voters' latent utility linearly, the effect of spending is larger in the primary election, almost doubling that in the general election. This is consistent with the benchmark implication and can be due to the reason that the campaign spending made by the candidates who participate in the general elections usually spend much more than the rest of the candidates. Thus when the latent utility is assumed to be a linear function where the spending dose not exhibit diminishing marginal effect, the marginal effect of the spending on the latent utility of voters is averaged out. Recall that in the benchmark model where the effect of campaign spending in the voter model is assumed to take the log form, the more spending candidates make, the smaller the marginal effect of spending entails. It

Table 1.12: Sensitivity analysis: alternative voter model specification

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

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

is important to compare the roles of the spending in the primary and general voter models given the same level of the campaign spending. The results from estimating the entry model are consistent with those of the benchmark model shown in Table 1.3. In this sensitivity analysis, the estimated standard deviations of the entry cost distributions for the incumbents and challengers are slightly larger than those in the benchmark model. The estimated coefficients on the explanatory variables have the same signs and similar magnitude as in the benchmark model.

Dynamic Change

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

I present the estimated results from the model with time fixed effects in Table 1.13. The estimates of the time-specific constants are not reported and will be discussed later in Figure 1.9. Note that since the voter model is unchanged, the results only contain the campaign spending model with the selection equation and the level equation, and the entry cost model.

As shown in Table 1.13, most of the estimates are consistent with those in Table 1.3. The estimated standard deviation of the error in the log of spending equation, σ_2, and the estimated covariance of the two errors in the selection equation and in the level equation of campaign spending, σ_{12}, are similar to those derived from estimating the benchmark model. Furthermore, in the specifications of campaign spending and entry cost distributions, the estimated coefficients on the covariates have expected signs and magnitudes. Compared with challengers, incumbents tend to spend non-zero and larger amount through campaigns in elections. Candidates in an open-seat election are more likely to

44 As a sensitivity analysis in Appendix 1.7.5, instead of using time-specific constants in the specifications, I also estimate an alternative model where a dummy variable being one for the period of post-2008 and zero for that of pre-2008 in order to capture the effect of time. The voter model estimates show no significant difference on how voters appreciate the campaign spending made by candidates before and after 2008, for both primary and general elections.

Table 1.13: Sensitivity analysis: estimation results with time fixed effects

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

Notes: Time fixed effects are included in the selection equation of campaign spending, the level equation of campaign spending, and the entry cost distribution. For explanatory variables, estimated coefficients in corresponding specifications are reported. The estimated results are obtained through the estimation method described in Section 1.4.2. Sd. errors are obtained through 500 nonparametric bootstrap at the election level.

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

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

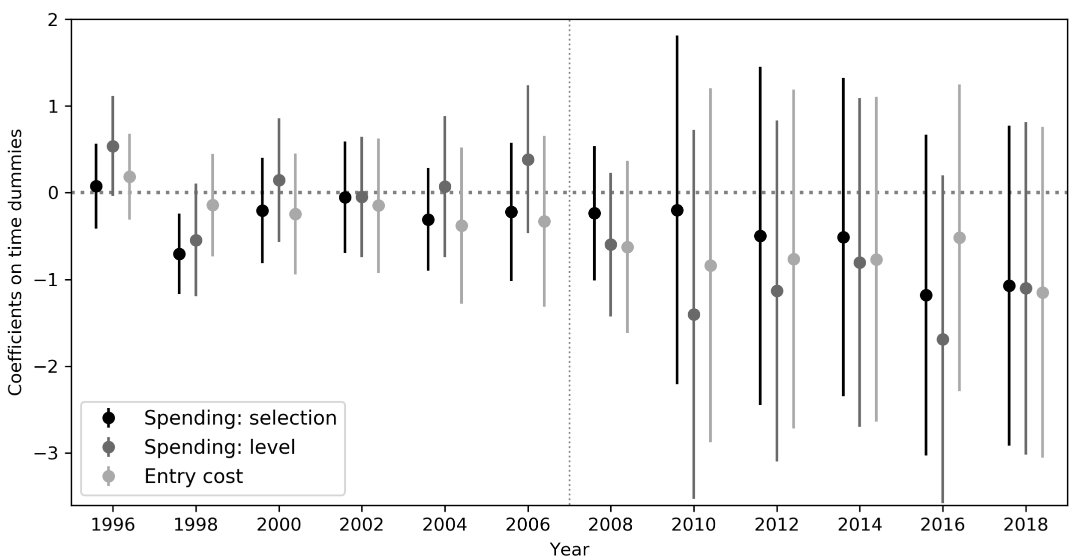


Figure 1.9: Coefficients on time dummies in multiple regressions with 90% confidence intervals

As reflected in Figure 1.9, most of the estimated time-specific effects are negative. If I divide the sample into two periods: pre-2008 and post-2008, the results before 2008 are mixed while those after 2008 are consistently negative with wider confidence intervals. Although all but one of the estimates are insignificant at the 90% significance level, the trend itself implies the devaluation of candidates for the Senate office across time, if any, because campaign spending is a strictly increasing function of the private value of holding the office for actual candidates. The entry costs of potential candidates also fall over the same period. Since the entry costs of running for office and the valuations of holding the office both decrease over time, it is hard to determine whether the net benefits of the Senate office decrease as well. The estimates correspond to an average entry probability of 0.8050 before 2008, and an average entry probability of 0.7838 after 2008. Therefore, with the falling entry costs of running for office and the falling valuations of holding office, the net benefits of office have a minuscule drop over the years, leading to slightly decreasing entry probabilities of potential candidates.

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

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

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

and Theriault and Rohde (2011)).

The Great-Recession Effect

In Section 1.7.5, I introduce the full set of time dummies into the benchmark model so that the dynamic change of the results can be taken into account. In this sensitivity analysis, I conduct an alternative exercise adding a dummy variable that is assigned one for the period of post-2008 and zero for that of pre-2008 to the specifications in the model. Figure 1.9 in Section 1.7.5 shows a persistent decrease of the propensity to spend through campaigns, the amount of positive campaign spending, and the entry cost of candidates after 2008, which may be due to the Great-Recession in 2008. Therefore, by introducing a dummy variable to indicate the period of post-2008, it may form additional evidence for the falling of the benefit of office by comparing the difference before and after 2008. Further, in this exercise, I allow the effects of campaign spending on voters to differ before and after 2008 in primary and general elections.

Table 1.14: Sensitivity analysis: difference over time

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

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

I present the results of this sensitivity analysis in Table 1.14. Most of the estimates have the same signs and similar magnitudes as those in the benchmark model shown in Table 1.3. Since I care about the effect of the Great-Recession in 2008, I focus on the coefficients on the variable "POST-2008" in the corresponding specifications for the spending and entry models, as well as the estimated parameters in the voter model. As in Section 1.7.5, the dummy variable indicating the period of post-2008 has a negative coefficient in both the selection equation and the level equation of the spending model, implying that after 2008 actual candidates are less likely to spend non-zero in campaigns, and make smaller amount of campaign spending. This dummy variable also has a negative coefficient in the entry cost specification, showing a possible decline of entry costs after 2008. However, these effects are not significant, as shown in Figure 1.9. For voter model, I observe a difference regarding the estimated effects of campaign spending on voters' utility before and after 2008, and this difference is more pronounced in general elections (but the difference is still not big). Voters tend to value how much candidates spend in the general election more after 2008. This can explain why I find null effect of time dummies in Section 1.7.5 and here. The devaluing of office may exist over time among candidates that should result in the decrease of the campaign spending in equilibrium, but because voters value the spending made by candidates through campaigns more over time, therefore in the end two conflicting effects on the campaign spending cancel out with each other and lead to no significant change of the campaign spending across time.

Chapter 2: Nonparametric Identification of Bayesian Games under Exclusion Restrictions

This chapter is adapted from the working paper "Nonparametric Identification of Bayesian Games under Exclusion Restrictions" and has been reproduced with the permission of my co-authors Tong Li and Jun Zhang.

2.1 Introduction

Bayesian games as a natural extension of games under complete information have been one of the most important models in modern economics. While it is natural to add incomplete information to most economic models, unfortunately, Bayesian games are difficult to analyze both from a theoretical perspective and using a structural approach because of the complexity arising from solving the equilibrium strategies and analyzing their properties. As a result, the structural analysis of Bayesian games has been quite limited, with a notable exception being the auction case which has been extensively studied since Paarsch (1992) especially thanks to the nonparametric identification and estimation method developed in Guerre, Perrigne, and Vuong (2000).[1] The reason for this exception is that due to the discontinuity in bidders' payoff functions, the first order conditions in characterizing bidders' Bayesian Nash equilibrium strategies are first order differential equations with well specified boundary conditions, while they are integral equations in general Bayesian games. Only a few special cases of Bayesian games with continuous payoffs and restrictive payoff structures have recently been studied. For example, He and Huang (2018) study nonparametric identification and estimation of the Tullock contest model with private information; Aryal and Zincenko (2019) study identification and estimation of the Cournot competition model under incomplete information; Aryal and Gabrielli (2020) estimate a competitive nonlinear pricing model with two firms; in all these papers the only model primitive that needs to be identified is the private type distribution.

In this paper, we consider an arbitrary Bayesian game within the private type paradigm and with an additive payoff function. This framework is general enough to cover many applications such as oligopoly competition (Raith (1996)), Diamond search (Diamond (1982)), public good provision (Bergstrom, Blume, and Varian (1986)), and Tullock contest (Tullock (1980)). To address the identification issue, we focus on the case where a part of the (additive) payoff function is unknown to the

1 For the fast growing literature in the empirical analysis of auction data, especially using the nonparametric approach, see surveys such as Athey and Haile (2007), Hendricks and Porter (2007), Hickman, Hubbard, and Sağlam (2012), Gentry, Hubbard, Nekipelov, and Paarsch (2018), and Li and Zheng (2020).

researcher, which can be more general than assuming the entire payoff functional form being known, and can find many applications. For example, the unknown part in the payoff function can be firms' cost functions in oligopoly competition games, searchers' cost functions in the Diamond search model, contributors' cost functions of private contribution in public good provision games, and contestants' cost of effort functions in Tullock contest model, to name a few.

We first show that the benchmark model is not nonparametrically identified from observed actions without further restrictions. We then impose the exclusion restriction in the form of an exogenous players' participation and establish nonparametric point or partial identification results of the model primitives, namely, the private type distribution and the unknown part of the payoff function. Specifically, we show that if the distributions of actions intersect with each other when the number of players varies, the private type distribution and the unknown structure are nonparametrically identified up to a scale. Otherwise, they are partially identified as they can be bounded nonparametrically. Our partial identification results are as important as the point identification results since they are applicable in many applications. Note that the distributions of actions can be nonparametrically estimated in a straightforward manner, thus whether they intersect or not when the number of players varies can be readily verified from the data.

Using exclusion restrictions as a means to identify structural econometric models in general and auction models in particular has been one of the key identification strategies; for examples in the empirical auction literature, see Athey and Haile (2002), Haile, Hong, and Shum (2003), Bajari and Hortaçsu (2005), Guerre, Perrigne, and Vuong (2009), Gentry and Li (2014), and Chen, Gentry, Li, and Lu (2019). In this regard, our identification strategy is related to Guerre, Perrigne, and Vuong (2009) who study nonparametric identification of first price auctions with unknown risk averse utility functions and establish nonparametric point identification of the model primitives, which are the private value distribution and the utility function, under the exclusion restriction. However, the Bayesian games under our consideration are more involved than the auction case considered in Guerre, Perrigne, and Vuong (2009). In the auction case, point identification can be attained because the equilibrium bid functions at different competition levels always cross at the lowest valuation if there is no binding reserve price or at the reserve price if there is, as a result of the boundary condition of the differential equation whose solution defines the Bayesian Nash equilibrium, while such a boundary condition does not exist for a general Bayesian game we consider.

Our identification approach starts with taking the ratio of the first order conditions corresponding to different numbers of players. This idea has also been used in Aryal, Grundl, Kim, and Zhu (2018) in studying identification of auctions with risk aversion and ambiguity, and Li and Zhu (2020) in studying identification of auctions with ambiguity and entry. However, the analysis that follows in these cases and ours takes a completely different approach from each other, due to the unique feature and complexity of the respective model. Furthermore, in a recent paper by D'Haultfœuille and Février (2020) that studies empirically the optimality of simple linear compensation contracts in a principal-agent model with both adverse selection and moral hazard using the contract data between the French National Institute of Statistics and Economics and the interviewers the Institute hired to conduct its surveys, the authors take the ratio of the first order conditions before and after the exogenous change of piece rate to obtain nonparametric partial identification results for the derivative of the cost function of the interviewers and the distribution of their types, respectively. In this regard, D'Haultfœuille and Février (2020) and our paper share similarities in terms of identification strategy, while in totally different contexts, using different exogenous variations, and dealing with different problems.

Our partial identification results are related to the partial identification literature in econometrics pioneered by Charles Manski; see Manski (2003) for a review of early contributions, and more recent work includes Manski and Tamer (2002), Tamer (2003), Chesher and Rosen (2013), Molinari (2008), among others, which are surveyed in Tamer (2010) and Molinari (2020). The partial identification approach has been used also in the structural analysis of auction data starting with Haile and Tamer (2003); see, e.g., Hortaçsu and McAdams (2010), Tang (2011), Aradillas-López, Gandhi, and Quint (2013), Gentry and Li (2014), and Chen, Gentry, Li, and Lu (2019), and also the papers surveyed in Li and Zheng (2020). To the best of our knowledge, our paper is the first one to study the nonparametric identification for a general class of Bayesian games.[2]

Our econometric approach is based on the theoretical literature on Bayesian games. The existence and/or uniqueness of a monotone pure strategy Nash equilibrium (MPSNE) have been established in several different frameworks that complement with each other using different approaches. First, Athey (2001) pioneers the literature and provides a central tool to establish the existence of MP-

2 It is worth noting that dealing with continuous actions in Bayesian games raises challenging identification problems that this paper attempts to address, while there has been some interesting work on identification of Bayesian games with discrete actions whose framework is vastly different from what we consider and requires different identification strategies/approaches. See, e.g., Sweeting (2009), Aradillas-Lopez (2010), Bajari, Hong, Krainer, and Nekipelov (2010), Tang (2010), de Paula and Tang (2012), Wan and Xu (2014), and Lewbel and Tang (2015), among others. On the other hand, for identification of games with complete information, see, e.g., Bajari, Hong, and Ryan (2010) for identification of discrete games of complete information, and Kline (2015) for complete information games that allow generalized interaction structures and generalized behavioral assumptions.

SNE for Bayesian games that satisfy the single crossing condition.[3] Second, Van Zandt and Vives (2007) establish the existence of a greatest and a least MPSNE for Bayesian games with strategic complimentarities. Polydoro (2011) further provides sufficient conditions for the uniqueness. Third, Mason and Valentinyi (2010) establish the existence and uniqueness of MPSNE using the contracting mapping method. Lastly, the existence and uniqueness of MPSNE in Tullock contest under incomplete information have been established separately by Fey (2008), Ryvkin (2010), and Ewerhart (2014). In our paper, we require the existence of strictly MPSNE. To accommodate as many applications as possible, we assume the existence of MPSNE, meaning that our nonparametric methodology can be applied to any of the above frameworks. We further impose strict supermodularity between one's own action and type to ensure that the existing MPSNE is strictly monotone.[4]

The rest of the paper is organized as follows. Section 2.2 introduces the Bayesian game with an unknown structure as well as the non-identification result. Section 2.3 introduces the exclusion restrictions and explores the functional comparative statics. Section 2.4 establishes the identification results. Section 2.5 presents the numerical illustration. Section 2.6 extends the model. Section 2.7 concludes. The proofs of our main results are in the Appendix 2.8, and all the other technical details are included in the Supplemental Appendix 2.9.

2.2 The Benchmark Model and Non-Identification

2.2.1 Preliminaries

There are $N \geq 2$ players engaging in a Bayesian game. Player $i \in \{1, \cdots, N\}$ has a private type t_i drawn from a distribution with CDF $F_i(\cdot)$ over the type space $\mathscr{T}_i \equiv [\underline{t}_i, \bar{t}_i] \subset \mathbb{R}$. Assume that $F_i(\cdot)$ is absolutely continuous and thus has an atomless density $f_i(\cdot)$. Types are drawn independently across players. All players choose actions simultaneously. Player i's action a_i is chosen from a compact action space $\mathscr{A}_i \subset \mathbb{R}$. Each player's strategy is a mapping from the type space to the action space, i.e., $s_i : \mathscr{T}_i \to \mathscr{A}_i$.

Player i's payoff, $\pi_i(\mathbf{a}, t_i)$, depends on all players' actions and her own type, and thus is a mapping $\pi_i : \mathbf{A} \times \mathscr{T}_i \to \mathbb{R}$, where $\mathbf{A} = \mathscr{A}_1 \times \cdots \mathscr{A}_N$ and $\mathbf{a} = (a_1, \cdots, a_N)$. We assume that $\pi_i(\mathbf{a}, t_i)$ is twice

3 McAdams (2003) extends Athey's model to allow partially ordered multidimensional type and action spaces. Reny (2011) further allows action spaces to be compact and locally complete metric semilattices, and type spaces to be partially ordered probability spaces.

4 There is a large literature that establishes the existence of pure strategy equilibrium but without the property of monotonicity, to name a few, see Vives (1990), Milgrom and Weber (1985), Barelli and Duggan (2015), Khan and Zhang (2014), and He and Sun (2019).

continuously differentiable with bounded derivatives.[5] In addition we assume that $\pi_i(\mathbf{a}, t_i)$ is strictly supermodular in (a_i, t_i). Thus, there exists a positive number M_π, such that

$$0 < \frac{\partial^2 \pi_i(a_i, \mathbf{a}_{-i}, t_i)}{\partial a_i \partial t_i} \leq M_\pi \textit{ almost everywhere}, \ \forall i, \tag{2.2.1}$$

where $\mathbf{a}_{-i}$ are the actions taken by all the players but player i.

2.2.2 Additive Payoff with an Unknown Structure

In order to explore the possibility of an unknown structure in the payoff function, throughout the rest of this paper, we focus on Bayesian games with an additive payoff structure, which takes one of the following three forms:

(1). $\pi_i(\mathbf{a}, t_i) = t_i x_i(\mathbf{a}) + y_i(a_i)$,

(2). $\pi_i(\mathbf{a}, t_i) = x_i(\mathbf{a}) + y_i(a_i)/t_i$,

(3). $\pi_i(\mathbf{a}, t_i) = t_i a_i + x_i(\mathbf{a}) + y_i(a_i)$,

where the structure of $x_i(\cdot)$ is known but the structure of $y_i(\cdot)$ is unknown by the researcher. Note that t_i can be replaced by a deterministic known function of t_i, $m_i(t_i)$. Many Bayesian games fall into one of the above three structures as shown in the following examples.[6]

Example 1 ***Differentiated Cournot competition:*** *The private type t_i is a firm's demand characteristic and the action a_i is a firm's production level.*[7] *Firm i's profit is*

$$\pi_i(a_i, \mathbf{a}_{-i}, t_i) = \left(t_i - \beta \sum_{j=1}^{N} a_j\right) a_i - c_i(a_i) = t_i a_i - \beta a_i \sum_{j=1}^{N} a_j - c_i(a_i),$$

This application fits Structure (3) with $x_i(\mathbf{a}) = \beta a_i \sum_{j=1}^{N} a_j$ (known) and $y_i(a_i) = -c_i(a_i)$ (unknown).

---

5 The assumption that π_i has bounded derivatives ensures the interchangeability of derivative and expectation (integral), through Lebesgue's dominated convergence theorem, see Bartle (1995). Strictly speaking, the Tullock contest model is not twice continuously differentiable. However, it has discontinuity only when all players choose zero efforts, which never happens in equilibrium. Therefore, the Tullock model can still be included in the class of games we study.

6 If $x_i(\mathbf{a})$ contains some parameters, we assume they can be identified in a preliminary step. For example, in the differentiated Cournot competition, if we impose the same exclusion restriction as in Section 2.3, the parameter β can be identified via taking difference of inverse demand functions of Cournot games with different numbers of players, because prices and productions of firms are known. The function $g(\cdot)$ in the Diamond's search model can be identified through the observed search intensity (measured by time) of each player, together with the observable matching rate. The function $k(\cdot)$ in the public good provision model can be identified through the observed contribution of each player, together with the amount of observable public goods. The parameter r for the Tullock contest model can be identified from each contestant's action and the result of the contest, both of which are observed.

7 A more standard and less general model is to assume firms have a constant private marginal cost. However, in that model, the payoff structure is known by researchers.

Example 2 ***Diamond search model:*** *The private type t_i is a player's valuation of a match and the action a_i is a player's search intensity. The payoff function for player i is equal to*

$$\pi_i(a_i, \mathbf{a}_{-i}, t_i) = t_i a_i g\left(\sum_{j\neq i} a_j\right) - c_i(a_i)$$

This application fits Structure (1) with $x_i(\mathbf{a}) = a_i g\left(\sum_{j\neq i} a_j\right)$ (known) and $y_i(a_i) = -c_i(a_i)$ (unknown).

Example 3 ***Public good provision:*** *The private type t_i is a player's value of the public good, and the action a_i is the individual contribution. The payoff function for player i is equal to*

$$\pi_i(a_i, \mathbf{a}_{-i}, t_i) = t_i k(a_i + \sum_{j\neq i} a_j) - c_i(a_i),$$

This application fits in Structure (1) with $x_i(\mathbf{a}) = k(a_i + \sum_{j\neq i} a_j)$ (known) and $y_i(a_i) = -c_i(a_i)$ (unknown).

Example 4 ***Tullock contest:*** *The private type t_i is a player's ability, and the action a_i is the expenditure in R&D contests, or fund raised in political campaign. Then the payoff function for player i is*

$$\pi_i(a_i, \mathbf{a}_{-i}, t_i) = V\frac{a_i^r}{a_i^r + \sum_{j\neq i} a_j^r} - \frac{c_i(a_i)}{t_i},$$

where V is the prize, $\frac{a_i^r}{a_i^r+\sum_{j\neq i} a_j^r}$ is the Tullock contest success function, and $c_i(a_i)$ is player i's cost function of action. This application fits in Structure (2) with $x_i(\mathbf{a}) = V\frac{a_i^r}{a_i^r+\sum_{j\neq i} a_j^r}$ (known) and $y_i(a_i) = -c_i(a_i)$ (unknown).

2.2.3 Strictly Monotone Pure Strategy Nash Equilibrium

For the ease of exposition, in the main context, we only examine form (1): $\pi_i(\mathbf{a}, t_i) = t_i x_i(\mathbf{a}) + y_i(a_i)$. All the results can be extended to the other two forms as we show in the Supplemental Appendix. The following assumption gives the regularity conditions for the unknown primitives of the Bayesian game, i.e., the distributions of the private types and the unknown structure in the payoff function.

Assumption 1 *Let $\mathscr{F}$ and $\mathscr{Y}$ be two classes of functions that satisfy the following conditions:*
(i) $\forall F(\cdot) \in \mathscr{F}$, $F(\cdot)$ satisfies: (1) $F(\cdot)$ is an absolutely continuous CDF with a compact support: $\mathscr{T} \equiv [\underline{t}, \bar{t}] \subset \mathbb{R}$; (2) $F(\cdot)$ has a bounded, continuous and atomless PDF $f(\cdot)$, which is strictly positive and bounded from infinity over $\mathscr{T}$, i.e.: $0 < f(\cdot) < \infty$.
(ii) $\forall y(\cdot) \in \mathscr{Y}$, $y(\cdot)$ satisfies: (1) $y(\cdot)$ is twice continuously differentiable (C^2) over a compact support:

$\mathscr{A} \subset \mathbb{R}$ with bounded derivatives; (2) the sign of $y'(\cdot)$ is known over the support $\mathscr{A}$; (3) $y(\cdot)$ is convex or concave over $\mathscr{A}$.

Condition (i) is the standard assumption on the latent type distribution $F_i(\cdot)$. Condition (ii) imposes some shape restrictions on the function $y_i(\cdot)$, which play an important role in the later nonparametric identification approach.[8] These conditions are not restrictive in empirical applications. For instance, it is natural to assume that the cost function is increasing and convex.

We assume that $x_i(\cdot)$ is twice continuously differentiable with bounded derivatives. Furthermore, the strict supermodularity of $\pi_i(\mathbf{a}, t_i)$ in (a_i, t_i) is also imposed. Thus together these assumptions lead to the following condition: there exists a positive number M_x such that

$$0 < \frac{\partial x_i(a_i, \mathbf{a}_{-i})}{\partial a_i} \leq M_x \textit{ almost everywhere}, \forall i. \tag{2.2.2}$$

In order to discuss the identification problem, we focus on the strictly MPSNE where player i adopts a strictly increasing equilibrium strategy $s_i(\cdot)$.[9] Given that all other players adopt the equilibrium strategy, i.e., $a_j = s_j(t_j)$ for $j \neq i$, player i's maximization problem, when having private type t_i, can be written as:

$$\max_{a_i} \Pi_i(a_i, t_i) \equiv \int_{\mathbf{t}_{-i} \in \mathscr{T}_{-i}} [t_i x_i(a_i, \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i) + y_i(a_i)] \, d\mathbf{F}_{-i}(\mathbf{t}_{-i}), \tag{2.2.3}$$

where $\mathscr{T}_{-i} = \cdots \times \mathscr{T}_{i-1} \times \mathscr{T}_{i+1} \times \cdots$, $\mathbf{s}_{-i}(\mathbf{t}_{-i}) = (\cdots, s_{i-1}(t_{i-1}), s_{i+1}(t_{i+1}), \cdots)$, $\mathbf{t}_{-i}$ are the types of the opponents of player i, and $d\mathbf{F}_{-i}(\mathbf{t}_{-i}) = \cdots dF_{i-1}(t_{i-1}) dF_{i+1}(t_{i+1}) \cdots$. Therefore the first order condition together with the equilibrium condition leads to that, $\forall i$ and $\forall t_i$,

$$t_i \cdot \int_{\mathbf{t}_{-i} \in \mathscr{T}_{-i}} \frac{\partial x_i(s_i(t_i), \mathbf{s}_{-i}(\mathbf{t}_{-i}))}{\partial a_i} d\mathbf{F}_{-i}(\mathbf{t}_{-i}) + y_i'(a_i) = 0. \tag{2.2.4}$$

where the partial derivative is with respect to the first argument of payoff function x_i.[10] The following proposition states the existence and regularity conditions of a strictly increasing equilibrium for the class of Bayesian games considered in the paper.

Proposition 2 *The equilibrium with $s_i(\cdot) : \mathscr{T}_i \to \mathscr{A}_i$ characterized by (2.2.4) for the specific class*

8 A known sign of the derivative function $y_i'(\cdot)$ over its support is necessary to attain point identification. Furthermore, for the partial identification result, we need this convexity or concavity restriction for all the three forms of payoff functions to bound the derivative function $y_i'(\cdot)$. Specifically, for the third form of payoff functions, the point identification result does not rely on (ii)-(2), but the partial identification result requires (ii)-(3) (see the Supplemental Appendix).

9 For Bayesian games with a strictly decreasing equilibrium strategy, we can simply redefine the type.

10 Here we focus on the interior solutions, and extend to the analysis for accommodating corner solutions in Section 2.6.1.

of Bayesian games described above with the structure $[F_i(\cdot), y_i(\cdot)] \in \mathscr{F} \times \mathscr{Y}$ *satisfies the following conditions:*

(i) $s_i(\cdot)$ *is strictly increasing, with the first derivative bounded away from infinity:* $0 < s_i'(\cdot) < \infty$,

(ii) $s_i(\cdot)$ *is continuously differentiable* (C^1) *with a compact image* $[\underline{a}_i, \bar{a}_i] = [s_i(\underline{t}_i), s_i(\bar{t}_i)] \subset \mathscr{A}_i$.

As discussed in the introduction, the existence and uniqueness of the MPSNE have been well established in the literature. The strict supermodularity (2.2.2) ensures that the equilibrium is strictly monotone. Note that the uniqueness is not essential here since with multiple equilibria, as long as players always act according to one of the equilibria, it does not affect our identification results below.

2.2.4 Non-Identification Results

The unknown model primitives are a pair of functions: $[F_i(\cdot), y_i(\cdot)]$. In order to illustrate the econometrics of this game, it is useful to represent condition (2.2.4) using the CDF of equilibrium actions, denoted as $G_i(\cdot)$, for each i. Note that $G_i(\cdot)$ has a compact support $[\underline{a}_i, \bar{a}_i] = [s_i(\underline{t}_i), s_i(\bar{t}_i)]$. For each $a \in [\underline{a}_i, \bar{a}_i]$, and each i, $G_i(a) = Pr(a_i \leq a) = Pr(t_i \leq s_i^{-1}(a)) = F_i(t)$, if we denote $a = s_i(t)$. Thus Equation (2.2.4) becomes

$$t_i \cdot \int_{\mathbf{a}_{-i} \in \mathscr{A}_{-i}} \frac{\partial x_i(a_i, \mathbf{a}_{-i})}{\partial a_i} d\mathbf{G}_{-i}(\mathbf{a}_{-i}) + y_i'(a_i) = 0. \tag{2.2.5}$$

Noting that the multiple integral in Equation (2.2.5) can be expressed by an expectation over player i's opponents' all equilibrium action profiles as $\mathbb{E}_{-i}[\frac{\partial x_i(a_i, \mathbf{a}_{-i})}{\partial a_i}]$, for simplicity. Equation (2.2.5) is thus simplified to

$$t_i \cdot \mathbb{E}_{-i}[\frac{\partial x_i(a_i, \mathbf{a}_{-i})}{\partial a_i}] + y_i'(a_i) = 0, \tag{2.2.6}$$

which is the key condition when discussing identification results in the rest of this paper. The following proposition gives properties of the CDF $G_i(\cdot)$.

Proposition 3 *Consider the specific class of Bayesian games described above with the structure* $[F_i(\cdot), y_i(\cdot)] \in \mathscr{F} \times \mathscr{Y}$. *The derived equilibrium action CDF* $G_i(\cdot)$ *satisfies the following properties:*

(i) $G_i(\cdot)$ *is an absolutely continuous CDF with a compact support:* $[\underline{a}_i, \bar{a}_i] = [s_i(\underline{t}_i), s_i(\bar{t}_i)] \subset \mathscr{A}_i$, *where* $s_i(\cdot)$ *is a strictly increasing equilibrium,*

(ii) $G_i(\cdot)$ *has a bounded, continuous and atomless PDF* $g_i(\cdot)$, *which is strictly positive and bounded from infinity over* $[\underline{a}_i, \bar{a}_i]$, *i.e.,* $0 < g_i(\cdot) < \infty$.

In the data, we observe the number of players N and the actions of players. Since the CDF of observed actions $G_i(\cdot)$ and its support $\mathscr{A}_i$ can be consistently estimated nonparametrically using the observed number of players and players' actions, they can be identified. Thus the identification problem reduces to whether the structure $[F_i(\cdot), y_i(\cdot)]$ is uniquely determined from the observables. The following assumption gives the conditions that the observed action CDF $G_i(\cdot)$ satisfies.

Assumption 2 *Let $\mathscr{G}$ be a class of functions that $G_i(\cdot)$ belongs to. Then $\mathscr{G}$ satisfies that: $\forall G(\cdot) \in \mathscr{G}$, $G(\cdot)$ satisfies properties (i) and (ii) in Proposition 3.*

The *identification* of the latent structure $[F_i(\cdot), y_i(\cdot)]$ from the known CDF $G_i(\cdot) \in \mathscr{G}$ with the number of players being N has to be unique, in the sense that a structure $[F_i(\cdot), y_i(\cdot)]$ is identified if there does not exist any other structure $[F_i^*(\cdot), y_i^*(\cdot)]$ leading to the same action CDF for the game with N players. In general, however, a structure $[F_i(\cdot), y_i(\cdot)] \in \mathscr{F} \times \mathscr{Y}$ is not identified, as shown in the following result.

Proposition 4 *For a structure $[F_i(\cdot), y_i(\cdot)] \in \mathscr{F} \times \mathscr{Y}$, there exists another observationally equivalent structure $[F_i^*(\cdot), y_i^*(\cdot)] \in \mathscr{F} \times \mathscr{Y}$, in the sense that both structures induce the same CDF of equilibrium actions $G_i(\cdot) \in \mathscr{G}$, given the number of players N, in the specific class of Bayesian games.*

The intuition of Proposition 4 can be better understood, by noting that two unknown functions $F_i(\cdot)$ and $y_i(\cdot)$ have to be recovered from only one condition, which is the first order condition of the model, given that $G_i(\cdot)$ is observed for each i. Therefore, by keeping $G_i(\cdot)$ unchanged, when adjusting one of the two functions, the other one can always be adapted accordingly. This leads to the non-uniqueness of the structure containing these two functions, to induce the same observed $G_i(\cdot)$.

In the next sections, targeting at the structure $[F_i(\cdot), y_i(\cdot)]$, this paper proposes the nonparametric (point or partial) identification approach through the exclusion restriction, to back out the unknown model primitives in the class of Bayesian games considered in this section.[11]

2.3 Functional Comparative Statics under Exclusion Restriction

To start with, we impose the symmetry restriction by assuming that each player draws her private type from the same distribution $F(\cdot|N) \in \mathscr{F}$ over the same compact support $[\underline{t}(N), \bar{t}(N)]$.[12] Besides, the unknown function $y_i(\cdot)$ is assumed to be identical for every player, denoted by $y(\cdot) \in \mathscr{Y}$ and

11 Without the exclusion restriction, it is possible to make other parametric or semiparametric assumptions to attain identification of the underlying model primitives. This is beyond the scope of this paper.

12 The extension to the asymmetric case will be discussed in Section 2.6.

independent of N. Furthermore, we assume that the known function $x_i(\mathbf{a};N)$ is independent of i and symmetric in the argument $\mathbf{a}_{-i}$ for each player i, thus is expressed as $x(a_i,\mathbf{a}_{-i};N)$.[13] Note that this function depends on N directly, since the argument is the whole vector of actions $\mathbf{a}$. Lastly, the action space $\mathscr{A}_i = \mathscr{A}(N)$ for each player i.

Due to symmetry, the focal point for the rest of the paper is the same equilibrium strategy and the same CDF of equilibrium actions for each player. Thus, the equilibrium $s(\cdot;N)$ satisfies Proposition 2; and $G(\cdot|N) \in \mathscr{G}$ has a compact and common support $[\underline{a}(N), \bar{a}(N)]$. For ease of exposition, in this and next sections, we omit the superscript i when necessary.

A key requirement for nonparametric identification for this class of Bayesian games is the *exclusion restriction* satisfying $F(\cdot|N) = F(\cdot)$ for all $N \geq 2 \in \mathbb{N}^+$, over the support $[\underline{t}, \bar{t}]$. Intuitively, this restriction requires that the CDF of latent types is independent of the number of players participating in the game, implying that participation is exogenous, see Guerre, Perrigne, and Vuong (2009) for the auction case. It is important to point out that, although the latent type distribution is independent of participation, the resulting equilibrium strategy and action distribution still depend on N.[14]

Specifically, let $2 \leq N_1 < N_2$ be two different numbers of players, where we use the index $j = 1,2$ to refer to the level of competition. We exploit the *functional comparative statics* under the exclusion restriction $F(\cdot|N_j) = F(\cdot)$, for $j = 1,2$. When N varies, the equilibrium action varies accordingly. For $j = 1,2$, denote the equilibrium strategy $s_j(\cdot) \equiv s(\cdot|N_j)$, leading to the equilibrium action denoted as a_j with support $[\underline{a}_j, \bar{a}_j]$; the corresponding action distribution is $G_j(\cdot) = G(\cdot|N_j)$.

For a given probability $\alpha \in [0,1]$, let $t(\alpha)$ be the α-th quantile of the private type distribution $F(\cdot)$, i.e. $F(t(\alpha)) = \alpha$. Note that $a_j(\alpha) = s_j(t(\alpha))$, and $G_j(a_j(\alpha)) = \alpha = F(t(\alpha))$ for all $\alpha \in [0,1]$ and $j = 1,2$. When the number of players increases from N_1 to N_2, the equilibrium strategy shifts from $s_1(t(\alpha))$ to $s_2(t(\alpha))$, for all $t \in [\underline{t}, \bar{t}]$, or equivalently for all $\alpha \in [0,1]$. This is why the comparative statics we investigate is in a functional sense. When the equilibrium strategy shifts from $s_1(\cdot)$ to $s_2(\cdot)$, the action distribution shifts accordingly from $G_1(\cdot)$ to $G_2(\cdot)$, for all $\alpha \in [0,1]$.

In short, Equation (2.2.6) can be expressed as a function of $\alpha \in [0,1]$, and, together with the symmetry assumption, shown as

13 For example, in the Tullock contest case, the contest success function is symmetric in $\mathbf{a}_{-i}$ for each contestant i.

14 Note that this exclusion restriction does not mean players' types are uncorrelated with the characteristics of the game; suppose the game has characteristics denoted by a vector $\mathbf{Z}$, which can be continuous or discrete, we allow that $F(\cdot|\mathbf{Z},N) = F(\cdot|\mathbf{Z})$ for all N.

$$t(\alpha) \cdot \mathbb{E}_{\mathbf{a}_{-N}}[\frac{\partial x(a(\alpha), \mathbf{a}_{-N}; N)}{\partial a}] + y'(a(\alpha)) = 0, \tag{2.3.1}$$

where $a(\alpha) = s(t(\alpha))$, for all $\alpha \in [0,1]$, $\mathbb{E}_{\mathbf{a}_{-N}}$ is the expectation over the $N-1$ players' all action profiles, and $\mathbf{a}_{-N}$ denotes the action profile of the $N-1$ players.

The analysis of functional comparative statics when the number of players increases is intractable, due to the generality of the class of Bayesian games, as well as the expectation in equation (2.3.1), which is essentially a multiple integral with an integrand of an arbitrary form taken over all possible action profiles of player i's opponents. When the number of players increases or the game becomes more competitive, there are two challenging features of evaluating this expectation. The first challenge is that the dimension of the joint distribution increases accordingly. The second challenge is that since the equilibrium strategy $s(\cdot)$ is implicitly contained in Equation (2.3.1) for all participating players without a closed-form solution in general, the changing direction and the magnitude of this expectation are even harder to decide analytically.

When the number of players increases from N_1 to N_2, we are interested in the shift of the whole equilibrium strategy function $s(t(\alpha))$ for all $\alpha \in [0,1]$. To utilize this functional comparative statics in the nonparametric identification approaches, we need to discuss every possible shifting pattern, due to the generality of the game. Therefore, graphs are useful in the discussion below.

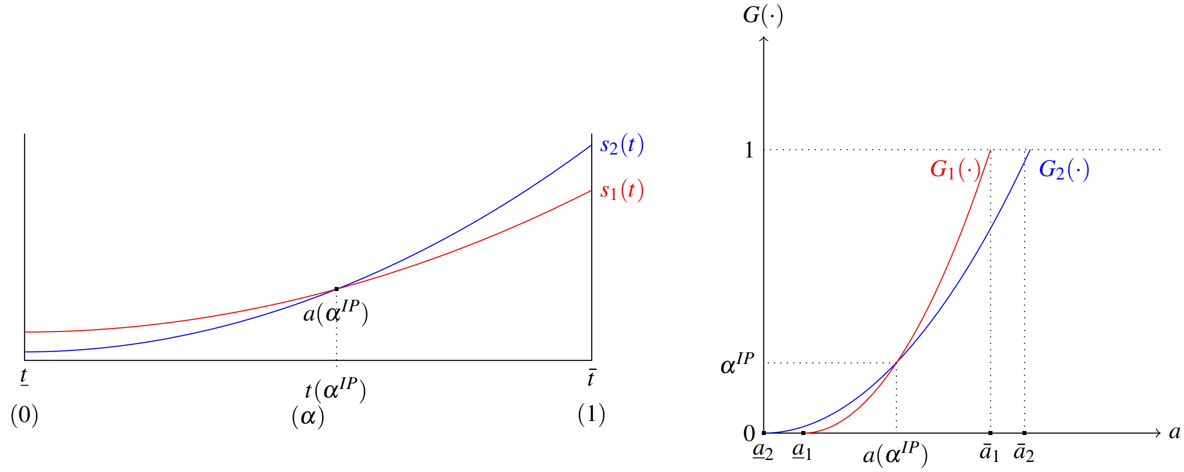
2.3.1 Single Intersection Point

We first consider the case where there is a single intersection point. Here, players with different types change the equilibrium actions towards different directions. Without loss of generality, we assume that the comparative statics correspond to the situation where players with higher types act more and players with lower types act less when the game becomes more competitive.

Figure 2.1a below shows the functional comparative statics in the single-intersection-point case. We draw the corresponding graph of the two action distributions, as shown in Figure 2.1b. Note that for $j = 1,2$, $G_j(a) = Pr(a_j \leq a|N_j) = Pr(t \leq s_j^{-1}(a)|N_j) = Pr(t \leq s_j^{-1}(a)) = F(s_j^{-1}(a))$. Thus when $s_1^{-1}(a) > s_2^{-1}(a)$, $G_1(a) > G_2(a)$, evaluated at one fixed a.

When the equilibrium functions intersect only once, the corresponding action distributions also intersect once at $(a(\alpha^{IP}), \alpha^{IP})$. In the identification section, this single-intersection-point case will be discussed first.

Figure 2.1: Equilibrium strategies and action CDFs with different numbers of players: single intersection point.



(a) Equilibrium strategies.

(b) Action CDFs.

Notes: In Figure 2.1a, the x-axis shows the support $\mathscr{T}$, which can be either labeled as the private type t over the support $[\underline{t}, \bar{t}]$, or α over the support $[0,1]$. Thus evaluated at the same α, the equilibrium actions are shown by the curves $a_1(\alpha)$ and $a_2(\alpha)$, respectively. The intersection of s_1 and s_2 is labeled as $a(\alpha^{IP})$ with the x-axis coordinate $t(\alpha^{IP})$. The action CDFs in Figure 2.1b come directly from Figure 2.1a. When two equilibrium strategy functions intersect once in Figure 2.1a, the corresponding CDFs intersect only once here.

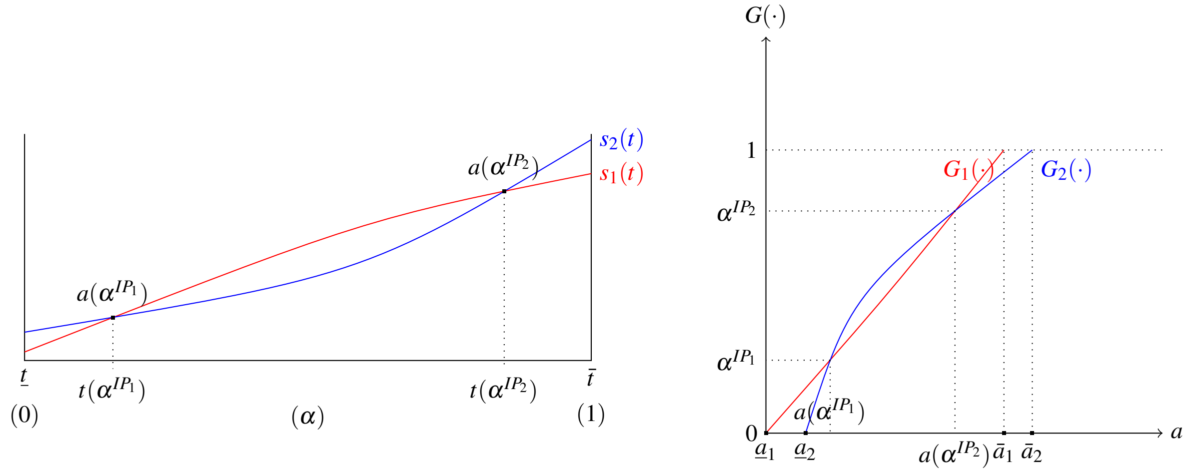
2.3.2 Multiple Intersection Points

The second case we consider is that there are finite $M > 1$ intersection points. Denote the intersection points as $\{\alpha^{IP_1}, \cdots, \alpha^{IP_M}\}$, where $0 \leq \alpha^{IP_1} < \cdots < \alpha^{IP_M} \leq 1$. Note that we do not allow the two equilibrium curves coincide on some open intervals as in this case, shown in Section 2.4.1, we can only identify the endpoints of these open intervals by continuity, but not the intervals themselves.

Figure 2.2a below presents a simple example of this functional comparative statics, where the two equilibrium action functions intersect twice. In this example, when the game becomes more competitive, lower and higher type players increase actions while players in between decrease actions. As a result, the corresponding action functions intersect twice at $(\alpha^{IP_1}, a(\alpha^{IP_1}))$ and $(\alpha^{IP_2}, a(\alpha^{IP_2}))$. Therefore when $\alpha \in [0, \alpha^{IP_1}) \cup (\alpha^{IP_2}, 1]$, $b_2(\alpha) > b_1(\alpha)$; when $\alpha \in (\alpha^{IP_1}, \alpha^{IP_2})$, $b_2(\alpha) < b_1(\alpha)$. Figure 2.2b below shows the corresponding action distributions.

As shown, the two action distributions intersect twice at $(a(\alpha^{IP_1}), \alpha^{IP_1})$ and $(a(\alpha^{IP_2}), \alpha^{IP_2})$. The identification problem in this case will be discussed as a direct application of the identification approach introduced for the single-intersection-point case.

Figure 2.2: Equilibrium strategies and action CDFs with different numbers of players: multiple intersection points.



(a) Equilibrium strategies.

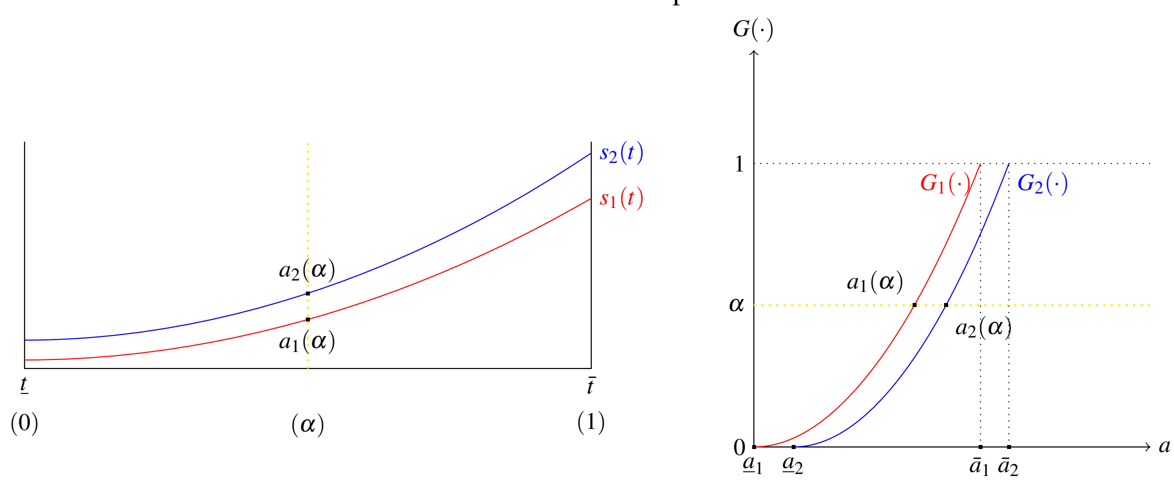
(b) Action CDFs.

Notes: In Figure 2.2a, there are two intersections of s_1 and s_2, denoted by $a(\alpha^{IP_1})$ and $a(\alpha^{IP_2})$, with the x-axis coordinates $t(\alpha^{IP_1})$ and $t(\alpha^{IP_2})$. The action CDFs in Figure 2.2b come directly from Figure 2.2a. When two equilibrium strategy functions intersect twice in Figure 2.2a, the corresponding CDFs intersect twice here.

2.3.3 No Intersection Point

The last shifting pattern of the equilibrium strategy function is when $s_1(\cdot)$ and $s_2(\cdot)$ do not intersect with each other over $\mathscr{T}$. Without loss of generality, we assume that more competition results in the upward shifting of the whole curve, as shown in Figure 2.3a. The corresponding action distributions $G_1(\cdot)$ and $G_2(\cdot)$ are shown in Figure 2.3b. There is no intersection of the two distributions.

Figure 2.3: Equilibrium strategies and action CDFs with different numbers of players: no intersection point.



(a) Equilibrium strategies.

(b) Action CDFs.

Notes: There is no intersection of s_1 and s_2 in Figure 2.3a. The action CDFs in Figure 2.3b come directly from Figure 2.3a. When two equilibrium strategy functions do not intersect in Figure 2.3a, the corresponding CDFs do not intersect here.

2.3.4 How to Determine the Pattern?

Since we deal with Bayesian games with a general structure, all of the three patterns can happen depending on a specific application under consideration. We first examine the patterns in the aforementioned examples.

Example 1' *Differentiated Cournot competition:* *Assume $c'(a) \geq 0$. The FOC yields:*

$$t_i - \beta[2a(t_i) + (N-1)\mathbb{E}(a(t))] - c'(a(t_i)) = 0, \forall t_i.$$

It is easy to see that when N changes, the solution should change, otherwise the FOC cannot hold at the same time for different N.[15] *For any pair of types, say t_i and t_j, $i \neq j$, we have*

$$t_i - 2\beta a(t_i) - c'(a(t_i)) = t_j - 2\beta a(t_j) - c'(a(t_j)).$$

Note that this equality does not depend on N directly. The LHS is strictly decreasing in $a(t_i)$ and the RHS is strictly decreasing in $a(t_j)$. Therefore, when N changes, the solution changes and $a(t)$ changes in the same direction for all t and there will not be crossing.

Example 2' *Diamond's search model:* *Assume $g(x) = x, c'(a) > 0, c''(a) > 0$, then the FOC yields:*

$$t_i(N-1)\mathbb{E}(a(t)) - c'(a(t_i)) = 0.$$

It is easy to see that when N changes, the solution should change, otherwise the FOC cannot hold at the same time for different N. For any pair of types, say t_i and t_j, $i \neq j$, we have

$$\frac{c'(a(t_i))}{t_i} = \frac{c'(a(t_j))}{t_j}.$$

Note that this equality does not depend on N directly. The LHS is strictly increasing in $a(t_i)$ and the RHS is strictly increasing in $a(t_j)$. Therefore, when N changes, the solution changes and $a(t)$ changes in the same direction for all t and there will not be crossing.

Example 3' *Public good provision:* *Suppose $k(a) = a^2/2$, and $c'(0) > 0$. Then the FOC yields:*

$$t_i[a(t_i) + (N-1)\mathbb{E}(a(t))] - c'(a(t_i)) = 0, \forall t_i.$$

SOC requires that

$$t_i - c''(a(t_i)) \leq 0, \ \forall t_i.$$

15 Formally, it is easy to see that $a(t) = 0, \forall t$ is not a solution and therefore $\mathbb{E}(a(t)) > 0$. Thus, when N changes, at least one type has to change the action, otherwise, the above FOC cannot hold at the same time under different N.

For any pair of types, say t_i and t_j, $i \neq j$, we have

$$\frac{c'(a(t_i)) - t_i a(t_i)}{t_i} = \frac{c'(a(t_j)) - t_j a(t_j)}{t_j}.$$

Note that this equality does not depend on N directly. In addition, if we assume that SOC holds strictly, the LHS is strictly decreasing in $a(t_i)$ and the RHS is strictly decreasing in $a(t_j)$. Therefore, when N changes, if the solution changes then $a(t)$ changes in the same direction for all t and there will not be crossing.

Example 4’ ***Tullock contest:*** *Assume $r = 1, c(a) = a$, t is uniformly distributed on $[0.5, 1.5]$. Wasser (2013) numerically shows that the action functions for $N = 2$ and $N = 3$ cross with each other, as well as $N = 2$ and $N = 5$. However, the action functions for $N = 3$ and $N = 5$ do not cross.*

While Examples 1’-3’ exhibit the no-intersection pattern, this does not necessary imply that no-intersection is more relevant in applications, since the results heavily rely on further assumptions on the model primitives. However, this suggests that the no-intersection case is important, if not more important than the intersection cases. From an econometric perspective, on the other hand, since the players’ actions are observed, the distribution of actions can be consistently estimated. Therefore the patterns of the action distributions can be inferred from the data. The next proposition establishes equivalence between the crossing patterns of two action strategies and their corresponding distributions.

Proposition 5 *The two equilibrium action strategies intersect with each other if and only if their action distributions intersect with each other.*

Proposition 5 implies that the crossing condition is straightforward to verify from the data, as it can be inferred from estimating the action distributions.

2.4 Contraction Sequence and Nonparametric Identification

In this section, we establish the nonparametric identification results of the structure $[F(\cdot), y(\cdot)]$ under the exclusion restriction $F(\cdot|N) = F(\cdot)$, in different cases of functional comparative statics as shown in Section 2.3, when the number of players increases from N_1 to N_2. While we assume independence between $F(\cdot)$ and N, the equilibrium action distribution G still depends on N. Therefore, Assumption 2 is still satisfied by $G(\cdot|N)$, given N.

Note that the expectation in Equation (2.3.1) is strictly positive for all $a \in [\underline{a}(N), \bar{a}(N)]$, due to Condition 2.2.2 that the known function $x(a(\alpha), \mathbf{a}_{-N}; N)$ has to satisfy. By rearranging terms, we get the inverse of the action strategy function $s(\cdot|N)$, satisfying the following equation for an interior solution

$$
\begin{aligned}
t(\alpha) &= -y'(a(\alpha)) \cdot \left\{ \mathbb{E}_{\mathbf{a}_{-N}}[\tfrac{\partial x(a(\alpha),\mathbf{a}_{-N};N)}{\partial a}] \right\}^{-1} \\
&\equiv \xi(a(\alpha); x, G, y | N),
\end{aligned} \tag{2.4.1}
$$

which expresses each participant's private type as a function ξ of the equilibrium action, the function x, the number of players, and the equilibrium action CDF G (all are known) as well as the function y (which is latent and needs to be recovered). Note that $\xi^{-1}(\cdot|N) = s(\cdot|N)$ for an interior solution, and Equation (2.4.1) holds for any probability $\alpha \in [0,1]$. We denote the corresponding ξ function as $\xi_j(\cdot)$ for $j = 1, 2$.

By stacking the key equations in (2.3.1) together when N increases from N_1 to N_2, we get the following system of two equations:

$$
\begin{aligned}
t(\alpha) \cdot \mathbb{E}_{\mathbf{a}_{-N_1}}[\tfrac{\partial x(a_1(\alpha),\mathbf{a}_{-N_1};N_1)}{\partial a_1}] + y'(a_1(\alpha)) = 0, \\
t(\alpha) \cdot \mathbb{E}_{\mathbf{a}_{-N_2}}[\tfrac{\partial x(a_2(\alpha),\mathbf{a}_{-N_2};N_2)}{\partial a_2}] + y'(a_2(\alpha)) = 0.
\end{aligned} \tag{2.4.2}
$$

A key step is to analyze the ratio of two key equations in the above system

$$
\frac{\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(\alpha),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]}{\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(\alpha),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]} = \frac{y'(a_1(\alpha))}{y'(a_2(\alpha))}, \tag{2.4.3}
$$

where the common term $t(\alpha)$, being the private type evaluated at one pre-determined α, is canceled out, under the exclusion restriction $F(t(\alpha)|N) = F(t(\alpha))$. On the left-hand side, we have the ratio of expected derivatives of function x when N varies, for the player whose private type is evaluated at the α-th quantile. On the right-hand side, we have the ratio of the derivatives of function y, evaluated at the actions for the same player when N varies. Since the ratio on the left hand side is known, the ratio on the right hand side is thus identified nonparametrically for all $\alpha \in [0,1]$. If $y'(\cdot)$ can be identified, the private types can be backed out from the first order condition (2.4.1).

It turns out whether $y'(\cdot)$ can be identified depends on whether the two equilibrium action strategies cross with each other. Proposition 5 implies that the crossing condition is straightforward to verify from the data. Thus the pattern of functional comparative statics can be decided from the data. We can characterize all the patterns into the two categories, namely, whether the two equilibrium action strategies intersect or not. In the rest of this section, we first derive the point identification result, through the construction of a *contraction sequence*, for the case with intersections. Then we shift to the no-intersection-point case, and establish the partial identification result. Throughout the discussion in this section, we maintain the following assumption.

Assumption 3 *Consider two numbers of players: $N_1 < N_2$. The corresponding conditional action CDFs satisfy: $G_j(\cdot) \in \mathscr{G}$ with the support $[\underline{a}_j, \bar{a}_j]$, for $j = 1, 2$.*

2.4.1 Intersection: Point Identification

Single Intersection

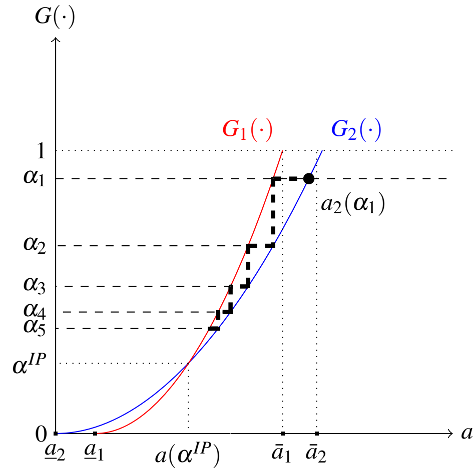
In the case where there is only one intersection point of the two equilibrium action strategies, we can follow Guerre, Perrigne, and Vuong (2009) to construct a contraction sequence which is a strictly monotonic sequence of probabilities converging to the intersection point, denoted as $\{\alpha_n\}_{n=1}^{\infty}$. The idea of nonparametric identification relies on exploiting the variations in the quantiles of two equilibrium action CDFs with two different numbers of players, while the quantiles of the CDF of private types do not change. Our identification strategy distinguishes from Guerre, Perrigne, and Vuong (2009) in the sense that we directly work on the action distributions instead of the equilibrium strategy functions, and we construct the sequence of quantiles in a different way. More importantly, their identification strategy only works for the benchmark case when the two equilibrium action strategies intersect, but not for the no-intersection-point case.

To see how the contraction sequence approach works, a graphic analysis is useful. Consider the case in Figure 2.1b. We choose an arbitrary starting probability $\alpha_1 > \alpha^{IP}$, and try to identify $y'(a_2(\alpha_1))$.[16] The identification approach is shown in the following figure.

Note that from Equation (2.4.3), the identification of the ratio $y'(a_2(\alpha_1))/y'(a_1(\alpha_1))$ is directly obtained through $\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(\alpha_1),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]/\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(\alpha_1),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]$. In order to disentangle $y'(a_2(\alpha_1))$, we construct a strictly decreasing contraction sequence $\{\alpha_n\}_{n=1}^{\infty}$, which satisfies $a_1(\alpha_n) = a_2(\alpha_{n+1})$, for $n > 1$. To begin with, note that $a_1(\alpha_1) = a_2(\alpha_2)$, thus $y'(a_1(\alpha_1)) = y'(a_2(\alpha_2))$. Evaluated at the new probability α_2, the ratio $y'(a_2(\alpha_2))/y'(a_1(\alpha_2))$ is obtained from Equation (2.4.3) by $\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(\alpha_2),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]/\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(\alpha_2),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]$. By repeating this procedure, for each n, the ratio $y'(a_2(\alpha_n))/y'(a_1(\alpha_n))$ is identified nonparametrically. Furthermore, when $n \to \infty$, $a_1(\alpha_n) \to a(\alpha^{IP})$

16 There is no special reason why we choose a starting point above the intersection point $(a(\alpha^{IP}), \alpha^{IP})$. The procedure is identical if we choose a starting point below it, leading to a strictly increasing sequence of probabilities.

Figure 2.4: Identification through contraction sequence in single-intersection-point case.



Notes: To identify $y'(a_2(\alpha_1))$, we construct a strictly decreasing sequence of probabilities $\{\alpha_n\}_{n=1}^{\infty}$, satisfying $a_1(\alpha_n) = a_2(\alpha_{n+1})$. As $n \to \infty$, $\alpha_n \to \alpha^{IP}$, $a_1(\alpha_n) \to a(\alpha^{IP})$, and $a_2(\alpha_n) \to a(\alpha^{IP})$.

and $a_2(\alpha_n) \to a(\alpha^{IP})$, thus the ratio $y'(a_2(\alpha_n))/y'(a_1(\alpha_n)) \to 1$. Notice that

$$\begin{aligned} y'(a_2(\alpha_1)) &= \tfrac{y'(a_2(\alpha_1))}{y'(a_1(\alpha_1))} \times \tfrac{y'(a_2(\alpha_2))}{y'(a_1(\alpha_2))} \times \cdots \times \tfrac{y'(a_2(\alpha_n))}{y'(a_1(\alpha_n))} \times \cdots \times y'(a(\alpha^{IP})), \\ &= y'(a(\alpha^{IP})) \times \Pi_{n=1}^{\infty} \tfrac{y'(a_2(\alpha_n))}{y'(a_1(\alpha_n))}. \end{aligned} \tag{2.4.4}$$

The above equation holds, because (1) the condition $y'(a_1(\alpha_n)) = y'(a_2(\alpha_{n+1}))$ leads to the cancellation of all intermediate terms, and (2) $y'(a_2(\alpha_n))/y'(a_1(\alpha_n)) \to 1$ as $n \to \infty$. Also, note that since the value $y'(a(\alpha^{IP}))$ is unknown, this point identification result holds only up to a scale. Knowing $y'(a(\alpha^{IP}))$ is equivalent to knowing $t(\alpha^{IP})$, since they are linked to each other through the first order condition (2.3.1).[17] Without loss of generality, we use $y'(a(\alpha^{IP}))$ as the scale for the point identification approach.

Another point worth noting is that the sign restriction imposed in Assumption 1-(ii)-(2) that the sign of $y'(\cdot)$ is known over its support plays a role in the identification of $y'(a_2(\alpha_1))$. Note that we utilize the system of equations (2.4.2) to cancel out the α_1-th quantile of the private type $t(\alpha_1)$, the same for both games. However, the supermodularity assumption regarding function $x(a(\alpha_1), \mathbf{a}_{-N}; N)$ shown in Condition 2.2.2 implies the strictly positive derivative of $x(a(\alpha_1), \mathbf{a}_{-N}; N)$ with respect to $a(\alpha_1)$, the first argument in x. This strictly positive derivative appears in the expectation on the right hand side

17 The scale can be defined through other values of quantile, by noting that from the function value $y'(a_j(\alpha))$ (or, equivalently $t(\alpha)$) at an arbitrary α for either $j = 1$ or 2, we can back out the value of $y'(a(\alpha^{IP}))$ and the identification approach goes through. Alternatively, we can assume that at α^{IP}, $t(\alpha^{IP})$ is normalized to be one, thus both $y'(a_1(\alpha^{IP}))$ and $y'(a_2(\alpha^{IP}))$ are known.

in the system of equations (2.4.2). Consequently, the sign of unchanged $t(\alpha_1)$ determines the sign of $y'(a_2(\alpha_1))$ (and $y'(a_1(\alpha_1))$). Since we do not restrict the sign of private type $t(\alpha_1)$, it is possible for $y'(a_2(\alpha_1))$ (and $y'(a_1(\alpha_1))$) to be negative, but by canceling out $t(\alpha_1)$, we always get strictly positive right hand side ratio in Equation (2.4.3). Further through the product of ratios in Equation (2.4.4), we always have strictly positive $y'(a_2(\alpha_1))$ from the benchmark contraction sequence approach, if $y'(a(\alpha^{IP}))$ is assumed to be positive. In this sense, the benchmark nonparametric identification we propose in this subsection can only identify $y'(a_2(\alpha_1))$, up to its absolute value. By assuming that we know the sign of $y'(\cdot)$, we know the sign of $y'(a_2(\alpha_1))$. As a result, the identification of $y'(a_2(\alpha_1))$ is obtained.[18] Moreover, the sign restriction plays a similar role in Sections 2.4.1 and 2.4.2.

By varying the starting probability α_1 given $\alpha_1 > \alpha^{IP}$, we can use the strictly decreasing contraction sequence approach to nonparametrically identify $y'(a_2(\cdot))$ over the support $(\alpha^{IP}, 1]$, up to the scale $y'(a(\alpha^{IP}))$. Through a similar procedure by constructing a strictly increasing contraction sequence, $y'(a_2(\cdot))$ is identified over $[0, \alpha^{IP})$, up to the scale $y'(a(\alpha^{IP}))$. In this way, $y'(a_2(\cdot))$ is nonparametrically identified over $[0,1]$; or equivalently, $y'(\cdot)$ is nonparametrically identified over $[\underline{a}_2, \bar{a}_2]$, up to the scale $y'(a(\alpha^{IP}))$.

Recall that, when evaluated at the same α, $y'(a_2(\alpha))/y'(a_1(\alpha))$ is identified. Together with the above argument that $y'(a_2(\cdot))$ is nonparametrically identified over $[0,1]$, we directly obtain the identification of $y'(a_1(\alpha))$ over the whole support of α, i.e. $[0,1]$, up to the scale $y'(a(\alpha^{IP}))$. Equivalently, $y'(\cdot)$ is nonparametrically identified over $[\underline{a}_1, \bar{a}_1]$. In summary, $y'(\cdot)$ is nonparametrically identified over $[min\{\underline{a}_1, \underline{a}_2\}, max\{\bar{a}_1, \bar{a}_2\}]$, up to the scale $y'(a(\alpha^{IP}))$.

Identifying $F(\cdot)$ over the support $[\underline{t}, \bar{t}]$ can be achieved through the inverse of the strategy function ξ_j stated in Equation (2.4.1), through which private types are recovered using either $j = 1$ or 2, up to the scale $t(\alpha^{IP})$. As a result, $F(\cdot)$ can be nonparametrically identified by noting that $F(t) = G_j(\xi_j^{-1}(t))$, for any $t \in [\underline{t}, \bar{t}]$ and for $j = 1$ or 2. Again, since $G_j(\xi_j^{-1}(t(\alpha^{IP})))$ is unknown, this identification holds up to a scale of $F(t(\alpha^{IP})) = G_j(\xi_j^{-1}(t(\alpha^{IP})))$, for $j = 1$ *or* 2.

Theorem 1 *Suppose Assumption 3 and the additional assumption hold: there is only one intersection point of two corresponding action distributions $G_1(\cdot)$ and $G_2(\cdot)$, denoted as: $(a(\alpha^{IP}), \alpha^{IP})$. Then under the exclusion restriction that $F(\cdot) = F(\cdot|N_1) = F(\cdot|N_2)$,*

18 In the Supplemental Appendix, where the corresponding assumptions and identification results of the second and third forms of the payoff function are given, we show that the identification of the second form also needs the restriction of sign. On the other hand, the identification of the third form does not need the sign restriction, implying that we can identify the exact derivative $y'(a_2(\alpha_1))$.

(1) $y'(\cdot)$ is nonparametrically identified over $[min\{\underline{a}_1,\underline{a}_2\},max\{\bar{a}_1,\bar{a}_2\}]$, up to the scale $y'(a(\alpha^{IP}))$.

(2) $F(\cdot)$ is nonparametrically identified over $[\underline{t},\bar{t}]$, with $F(\cdot) = G_1(\xi_1^{-1}(\cdot)) = G_2(\xi_2^{-1}(\cdot))$, up to the scale $F(t(\alpha^{IP})) = G_1(\xi_1^{-1}(t(\alpha^{IP}))) = G_2(\xi_2^{-1}(t(\alpha^{IP})))$.

Note that in the single-intersection-point case, the contraction sequence $\{\alpha_n\}_{n=1}^{\infty}$ is unique. Thus $y'(\cdot)$ and $F(\cdot)$ are uniquely identified over their identifiable supports, up to a scale.

Multiple Intersections

The identification in this subsection is a suitable adaption of the contraction sequence approach adopted in the single-intersection-point case. Since in the multiple-intersection-point case, there are always intersection points of two equilibrium action CDFs, we can obtain the point identification result as in Section 2.4.1.

As in Section 2.3.2, we do not allow the two action distributions to coincide on some open intervals; instead, they intersect at $M > 1$ finite points. Denote the intersection points as $\{\alpha^{IP_1}, \cdots, \alpha^{IP_M}\}$, where $0 \leq \alpha^{IP_1} < \cdots < \alpha^{IP_M} \leq 1$. The M intersection points divide the 2-dimensional space containing the two action distributions into a finite number of parts. Note that while in the simple example in Section 2.3.2 shown in Figure 2.2b, the two intersection points divide the space into 4 parts, and the number of the parts is not definitely $M+2$, since we allow the first and the last intersection points satisfying $\alpha^{IP_1} = 0$ or $\alpha^{IP_M} = 1$. For nonparametric identification, we then show how to identify $y'(\cdot)$ (up to a scale) over $[\underline{a}_2, \bar{a}_2]$ in this case. With the first intersection point α^{IP_1}, the point identification result holds with $t(\alpha^{IP_1})$ normalized to be one and thus $y'(a_1(\alpha^{IP_1}))$ and $y'(a_2(\alpha^{IP_1}))$ both known. Without loss of generality, we use this normalization.

Note that if we choose an arbitrary part in the 2-dimensional space containing the two distribution curves, the following four different possibilities can occur: (1) G_2 is above G_1, and both endpoints are intersection points, (2) G_2 is above G_1, and only one endpoint is intersection point, (3) G_2 is below G_1, and both endpoints are intersection points, and (4) G_2 is below G_1, and only one endpoint is intersection point. (2) and (4) can only happen in the first and the last parts. (1) and (3) can happen in any part. In the example shown in Figure 2.2b, when $a \in [\underline{a}_2, a(\alpha^{IP_1})]$, $G_2(a) < G_1(a)$, only the right endpoint is intersection point; when $a \in [a(\alpha^{IP_1}), a(\alpha^{IP_2})]$, $G_2(a) > G_1(a)$, both endpoints are intersection points; when $a \in [a(\alpha^{IP_2}), \bar{a}_2]$, $G_2(a) < G_1(a)$, the left endpoint is intersection point.

We can use the contraction sequence approach in the benchmark case in Section 2.4.1 to identify

$y'(\cdot)$. Choose any part from the x-axis in the action CDF graph corresponding to $[a(\alpha^{IP_m}), a(\alpha^{IP_{m+1}})]$, which satisfies one of the four possibilities depicted above. Notice that here we allow that $a(\alpha^{IP_m}) = min\{\underline{a}_1, \underline{a}_2\}$, or $a(\alpha^{IP_{m+1}}) = max\{\bar{a}_1, \bar{a}_2\}$. In order to identify $y'(\cdot)$ over $(a(\alpha^{IP_m}), a(\alpha^{IP_{m+1}}))$, we choose any a in this open interval. The identification then relies on which event happens:

(i) First, the contraction sequence of probabilities $\{\alpha_n\}_{n=0}^{\infty}$ for identification is chosen depending on the location of the intersection point. When only one endpoint is the intersection point, the right endpoint, whose corresponding $\alpha^{IP_{m+1}}$ is the limit of the sequence, implies that the sequence is increasing; the left endpoint, whose corresponding, α^{IP_m} is the limit of the sequence, implies that the sequence is decreasing. When both endpoints are intersection points, one can choose either endpoint as the limit of the sequence. The identification of $y'(a)$ is unique, because we rely on Equation (2.4.4) which is an identity where the intermediate ratios are canceled out.

(ii) When G_2 is above G_1, and we choose the right endpoint as the limit of the contraction sequence of probabilities $\{\alpha_n\}_{n=0}^{\infty}$, the identification of $y'(a_2(\alpha_0))$ is determined by Equation (2.4.4) directly, without any adjustment. Instead, if we choose the left endpoint as the limit, Equation (2.4.4) is applied with a minor adaption, by noting that $y'(a_2(\alpha_0)) = y'(a_1(\alpha_1))$ since $a_2(\alpha_0) = a_1(\alpha_1)$ and $a_1(\alpha_1)$ is right below $a_2(\alpha_0)$. Therefore, $y'(a_1(\alpha_1))$ is identified, and $y'(a_2(\alpha_0))$ is the same as $y'(a_1(\alpha_1))$.

(iii) When G_2 is below G_1, choosing the right endpoint as the limit requires the minor adaption mentioned above to identify $y'(a_2(\alpha_0))$ and $a_1(\alpha_1)$ is right above $a_2(\alpha_0)$; and choosing the left endpoint as the limit gives the direction identification of $y'(a_2(\alpha_0))$.

These discussions show that we identify $y'(\cdot)$ over all open intervals. Lastly, the endpoints of these open intervals can be determined by continuity. The next proposition summarizes the identification results of $y'(\cdot)$ and $F(\cdot)$ in the multiple-intersection-point case.

Proposition 6 *Suppose Assumption 3 and the additional assumption hold: there are $M > 1$ finite intersection points of two corresponding distributions $G_1(\cdot)$ and $G_2(\cdot)$. Then*

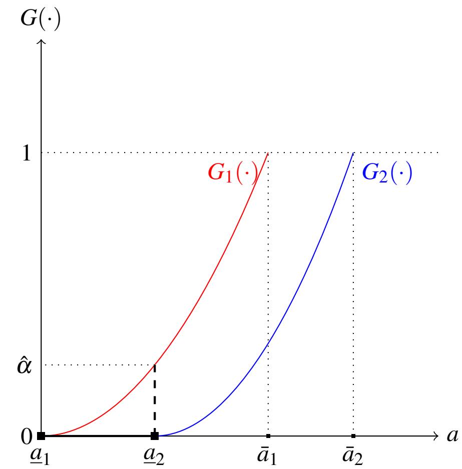
(1) $y'(\cdot)$ is nonparametrically identified over $[min\{\underline{a}_1, \underline{a}_2\}, max\{\bar{a}_1, \bar{a}_2\}]$, up to a scale.

(2) $F(\cdot)$ is nonparametrically identified over $[\underline{t}, \bar{t}]$ with $F(\cdot) = G_2(\xi_2^{-1}(\cdot)) = G_1(\xi_1^{-1}(\cdot))$, up to a scale.

2.4.2 No-Intersection: Partial Identification

When two action CDFs do not intersect with each other, as shown in Figure 2.3b, in general we do not have point identification even under the exclusion restriction. We first illustrate it using the following graph.

Figure 2.5: No point identification in no-intersection-point case.



Notes: Any functional form of $y'(\cdot)$ over the support $[\underline{a}_1, \underline{a}_2]$ can be chosen, but still leading to the same ratios $y'(a_1(\alpha))/y'(a_2(\alpha))$ for all $\alpha \in [0,1]$ and the same CDFs G_1 and G_2, which are known. Thus $y'(\cdot)$ is not pointly-identified over either $[\underline{a}_1, \bar{a}_1]$ or $[\underline{a}_2, \bar{a}_2]$, under the exclusion restriction $F(\cdot) = F(\cdot|N_1) = F(\cdot|N_2)$.

As shown in Figure 2.5, evaluated at any action a between $[\underline{a}_1, \underline{a}_2]$, we can assign any value to the function $y'(\cdot)$, which corresponds to the function value $y'(a_1(\alpha))$ at $\alpha \in [0, \hat{\alpha}]$ along the known CDF G_1. Then, we can derive $y'(a_2(\alpha))$ through the known ratio $y'(a_1(\alpha))/y'(a_2(\alpha))$ which is known as it is determined by the ratio $\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(\alpha), \mathbf{a}_{-N_1}; N_1)}{\partial a_1}]/\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(\alpha), \mathbf{a}_{-N_2}; N_2)}{\partial a_2}]$ from Equation (2.4.3), at the same quantile, along the known CDF G_2. Repeat this procedure, we can finally determine all function value $y'(\cdot)$ over $[\underline{a}_2, \bar{a}_2]$, by choosing *any arbitrary function* of $y'(\cdot)$ over $[\underline{a}_1, \underline{a}_2]$. This implies that $y'(\cdot)$ is not pointly identified, since we can essentially derive the same known information with the freedom of picking any function $y'(\cdot)$ over the continuum $[\underline{a}_1, \underline{a}_2]$. This non-identification result is given in the following proposition.

Proposition 7 *Suppose Assumption 3 and the additional assumption hold: there is no intersection point of two corresponding action distributions $G_1(\cdot)$ and $G_2(\cdot)$. Then under the exclusion restriction that $F(\cdot) = F(\cdot|N_1) = F(\cdot|N_2)$, $y'(\cdot)$ is not pointly identified over $[min\{\underline{a}_1, \underline{a}_2\}, max\{\bar{a}_1, \bar{a}_2\}]$. Specifically, for a function $y'(\cdot)$, there exists another function $y'^*(\cdot)$ observationally equivalent to $y'(\cdot)$, in the sense that both functions induce the same CDFs $G_1(\cdot)$ and $G_2(\cdot)$ over their corresponding supports $[\underline{a}_1, \bar{a}_1]$ and $[\underline{a}_2, \bar{a}_2]$, as well as the same ratios $y'(a_1(\alpha))/y'(a_2(\alpha))$ for any $\alpha \in [0,1]$.*

However, we can get bounds for the function $y'(\cdot)$, thus deriving partial identification result (up to a scale). We first assume that there exists one α^* whose corresponding private type $t(\alpha^*)$ is normalized to be one, from which the values $y'(a_1(\alpha^*))$ and $y'(a_2(\alpha^*))$ can be known from the system of equations (2.4.2).[19] Starting from this normalized α^*, we obtain a finite set of probabilities, whose corresponding function values of $y'(a_1(\cdot))$ and $y'(a_2(\cdot))$ are known, denoted as $\{q_1, q_2, \cdots, q_I\}$, designed to be a strictly increasing sequence, meaning that $q_i < q_{i+1}$, and satisfy the condition: $a_1(q_{i+1}) = a_2(q_i)$ for $i < I$. We consider the pairs $(y'(a_1(q_1)), y'(a_2(q_1)))$, $\cdots$, and $(y'(a_1(q_I)), y'(a_2(q_I)))$, where we express every value as a function of $y'(a_1(q_1))$ in the following fashion:

$$\begin{aligned}
& y'(a_2(q_1)) = \frac{\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(q_1),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]}{\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(q_1),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]} y'(a_1(q_1)), \\
& y'(a_1(q_2)) = y'(a_2(q_1)) = \frac{\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(q_1),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]}{\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(q_1),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]} y'(a_1(q_1)), \\
& \cdots, \\
& y'(a_1(q_I)) = y'(a_2(q_{I-1})) = \prod_{i=1}^{I-1} \frac{\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(q_i),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]}{\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(q_i),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]} y'(a_1(q_1)), \\
& y'(a_2(q_I)) = \prod_{i=1}^{I} \frac{\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(q_i),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]}{\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(q_i),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]} y'(a_1(q_1)),
\end{aligned} \tag{2.4.5}$$

due to the ratio of the two first order conditions as in (2.4.3) and the properties of the finite set of probabilities $\{q_1, q_2, \cdots, q_I\}$.

Without loss of generality, we assume that $y(\cdot)$ is a strictly decreasing and concave function, which implies $y'(a_2(q_i)) \le y'(a_1(q_i))$ for each i. Thus we have I bounds from these paired function values: $[y'(a_2(q_1)), y'(a_1(q_1))]$, $\cdots$, $[y'(a_2(q_I)), y'(a_1(q_I))]$. The identified sets are constructed from these bounds as follows.

Given the finite set $\{q_1, q_2, \cdots, q_I\}$, the whole support $[0,1]$ is divided into J parts, where J can take three possible values: $I-1$, I, and $I+1$, which translates into four different cases: (1) when $J = I-1$, and $q_1 = 0$, $q_I = 1$, (2.a) when $J = I$, and $q_1 = 0$, $q_I < 1$, (2.b) when $J = I$, and $q_1 > 0$, $q_I = 1$, (3) when $J = I+1$, and $q_1 > 0$, $q_I < 1$.

Consider another strictly increasing sequence of probabilities $\{\alpha_1, \cdots, \alpha_J\}$, with each element falling into one of the J parts. Next we will construct bounds for each α_j, for $j = 1, \cdots, J$. Naturally,

19 This pre-specified probability serves as the scale in our partial identification approach. As in the point identification for the single-intersection-point case, we can choose an arbitrary probability for normalization. But at the normalized probability, the private type value along with the function values of y' are unknown; thus it is impossible to compare bounds derived from different normalized probabilities, because this entails comparing the unknown function values at different normalized probabilities. Therefore, to fix the idea, we choose the probability to be normalized first, then derive bounds based on this normalized probability.

an arbitrary α_j will have I bounds derived from the I pairs of known values, and then we can pick the tightest bound.

We now discuss the first case as an example for illustration, and leave the discussion on the remaining cases to the Supplemental Appendix. For ease of illustration, we adopt the following notation:

$$\frac{\mathbb{E}_1(a_1(\alpha))}{\mathbb{E}_2(a_2(\alpha))} \equiv \frac{\mathbb{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_1(\alpha),\mathbf{a}_{-N_1};N_1)}{\partial a_1}]}{\mathbb{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_2(\alpha),\mathbf{a}_{-N_2};N_2)}{\partial a_2}]},$$

which is always strictly positive due to the strict supermodularity condition 2.2.2.

Case 1: $J = I - 1$

For an arbitrary $\alpha_j \in (q_j, q_{j+1})$, bounds using $q_{j+1}, \cdots, q_{J+1}$ come from bounding $y'(a_2(\alpha_j)), \cdots, y'(a_2(\alpha_J))$, while bounds using $q_1, \cdots, q_j$ come from bounding $y'(a_1(\alpha_1)), \cdots, y'(a_1(\alpha_j))$. Note that we can always express $y'(a_2(\alpha_j))$ in the following two ways:

$$\begin{aligned}
y'(a_2(\alpha_j)) =& y'(a_1(\alpha_{j+1})) \\
=& \frac{\mathbb{E}_1(a_1(\alpha_{j+1}))}{\mathbb{E}_2(a_2(\alpha_{j+1}))} y'(a_2(\alpha_{j+1})) \\
=& \frac{\mathbb{E}_1(a_1(\alpha_{j+1}))}{\mathbb{E}_2(a_2(\alpha_{j+1}))} \frac{\mathbb{E}_1(a_1(\alpha_{j+2}))}{\mathbb{E}_2(a_2(\alpha_{j+2}))} y'(a_2(\alpha_{j+2})) \\
=& \cdots \\
=& \frac{\mathbb{E}_1(a_1(\alpha_{j+1}))}{\mathbb{E}_2(a_2(\alpha_{j+1}))} \frac{\mathbb{E}_1(a_1(\alpha_{j+2}))}{\mathbb{E}_2(a_2(\alpha_{j+2}))} \cdots \frac{\mathbb{E}_1(a_1(\alpha_J))}{\mathbb{E}_2(a_2(\alpha_J))} y'(a_2(\alpha_J)),
\end{aligned}$$

or

$$\begin{aligned}
y'(a_2(\alpha_j)) =& \frac{\mathbb{E}_2(a_2(\alpha_j))}{\mathbb{E}_1(a_1(\alpha_j))} y'(a_1(\alpha_j)) \\
=& \frac{\mathbb{E}_2(a_2(\alpha_j))}{\mathbb{E}_1(a_1(\alpha_j))} \frac{\mathbb{E}_2(a_2(\alpha_{j-1}))}{\mathbb{E}_1(a_1(\alpha_{j-1}))} y'(a_1(\alpha_{j-1})) \\
=& \cdots \\
=& \frac{\mathbb{E}_2(a_2(\alpha_j))}{\mathbb{E}_1(a_1(\alpha_j))} \frac{\mathbb{E}_2(a_2(\alpha_{j-1}))}{\mathbb{E}_1(a_1(\alpha_{j-1}))} \cdots \frac{\mathbb{E}_2(a_2(\alpha_1))}{\mathbb{E}_1(a_1(\alpha_1))} y'(a_1(\alpha_1)).
\end{aligned}$$

Thus, the first group of bounds are constructed as follows:

The bound using q_{j+1} has a coefficient $y'(a_1(q_1)) \prod_{l=1}^{j} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}$ on $[1, \frac{\mathbb{E}_2(a_2(q_{j+1}))}{\mathbb{E}_1(a_1(q_{j+1}))}]$.

The bound using q_{j+2} has a coefficient $y'(a_1(q_1)) \prod_{k=j+1}^{j+1} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}) \prod_{l=1}^{j+1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}$ on $[1, \frac{\mathbb{E}_2(a_2(q_{j+2}))}{\mathbb{E}_1(a_1(q_{j+2}))}]$.

Finally, the bound using q_{J+1} has a coefficient $y'(a_1(q_1)) \prod_{k=j+1}^{J} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}) \prod_{l=1}^{J} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}$ on

$[1, \frac{\mathbb{E}_2(a_2(q_{J+1}))}{\mathbb{E}_1(a_1(q_{J+1}))}]$.

Therefore, if we define $\prod_{k=j+1}^{s} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} = 1$ for $s < j+1$, we can use a unified way to represent the bound from q_i to be

$$y'(a_1(q_1)) \prod_{k=j+1}^{i-1} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} \prod_{l=1}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} [1, \frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

for all $i = j+1, \cdots, J+1$. Then, the second group of bounds are constructed as follows:
The bound using q_j has a coefficient $y'(a_1(q_1)) \prod_{k=j}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}) \prod_{l=1}^{j-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}$ on $[1, \frac{\mathbb{E}_2(a_2(q_j))}{\mathbb{E}_1(a_1(q_j))}]$.
The bound using q_{j-1} has a coefficient $y'(a_1(q_1)) \prod_{k=j-1}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}) \prod_{l=1}^{j-2} \frac{E_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}$ on $[1, \frac{\mathbb{E}_2(a_2(q_{j-1}))}{\mathbb{E}_1(a_1(q_{j-1}))}]$.
Finally, the bound using q_1 has a coefficient $y'(a_1(q_1)) \prod_{k=1}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))})$ on $[1, \frac{\mathbb{E}_2(a_2(q_1))}{\mathbb{E}_1(a_1(q_1))}]$.

Therefore, if we define $\prod_{l=0}^{0} \frac{\mathbb{E}_2(b_2(q_l))}{\mathbb{E}_1(b_1(q_l))} = 1$, we can use a generic way to represent the bound from q_i to be

$$y'(a_1(q_1)) \prod_{k=i}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} [1, \frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

for all $i = 1, \cdots, j$.

Finally, the upper bound through the above procedure is thus $y'(a_1(q_1))$ times

$$max\left(max\{\prod_{k=i}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j}, max\{\prod_{k=j+1}^{i-1} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} \prod_{l=1}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j+1}^{J+1} \right),$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\left(min\{\prod_{k=i}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j}, min\{\prod_{k=j+1}^{i-1} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} \prod_{l=1}^{i} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j+1}^{J+1} \right).$$

As a result, for any $\alpha \in (q_1, q_2) \cup (q_2, q_3) \cup \cdots \cup (q_{I-1}, q_I)$, we adopt the approach discussed above, and derive the lower and upper bounds for the function value $y'(a_2(\alpha))$, denoted as $lb[y'(a_2(\alpha))]$ and $ub[y'(a_2(\alpha))]$. Thus the identified set of the function $y'(\cdot)$ is given as follows:

$$\begin{aligned}\mathscr{Y}'^{*} \equiv \{y' \in \mathscr{Y}' : \; & y'(a_2(\alpha)) \in \left[lb[y'(a_2(\alpha))], ub[y'(a_2(\alpha))]\right] \; when \; \alpha \in \cup_{i=1}^{I-1}(q_i, q_{i+1}), \\ & y'(a_2(\alpha)) \; is \; pointly \; determined \; through \; the \; system \; (2.4.5) \; when \; \alpha \in \{q_1, \cdots, q_I\}\},\end{aligned} \tag{2.4.6}$$

where $\mathscr{Y}'$ is defined following the regularity conditions in Assumption 1-(ii) of $\mathscr{Y}$, thus is omitted here. One restriction worth noting here is that the function $y(\cdot)$ is assumed to be concave, thus the derivative function $y'(\cdot)$ is decreasing. Although the bounds cannot be shown to be monotonic in α, this shape restriction is already included in the assumption of $\mathscr{Y}$, thus in $\mathscr{Y}'$, and one needs to take it into consideration when constructing bounds and the identified set. This identified set is sharp in the sense that any function $y'(\cdot)$ that belongs to this set would induce the same data, i.e., the action CDFs G_1 and G_2 and the ratios $y'(a_1(\alpha))/y'(a_2(\alpha))$ for any $\alpha \in [0,1]$.

From the bounds of $y'(a_2(\alpha))$, we can easily recover the bounds for $y'(a_1(\alpha))$ for any α from $\cup_{i=1}^{I-1}(q_i, q_{i+1})$. The lower and upper bounds can be shown to be continuously differentiable. Therefore, we can use these bounds to recover the bounds of private types via the inverse of the strategy function as in Section 2.4.1, from which the bounds on the private type CDF $F(\cdot)$ can then be recovered.

Theorem 2 *Suppose Assumption 3 and the additional assumption hold: there is no intersection point of two corresponding action distributions $G_1(\cdot)$ and $G_2(\cdot)$. Then under the exclusion restriction that $F(\cdot) = F(\cdot|N_1) = F(\cdot|N_2)$,*
(1) $y'(\cdot)$ is nonparametrically partially identified over $[min\{\underline{a}_1, \underline{a}_2\}, max\{\bar{a}_1, \bar{a}_2\}]$, up to a scale. The identified set is given by $\mathscr{Y}'^*$ in (2.4.6).
(2) $F(\cdot)$ is nonparametrically partially identified over $[\underline{t}, \bar{t}]$, with $F(\cdot) = G_1(\xi_1^{-1}(\cdot)) = G_2(\xi_2^{-1}(\cdot))$, up to a scale.

2.4.3 Discussion on the Identification Results

We obtain point identification results for the single-intersection-point case and the multiple-intersection-point case, but only partial identification result for the no-intersection-point case. Although the point identification result is not always guaranteed, we emphasize that the patterns of functional comparative statics discussed in Section 2.3 are testable through checking the conditional action CDFs G_1 and G_2. Therefore, whether the point or partial identification result holds can be verified from the data. If the data shows that G_1 and G_2 cross at least once, the identification approaches in Section 2.4.1 and Section 2.4.1 can be adopted, thus leading to point identification result. If G_1 and G_2 do not cross, the partial identification result in Section 2.4.2 applies.

The point and partial identification results are actually connected to each other. Consider the case of no-intersection-point. Since we impose the convexity or concavity restriction of $y(\cdot)$ in Assumption 1-(ii), the relative locations of G_1 and G_2 give more information about the bounds. If G_1 and G_2 are closer to each other, the bounds from known and finite pairs of function values become tighter, because the derivative function $y'(\cdot)$ is increasing or decreasing. Specifically, when G_1 and G_2 are close enough such that they intersect once, the bounds from known pairs of function values will be infinite, thus collapsing to point identification, because at the intersection point, the bound becomes a singleton.

Although the discussion about the identification problem focuses on the game with incomplete information, the game with complete information can also be included in our framework. Recall that the key of identification relies on the ratio of the two first order conditions as in Equation (2.4.3). The left-hand side ratio of expected derivatives of function x can be replaced by the ratio of the derivatives of function x, without the expectation. This does not affect the identification approach, in the sense that this derivative of function x is known, and varies with the number of players N. Thus the point or partial identification result still holds.

If N can take more than two values, the over-identification of the problem might occur. However, since the model is general, it is hard to know whether knowing a third N can help tighten the bounds in the case of partial identification, which would depend on how far away the third action CDF lies from the first two action CDFs.

2.5 Numerical Illustration

In this section, we use the Tullock contest model to illustrate the nonparametric identification results. We further assume in Example 4 that $V = 1$, $r = 1$ and all contestants are ex-ante symmetric. Thus, the payoff function becomes

$$\pi_i(a_i, \mathbf{a}_{-i}, t_i) = \frac{a_i}{a_i + \sum_{j \neq i} a_j} - \frac{c(a_i)}{t_i}, \tag{2.5.1}$$

According to Wasser (2013) and Ewerhart (2014), the equilibrium is interior and can be reduced to

$$t_i \cdot \mathbb{E}_{-i}\left[\frac{\sum_{j \neq i} a_j}{(a_i + \sum_{j \neq i} a_j)^2}\right] = c'(a_i), \tag{2.5.2}$$

Denote $p = 1/t$, which can be interpreted as the private cost. We specify the distribution of private costs in this numerical illustration following Wasser (2013).

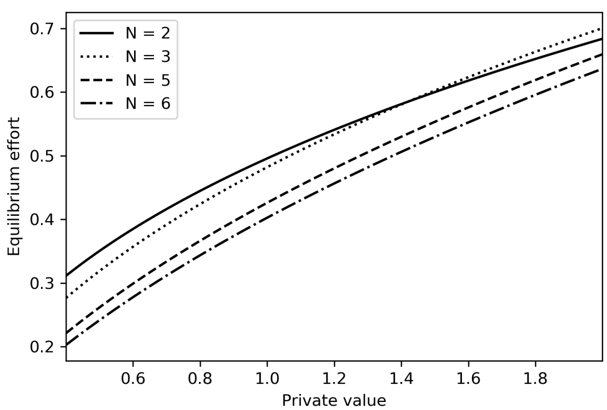
2.5.1 When the Exclusion Restriction Holds

We first consider the scenario when the exclusion restriction holds. The true distribution of private cost p is assumed to be a uniform distribution on $[0.5, 2.5]$. The true cost function is assumed to be a quadratic one: $c(x) = (1/2)x^2$, which leads to an identity cost derivative function: $c'(x) = x$. We approximate $s(t) = \nu(p)$ numerically by a discrete function on a grid of points in $[0.5, 2.5]$. The size of the grid is set to be $g = 1000$, and the set of points is denoted as:

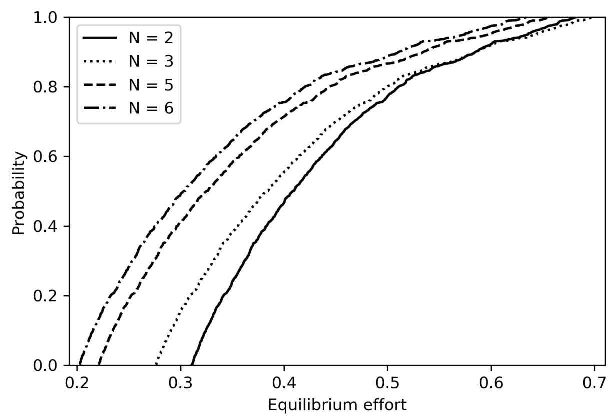
$$\hat{\mathbf{p}} = \{p^1, p^2, \cdots, p^g\}.$$

When the exclusion restriction holds, we consider the same $\hat{\mathbf{p}}$ for different numbers of contestants: $N = 2$, $N = 3$, $N = 5$, and $N = 6$. The numerical equilibrium strategy functions and corresponding equilibrium efforts distributions are shown in Figures 2.6a and 2.6b, respectively. When N increases from 2 to 3, the change corresponds to the first pattern of functional comparative statics: single-intersection-point; when N increases from 3 to 5 and 6, the last pattern of functional comparative statics emerges: no-intersection-point. We apply the point identification approach to the single-intersection-point pattern, e.g., $N = 2$ and $N = 3$ to identify the cost derivative function over the support of equilibrium efforts when $N = 3$. We also apply the partial identification result to two no-intersection-point patterns, e.g., $N = 3$ and $N = 5$, as well as $N = 3$ and $N = 6$, in order to identify the cost derivative function over the support of efforts when $N = 3$ as well.

Figure 2.6: When the exclusion restriction holds: equilibrium effort functions and effort distributions.



(a) Equilibrium effort functions with different N.

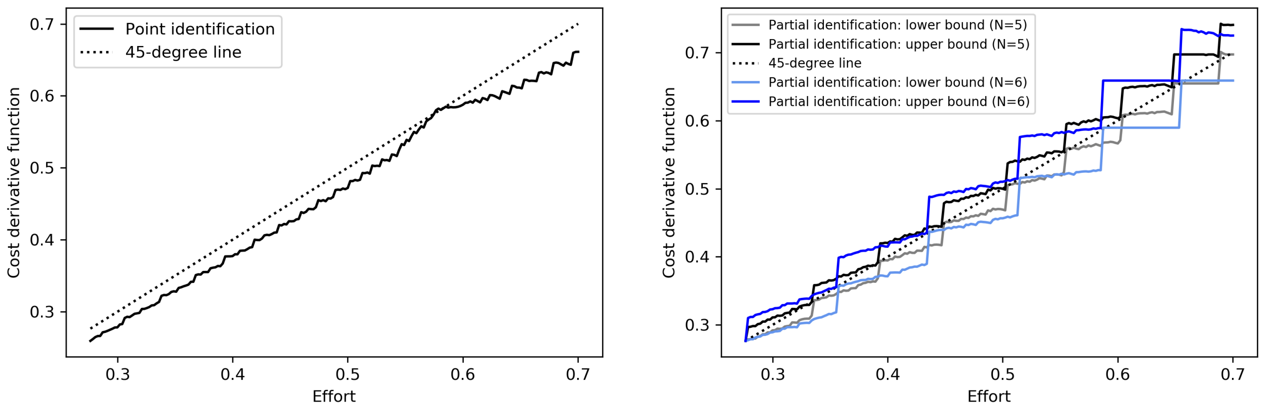


(b) Effort distributions with different N.

For point identification, we normalize the cost derivative at the intersection point to be the effort itself, since the true cost derivative function is an identity function, which is represented by the 45-degree line. As shown in Figure 2.7a, the identified cost derivative function is more accurate as the effort value being evaluated is closer to the intersection point. The reason is that the point identification result relies on the construction of a contraction sequence that converges to the intersection point. The infinite nature of this sequence can cause the numerical error induced by calculating the empirical CDF and the empirical quantile to accumulate. As the effort being identified gets further away from the intersection point, the cumulative error gets larger.

The partial identification results are shown in Figure 2.7b, where the upper and lower bounds of the cost derivative function are displayed. First of all, starting from the lower support of equilibrium efforts when $N = 3$, there is a finite set of effort values evaluated at which the cost derivatives are pointly identified. The bounds are thus derived for each section between two pointly identified values. Compared to $N = 6$, the effort distribution with $N = 5$ is closer to that with $N = 3$. Therefore, there are more effort values at which the cost derivatives are pointly identified resulting in tighter bounds.

Figure 2.7: When the exclusion restriction holds: point and partial identification results.



(a) Point identification of the cost derivative function.

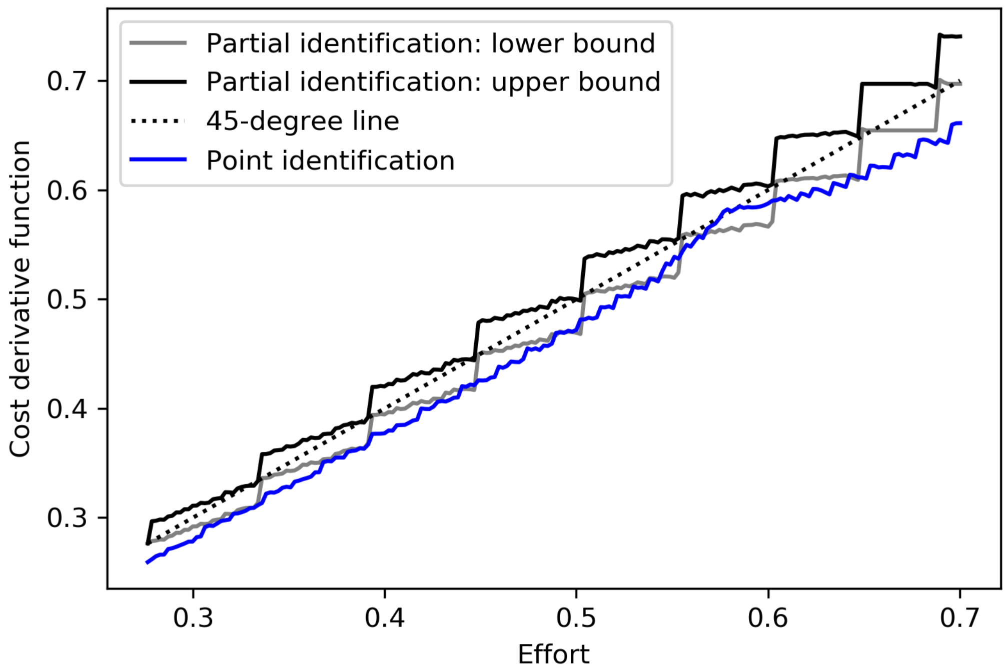
(b) Partial identification of the cost derivative function.

A few remarks regarding the results from the point identification and the partial identification approaches are in order. First, the point identification results come from the product of infinitely many ratios, which lead to the accumulation of numerical errors. In contrast, the bounds in the partial identification case are derived from the pairs of point identified values that are finite; hence the numerical error seems smaller relative to the case of point identification. Second, in the point identification case the normalized point is chosen to be the intersection point; while in the partial identification case, the normalized point is always the lower support of equilibrium efforts when $N = 3$. Figure 2.8 shows the results from the point identification ($N = 2$ and $N = 3$) and the partial identification ($N = 3$ and $N = 5$). Roughly speaking, the results are compatible although the point identification result is further away from the true function (45-degree line) as moving further away from the intersection point due to the cumulative numerical error. Nevertheless and importantly, in a local area around the intersection point, the point identification result is always inside the derived bounds.

2.5.2 When Exclusion Restriction Does Not Hold

When the exclusion restriction does not hold, the private cost distribution varies as N increases. We assume that when $N = 2$, the cost is distributed as a truncated standard normal distribution in $[0.5, 2.5]$;

Figure 2.8: Comparison of the point and partial identification when the exclusion restriction holds.

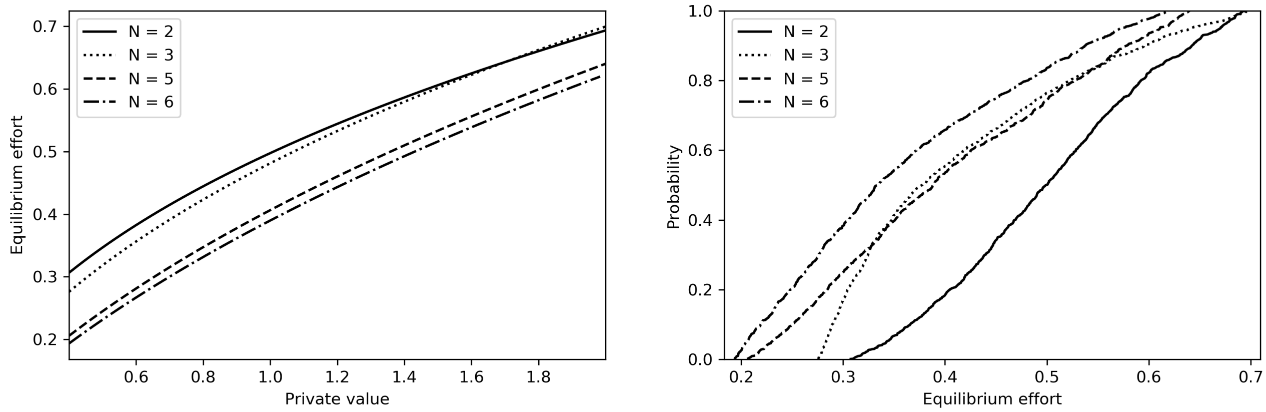


when $N = 3$, the cost is distributed as a uniform distribution in $[0.5, 2.5]$; when $N = 5$, the cost is distributed as a truncated exponential distribution in $[0.5, 2.5]$; when $N = 6$, the cost is distributed as a truncated logistic distribution in $[0.5, 2.5]$.

The numerical equilibrium strategy functions are shown in Figure 2.9a, and the distributions of equilibrium efforts are shown in Figure 2.9b. Clearly the patterns presented in Figure 2.9a do not transmit to Figure 2.9b, because the exclusion restriction does not hold. We apply the point identification approach to the single-intersection-point pattern, e.g., $N = 2$ and $N = 3$ to identify the cost derivative function over the support of equilibrium efforts when $N = 2$ (Figure 2.10a). We also apply the partial identification result to two no-intersection-point patterns, e.g., $N = 2$ and $N = 5$, as well as $N = 2$ and $N = 6$, in order to identify the cost derivative function over the support of efforts when $N = 2$ as well (Figure 2.10b).

The point identification result shown in Figure 2.10a is substantially different from the true function (not even increasing). While the bounds from the partial identification shown in Figure 2.10b seem closer to the true function than the point identification result, the 45-degree line lies outside the identified sets. When the exclusion restriction does not hold, the bounds from different Ns are not compatible with each other. The derived bounds from $N = 2$ and $N = 5$ are not tighter and do not

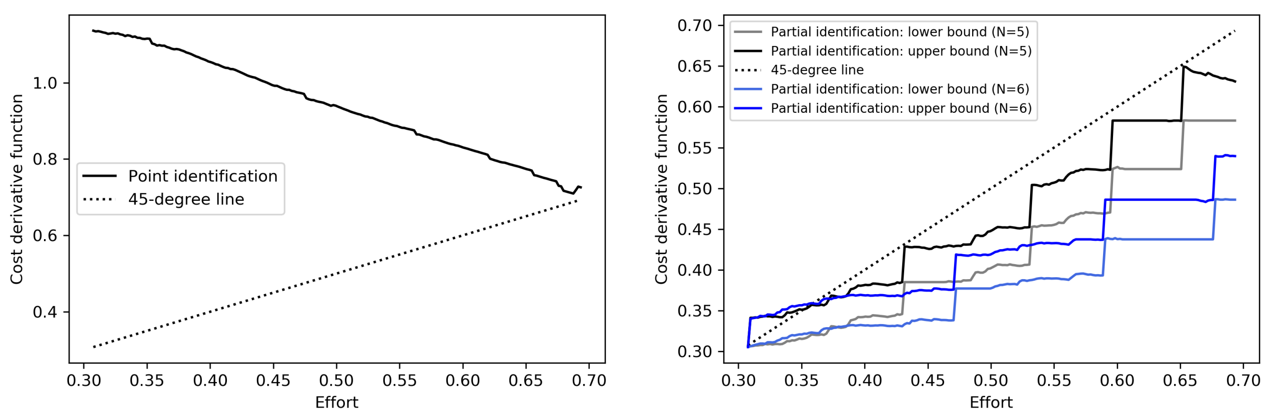
Figure 2.9: When the exclusion restriction does not hold: equilibrium effort functions and effort distributions.



(a) Equilibrium effort functions with different N.
(b) Effort distributions with different N.

lie within those from $N = 2$ and $N = 6$, which is in sharp contrast with the results in Figure 2.7b. As a result, this can be used as a basis to propose a test for the exclusion restriction, provided that nonparametric set inference methods are developed.

Figure 2.10: When the exclusion restriction does not hold: point and partial identification results.



(a) Point identification of the cost derivative function.
(b) Partial identification of the cost derivative function.

2.6 Extensions

2.6.1 Corner Solutions

In the previous sections, we only consider the interior solution case, where the first order condition yields Equation (2.3.1). However, in some empirical applications, *corner solutions* may occur.

Example 4 illustrates the payoff function in a Tullock contest that fits in the framework of Bayesian games in this paper as in Section 2.5. Note that unlike Example 4, here we let the cost function $c(\cdot)$ be the same for every contestant and assume that r is known to researchers. In this subsection, we use this contest model to discuss the nonparametric identification of the unknown structure $[F, c']$, given the exclusion restriction $F(\cdot|N) = F(\cdot)$, when the contestant's equilibrium effort is allowed to be zero.[20] The first order condition is now characterized by the Krush-Kuhn-Tucker (KKT) conditions:

$$t_i \cdot V \cdot \mathbb{E}_{-i}[\frac{ra_i^{r-1}(\sum_{j\neq i} a_j^r)}{(a_i^r + \sum_{j\neq i} a_j^r)^2}] = c'(a_i), \ if \ a_i > 0, \tag{2.6.1}$$

$$t_i \cdot V \cdot \mathbb{E}_{-i}[\frac{ra_i^{r-1}(\sum_{j\neq i} a_j^r)}{(a_i^r + \sum_{j\neq i} a_j^r)^2}] \leq c'(a_i), \ if \ a_i = 0. \tag{2.6.2}$$

Therefore, there exist both "inactive" contestants who make zero effort and obtain zero expected payoffs, as well as "active" ones whose efforts are strictly positive. Ewerhart (2014) shows that there exists a unique, symmetric, and monotone equilibrium strategy $s(\cdot)$ in the above Tullock contest, and at least one contestant remains active in equilibrium. For active contestants, $s(\cdot)$ is strictly increasing since the strict supermodularity condition is satisfied.

To apply the contraction sequence approach of nonparametric identification proposed in Section 2.4, we can classify the cases of comparative statics in the same way as before, by looking at how the two effort distributions behave for the private types who are active when N varies.

It is worth noting that in the corner solution case, $c'(\cdot)$ and $F(\cdot)$ can only be pointly or partially identified over a restricted support, due to the reason that the identification approaches can only be adopted for strictly positive efforts in Equation (2.6.1). For instance, in the Tullock contest, let α_1^* and α_2^* be two probabilities, which correspond to the upper boundaries of zero actions in the two contests (also the lower boundaries of non-zero actions in the two contests). Then $c'(\cdot)$ is only identified over the restricted support $[min\{a_1(\alpha_1^*), a_2(\alpha_2^*)\}, max\{\bar{a}_1, \bar{a}_2\}]$, and $F(\cdot)$ is identified over the restricted support $[min\{t(\alpha_1^*), t(\alpha_2^*)\}, \bar{t}]$, with $F(\cdot) = G_j(\xi_j^{-1}(\cdot))$ for $j = 1, 2$.

2.6.2 Asymmetric Private Type Distribution

We now consider the asymmetric case where each player i has her private type drawn from asymmetric $F_i(\cdot)$ over the support $[\underline{t}_i, \bar{t}_i]$, satisfying Assumption 1-(i). We maintain the exclusion restriction. The strictly MPSNE of player i is denoted by $s_i(t_i) = a_i$, with equilibrium action CDF $G_i(\cdot|N)$ over the support $[\underline{a}_i(N), \bar{a}_i(N)]$, which satisfies Assumption 2. The first order condition (an

20 Here we focus on identification of $c'(\cdot)$, the derivative of the cost function.

adaption of Equation (2.3.1)) now becomes:

$$t_i(\alpha)\cdot \mathbf{E}_{\mathbf{a}_{-N}}[\frac{\partial x(a_i(\alpha),\mathbf{a}_{-N};N)}{\partial a_i}]+y'(a_i(\alpha))=0, \tag{2.6.3}$$

where the bolded $\mathbf{E}_{-N}$ denotes the expectation over the joint action distribution except player i: $\mathbf{G}_{-i}(\cdot,\cdots,\cdot|N)=\Pi_{j\neq i}G_j(\cdot|N)$. We consider the interior solution case.

Consider two numbers of players $N_1<N_2$. Our identification problem is to recover the structure $[F_1(\cdot),\cdots,F_{N_j}(\cdot);y'(\cdot)]$, given the equilibrium action vector $\{a_{j,1},\cdots,a_{j,N_j}\}$, and action CDFs $\{G_{j,1}(\cdot),\cdots,G_{j,N_j}(\cdot)\}$ for $j=1,2$.

Under the exclusion restriction, it is crucial to cancel out $t_i(\alpha)$ in Equation (2.6.3) above, as in the beginning of Section 2.4. Therefore, one necessary constraint to apply the identification approach is to assume that there are N_0 common players with $1\leq N_0\leq N_1$ when N varies. We assume that these common players are indexed as the first N_0 players. The identification procedure is thus as follows (either pointly or partially):

Step 1: Choose an arbitrary common player i, with private type distribution $F_i(\cdot)$. By canceling out $t_i(\alpha)$, we get the following equation

$$\frac{\mathbf{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_{1,i}(\alpha),\mathbf{a}_{-N_1};N_1)}{\partial a_{1,i}}]}{\mathbf{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_{2,i}(\alpha),\mathbf{a}_{-N_2};N_2)}{\partial a_{2,i}}]}=\frac{y'(a_{1,i}(\alpha))}{y'(a_{2,i}(\alpha))}.$$

Step 2: Applying the identification approach in Section 2.4, $y'(\cdot)$ is nonparametrically identified over $[min\{\underline{a}_{1,i},\underline{a}_{2,i}\},max\{\bar{a}_{1,i},\bar{a}_{2,i}\}]$, and $F_i(\cdot)$ is nonparametrically identified over $[\underline{t}_i,\bar{t}_i]$, with $F_i(\cdot)=G_{j,i}(\xi_{j,i}^{-1}(\cdot))$ for $j=1,2$.

Step 3: Varying the index i within the group of common players, $y'(\cdot)$ is identified over $[\min_{1\leq i\leq N_0}\left\{min\{\underline{a}_{1,i},\underline{a}_{2,i}\}\right\},\max_{1\leq i\leq N_0}\left\{max\{\bar{a}_{1,i},\bar{a}_{2,i}\}\right\}]$.[21]

Step 4: Choosing an arbitrary remaining contestant k, with private type distribution $F_k(\cdot)$. Thus when $a_{j,k}\in[\min_{1\leq i\leq N_0}\left\{min\{\underline{a}_{1,i},\underline{a}_{2,i}\}\right\},\max_{1\leq i\leq N_0}\left\{max\{\bar{a}_{1,i},\bar{a}_{2,i}\}\right\}]$, when $j=1$ *or* 2, the corresponding private type t_k can be recovered. As a result, $F_k(\cdot)$ is identified over such values of t_k.

2.6.3 Asymmetric Function y

We now consider the asymmetric case where N players draw their private types from a common CDF $F(\cdot)$ over $[\underline{t},\bar{t}]$, but have different functions $y_i(\cdot)$ satisfying Assumption 1. Although the latent

21 If the support $[min\{\underline{a}_{1,i},\underline{a}_{2,i}\},max\{\bar{a}_{1,i},\bar{a}_{2,i}\}]$ of different common player i overlaps, varying the index i may potentially tighten the bounds of $y'(\cdot)$ in the case of partial identification.

private type distribution is common, each player i can have a different equilibrium strategy $s_i(t_i) = a_i$ with different equilibrium action CDF $G_i(\cdot|I)$ over $[\underline{a}_i(N), \bar{a}_i(N)]$, due to the different function y_i.

Consider two numbers of players $N_1 < N_2$, and assume that there are N_0 common players where $1 \leq N_0 \leq N_1$. We assume that these common players are indexed as the first N_0 players. The identification approach is as follows (either pointly or partially):

Step 1: Choose an arbitrary common player i, with the function $y_i(\cdot)$. By canceling out $t(\alpha)$, we get the following equation

$$\frac{\mathbf{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_{1,i}(\alpha),\mathbf{a}_{-N_1};N_1)}{\partial a_{1,i}}]}{\mathbf{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_{2,i}(\alpha),\mathbf{a}_{-N_2};N_2)}{\partial a_{2,i}}]} = \frac{y_i'(a_{1,i}(\alpha))}{y_i'(a_{2,i}(\alpha))}.$$

Step 2: Applying the identification approach in Section 2.4, $y_i'(\cdot)$ is nonparametrically identified over $[min\{\underline{a}_{1,i}, \underline{a}_{2,i}\}, max\{\bar{a}_{1,i}, \bar{a}_{2,i}\}]$, and $F(\cdot)$ is nonparametrically identified over $[\underline{t}, \bar{t}]$, with $F(\cdot) = G_{j,i}(\xi_{j,i}^{-1}(\cdot))$ for $j = 1, 2$.[22]

Step 3: Choose an arbitrary remaining contestant k when $N = N_j$, $j = 1$ *or* 2. Thus from player k's first order condition $t(\alpha) \cdot \mathbf{E}_{\mathbf{a}_{-N_j}}[\partial x(a_{j,k}(\alpha), \mathbf{a}_{-N_j}; N_j)/\partial a_{j,k}] = y_k'(a_{j,k}(\alpha))$, $y_k'(\cdot)$ is identified over $[\underline{a}_{j,k}, \bar{a}_{j,k}]$.

2.6.4 Asymmetric Private Type Distribution and Function y

We now consider the asymmetric Bayesian game, where player i has own private type CDF $F_i(\cdot)$ over $[\underline{t}_i, \bar{t}_i]$ and own function $y_i(\cdot)$ satisfying Assumption 1. Each player i has a different equilibrium action strategy $s_i(t_i) = a_i$ with a different action distribution $G_i(\cdot|N)$ over $[\underline{a}_i(N), \bar{a}_i(N)]$.

Consider two numbers of players $N_1 < N_2$. assume that there are N_0 common players where $1 \leq N_0 \leq N_1$. We assume that these common players are indexed as the first N_0 players. The identification approach is as follows (either pointly or partially):

Step 1: Choose an arbitrary common player i, with private type distribution $F_i(\cdot)$. By canceling out $t_i(\alpha)$, we get the following equation

$$\frac{\mathbf{E}_{\mathbf{a}_{-N_1}}[\frac{\partial x(a_{1,i}(\alpha),\mathbf{a}_{-N_1};N_1)}{\partial a_{1,i}}]}{\mathbf{E}_{\mathbf{a}_{-N_2}}[\frac{\partial x(a_{2,i}(\alpha),\mathbf{a}_{-N_2};N_2)}{\partial a_{2,i}}]} = \frac{y'(a_{1,i}(\alpha))}{y'(a_{2,i}(\alpha))}.$$

Step 2: Applying the identification approach in Section 2.4, $y'(\cdot)$ is nonparametrically identified over $[min\{\underline{a}_{1,i}, \underline{a}_{2,i}\}, max\{\bar{a}_{1,i}, \bar{a}_{2,i}\}]$, and $F_i(\cdot)$ is nonparametrically identified over $[\underline{t}_i, \bar{t}_i]$, with $F_i(\cdot) =$

22 Varying the index i may potentially tighten the bounds of $F(\cdot)$ in the case of partial identification.

$G_{j,i}(\xi_{j,i}^{-1}(\cdot))$ for $j = 1, 2$.

Note that in this setup, only the private type CDFs $\{F_1(\cdot), \cdots, F_{N_0}(\cdot)\}$, and functions $\{y_1(\cdot), \cdots, y_{N_0}(\cdot)\}$ of common players can be nonparametrically identified over their corresponding identified supports.

2.6.5 Asymmetric Function x

We now consider the extension to allow for each player i to have different $x_i(a_i, \mathbf{a}_{-i}; N)$. It is important to note that this function has to vary with N, which implies that for player i, when the number of participants change, x_i has to change as well, in order to cause variation of the expectation expressed as a multiple integral.

Since x_i is assumed to be known to the researchers, its asymmetry does not add any more difficulty to the identification problem as it does not influence the cancellation of private type for common players when the first order conditions for common players are stacked, which is the key insight we use in our identification strategy. Specifically, when asymmetric x_i appears together with asymmetric private type CDF F_i, the procedure in Section 2.6.2 can be adopted. Second, when both x_i and y_i are asymmetric, we can utilize the procedure in Section 2.6.3. Lastly, when the functions x_i, F_i and y_i are all asymmetric, the procedure in Section 2.6.4 can be applied.

Furthermore, as mentioned in Section 2.2.2, a deterministic function of t_i appearing in payoff functions in all three forms, i.e. $m_i(t_i)$, can be relaxed to be asymmetric, for the same reason as in the asymmetric x_i case. As a result, asymmetric m_i can be combined with asymmetric F_i or asymmetric y_i (along with asymmetric x_i), which will not affect identification results in this section.

2.6.6 Unobserved Heterogeneity

In this subsection, we discuss the extension of the benchmark model to the case with unobserved game heterogeneity, which is modeled as a random variable U drawn independently from the CDF $F_U(\cdot)$ with the support denoted by $\mathscr{U}$. The realization of the unobserved heterogeneity u is common knowledge among players in the Bayesian game, but as researchers we do not observe this u. Thus conditional on this u, players act as in the setup of our benchmark model, in the sense that their private types are independent conditionally. As a consequence, the existence of strictly MPSNE and its properties follow Proposition 2, given the conditional private type distribution $F(\cdot|u) \in \mathscr{F}$ defined in Assumption 1-(i). Therefore, conditional on this u, the distribution of equilibrium actions denoted as $G(\cdot|u)$ satisfies the properties in Proposition 3, i.e. $G(\cdot|u) \in \mathscr{G}$. Thus, the identification problem now becomes how to identify the unknown structure conditional on the realization $U = u$, together with the unknown distribution of the unobserved heterogeneity, i.e. $F_U(\cdot)$. Note that now the unknown private type's CDF is conditional on this u, thus $F(\cdot|u)$, while the other unknown function $y(\cdot)$ is unrelated to u.

In order to discuss the identification problem, we impose the following assumptions, which hold throughout this section.

Assumption 4

(i) Conditional exclusion restriction: for all $N \geq 2 \in \mathbb{N}^+$, $F(\cdot|u,N) = F(\cdot|u)$.

(ii) Stochastic monotonicity restriction: conditional on u, the private type is strictly increasing in u with respect to the first order.

There are two challenges in nonparametric identification of the model primitives under the unobserved heterogeneity. First, we need to identify the distribution of actions conditional on the unobserved heterogeneity and the distribution of the unobserved heterogeneity, given the unconditional distribution of actions identified in the data, for each N. Second, after identifying the above distributions, we have to know how to match the conditional distributions of actions for N_1 and N_2, denoted as $G_j(\cdot|u)$ for $j = 1,2$, on the realization of the unobserved heterogeneity, or when $G_1(\cdot|u)$ and $G_2(\cdot|u)$ correspond to the same realization u, in order to apply the identification approaches described in Section 2.4.

Assumption 4-(ii) is needed for both discrete and continuous u in order to exploit support variation, see Hu, McAdams, and Shum (2013), Gentry and Li (2014), and Grundl and Zhu (2019) in the auction context. Particularly, the strict inequality is assumed to hold at the upper boundary of t, i.e., $\bar{t}(u)$ is strictly increasing in u. In general, that the private type is strictly FOSD-increasing with respect to the unobserved heterogeneity u does not necessarily imply that the equilibrium action is also FOSD-increasing in u.[23] If the Bayesian game under consideration is a game with strategic substitutes, or the payoff function satisfies

$$\frac{\partial^2 \pi(\mathbf{a},t)}{\partial a_i \partial a_j} \leq 0, \textit{ almost everywhere}, \forall i \neq j,$$

this stochastic monotonicity restriction on the private type with respect to the unobserved heterogeneity

23 An example can be found in Aryal and Zincenko (2019), who consider a Cournot-oligopoly model with private information of firms and linear demand function faced by firms. They establish identification of the model primitives, including the distribution of the unobserved heterogeneity. However, since their Cournot model does not have an unknown function like the unknown cost function $q(a_i)$ in Example 1, our models are different and it is unclear whether the strict monotonicity of equilibrium strategies still holds in our case.

implies that the equilibrium action at a fixed α: $a(\alpha;u) = s(t(\alpha);u)$ as a function of u, is increasing in u. Specifically, at the upper boundary, or $\alpha = 1$, $\bar{a}(u)$ is strictly increasing in u. After establishing the monotonicity of the equilibrium action with respect to the unobserved heterogeneity, we can proceed to discuss the identification problem; we need to distinguish between the cases of discrete and continuous u, in order to extend Grundl and Zhu (2019) who deals with first-price auctions with risk aversion and unobserved auction heterogeneity to our setting.

If u is discrete, with a finite support, i.e., the support of u is $1, 2, \cdots, K$ with $K < \infty$, we need two numbers of players N_1 and N_2; for each one, two randomly selected actions are observed. For different u, the upper boundaries of observed actions are distinguishable. Thus, we can identify the K action distributions conditional on one value of u for each N. Furthermore, the action distribution for each N can be sorted by the upper boundaries of their supports. If two different numbers of players correspond to the same u, they should have the same rank. Therefore, they should have the same private type distributions conditional on this u. After matching the games on the value of u, we can apply the identification approaches in the previous section to identify the private type distributions. By doing this for each u, we can obtain the final identification.

If u is continuous, for each N we need three randomly selected actions. Besides, in order to use the results on identification of triangular non-separable models, we need to impose an additional assumption that the lower boundary of the private type $\underline{t}(u)$ is also strictly increasing in u, together with a normalization condition such as $\underline{t}(u) = u$. With these extra conditions, we ensure that the lower and upper boundaries of equilibrium actions are both strictly increasing with u, for each N. As a result, we can adopt the support variation approach to identify the distribution of actions conditional on u, together with the distribution of U. Moreover, it guarantees that $G_1(\cdot|u)$ and $G_2(\cdot|u)$ with different numbers of players are conditional on the same u if their supports have the same lower boundary, which fits into the identification approach proposed in the previous section.

2.6.7 Endogenous Participation

Our identification strategy has replied on the exclusion restriction in the form of a players' exogenous participation, i.e., $F(\cdot|N) = F(\cdot)$. However, this restriction is no longer valid if players' private types are influenced by the competition level of the game, i.e., the number of players. In this subsection, we consider the case when there is unobserved heterogeneity ε at the game level that affects players' participation decisions, leading to the endogenous participation.[24] Specifically, we assume $N = N(\mathbf{Z}, \varepsilon)$, where $\mathbf{Z}$ is a vector of observed characteristics of the game. We discuss two empirically-related scenarios regarding the correlation between the observed $\mathbf{Z}$ and the unobserved ε.

The first scenario is when $t \perp\!\!\!\perp \varepsilon | \mathbf{Z}$, thus the unobserved heterogeneity ε only affects players' participation, but not their private values, because the distribution of private types $F(\cdot|\mathbf{Z}, \varepsilon) = F(\cdot|\mathbf{Z})$. Thus for any given $\mathbf{Z}$, we obtain the exclusion restriction $F(\cdot|N, \mathbf{Z}) = F(\cdot|\mathbf{Z})$, but in the equilibrium, the action distribution $G(\cdot|N, \mathbf{Z})$ still depends on N. As a consequence, we can exploit the variation from two numbers of players $N_1 < N_2$ as before, but note that now they have the same characteristics $\mathbf{Z}$. Our point and partial identification approaches can thus be adopted in a straightforward manner.

The second scenario is more involved, when the unobserved heterogeneity ε is correlated with the private type.[25] We assume the availability of instrumental variable W such that $N = N(\mathbf{Z}, W, \varepsilon)$, and $t \perp\!\!\!\perp W | (\mathbf{Z}, \varepsilon)$. As a result, the private type distribution satisfies $F(\cdot|\mathbf{Z}, W, \varepsilon) = F(\cdot|\mathbf{Z}, \varepsilon)$. Thus for any fixed $(\mathbf{Z}, \varepsilon)$, we want to utilize the exclusion restriction $F(\cdot|N, \mathbf{Z}, \varepsilon) = F(\cdot|\mathbf{Z}, \varepsilon)$, in order to use the identification approaches, because in equilibrium, the action distribution is $G(\cdot|N, \mathbf{Z}, \varepsilon)$. We follow Guerre, Perrigne, and Vuong (2009) to impose the following two conditions:

(i) $\varepsilon = N - \mathbb{E}[N|\mathbf{Z}, W]$.

(ii) $W = h(X, \varepsilon)$, where $h(\cdot, \cdot)$ is strictly increasing in ε, $X \perp\!\!\!\perp \varepsilon$, and X cannot be a subset of $\mathbf{Z}$.

Under either one of the above conditions, we can identify ε as a first step, in order to identify the other structure of the model. When condition (i) holds, N is known and $E[N|\mathbf{Z}, W]$ is identified. Therefore, ε can be treated as the error term and is identifiable. When condition (ii) holds, the function $h(\cdot, \cdot)$ can be identified following Matzkin (2003), and ε can be identified as the inverse of the function, i.e., $\varepsilon = h^{-1}(X, W)$. After ε is recovered, for any given $(\mathbf{Z}, \varepsilon)$, under the exclusion restriction $F(\cdot|N, \mathbf{Z}, \varepsilon) = F(\cdot|\mathbf{Z}, \varepsilon)$, our identification approaches are applicable by noting that $G(\cdot|N, \mathbf{Z}, \varepsilon)$ depends on N.

---

24 In auction models, bidders' entry behavior has been studied both theoretically (see Samuelson (1985), McAfee and McMillan (1987), and Levin and Smith (1994), among others), and empirically (see Li and Zheng (2009), Athey, Levin, and Seira (2011), Krasnokutskaya (2011), and Compiani, Haile, and Sant'Anna (2019), among others). For Bayesian game models, however, both theoretical and empirical studies on players' participation have been limited. To the best of our knowledge, the two theoretical papers (Fu, Jiao, and Lu (2015) and Gu, Hehenkamp, and Leininger (2019)) that study players' entry in contest models consider the complete information framework, thus not a Bayesian game. On the empirical side, Kawai and Sunada (2015) estimate a dynamic model to analyze the campaign finance of electoral candidates, together with the challenger's entry decision.

25 Note that this situation is different from the case of unobserved heterogeneity which enters directly players' private type distribution in Section 2.6.6. Here we consider the scenario when the unobserved heterogeneity affects private types through players' participation.

2.7 Conclusions

In this paper, we explore a general approach to nonparametric identification of general Bayesian games with separable payoff functions, which include many interesting applications. Using the exclusion restriction in the form of an exogenous players' participation, we characterize conditions under which we can establish either point or partial identification of the model primitives, which are the payoff function and the private type distribution. We make two core contributions to the related literature. First, for the benchmark model we show that if the distributions of actions intersect with each other when the number of players varies, the model primitives can be nonparametrically identified. Otherwise, they can be bounded. Whether the distributions of actions intersect with each other as the level of competition changes can be verified readily from the data. As a result, while the equilibrium strategies in Bayesian games we consider are challenging to analyze, we are able to provide empirically verifiable conditions and corresponding identification results. Second, we extend these results to accommodate corner solutions, asymmetric players, unobserved heterogeneity, and endogenous participation, thus making our results applicable to a broad class of empirically relevant Bayesian games.

While we focus on studying nonparametric identification of Bayesian games, in a similar spirit to the papers in addressing identification issues in structural IO models such as auction models (Athey and Haile (2002), Guerre, Perrigne, and Vuong (2009) and Gentry and Li (2014)) and the models of differentiated products markets (Berry and Haile (2014)), our identification results can also provide a basis for nonparametric estimation and inference (including testing for the exclusion restriction as discussed in Section 2.5.2) in the class of Bayesian games we consider. In the point identification case when the distributions of actions intersect with each other as the number of players varies, the estimation method proposed in Zincenko (2018) in estimating the model primitives in the first price auction model with risk averse bidders can be extended to our case using our point identification results. On the other hand, in the partial identification case when the distributions of actions do not intersect as the number of players changes, the nonparametric bounds can be consistently estimated by replacing the bounds by the nonparametric estimators of the point identified quantities. However, it raises new and challenging questions about inference for nonparametrically partially identified models, as the recent advances in inference on incomplete or partially identified econometric models have mainly focused on the parametric or semiparametric framework (e.g. Chernozhukov, Hong, and Tamer (2007), Beresteanu, Molchanov, and Molinari (2011), Ciliberto and Tamer (2009), Galichon and Henry (2011), and Chesher and Rosen (2017), among others), thus will be left for future research.[26]

2.8 Appendix: Proofs of Main Results

Proof of Proposition 2:

We prove the result for the general payoff $\pi_i(a_i, \mathbf{a}_{-i}, t_i)$ that satisfies the strict supermodularity condition 2.2.1. Then Proposition 2 for the additive payoff case follows directly. The F.O.C. is

$$\int_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} \frac{\partial \pi_i(s_i(t_i), \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i)}{\partial a_i} d\mathbf{F}_{-i}(\mathbf{t}_{-i}) = 0, \tag{2.8.1}$$

where we define the existing MPSNE as $s_i(\cdot)$ for each player i. Differentiating Equation (2.8.1) with respect to t_i yields:

$$s_i'(t_i) \cdot \int_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} \frac{\partial^2 \pi_i(a_i, \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i)}{\partial a_i^2} d\mathbf{F}_{-i}(\mathbf{t}_{-i}) + \int_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} \frac{\partial^2 \pi_i(a_i, \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i)}{\partial a_i \partial t_i} d\mathbf{F}_{-i}(\mathbf{t}_{-i}) = 0. \tag{2.8.2}$$

Note that the existence of equilibrium implies that the second order necessary condition for this optimization is valid, i.e.

$$\int_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} \frac{\partial^2 \pi_i(a_i, \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i)}{\partial a_i^2} d\mathbf{F}_{-i}(\mathbf{t}_{-i}) \leq 0.$$

Due to the strict supermodularity condition, we have $s_i'(t_i) > 0$ almost everywhere.

The property that $\frac{\partial^2 \pi_i(a_i, \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i)}{\partial a_i \partial t_i}$ is bounded away from infinity, together with Equation (2.8.2), immediately implies that $s_i'(\cdot)$ is bounded away from infinity. Thus $0 < s_i'(\cdot) < \infty$. Also the existence of equilibrium implies at the upper and lower boundaries of the equilibrium actions exist, denoted as $\underline{a}_i = s_i(\underline{t}_i)$ and $\bar{a}_i = s_i(\bar{t}_i)$, thus forming the image of the equilibrium strategy $s_i(\cdot)$, i.e. $[\underline{a}_i, \bar{a}_i] \subset \mathscr{A}_i$.

Next we show the smoothness condition of the equilibrium strategy $s_i(\cdot)$. Note that $s_i(\cdot)$ has a bounded first derivative. Suppose that there exists a positive number M_s such that $0 < s_i'(\cdot) \leq M_s$. Choose any point $t \in [\underline{t}_i, \bar{t}_i]$, for any $\varepsilon > 0$, then there exists a positive $\delta = \varepsilon / M_s$ such that $|s_i(t_i) - s_i(t)| < \varepsilon$ as long as $|t_i - t| < \delta$. This holds because by the mean value theorem, there exists a $\tilde{t}$ between t_i and t such that $|s_i(t_i) - s_i(t)| = |s_i'(\tilde{t})||t_i - t|$. Furthermore, by Lebesgue's dominated convergence theorem, through Equation (2.8.2) above, since the two integrals $\int_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} \frac{\partial^2 \pi_i(a_i, \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i)}{\partial a_i^2} d\mathbf{F}_{-i}(\mathbf{t}_{-i})$ and $\int_{\mathbf{t}_{-i} \in \mathbf{T}_{-i}} \frac{\partial^2 \pi_i(a_i, \mathbf{s}_{-i}(\mathbf{t}_{-i}), t_i)}{\partial a_i \partial t_i} d\mathbf{F}_{-i}(\mathbf{t}_{-i})$ are continuous, the first derivative of $s_i(\cdot)$ is continuous. In conclusion, $s_i(\cdot)$ is continuously differentiable (C^1). $\square$

26 D'Haultfœuille and Février (2020) establish consistency of their nonparametric estimator of the bounds, and suggest to use bootstrap to make inference. They then use their nonparametric bound estimates to specify parametric forms for the underlying structural elements and conduct structural estimation using the parametric approach.

Proof of Proposition 3:

First of all, the equilibrium action has a compact support $[\underline{a}_i, \bar{a}_i] = [s_i(\underline{t}_i), s_i(\bar{t}_i)]$, which is also the support for the derived CDF of equilibrium action for player i. For any $a \in [\underline{a}_i, \bar{a}_i]$ and each i, note that $G_i(a) = Pr(a_i \leq a) = Pr(t_i \leq s_i^{-1}(a)) = F_i(t)$, by denoting $a = s_i(t)$. Since $F_i(\cdot) \in \mathscr{F}$ in Assumption 1-(i), $F_i(\cdot)$ is absolutely continuous; thus $G_i(\cdot)$ is absolutely continuous in its support $[\underline{a}_i, \bar{a}_i]$.

Since for any $a \in [\underline{a}_i, \bar{a}_i]$, let $a = s_i(t)$, we have $G_i(s_i(t)) = F_i(t)$. By taking derivatives on both sides, we have

$$g_i(a)s_i'(t) = f_i(t),$$

which leads to $g_i(a) = f_i(t)/s_i'(t)$. Due to the assumption that $f_i(\cdot)$ is bounded, continuous and atomless, and the result in Proposition 2 that $s_i'(\cdot)$ is bounded and continuous, $g_i(\cdot)$ is therefore bounded, continuous and atomless. In addition, since both $f_i(\cdot)$ and $s_i'(\cdot)$ are strictly positive, and bounded from infinity, we have that $g_i(\cdot)$ is strictly positive and bounded from infinity, i.e.: $0 < g_i(\cdot) < \infty$. □

Proof of Proposition 4:

We provide a counter example to show this nonidentification result. First, for any i, consider a structure $[F_i(\cdot), y_i(\cdot)] \in \mathscr{F} \times \mathscr{Y}$ in the class of Bayesian games under investigation, which leads to an observed equilibrium action CDF $G_i(\cdot) \in \mathscr{G}$. Specifically, we assume that $y_i(\cdot)$ is strictly increasing and convex. As a result, this structure satisfies the following first order condition, for any i:

$$t_i \cdot \int_{\mathbf{a}_{-i} \in \mathscr{A}_{-i}} \frac{\partial x_i(a_i, \mathbf{a}_{-i})}{\partial a_i} d\mathbf{G}_{-i}(\mathbf{a}_{-i}) + y_i'(a_i) = 0. \tag{2.8.3}$$

Consider another function $y_i^*(\cdot) = \phi_i(y_i(\cdot))$ for each player i where $\phi_i(\cdot)$ is twice continuously differentiable, strictly increasing, and convex, thus $y_i^*(\cdot) \in \mathscr{Y}$. To derive the same G_i, it suffices is to show there exists another CDF $F_i^*(\cdot)$ that produces the private type $t_i^* = \phi_i'(y_i(a_i))t_i$. We now rewrite the first order condition (2.8.3) through quantiles. For any $\alpha \in [0, 1]$, Equation (2.8.3) is equivalent to

$$F_i^{-1}(\alpha) \cdot \int_{\mathbf{a}_{-i} \in \mathscr{A}_{-i}} \frac{\partial x_i(G_i^{-1}(\alpha), \mathbf{a}_{-i})}{\partial a_i} d\mathbf{G}_{-i}(\mathbf{a}_{-i}) + y_i'(G_i^{-1}(\alpha))) = 0. \tag{2.8.4}$$

Thus if we define a new quantile function $Q_i^*(\alpha) = \phi_i'(y_i(G_i^{-1}(\alpha)))F_i^{-1}(\alpha)$, the corresponding CDF $F_i^*(t) = Pr(t_i^* \leq Q_i^*(\alpha))$ satisfies that $F_i^*(\cdot) \in \mathscr{F}$. Therefore, the structure $[F_i^*(\cdot), y_i^*(\cdot)]$ induces the same observed equilibrium action CDF $G_i(\cdot)$ as $[F_i(\cdot), y_i(\cdot)]$. □

Proof of Proposition 5:

Without loss of generality, we assume that the two equilibrium action strategies intersect with each other at the intersection point $(\alpha^{IP}, a(\alpha^{IP}))$ as in Figure 2.1a. Note that for $j = 1,2$, $G_j(a) = Pr(a_j \leq a|N_j) = Pr(t \leq s_j^{-1}(a)|N_j) = Pr(t \leq s_j^{-1}(a)) = F(s_j^{-1}(a))$. Thus when $s_1^{-1}(a) > s_2^{-1}(a)$, $G_1(a) > G_2(a)$, evaluated at one fixed $a > a(\alpha^{IP})$; when $s_1^{-1}(a) < s_2^{-1}(a)$, $G_1(a) < G_2(a)$, evaluated at one fixed $a < a(\alpha^{IP})$. Therefore, when the two equilibrium strategy functions intersect once, the two action CDFs intersect once, too.

On the other hand, if the two action CDFs intersect once at $(a(\alpha^{IP}, \alpha^{IP})$ as in Figure 2.1b, with the same logic as above, the two equilibrium strategy functions intersect once, too. $\square$

Proof of Theorem 1:

The assumptions in Assumption 3 and the additional assumption in Theorem 1 hold. Focusing on N_2, by constructing the contraction sequence as in the main text of Section 2.4.1, $y'(\cdot)$ is nonparametrically identified over $[\underline{a}_2, \bar{a}_2]$ up to a scale $y'(a(\alpha^{IP}))$. Due to the same logic regarding N_1, $y'(\cdot)$ is nonparametrically identified over $[\underline{a}_1, \bar{a}_1]$ up to the same scale. Combining them together yields the identification of $y'(\cdot)$ over $[min\{\underline{a}_1, \underline{a}_2\}, max\{\bar{a}_1, \bar{a}_2\}]$ up to the scale $y'(a(\alpha^{IP}))$.

Next, since $y'(\cdot)$ is nonparametrically identified over $[\underline{a}_2, \bar{a}_2]$, we use N_2 to recover the private type over the whole support $[\underline{t}, \bar{t}]$, through Equation (2.4.1), i.e. $t(\alpha) = \xi(a_2(\alpha); x, G, y|N_2) = \xi_2(a_2(\alpha))$, for any $\alpha \in [0,1]$. From this, $F(\cdot)$ is nonparametrically identified over $[\underline{t}, \bar{t}]$, up to a scale $F(t(\alpha^{IP})) = G_2(\xi_2^{-1}(t(\alpha^{IP})))$. Using N_1 produces the same result, thus $F(\cdot) = G_j(\xi_j^{-1}(\cdot))$, since $a_j(\alpha) = \xi_j^{-1}(t(\alpha))$ for any $\alpha \in [0,1]$ and for $j = 1,2$.

Lastly, since the contraction sequence is unique in this case, the identification results above are all unique. $\square$

Proof of Proposition 7:

The assumptions in Assumption 3 and the additional assumption in Proposition 7 hold. Assume that the conditional action CDFs G_1 and G_2 are known; furthermore, the ratios $y'(a_1(\alpha))/y'(a_2(\alpha))$ are known for each $\alpha \in [0,1]$ through the ratio of two first order conditions. In order to show non-identification, we need to show that different functions of $y'(\cdot)$ can lead to the same G_1, G_2, and $y'(a_1(\alpha))/y'(a_2(\alpha))$ for any $\alpha \in [0,1]$.

Choosing any fixed $\breve{\alpha}$, we show this non-identification result. Suppose we have one arbitrary function $y'(\cdot)$ over $[a_1(\breve{\alpha}), a_2(\breve{\alpha})]$ that are known, which corresponds to the function values $y'(a_1(\cdot))$ over $[\breve{\alpha}, min\{1, \hat{\alpha}\}]$ along the known CDF G_1, where $\hat{\alpha}$ satisfies $a_1(\hat{\alpha}) = a_2(\breve{\alpha})$. Thus, from the known ratios $y'(a_1(\alpha))/y'(a_2(\alpha))$, we can derive $y'(a_2(\cdot))$ over the same support along the known CDF G_2. Repeat this procedure, we can determine all function values of $y'(\cdot)$ over the whole support $[\underline{a}_1, \bar{a}_2]$, which accord with the known information about G_1, G_2, and $y'(a_1(\alpha))/y'(a_2(\alpha))$ over $\alpha \in [0,1]$. This procedure does not depend on what function we choose for $y'(\cdot)$ over $[a_1(\check{\alpha}), a_2(\check{\alpha})]$. This arbitrariness of the function $y'(\cdot)$ over a continuum of its support implies that $y'(\cdot)$ is not point identified. □

Proof of Theorem 2:

The assumptions as in Proposition 7 hold. Focusing on N_2, the bounds of $y'(a_2(\alpha))$ are constructed as in the main text, resulting in the identified set (2.4.6). Thus $y'(\cdot)$ is partially identified over $[\underline{a}_2, \bar{a}_2]$, up to a scale. Due to the same logic regarding N_1, $y'(\cdot)$ is partially identified over $[\underline{a}_1, \bar{a}_1]$, up to the same scale. Therefore, $y'(\cdot)$ is partially identified over $[\underline{a}_1, \bar{a}_2]$, up to the same scale.

After the bounds on $y'(\cdot)$ are derived, we can use these bounds to recover the bounds on private types over the whole support $[\underline{t}, \bar{t}]$, from which, the bounds on the private type CDF $F(\cdot)$ are then recovered, because both the inverse strategy functions $\xi_j(\cdot)$ for $j = 1, 2$ and $F(\cdot)$ are strictly increasing over their supports. Denote the bounds of $F(t(\alpha))$ as $lb[F(t(\alpha))]$ and $ub[F(t(\alpha)]$, for any $\alpha \in \cup_{i=1}^{I-1}(q_i, q_{i+1})$ (here, we use the same notations as in Section 2.4.2). The identified set of $F(\cdot)$ is thus:

$$\mathscr{F}^* \equiv \{F \in \mathscr{F} : F(t(\alpha)) \in [lb[F(t(\alpha))], ub[F(t(\alpha))]] \text{ when } \alpha \in \cup_{i=1}^{I-1}(q_i, q_{i+1}), \tag{2.8.5}$$

$$F(t(\alpha)) \text{ is pointly determined when } \alpha \in \{q_1, \cdots, q_I\}\}. \tag{2.8.6}$$

□

2.9 Supplemental Appendix

This supplemental appendix collects the partial identification of Case 2.a, Case 2.b, and Case 3 when the action distributions do not cross, and the identification of the second and third forms of the additive payoff structure for the Bayesian game under investigation in the main text.

2.9.1 Partial Identification of Case 2.a, Case 2.b, and Case 3

In Section 2.4.2, given a finite set of probabilities $\{q_1, q_2, \cdots, q_I\}$, the whole support $[0,1]$ is divided into J parts. In the main text, we discuss the partial identification approach regarding the first case (1) when $J = I - 1$, and $q_1 = 0$, $q_I = 1$. Here we focus on the remaining cases: (2.a) when $J = I$, and $q_1 = 0$, $q_I < 1$, (2.b) when $J = I$, and $q_1 > 0$, $q_I = 1$, (3) when $J = I + 1$, and $q_1 > 0$, $q_I < 1$.

Case 2.a: $J = I$, $q_1 = 0$

For this case, we need to separate $j = 1, \cdots, J-1$, and $j = J$, because for $j = J$ we utilize the bounds in only one form instead of two. For any $j = 1, \cdots, J-1$, the bound from q_i, for $i = j+1, \cdots, J$ takes the following form:

$$y'(a_1(q_1)) \prod_{k=j+1}^{i-1} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} \prod_{l=1}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} [1, \frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{k=j+1}^{s} \frac{E_1(a_1(\alpha_k))}{E_2(a_2(\alpha_k))} = 1$ for $s < j+1$. The bound from q_i, for $i = 1, \cdots, j$ takes the following form:

$$y'(a_1(q_1)) \prod_{k=i}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} [1, \frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{l=0}^{0} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} = 1$. Thus, we adopt the same procedure as in Case 1 in the main text to get the final lower and upper bound for $y'(a_2(\alpha_j))$. The upper bound is $y'(a_1(q_1))$ times

$$max\Big(max\{\prod_{k=i}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j}, max\{\prod_{k=j+1}^{i-1} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} \prod_{l=1}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j+1}^{J}\Big),$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\Big(min\{\prod_{k=i}^{j} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j}, min\{\prod_{k=j+1}^{i-1} \frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} \prod_{l=1}^{i} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j+1}^{J}\Big).$$

For $j = J$, the bound from q_i, for $i = 1, \cdots, J$ takes the following form:

$$y'(a_1(q_1)) \prod_{k=i}^{J} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} [1, \frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{l=0}^{0} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} = 1$. One subtle difference for $j = J$ is that, from the monotonic property of the function $y'(\cdot)$, we know $y'(a_2(\alpha_J))$ has a natural upper bound, i.e., $y'(a_2(q_J))$, which takes the form

$$y'(a_1(q_1)) \prod_{l=0}^{J} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))},$$

and has not been considered, thus this lower bound should be compared with the lower bounds from the q_i, for $i = 1, \cdots, J$ to further tighten the identified set.

For $j = J$, the upper bound is $y'(a_1(q_1))$ times

$$max\Big(max\{\prod_{k=i}^{J} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{J}, \prod_{l=0}^{J} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\Big),$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\{\prod_{k=i}^{J}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{J}.$$

Case 2.b: $J = I$, $q_I = 1$

For this case, we need to separate $j = 2,\cdots,J$, and $j = 1$, because for $j = 1$ we utilize the bounds in only one form instead of two. For any $j = 2,\cdots,J$, the bound from q_i, for $i = j,\cdots,J$ takes the following form:

$$y'(a_1(q_1))\prod_{k=j+1}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}[1,\frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{k=j+1}^{s}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} = 1$ for $s < j+1$. The bound from q_i, for $i = 1,\cdots,j-1$ takes the following form:

$$y'(a_1(q_1))\prod_{k=i+1}^{j}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}[1,\frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{l=0}^{0}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} = 1$. Thus, we adopt the same procedure as in Case 1 in the main text to get the final lower and upper bound for $y'(a_2(\alpha_j))$. The upper bound is $y'(a_1(q_1))$ times

$$max\Big(max\{\prod_{k=i+1}^{j}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j-1},max\{\prod_{k=j+1}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j}^{J}\Big),$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\Big(min\{\prod_{k=i+1}^{j}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j-1},min\{\prod_{k=j+1}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j}^{J}\Big).$$

For $j = 1$, the bound from q_i, for $i = 1,\cdots,J$ takes the following form:

$$y'(a_1(q_1))\prod_{k=2}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}[1,\frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{k=2}^{s}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} = 1$ for $s < 2$.

For $j = 1$, the upper bound is $y'(a_1(q_1))$ times

$$max\{\prod_{k=2}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{J},$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\{\prod_{k=2}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{J}.$$

Case 3: $J = I + 1$

For this case, we need to separate $j = 2, \cdots, J-1$, and $j = 1$, $j = J$, because for $j = 1$ and $j = M$ we utilize the bounds in only one form instead of two. For any $j = 2, \cdots, J-1$, the bound from q_i, for $i = j, \cdots, J-1$ takes the following form:

$$y'(a_1(q_1))\prod_{k=j+1}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}[1,\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}],$$

if we define $\prod_{k=j+1}^{s}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} = 1$ for $s < j+1$. The bound from q_i, for $i = 1, \cdots, j-1$ takes the following form:

$$y'(a_1(q_1))\prod_{k=i+1}^{j}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}[1,\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}],$$

if we define $\prod_{l=0}^{0}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} = 1$.

Thus, for any $j = 2, \cdots, J-1$, the upper bound is $y'(a_1(q_1))$ times

$$max\Big(max\{\prod_{k=i+1}^{j}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j-1}, max\{\prod_{k=j+1}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j}^{J-1}\Big),$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\Big(min\{\prod_{k=i+1}^{j}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{j-1}, min\{\prod_{k=j+1}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=j}^{J-1}\Big).$$

For $j = 1$, the bound from q_i, for $i = 1, \cdots, J-1$ takes the following form:

$$y'(a_1(q_1))\prod_{k=2}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}[1,\frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{k=2}^{s}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))} = 1$ for $s < 2$.

For $j = 1$, the upper bound is $y'(a_1(q_1))$ times

$$max\{\prod_{k=2}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{J-1},$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\{\prod_{k=2}^{i}\frac{\mathbb{E}_1(a_1(\alpha_k))}{\mathbb{E}_2(a_2(\alpha_k))}\prod_{l=1}^{i}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{J-1}.$$

For $j = J$, the bound from q_i, for $i = 1, \cdots, J-1$ takes the following form:

$$y'(a_1(q_1))\prod_{k=i+1}^{J}\frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))}\prod_{l=0}^{i-1}\frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}[1,\frac{\mathbb{E}_2(a_2(q_i))}{\mathbb{E}_1(a_1(q_i))}],$$

if we define $\prod_{l=0}^{0} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))} = 1$. One subtle difference for $j = J$ is that, from the monotonic property of the function $y'(\cdot)$, we know $y'(a_2(\alpha_J)$ has a natural upper bound, i.e., $y'(a_2(q_{J-1}))$, which takes the form

$$y'(a_1(q_1)) \prod_{l=0}^{J-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))},$$

and has not been considered, thus this lower bound should be compared with the lower bounds from the q_i, for $i = 1, \cdots, J-1$ to further tighten the identified set.

For $j = J+1$, the upper bound is $y'(a_1(q_1))$ times

$$max\Big(max\{\prod_{k=i+1}^{J} \frac{\mathbb{E}_2(a_2(\alpha_k))}{\mathbb{E}_1(a_1(\alpha_k))} \prod_{l=0}^{i-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\}_{i=1}^{J-1}, \prod_{l=0}^{J-1} \frac{\mathbb{E}_2(a_2(q_l))}{\mathbb{E}_1(a_1(q_l))}\Big),$$

and the lower bound is thus $y'(a_1(q_1))$ times

$$min\{\prod_{k=i+1}^{J} \frac{E_2(a_2(\alpha_k))}{E_1(a_1(\alpha_k))} \prod_{l=0}^{i} \frac{E_2(a_2(q_l))}{E_1(a_1(q_l))}\}_{i=1}^{J-1}.$$

2.9.2 Identification of the Second and Third Forms of Payoff

Identification of the Second Form of Payoff

Recall that the second form of payoff is: (2). $\pi_i(\mathbf{a}, t_i) = x_i(\mathbf{a}) + y_i(a_i)/t_i$. In this appendix subsection, we first provide the assumptions that all functions have to satisfy. Then, we show that the nonparametric identification approach in the main text can be applied to this form of payoff function. Therefore, all proofs in this subsection are omitted, because they follow the proofs of the main results given in the main text Appendix.

As in the main text, we impose the restriction that the function $x_i(\cdot)$ is known. Unlike in the main text, however, we do not require the strictly positive sign of its derivative with respect to a_i. The unknown model primitives consist of a pair of functions: $[F_i(\cdot), y_i(\cdot)]$. The following assumption gives the regularity conditions that this pair must satisfy.

Assumption 5 *Let $\mathscr{F}$ and $\mathscr{Y}$ be two classes of functions that $F_i(\cdot)$ and $y_i(\cdot)$ belong to. Then $\mathscr{F}$ and $\mathscr{Y}$ satisfy the following conditions:*
(i) $\forall F(\cdot) \in \mathscr{F}$, $F(\cdot)$ satisfies: (1) $F(\cdot)$ is an absolutely continuous CDF with a compact support: $\mathscr{T} \equiv [\underline{t}, \bar{t}] \subset \mathbb{R}$, where $0 \notin \mathscr{T}$; (2) $F(\cdot)$ has a bounded, continuous and atomless PDF $f(\cdot)$, which is strictly positive and bounded from infinity over $\mathscr{T}$, i.e.: $0 < f(\cdot) < \infty$.
(ii) $\forall y(\cdot) \in \mathscr{Y}$, $y(\cdot)$ satisfies: (1) $y(\cdot)$ is twice continuously differentiable (C^2) over a compact support:

$\mathscr{A} \subset \mathbb{R}$; (2) there exists a positive number M_y, such that the first derivative $y'(\cdot)$ satisfies: $0 < -y'(\cdot) \leq M_y$ almost everywhere; (3) $y(\cdot)$ is convex or concave over $\mathscr{A}$.

Most conditions of Assumption 5 are the same as those in Assumption 1 for the first form of payoff. The only condition that needs to be noted is (ii)-(2), which is assumed to ensure that the payoff function $\pi_i(\mathbf{a}, t_i)$ satisfies the strict supermodularity condition.

The existence and the regularity of strictly MPSNE, denoted as $s_i(\cdot)$, in this case is shown in Proposition 2. The properties of the derived CDF $G_i(\cdot)$ are the same as in Proposition 3. The assumption, with conditions that the observed action CDF $G_i(\cdot)$ satisfies, is equivalent to Assumption 2, and we still call the class of such functions $\mathscr{G}$ for convenience. All of the three can be found in the main text for the first form of payoff, thus are suppressed here.

To obtain nonparametric identification in this case, we again introduce the exclusion restriction $F_i(\cdot|N) = F_i(\cdot)$ for each $N \geq 2 \in \mathbb{N}^+$. The nonparametric identification approach then follows directly Section 2.4 for symmetric equilibrium, and Section 2.6 for asymmetric equilibrium.

Taking symmetric equilibrium as an example, one only needs to note that the first order condition for this payoff form, which corresponds to Equation (2.3.1), becomes:

$$\mathbb{E}_{\mathbf{a}_{-N}}[\frac{\partial x(a(\alpha), \mathbf{a}_{-N}; N)}{\partial a}] + \frac{y'(a(\alpha))}{t(\alpha)} = 0. \tag{2.9.1}$$

As the final remark, we emphasize the crucial role of sign restriction as in Assumption 1-(ii)-(2) and Assumption 5-(ii)-(2) in nonparametric identification. The difference is that here since we have to assume the supermodularity condition, with this form of payoff, function $y(\cdot)$ has to be strictly decreasing; as a result, the sign restriction is satisfied naturally.

Since by stacking the two first order conditions as in Equation (2.9.1) for two numbers of players $N_1 < N_2$, we cancel out $t(\alpha)$ evaluated at one specific α, whose sign is unrestricted. Furthermore, the sign of expectation $\mathbb{E}_{\mathbf{a}_{-N}}[\partial x(a(\alpha), \mathbf{a}_{-N}; N)/\partial a]$ is unrestricted, unlike in the main text. Therefore, it is essential to know the sign of $y'(\cdot)$ in advance, since taking ratio sometimes causes the confusion of the sign.

Identification of the Third Form of Payoff

Recall that the third form of payoff is: (3). $\pi_i(\mathbf{a}, t_i) = t_i a_i + x_i(\mathbf{a}) + y_i(a_i)$. In this appendix subsection, we first provide the assumptions that all functions have to satisfy. Then we propose the nonparametric identification approach for this form of payoff function, which is an adaption of the identification approach. Besides, all proofs in this subsection are omitted, because they follow the proofs of the main results given in Appendix 2.8, with minor adjustment.

As in the main text, we impose the restriction that function $x_i(\cdot)$ is known. Also as in the second form of payoff, we do not require the strictly positive sign of its derivative with respect to a_i, because the strict supermodularity condition is automatically satisfied. The unknown model primitives consists of a pair of functions: $[F_i(\cdot), y_i(\cdot)]$. The following assumption gives the regularity conditions that this pair needs to satisfy.

Assumption 6 *Let $\mathscr{F}$ and $\mathscr{Y}$ be two classes of functions that $F_i(\cdot)$ and $y_i(\cdot)$ belong to. Then $\mathscr{F}$ and $\mathscr{Y}$ satisfy the following conditions:*
(i) $\forall F(\cdot) \in \mathscr{F}$, $F(\cdot)$ satisfies: (1) $F(\cdot)$ is an absolutely continuous CDF with a compact support: $\mathscr{T} \equiv [\underline{t}, \bar{t}] \subset \mathbb{R}$; (2) $F(\cdot)$ has a bounded, continuous and atomless PDF $f(\cdot)$, which is strictly positive and bounded from infinity over $\mathscr{T}$, i.e.: $0 < f(\cdot) < \infty$.
(ii) $\forall y(\cdot) \in \mathscr{Y}$, $y(\cdot)$ satisfies: (1) $y(\cdot)$ is twice continuously differentiable (C^2) over a compact support: $\mathscr{A} \subset \mathbb{R}$ with bounded derivatives; (2) $y(\cdot)$ is convex or concave over $\mathscr{A}$.

Note that, unlike the main text or the second form of payoff, here Assumption 6-(ii) does not require the sign restriction, either for the strict supermodularity condition, or for the confusion caused by taking ratio of first order conditions. This is because the identification approach adopted here is not taking ratio, but *taking difference* of first order conditions.

The existence and the regularity of strictly MPSNE, denoted as $s_i(\cdot)$, in this case is shown in Proposition 2. The properties of the derived CDF $G_i(\cdot)$ are the same as in Proposition 3. Also the assumption, with conditions the observed action CDF $G_i(\cdot)$ satisfies, is equivalent to Assumption 2, and we still call the class of such functions $\mathscr{G}$ for convenience. All of the three can be found in the main text for the first form of payoff, thus are suppressed here.

Turn to the nonparametric identification approach. First note that corresponding to the key first order equation as Equation (2.9.1) in the above subsection, for the third form of payoff, we have the following first order condition, expressed through the expectation over all possible opponents' action profile, when considering symmetric equilibrium:

$$t(\alpha) + \mathbb{E}_{\mathbf{a}_{-N}}[\frac{\partial x(a(\alpha), \mathbf{a}_{-N}; N)}{\partial a}] + y'(a(\alpha)) = 0. \tag{2.9.2}$$

Given the exclusion restriction that $F(\cdot|N) = F(\cdot)$ for all $N \geq 2 \in \mathbb{N}^{+}$, in the case of symmetric equilibrium as in Section 2.4 in the main text, we again consider two numbers of players, i.e. $N_1 < N_2$.

The system of first order conditions is thus

$$\begin{aligned} t(\alpha) + \mathbb{E}_{\mathbf{a}_{-N_1}}\left[\frac{\partial x(a_1(\alpha), \mathbf{a}_{-N_1}; N_1)}{\partial a_1}\right] + y'(a_1(\alpha)) = 0, \\ t(\alpha) + \mathbb{E}_{\mathbf{a}_{-N_2}}\left[\frac{\partial x(a_2(\alpha), \mathbf{a}_{-N_2}; N_2)}{\partial a_2}\right] + y'(a_2(\alpha)) = 0, \end{aligned} \tag{2.9.3}$$

evaluated at a pre-specified $\alpha \in [0,1]$.

The key idea of nonparametric identification now relies on the *difference* of two equations in the above system:

$$\mathbb{E}_{\mathbf{a}_{-N_2}}\left[\frac{\partial x(a_2(\alpha), \mathbf{a}_{-N_2}; N_2)}{\partial a_2}\right] - \mathbb{E}_{\mathbf{a}_{-N_1}}\left[\frac{\partial x(a_1(\alpha), \mathbf{a}_{-N_1}; N_1)}{\partial a_1}\right] = y'(a_1(\alpha)) - y'(a_2(\alpha)). \tag{2.9.4}$$

The contraction sequence $\{\alpha_n\}_{n=0}^{\infty}$ and the bounds from a finite set of probabilities at which $y'(a_1(\cdot))$ and $y'(a_2(\cdot))$ are known play exactly the same roles in the point and partial identification approaches as in Sections 2.4.1 and 2.4.2, depending on the pattern of functional comparative statics. However, for the third form of payoff, due to the difference in Equation (2.9.4), the identity equation that resembles Equation (2.4.4) is thus:

$$\begin{aligned} y'(a_2(\alpha_0)) = & \left[y'(a_2(\alpha_0)) - y'(a_1(\alpha_0))\right] + \left[y'(a_2(\alpha_1)) - y'(a_1(\alpha_1))\right] \\ & + \cdots + \left[y'(a_2(\alpha_n)) - y'(a_1(\alpha_n))\right] + \cdots + y'(a(\alpha^{IP})), \\ = & y'(a(\alpha^{IP})) + \sum_{n=0}^{\infty} \left[y'(a_2(\alpha_n)) - y'(a_1(\alpha_n))\right]. \end{aligned} \tag{2.9.5}$$

Note that this identity holds because (1) the condition $y'(a_1(\alpha_n)) = y'(a_2(\alpha_{n+1}))$ leads to the cancellation of all intermediate terms, and (2) $y'(a_2(\alpha_n)) - y'(a_1(\alpha_n)) \to 0$ as $n \to \infty$. Also, note that since the value $y'(a(\alpha^{IP}))$ is unknown, the identification results in the following hold up to a *location*.

Each difference term in the identity equation (2.9.5) is identified through the left hand side difference in Equation (2.9.4). The point and partial identification approaches thus follow the main text, including those for symmetric equilibrium in Section 2.4 and those for asymmetric equilibrium in Section 2.6.

Note that for this form of payoff there is no sign restriction, since we take difference of two first order conditions like in Equation (2.9.4). By taking difference instead of ratio, there is no ambiguity of signs on $y'(\cdot)$.

Chapter 3: Testing for Collusion in Procurement Auctions

This chapter is adapted from the working paper "Testing for Collusion in Procurement Auctions" and has been reproduced with the permission of my co-authors Tong Li and Jun Nakabayashi.

3.1 Introduction

Collusion has been a serious antitrust issue in procurement auctions and served as a threat to the integrity of procurement auctions. The basic purpose of a procurement auction, which comprises purchasing of goods or services required for activities, is to secure the best value for the buyer's money. Bid rigging, as the typical mechanism of collusion in procurement contracts, distorts the efficient functioning of procurement auctions, because bidders may determine among themselves how to arrange the bids and alter the winner, for example, by bid rotation scheme, phantom bidding with side-payments, or geographic territory. Therefore, it is important for investigators, prosecutors, auditors, and other anti-collusion professionals to detect and prove bid rigging.

In a seminal paper, Porter and Zona (1993) propose a test based on the rank distribution of the cartel and non-cartel bids and use their test to analyze the Long Island highway construction data, and find the evidence of prevalent existence of cartel and bid rigging schemes. Their test is based on the assumption that the phantom bidding scheme was used (as supported by the empirical observation) and thus there were fundamental differences between the ordering of the cartel and competitive bids. They assume bidders' private costs are independent, and their rank-based test is derived from a multinomial logit (MNL) model that exhibits the Independence of Irrelevant Alternatives (IIA) property. However, in auctions in general and in procurement auctions in particular, bidders' private values or costs are likely to be positively correlated, or affiliated through some common factors.[1] For example, in the bridge construction setting considered in this paper, bidders' private costs can be correlated through uncertainties they all face in assessing soil conditions when relaying or removing existing bridges (e.g. De Silva, Dunne, Kankanamge, and Kosmopoulou (2008)). Moreover, in procurement auctions, firms often use the same subcontractors and thus have dependence among their private costs through

1 See the discussion in "Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2020" by the Committee for the Prize in Economic Sciences in Memory of Alfred Nobel (The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel (2020)), on the importance of the affiliation concept introduced in Milgrom and Weber (1982) and its empirical relevance.

their shared subcontractors (Rosa (2019)). The structural analysis of the affiliated private value (APV) model has been conducted in the literature (e.g. Li, Perrigne, and Vuong (2000), Li, Perrigne, and Vuong (2002), Li, Perrigne, and Vuong (2003), Campo, Perrigne, and Vuong (2003), Hubbard, Li, and Paarsch (2012) and Li and Zhang (2015)), and several tests for affiliation have been proposed (de Castro and Paarsch (2010), Li and Zhang (2010), and Jun, Pinkse, and Wan (2010)) that also offer empirical support for the validity of the affiliation assumption in real world auctions . However, most of the analyses of cartels and bid rigging and of the structural analyses of procurement auctions have been conducted within the independent private value (IPV) paradigm.[2]

In this paper, we extend the rank-based test in Porter and Zona (1993) to allow for affiliated private costs among bidders. Toward this end, we assume that the log of bids (conditional on auction heterogeneity and bidder-specific covariates) follow a joint normal distribution. As such, the probabilities of rankings of bids can be modeled as a multivariate probit model instead of the MNL as in Porter and Zona (1993). To deal with the multiple integrals in the likelihood function, we propose a simulation-based method using the GHK simulator developed by Geweke (1991), Börsch-Supan and Hajivassiliou (1993), and Keane (1994). We use a Wald test to test for collusion. In a similar spirit to Porter and Zona (1993), our rank-based test is designed to detect differences of the determinants for the bidding rules between the bidder delivering the lowest bid and all the other bidders with higher bids. Our test can also detect more sophisticated collusion schemes such as the one where the cartel members all inflate their bids above costs by the same percentage. Moreover, as in Porter and Zona (1993), we are able to control for the unobserved auction heterogeneity in conducting the empirical test. Controlling for unobserved auction heterogeneity is an important econometric issue that has recently drawn attention in the structural analysis of auction data; see, e.g., Li and Zheng (2009), Li and Zheng (2012), Athey, Levin, and Seira (2011), Krasnokutskaya (2011), Li and Zhang (2010), Li and Zhang (2015), Roberts (2013), Roberts and Sweeting (2013), Bhattacharya, Roberts, and Sweeting (2014), and Compiani, Haile, and Sant'Anna (2019). In our setting where the bidders' private costs are affiliated, the importance of controlling for unobserved heterogeneity is more pronounced. This is because the dependence among the observed bids (conditional on the observables) can be attributed

2 The exception for the former is Chassang, Kawai, Nakabayashi, and Ortner (2019) who propose a way of detecting collusion based on a general information structure that includes the APV paradigm. The exceptions for the latter are Hong and Shum (2003) who study competition and winner's curse in procurement auctions using a common value model, Rosa (2019) and De Silva, Li, and Zhao (2020) who study a residence preference program and a bidder training program, respectively, in procurement auctions within the APV paradigm.

to two sources, namely, the unobserved auction heterogeneity that is observed by the bidders when they make bidding decisions but unobserved by the econometrician, and the affiliation that is caused by some common factors unobserved by the bidders when they submit bid. Only when unobserved heterogeneity is controlled for, can one isolate the affiliation level from the dependence among bids.

We use our test to examine possible bid rigging occurrence in Japanese procurement auctions for bridge construction projects let by the Ministry of Infrastructure Land and Transportation (MILT). It had been the largest bid rigging case in the history of the nation. The Japan Fair Trade Commission (JFTC) ordered 44 firms to pay a total fine of approximately 130 million dollars. Our testing results indicate evidence of prevalent collusion, not only among the firms who were convicted by the JFTC, but also among some of those who were not convicted. To see affiliation matters in testing collusion, we also conduct the MNL test used in Porter and Zona (1993). We find that the estimates based on the MNL with the IPV assumption are quite different from the estimates we obtain assuming APV. Moreover, the test for collusion with the IPV assumption yields less plausible results.

Our paper contributes to the literature on bid rigging detection. For the literature in detecting collusion in first price auctions, see Porter and Zona (1999), Pesendorfer (2000), Bajari and Ye (2003), List, Millimet, and Price (2007), Ishii (2008), Price (2008), Aryal and Gabrielli (2013), Schurter (2020), and Chassang and Ortner (2019). For detecting collusion in English auctions, see Baldwin, Marshall, and Richard (1997), Athey, Levin, and Seira (2011), Marmer, Shneyerov, and Kaplan (2017), and Marshall, Richard, and Shen (2019). Moreover, for detecting collusion in non-standard auction formats, such as the two-stage model with the first stage being knockout auction (first-price sealed auction with side-payments) among collusive bidders and the second stage being English auction and the average-bid auction, respectively; see Asker (2010) for the former and Conley and Decarolis (2016) for the latter. Note that all these papers study or detect collusions in an IPV paradigm. See Porter (2005) and Marshall and Marx (2012) for the thorough discussion of the issues with detecting collusion and review of the related literature.

The paper is organized as follows. Section 3.2 depicts the data and the background. The estimation method and the testing procedure are proposed in Section 3.3. In Section 3.4, we apply our test to analyze the data. Section 3.5 concludes.

3.2 Data and Background

The data we study consists of 709 procurement auctions for bridge superstructure construction let by the MILT in Japan from April 2003 through December 2006. The format of the auctions used by the MILT is a first-price sealed-bid auction with a secret reserve price.[3] These auctions can be separated into two groups, namely, "open" and "invited," depending on how the MILT imposes on the participation restriction. In an open auction, any qualified bidder can participate. In an invited-bidder auction, the government typically chooses 10 firms from a list of qualified contractors and solicits bids only from these firms.[4]

The data includes auctions for a major collusion case: one of the largest scales of collusion the country had ever experienced. The JFTC ruling documents that the ring consists of 47 manufactures of steel bridge upperstructrue, such as IHI, Nippon Steel Corporation, and Mitsubishi Heavy Industries, and was implicated to keep colluding for a long time. In October 2004, the JFTC begun on-site inspections of thirty firms. The JFTC first criminalized the prosecutor's office for collusion with eight rigging companies in May 2005 and with additional 18 companies on June 15 in 2005. On the same day, the Tokyo High Public Prosecutor's Office indicted these 26 companies. Perhaps the most publicly noticed event in the bid rigging scandal was the prosecution's arrest of the vice president of the Japan Highway Public Corporation on July 12, 2005, on suspicion of supporting rigging and violating the antitrust law. In response to these events, the MLIT announced on July 29th a thorough investigation of the bid rigging case and a drastic revision of the bidding system. Given these events, we split our data into two periods, one includes auctions that occurred on July 31, 2005 and earlier (referred to as before mid-2005 hereinafter), and the other includes the remaining auctions (referred to as after mid-2005 hereinafter). Figure 3.1 illustrates the winning bids (relative to the reserve prices) and the share of open auctions (with the solid line) in our sample. The vertical line is July 31, 2005. We see that winning bids fall across the vertical line and that the share of open auctions increases significantly after July 31. The data also has the index for collusive firms who were convicted by the JFTC; for

3 The government agencies reauction projects if no bid falls below the secret reserve price. Except for this, the auction format is exactly the same as the first-price sealed-bid procurement as long as in the first round the lowest bid is below the secret reserve price, and we restrict our attention to the first round of bidding. Kawai and Nakabayashi (2014) exploit the reauction feature and develop the method similar to regression discontinuity by focusing on the projects that were reauctioned.

4 The MILT acknowledges that invited bidders had been facilitating collusion, which can explain why the government changed the policy of participation restriction to encourage "open" (free) participation in auctions as part of the government's initiative in adopting a general competition method with high objectivity, transparency, and competitiveness.

these firms that were not convicted by the JFTC, we call them "competitive" firms.

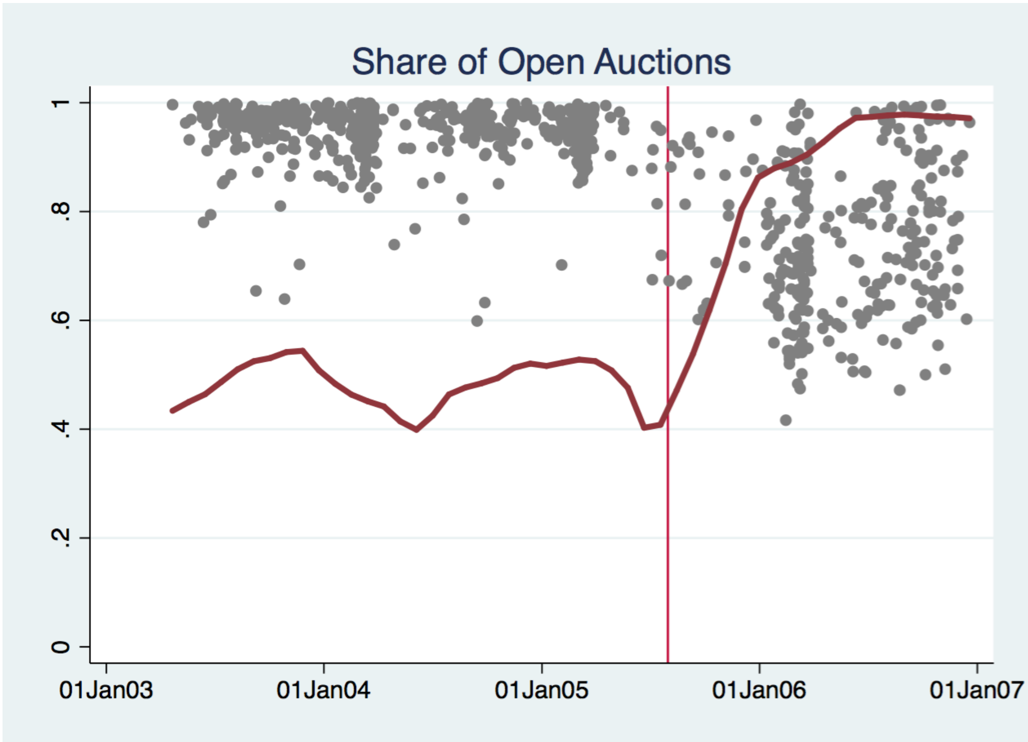


Figure 3.1: The increasing share of open auctions and the decreasing percent of winning bids normalized by the reserve prices.

Table 3.1 presents the auction-specific summary statistics. Regardless of the auction type, the average number of bidders drops after the institutional change on July 31, 2005. Both the winning bids and the reserve prices in the after mid-2005 subsamples are smaller than those in the before mid-2005 subsamples on average, and the magnitudes of decreases are larger invited auctions relative to open auctions. After normalized by the reserve prices, the average winning bids fall over time, specifically in open auctions, which is consistent with the patterns shown in Figure 3.1.

To capture the asymmetry among the bidders and for the purpose of testing for the difference between the lowest bid and the other bids in one auction, which is the key idea behind our test, we follow Porter and Zona (1993) to construct bidder-specific covariates. The variable "backlog" is defined as the sum of the amounts of contract awarded within one year of the auction in which the firm participates. The variable "Capacity" is a firm's maximum backlog during the sample period. A firm's utilization rate ("Util") measures the firm's backlog at a given time and is constructed by dividing the backlog by the capacity. We include the squared terms of the capacity and the utilization rate ("Capacity Sq" and "Util Sq") to account for possible nonlinearities. Following Porter and Zona (1993), we create a dummy variable "Noback" for "competitive" bidders, which is equal to one if the

Table 3.1: Auction-Specific Summary Statistics

	Before Mid-2005		After Mid-2005	
	Open Auctions	Invited Auctions	Open Auctions	Invited Auctions
No. of Auctions	210	275	44	18
No. of Bidders	10.862	9.7564	5.8182	6.3333
	(2.646)	(1.762)	(4.161)	(3.087)
Winning Bid	4.1796	1.4959	2.0136	0.4111
	(3.273)	(1.404)	(1.879)	(0.262)
Reserve Price	4.3986	1.5355	2.9209	0.4797
	(3.569)	(1.434)	(3.014)	(0.301)
Winning Bid / Reserve Price	0.9639	0.9751	0.7618	0.8785
	(0.063)	(0.094)	(0.133)	(0.143)

Notes: The means are reported for variables of interest, with the standard deviations shown in parentheses. The means of "Winning bid" and "Reserve price" and the associated standard deviations are scaled down by 10^8.

"competitive" bidder never won during the same period.

Seven different datasets are used in the application. The first and the second consist of the ranking of "competitive" bids in the open and invited auctions before mid-2005, respectively, where we keep the auctions that receive two or more "competitive" bids. The third and the fourth consist of the ranking of cartel bids in the open and invited auctions before mid-2005, respectively, and the auctions with one or no cartel bids are dropped. The fifth and the sixth consist of the ranking of "competitive" bids in the open and invited auctions after mid-2005, respectively, where auctions with two or more "competitive" bids are included. Since all the invited auctions after mid-2005 have fewer than two cartel bids, we only keep the open auctions with two or more cartel bids after mid-2005 as our last subset.

Table 3.2 gives the bidder-specific summary statistics of the data. The mean "competitive" bids for the open and invited auctions before mid-2005 are, approximately, 292 million and 51 million Yen, respectively. The mean cartel bids for the open and invited auctions before mid-2005 are 471 million and 187 million Yen, respectively. After mid-2005, the mean "competitive" bids for open and invited auctions are 292 million and 50 million Yen, respectively. The mean cartel bids for the open auctions after mid-2005 is 306 million Yen. Regardless of the time period and the type of the procurement auction, the bids submitted by by "competitive" and cartel bidders exhibit similar patterns: cartel bidder's bids are relatively larger than those submitted by "competitive" bidders on average, meaning that cartel bidders tend to be less aggressive or bid on larger projects.

Table 3.2: Bidder-Specific Summary Statistics

	Before Mid-2005				After Mid-2005		
	"Competitive" Ranks		Cartel Ranks		"Competitive" Ranks		Cartel Ranks
	Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions
No. of Auctions	72	90	209	235	32	17	19
No. of Bids	225	521	2021	2105	126	110	112
Bid	2.9163	0.5139	4.7089	1.8676	2.9184	0.4966	3.0575
	(2.428)	(0.467)	(3.586)	(1.539)	(2.257)	(0.276)	(2.277)
Util	0.2282	0.1391	0.4545	0.4778	0.4389	0.2299	0.3970
	(0.309)	(0.283)	(0.324)	(0.326)	(0.409)	(0.387)	(0.300)
Util Sq	0.1473	0.0991	0.3113	0.3345	0.3590	0.2016	0.2468
	(0.297)	(0.252)	(0.320)	(0.326)	(0.410)	(0.378)	(0.279)
Capacity	3.7330	1.5009	28.586	23.592	4.8399	1.2524	20.885
	(3.397)	(2.914)	(16.12)	(15.73)	(5.112)	(2.091)	(15.76)
Capacity Sq	0.2542	0.1073	10.767	8.0374	0.4935	0.0590	6.8229
	(0.414)	(0.392)	(11.99)	(10.76)	(0.905)	(0.232)	(9.996)
Noback	0.2222	0.5048			0.1190	0.2909	
	(0.417)	(0.500)			(0.325)	(0.456)	

Notes: The means are reported for variables of interest, with the standard deviations shown in parentheses. The means of "Bid" and "Capacity" and the associated standard deviations are scaled down by 10^8. The means and the standard deviations of "Capacity Sq" are scaled down by 10^{18}.

3.3 The Collusion Test

We consider a first-price sealed-bid procurement auction in which a single project is auctioned off to n risk-neutral bidders within the APV paradigm, extending the IPV paradigm considered in Porter and Zona (1993).[5] Affiliation is a terminology describing the dependence among random variables, which is introduced in Milgrom and Weber (1982) as follows:

Definition 1 *Let y and y' be any two values of a random vector $Y \subseteq \mathbb{R}^n$ with a density $f(\cdot)$. Then all elements of Y are affiliated if $f(y \vee y')f(y \wedge y') \geq f(y)f(y')$, where $y \vee y' = (\max(y_1, y'_1), \ldots \max(y_n, y'_n))$, and $y \wedge y' = (\min(y_1, y'_1), \ldots \min(y_n, y'_n))$.*

Intuitively speaking, affiliation means that large values for some of the components in Y make other components more likely to be large than small. In the case where the bidders' private costs follow a log-normal distribution, the (positive) correlation of the normal distribution measures the level of affiliation.

We assume that we observe a $1 \times p_{\mathbf{x}}$ auction-specific covariate vector, denoted by $\mathbf{x}_l$, $l = 1, \cdots, L$, where L is the number of auctions in the data. For auction l with n_l bidders, we define b_{li} as the bid submitted by bidder i, for $i = 1, \cdots, n_l$. Since we deal with procurement auctions, the bids are

5 We follow Porter and Zona (1993) to model procurement auctions as private value auctions. As Porter (2005) points out, collusion often occurs in private value auctions in which bidders differ in their idiosyncratic willingness to pay, rather than differences in information concerning common components of valuations; procurement auctions are a primary example of such.

arranged in the descending order: $b_{l1} > b_{l2} > \cdots > b_{ln_l}$, with the last bid b_{ln_l} representing the lowest (winning) bid of the auction. Bidder-specific covariate vector denoted by $\mathbf{z}_{li}$ is assumed to be $1 \times p_{\mathbf{z}}$, for $i = 1, \cdots, n_l$. Following Porter and Zona (1993), we suppose that the equilibrium behavior can be approximated by the log-linear bidding rule:

$$log(b_{li}) = \mathbf{x}_l\beta + \mathbf{z}_{li}\gamma_i + \eta_l + \varepsilon_{li}, \tag{3.3.1}$$

where η_l is the unobserved auction heterogeneity while ε_{li} denotes the unobserved idiosyncratic error at the bidder level, independent of $\mathbf{x}_l$, $\mathbf{z}_{li}$, and η_l. Denote the vector of all γ_i, $i = 1, ..., n_l$, $l = 1, \cdots, L$ as γ, which capture the asymmetry among bidders. Note that here we allow the coefficient of z_{li} to depend on bidder i, while it is assumed to be a constant across all bidders in Porter and Zona (1993).

In Porter and Zona (1993), they assume $-\varepsilon_{l1}, \cdots, -\varepsilon_{ln_l}$ are i.i.d and each follows a type I extreme value distribution, consistent with the IPV paradigm they consider. To accommodate affiliation, instead we assume that $\varepsilon_l = (\varepsilon_{l1}, \cdots, \varepsilon_{ln_l})'$ follows a joint normal distribution with mean $\mu = (0, \cdots, 0)'$ and covariance

$$\Sigma_{n_l} = \begin{bmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{bmatrix},$$

implying that ε_{li} and ε_{lj} have a correlation denoted by ρ, the same across all the bidders/auctions. Therefore, $\rho \geq 0$ represents the affiliation in our APV paradigm.[6]

The likelihood of observing the ranking of bids for auction l, denoted by $p_l(\gamma, \rho)$ is as follows:

$$\begin{aligned} & p_l(\gamma, \rho) \\ =& P\big(b_{l1} > b_{l2} > \cdots > b_{ln_l}\big) \\ =& P\big(log(b_{l1}) > log(b_{l2}) > \cdots > log(b_{ln_l})\big) \\ =& \int_{-\infty}^{+\infty} \int_{-\infty}^{log(b_{l1})} \cdots \int_{\infty}^{log(b_{ln_l - 1})} dG\big(log(b_{l1}), \cdots, log(b_{ln_l})\big), \end{aligned}$$

where $G\big(log(b_{l1}), \cdots, log(b_{ln_l})\big)$ denotes the joint distribution of $log(b_{li})$'s induced by η_l and ε_{li}'s.

6 As discussed earlier in Section 3.1 and can be clearly seen in (3.3.1), the dependence among the observed bids in an auction can be attributed to the two sources, one from the unobserved auction heterogeneity, and the other from the (positive) dependence among the idiosyncratic error terms, which is the affiliation among the bidders' private costs.

Note that this likelihood can also be represented as following:

$$
\begin{aligned}
& P\big(b_{l1} > b_{l2} > \cdots > b_{ln_l}\big) \\
=& P\big((z_{l1}\gamma_1 + \varepsilon_{l1}) > (z_{l2}\gamma_2 + \varepsilon_{l2}) > \cdots > (z_{ln_l}\gamma_{n_l} + \varepsilon_{ln_l})\big) \\
=& \int_{-\infty}^{+\infty} \int_{-\infty}^{z_{l1}\gamma_1 - z_{l2}\gamma_2 + \varepsilon_{l1}} \cdots \int_{\infty}^{z_{ln_l-1}\gamma_{n_l-1} - z_{ln_l}\gamma_{n_l} + \varepsilon_{ln_l-1}} dF(\varepsilon_{l1}, \cdots, \varepsilon_{ln_l}) \\
=& \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} 1\{\varepsilon_l \in A_l\} dF(\varepsilon_{l1}, \cdots, \varepsilon_{ln_l}),
\end{aligned}
$$

where the auction-specific covariate vector $\mathbf{x}_l$ and the auction unobserved heterogeneity η_l are cancelled out, and A_l denotes the corresponding set of ε_l induced by the ranking of bids. The dimension of the integral above is equal to n_l, the number of bidders in auction $l = 1, \cdots, L$, which renders the integral difficult to evaluate as the average number of bidders in the data is about 8 with the largest being 22.

3.3.1 A Smoothly Simulated Maximum Likelihood Estimator

Let $\theta = (\gamma, \rho)$ denote the unknown parameters. The likelihood of observing the ranking of bids for auction l can be written as $p_l(\theta)$. The joint log-likelihood function for all auctions in the data is defined as

$$
L(\mathbf{b}, \mathbf{z}; \theta) = \sum_{l=1}^{L} \ln p_l(\theta),
$$

where $\mathbf{b}_l = (b_{l1}, \cdots, b_{ln_l})'$, $\mathbf{z}_l = (z_{l1}, \cdots, z_{ln_l})'$, and $\mathbf{b} = (\mathbf{b}_1, \cdots, \mathbf{b}_L)'$, $\mathbf{z} = (\mathbf{z}_1, \cdots, \mathbf{z}_L)'$.

To overcome the computational burden associated with the multiple integral in $p_l(\theta)$, we adopt a simulation-based method to approximate the above log-likelihood function. Also with the use of the simulation-based method, the objective function may become non-smooth due to the index function $1\{\varepsilon_l \in A_l\}$ in the integrand. Toward this end, we use a smooth and unbiased simulator that is called the GHK simulator after Geweke, Hajivassiliou, and Keane, which can address both issues; see Hajivassiliou, McFadden, and Ruud (1996).[7] Specifically, there exists a lower triangular matrix h_l such that $h_l h_l' = \Sigma_{n_l}$. Then ε_l can be written as $\varepsilon_l = h_l \xi_l$, where $\xi_l = (\xi_{l1}, \cdots, \xi_{ln_l})'$ follows an n_l-variate standard normal distribution. Using the Choleski decomposition, we have the following recursive formula for ε_l:

$$
\begin{aligned}
\varepsilon_{l1} &= h_{l,11}\xi_{l1} \\
\varepsilon_{l2} &= h_{l,21}\xi_{l1} + h_{l,22}\xi_{l2}
\end{aligned}
$$

7 See also Li and Zhang (2010) for using the GHK simulator to test for affiliation using bidders' entry behavior.

$$\cdots$$

$$\varepsilon_{ln_l} = h_{l,n_l 1}\xi_{l1} + \cdots + h_{l,n_l n_l}\xi_{ln_l},$$

where $h_{l,ij} = h_l(i,j)$. Then we can make another transformation from ε_l to ξ_l to calculate the likelihood $P(b_{l1} > b_{l2} > \cdots > b_{ln_l})$ as follows:

$$\begin{aligned}
&P\big(b_{l1} > b_{l2} > \cdots > b_{ln_l}\big)\\
=&\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} 1\{\varepsilon_l \in A_l\} dF(\varepsilon_{l1}, \cdots, \varepsilon_{ln_l})\\
=&\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} 1\{\xi_l \in B_l\}\phi(\xi_{l1})\cdots\phi(\xi_{ln_l}) d\xi_{l1}\cdots d\xi_{ln_l}\\
=&\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} q(\xi_l) 1\{\xi_l \in B_l\}\frac{\phi(\xi_{l1})\cdots\phi(\xi_{ln_l})}{q(\xi_l)} d\xi_{l1}\cdots d\xi_{ln_l}\\
=&E_{\xi_l}[q(\xi_l)],
\end{aligned}$$

where B_l denotes the corresponding set of ξ_l induced by A_l; and

$$\begin{aligned}
q(\xi_l) =&\Phi\Big(\frac{z_{l1}\gamma_1 - z_{l2}\gamma_2 + (h_{l,11} - h_{l,21})\xi_{l1}}{h_{l,22}}\Big) \times \Phi\Big(\frac{z_{l2}\gamma_2 - z_{l3}\gamma_3 + (h_{l,21} - h_{l,31})\xi_{l1} + (h_{l,22} - h_{l,32})\xi_{l2}}{h_{l,33}}\Big)\\
&\times \cdots \times \Phi\Big(\frac{z_{ln_l-1}\gamma_{n_l-1} - z_{ln_l}\gamma_{n_l} + (h_{l,n_l-11} - h_{l,n_l1})\xi_{l1} + \cdots + (h_{l,n_l-1n_l-1} - h_{l,n_ln_l-1})\xi_{ln_l-1}}{h_{l,n_ln_l}}\Big);
\end{aligned}$$

by noting that $1\{\xi_l \in B_l\}\frac{\phi(\xi_{l1})\cdots\phi(\xi_{ln_l})}{q(\xi_l)}$ is actually a density function of ξ_l, the last equality holds.

Thus, $P(b_{l1} > b_{l2} > \cdots > b_{ln_l})$ can be simulated by $\frac{1}{R}\sum_{r=1}^{R} q(\xi_l^r)$, where ξ_l^r is simulated by the following procedure.

Step 1: Draw a realization of ξ_{l1}, denoted by ξ_{l1}^r from $N(0,1)$.

Step 2: Calculate $P(\xi_{l2} < \frac{z_{l1}\gamma_1 - z_{l2}\gamma_2 + (h_{l,11} - h_{l,21})\xi_{l1}}{h_{l,22}} | \xi_{l1} = \xi_{l1}^r) = \Phi(\frac{z_{l1}\gamma_1 - z_{l2}\gamma_2 + (h_{l,11} - h_{l,21})\xi_{l1}^r}{h_{l,22}})$.

Step 3: Draw a realization of ξ_{l2}, denoted by ξ_{l2}^r from a truncated standard normal truncated at $\frac{z_{l1}\gamma_1 - z_{l2}\gamma_2 + (h_{l,11} - h_{l,21})\xi_{l1}^r}{h_{l,22}}$ from above.

$\vdots$

Step 4: Draw a realization of ξ_{ln_l-1}, denoted by $\xi_{ln_l-1}^r$ from a truncated standard normal truncated at $\frac{z_{ln_l-2}\gamma_{n_l-2} - z_{ln_l-1}\gamma_{n_l-1} + (h_{1,n_l-21} - h_{l,n_l-11})\xi_{l1}^r + \cdots + (h_{l,n_l-2n_l-2} - h_{l,n_l-1n_l-2})\xi_{l_{n_l-2}}^r}{h_{l,n_l-1n_l-1}}$ from above.

Step 5: Calculate $\Phi(\frac{z_{ln_l-1}\gamma_{n_l-1} - z_{ln_l}\gamma_{n_l} + (h_{l,n_l-11} - h_{l,n_l1})\xi_{l1}^r + \cdots + (h_{l,n_l-1n_l-1} - h_{l,n_ln_l-1})\xi_{ln_l-1}^r}{h_{l,n_ln_l}})$.

Step 6: Calculate corresponding $q(\xi_l^r)$ given the realization of ξ_l.

For each auction denoted by l, the simulated likelihood can be written as $\tilde{p}_l(\theta) = \frac{1}{R}\sum_{r=1}^{R} q(\xi_l^r)$. A smoothly simulated maximum likelihood estimator (SSMLE) for θ is proposed as

$$\widehat{\theta}_{SSMLE} = \arg\max_{\theta} \tilde{L}(\mathbf{b}, \mathbf{z}; \theta) = \arg\max_{\theta} \sum_{l=1}^{L} \ln \tilde{p}_l(\theta).$$

As a simulated maximum likelihood estimator, $\widehat{\theta}_{SSMLE}$ has the same asymptotic normal distribution as that of the usual MLE, as the number of auctions $L \to \infty$, the number of simulations $R \to \infty$, and $R/\sqrt{L} \to \infty$. For the purpose of conducting inference, we will use the nonparametric bootstrap method.

3.3.2 The Wald Test for Collusion

The key idea behind the rank-based test in Porter and Zona (1993) is that one can remain agnostic about how the low cartel bidder is selected and the test is designed to detect differences in the order between competitive and cartel bids conditional on the observed data. Under the null hypothesis that the model is correctly specified (and thus no collusion), due to the IIA property, the model parameters can be consistently estimated by any subset of the data using maximum likelihood estimation. Therefore, they construct a likelihood ratio (LR) test statistic by comparing the likelihood of observing the low bidders from all auctions and the likelihood of observing all other rankings of bids from these auctions. Their approach, however, cannot be adopted in our case, because the IIA property is no longer valid when we allow bidders' private costs to be affiliated and we use the joint normal distribution to model the joint distribution of log of bids. Nevertheless, we can still build on the key idea of the rank-based test in Porter and Zona (1993) to detect the discrepancy between the bidding rules for the bidder who submits the lowest bid and the rest of the bidders in a procurement auction. Toward this end, we model the bidding rules for the bidder who submits the lowest bid and the rest of the bidders differently for auction l, as follows:

$$\begin{aligned} log(b_{li}) &= \mathbf{x}_l\beta + \mathbf{z}_{li}\gamma_R + \eta_l + \varepsilon_{li}, \quad for\ i = 1, \cdots, n_l - 1, \\ log(b_{ln_l}) &= \mathbf{x}_l\beta + \mathbf{z}_{ln_l}\gamma_{LOW} + \eta_l + \varepsilon_{ln_l}, \end{aligned} \tag{3.3.2}$$

(3.3.2) is a special case of (3.3.1). Thus we can construct the likelihood of observing a particular ranking in an auction based on (3.3.2) and use the SSMLE procedure proposed in Section 3.3.1 in estimating the model parameters. We construct the following hypothesis to test for collusion:

$$H_0 : \gamma_R = \gamma_{LOW},$$

$$H_1 : \gamma_R \neq \gamma_{LOW}.$$

While Wald, Lagrange multiplier, and LR tests can be implemented to test for H_0 against H_1, we use the Wald test, which is relatively straightforward to be implemented in our case. A by-product of our setting is that the affiliation among bidders can be tested by testing $\rho \geq 0$.

3.4 Collusion in Japanese Procurement Auctions

We test for collusive behavior in Japanese procurement auctions by applying the test proposed in the previous section to the data.

As aforementioned, we split data into two periods, before and after July 31, 2005. In the subset before mid-2005, we apply the test for the "competitive" bidders who were not convicted by the JFTC and the cartel bidders who were convicted by the JFTC, respectively. We separate open and invited auctions, because they entail different formats of governmental regulation of bidder participation. Thus four different sub-datasets are considered. For the period after mid-2005, there is no auction left for the invited auction subset after excluding those auctions with fewer than two cartel bids. Hence we have three different sub-datasets for this period.

Table 3.3: Estimation Results: Same ρ

		Before Mid-2005				After Mid-2005		
		"Competitive" Ranks		Cartel Ranks		"Competitive" Ranks		Cartel Ranks
		Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions
No. of Auctions		72	90	209	235	32	17	19
No. of Bids		225	521	2021	2105	126	110	112
Log Likelihood		-122.0	-717.3	-3779	-3778	-118.6	-142.1	-198.1
Wald Statistic		1.3240	9.5394	12.024	21.392	8.7618	5.6842	1.1365
P Value		0.9324	0.0894	0.0172	0.0003	0.1190	0.3382	0.8884
ρ		0.0018	0.0737	0.0913	0.0914	0.0560	0.0555	0.0990
		(0.0221)	(0.0032)	(0.0051)	(0.0066)	(0.0100)	(0.0037)	(0.0082)
Low Ranks	Util	-2.1042	-0.2564	1.6995	1.0988	-1.8277	1.0187	2.4124
		(1.327)	(1.132)	(0.382)	(0.513)	(1.392)	(0.765)	(2.574)
	Util Sq	1.5171	0.4061	-2.0174	-1.3376	0.6615	-1.0114	-0.9918
		(2.550)	(1.336)	(0.482)	(0.542)	(1.684)	(0.814)	(2.908)
	Capacity	-0.3690	0.4191	1.3553	1.2065	0.3334	0.3321	1.2204
		(0.524)	(0.113)	(0.222)	(0.595)	(0.730)	(0.186)	(1.655)
	Capacity Sq	0.0119	-0.0410	-0.3505	-0.3393	0.0002	-0.0234	-0.3189
		(0.135)	(0.012)	(0.097)	(0.108)	(0.156)	(0.044)	(0.946)
	Noback	0.2049	-0.3638			-0.9004	0.1603	
		(0.484)	(0.215)			(0.971)	(0.200)	
Rest Ranks	Util	-2.5024	0.9103	1.1617	0.4436	0.3660	0.4770	0.6179
		(0.959)	(0.701)	(0.289)	(0.404)	(0.907)	(0.682)	(0.883)
	Util Sq	2.7807	-0.8658	-1.0188	-0.4640	-0.5250	-0.4056	-0.2018
		(1.090)	(0.711)	(0.240)	(0.323)	(0.894)	(0.705)	(0.941)
	Capacity	0.3851	0.2623	1.0633	0.0376	-0.1258	0.6682	0.4863
		(0.481)	(0.082)	(0.126)	(0.098)	(0.370)	(0.150)	(0.446)
	Capacity Sq	-0.1189	-0.0212	-0.2912	-0.0168	0.0610	-0.1092	-0.0812
		(0.117)	(0.007)	(0.041)	(0.029)	(0.088)	(0.027)	(0.088)
	Noback	0.7856	0.0770			0.1100	0.3182	
		(0.362)	(0.132)			(0.453)	(0.191)	

Notes: The estimated results are obtained through the SSMLE, with 100 simulations. The standard errors are in parentheses, which are calculated through 100 nonparametric bootstrap samples. The variable "Capacity" is normalized via dividing by the sample mean, and the variable "Capacity Sq" is constructed by squaring the normalized "Capacity". The correlation coefficient ρ is the same for all bidders.

Table 3.3 shows the results of the SSMLE obtained with 100 simulations, together with the Wald test results, when the correlation of the unobserved idiosyncratic errors is assumed to be the same across all bidders within one procurement auction. We report the results for the two different periods, before mid-2005 and after mid-2005, in the first four columns and the last three columns, respectively. The first two columns correspond to the ranking of "competitive" bids, and the middle two columns correspond to the ranking of cartel bids, both before mid-2005. The last three columns present the results of "competitive" bids for both types of auctions and of cartel bids for the open auctions only, after mid-2005.

Our primary goal is to test for collusion. Thus the results of the Wald test are of main interest. For cartel ranks before mid-2005, the Wald statistic is distributed asymptotically as a χ^2 with four degrees of freedom under the null hypothesis. For the open and invited auctions, the Wald statistics are 12.02 and 21.39, respectively, both being significant at the 5 percent level. Therefore, we can reject the null hypothesis in the case of cartel ranks, which means the existence of the collusion behavior among cartel firms as expected. For "competitive" ranks before mid-2005, the Wald statistics for the open and invited auctions are 1.32 and 9.54, respectively, which are distributed asymptotically as a χ^2 with five degrees of freedom under the null hypothesis. For open auctions, the associated p-value is 0.93, thus we cannot reject the null hypothesis of the equality in bidding behavior between the lowest and other bidders. In contrast, for invited auctions, the p-value is 0.09, indicating that the test statistic is significant at the 10 percent level and that we can reject the null hypothesis of competitive bidding. This suggests that there may be collusion even among the "competitive" firms in the invited auctions. That these "competitive" firms may collude yet were not convicted by the government may be because implementing antitrust law requires hard evidence and that it is not unusual that some collusive firms do not leave any hard evidence of collusion. As a result, we conclude that before mid-2005, there is evidence of collusion regardless of "competitive" or cartel firms in the invited auctions, and there is evidence of collusion among cartel firms in the open auctions.

Next we turn to the period after mid-2005. The Wald statistics for "competitive" ranks in the open and invited auctions are 8.76 and 5.68, with the corresponding p-values 0.12 and 0.34, respectively, and therefore, we cannot reject the null hypothesis of coefficient stability across ranks. For the cartel ranks in the open auctions in the same period, the Wald statistic is 1.14, with a p-value being 0.89, implying that we cannot reject the null hypothesis of competitive behavior. As a result, we find no evidence of collusive behavior for both types of bidders after mid-2005.

We also find evidence from the estimates of ρs that cartel firms and "competitive" firms collude with different information sets. For open auctions and "competitive" bidders, the estimated ρ is 0.002, being insignificant with standard error 0.022. However, the estimated ρ for "competitive" bidders in the invited auctions is 0.074, with standard error around 0.003. This result suggests that ρ is significantly positive, thus the "competitive" bidders are affiliated in the invited auctions. For the cartel bidders in the open and invited auctions, the estimates for ρs are both 0.091, with standard errors 0.005 and 0.007 respectively, both corresponding to a larger affiliation level among the cartel firms, compared to the results of the "competitive" firms. This suggests that the cartel firms may have a different information set and share more information relative to the "competitive" firms. After mid-2005, the affiliation levels for the "competitive" bidders in the open and invited auctions are estimated as significantly positive values, both being 0.056, with standard errors around 0.010 and 0.004 respectively. For the cartel bidders in the open auctions after mid-2005, the estimated ρ is 0.099, with standard error 0.008. This means that after mid-2005, the affiliation level among the cartel bidders is also larger than that among the "competitive" bidders.

Most of the estimates have the expected signs. The utilization rate as a measure of the job backlog affects the subsets before mid-2005 significantly, and the effects on "competitive" and cartel bidders are different. For the "competitive" bids in the open auctions, the rest of the bids first decrease then increase as "Util" increases. In contrast, for the cartel bids in the open auctions, bids first increase then decrease as "Util" increases, regardless of the bid ranking. For the cartel bids in the invited auctions, "Util" only has a significant and inverted-U-shaped effect on the lowest bids, but not the rest of the bids. Taking the subset of cartel bids in the open auctions before mid-2005 as an example, at a utilization rate of 42 percent, the lowest bids are about 26 percent higher than usual; at a full utilization rate of 100 percent with no usable capacity left, bids are about 32 percent lower.

The capacity has a consistent and inverted-U-shaped effect on the log of bids across subsets. Specifically, "Capacity" affects the lowest bids for the "competitive" bids in the invited auctions regardless of the period and the cartel bids in both the open and invited auctions before mid-2005 significantly. "Capacity" also has significant effects on the rest of the bids for the "competitive" bids in the invited auctions regardless of the period and the cartel bids in the open auctions. In the aforementioned specifications, bids are first an increasing function of "Capacity," and then decreasing. For "competitive" bids, the dummy variable "Noback" has a significant and positive effect on the log of the lowest bids in the open auctions before mid-2005. Besides, "Noback" affects the rest "competitive" bids in the invited auctions after mid-2005 significantly.

The results above are obtained based on the assumption of equal affiliation level between the lowest and other bidders. For robustness, we conduct our test assuming that the correlation between the unobserved idiosyncratic errors from the lowest bids and from the rest of the bids is different from that among the unobserved idiosyncratic errors of the rest of the bids. The results are very similar to these reported in Table 3.3, and are not reported here. Moreover, the estimated correlation between the unobserved idiosyncratic errors from the lowest bids and from the rest of the bids has similar magnitude relative to that among the unobserved idiosyncratic errors of the rest of the bids.

Table 3.4: Estimation Results: Porter and Zona (1993)

		Before Mid-2005				After Mid-2005		
		"Competitive" Ranks		Cartel Ranks		"Competitive" Ranks		Cartel Ranks
		Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions
No. of Auctions		72	90	209	235	32	17	19
No. of Bids		225	521	2021	2105	126	110	112
Likelihood Ratio Statistic		4.1906	18.400	15.826	48.903	12.734	12.875	1.2981
P Value		0.5223	0.0025	0.0033	0.0000	0.0260	0.0246	0.8617
Low Ranks	Util	-1.9902	0.0795	0.6679	1.4111	2.2374	-148.69	4.7251
		(2.220)	(1.868)	(1.117)	(1.042)	(3.039)	(569.3)	(3.608)
	Util Sq	-0.0724	-0.5227	-1.8628	-2.6487	-5.2908	148.62	-4.4486
		(2.314)	(2.028)	(1.048)	(0.978)	(3.531)	(569.4)	(3.745)
	Capacity	0.3420	-0.0167	0.3255	-1.2163	0.8133	-0.6173	-1.0486
		(0.777)	(0.231)	(0.452)	(0.344)	(1.073)	(0.996)	(1.211)
	Capacity Sq	-0.3082	0.0118	-0.1154	0.3328	-0.2274	0.4967	0.5178
		(0.388)	(0.100)	(0.217)	(0.173)	(0.568)	(0.558)	(0.623)
	Noback	-0.1938	-1.6090			-0.7878	-1.1201	
		(0.720)	(0.440)			(1.193)	(0.909)	
	Observations	72	90	209	235	32	17	19
	Log Likelihood	-72.25	-128.8	-456.1	-483.0	-29.00	-23.24	-27.84
Rest Ranks	Util	-0.8861	2.4102	-0.9469	-0.1132	1.9017	3.2430	3.8862
		(2.028)	(1.338)	(0.426)	(0.442)	(1.856)	(1.960)	(1.712)
	Util Sq	-0.2446	-2.5986	0.8040	-0.1146	-2.5183	-3.1182	-4.5490
		(2.022)	(1.395)	(0.379)	(0.387)	(1.833)	(2.024)	(1.999)
	Capacity	-1.3914	-0.0148	0.2144	0.4821	-0.2359	0.9228	0.2475
		(0.837)	(0.160)	(0.172)	(0.143)	(0.571)	(0.572)	(0.551)
	Capacity Sq	0.6483	0.0142	-0.0322	-0.1966	0.1302	-0.7883	-0.1119
		(0.409)	(0.065)	(0.085)	(0.068)	(0.321)	(0.584)	(0.251)
	Noback	-1.0478	0.1401			0.5665	0.5956	
		(0.743)	(0.214)			(0.658)	(0.458)	
	Observations	153	431	1812	1870	94	93	93
	Log Likelihood	-68.92	-503.3	-2596	-2528	-65.96	-106.2	-114.0
All Ranks	Util	-1.6670	1.7755	-0.7790	0.0323	1.8423	1.8627	3.9158
		(1.482)	(1.080)	(0.397)	(0.406)	(1.506)	(1.777)	(1.522)
	Util Sq	0.1486	-2.0848	0.5135	-0.4085	-2.8583	-2.2922	-4.4354
		(1.499)	(1.140)	(0.355)	(0.358)	(1.521)	(1.800)	(1.746)
	Capacity	-0.4402	-0.0019	0.2233	0.2213	0.0423	0.1134	0.0328
		(0.554)	(0.130)	(0.161)	(0.131)	(0.493)	(0.288)	(0.496)
	Capacity Sq	0.1335	0.0082	-0.0418	-0.1129	0.0615	0.0759	-0.0225
		(0.269)	(0.054)	(0.079)	(0.063)	(0.271)	(0.183)	(0.233)
	Noback	-0.5750	-0.1384			0.1953	0.0102	
		(0.510)	(0.184)			(0.552)	(0.355)	
	Observations	225	521	2021	2105	126	110	112
	Log Likelihood	-143.3	-641.3	-3060	-3036	-101.3	-135.8	-142.5

Notes: The estimated results are obtained through the MNL as in Porter and Zona (1993). The standard errors are in parentheses.

We also apply the test in Porter and Zona (1993) to our data. The results are shown in Table 3.4.

We obtain quite different conclusions from the Wald test results in Table 3.3. Out of seven subsets, for only two subsets we cannot reject the null hypothesis of no bid rigging. They are the "competitive" bids in the open auctions before mid-2005 and the cartel bids in the open auctions after mid-2005, with the likelihood ratio statistics being 4.19 and 1.30 and the p-values being 0.52 and 0.86, respectively. This means for "competitive" bidders, using the MNL assuming independence of bidders' private costs, they would be deemed as non-collusive before mid-2005 and collusive after, which is implausible and means the test could have both size and power problems. Moreover, most of the estimated coefficients are not significant using the full data. If any, "Capacity" has a significant and inverted-U-shaped effect on the log of bids for the cartel bids in the invited auctions before mid-2005, and "Util" has a significant and inverted-U-shaped effect for the cartel bids in the open auctions after mid-2005. This means that it is important to take into account the affiliation among bidders' private costs when testing for the possible collusive behavior in procurement auctions.

Table 3.5: Estimation Results: Marginal Bids

		Before Mid-2005				After Mid-2005		
		"Competitive" Ranks		Cartel Ranks		"Competitive" Ranks		Cartel Ranks
		Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions	Invited Auctions	Open Auctions
No. of Auctions		69	82	202	231	21	10	15
No. of Bids		210	379	1849	1940	55	35	38
Log Likelihood		-84.45	-458.0	-3282	-3341	-31.73	-25.76	-28.32
Wald Statistic		5.2491	15.167	7.8118	20.639	5.3525	5.3043	0.539
P Value		0.3862	0.0097	0.0987	0.0004	0.3744	0.3799	1.2928
ρ		-0.0039	0.0649	0.1004	0.0928	0.0544	0.0560	0.0991
		(0.0421)	(0.0038)	(0.0057)	(0.0089)	(0.0093)	(0.0033)	(0.0285)
Low Ranks	Util	-2.6138	0.4386	1.6967	1.1339	-1.8400	1.0164	2.2607
		(1.327)	(1.243)	(0.402)	(0.529)	(2.864)	(1.075)	(3.195)
	Util Sq	1.5301	-0.2673	-1.9826	-1.3614	0.6659	-1.0091	-1.0263
		(3.637)	(1.398)	(0.497)	(0.569)	(2.859)	(1.316)	(3.345)
	Capacity	-0.4619	0.5216	1.2763	1.1638	0.3315	0.3313	1.1851
		(0.528)	(0.189)	(0.231)	(0.961)	(0.928)	(0.529)	(1.891)
	Capacity Sq	0.0001	-0.0533	-0.3315	-0.3311	0.0002	-0.0232	-0.3188
		(0.141)	(0.036)	(0.103)	(0.170)	(0.246)	(0.226)	(0.881)
	Noback	0.2951	-0.3414			-0.8998	0.1599	
		(0.445)	(0.239)			(0.803)	(0.264)	
Rest Ranks	Util	-1.2906	1.1311	1.1631	0.5421	0.3669	0.7185	0.6258
		(1.170)	(0.777)	(0.295)	(0.417)	(0.975)	(1.006)	(1.300)
	Util Sq	3.4547	-1.1769	-1.0570	-0.5463	-0.5285	-0.7334	-0.2035
		(1.143)	(0.840)	(0.246)	(0.349)	(0.696)	(1.054)	(1.290)
	Capacity	0.6728	0.2473	1.0361	0.0544	-0.1267	0.3734	0.4848
		(0.573)	(0.109)	(0.127)	(0.117)	(0.252)	(0.460)	(0.761)
	Capacity Sq	-0.0738	-0.0204	-0.2791	-0.0229	0.0619	-0.0220	-0.0826
		(0.117)	(0.012)	(0.041)	(0.031)	(0.067)	(0.198)	(0.200)
	Noback	1.0208	-0.0058			0.1107	0.1719	
		(0.383)	(0.139)			(0.352)	(0.250)	

Notes: The estimated results are obtained through the SSMLE, with 100 simulations. The standard errors are in parentheses, which are calculated through 100 nonparametric bootstrap samples. The variable "Capacity" is normalized via dividing by the sample mean, and the variable "Capacity Sq" is constructed by squaring the normalized "Capacity". The correlation coefficient ρ is the same for all bidders.

It is possible that bidders with relatively higher private costs do not submit "serious" bids in the sense that they do not follow the equilibrium bidding rule. To rule out this possibility, we only keep the bids within 10% range of the lowest bid in each auction and re-conduct our test. The results are reported in Table 3.5, which remain consistent with those reported in Table 3.3 with the full data.[8] We find that for the same subsets as in Table 3.3 before mid-2005 (the "competitive" bids in the invited auctions and the cartel bids in both open and invited auctions), we reject the null hypothesis of competitive behavior, while for the subset of the "competitive" bids in the open auctions and the subsets after mid-2005 we cannot reject the null hypothesis. Note that in this analysis, some of the estimates have greater standard errors than those in Table 3.3, because the sample size is smaller here. It is worth noting that, for the four subsets before mid-2005, trimming the bids outside the range of 10% of the lowest bid in each auction does not exclude most bids, which keeps 93% and 73% "competitive" bids in the open and invited auctions respectively, and 91% and 92% cartel bids in the open and invited auctions respectively. In contrast, after mid-2005, we keep only 43% and 32% "competitive" bids in the open and invited auctions, respectively, and 34% cartel bids in the open auctions. Such a striking difference can be attributed to that for the "competitive" bidders in the invited auctions and the cartel bidders in both open and invited auctions before mid-2005, the lowest bid is too high, or bidders who do not have the lowest bid and may try to collude in auctions submit "phantom" bids to create an illusion of competition; however, they may bid too aggressively compared to when they submitted "competitive" bids according to their moderate or high private costs. This also offers a piece of evidence showing that collusion was widespread during this period.

3.5 Conclusion

In this paper, we extend the rank-based test in Porter and Zona (1993) who assume independence among bidders' private costs in procurement auctions to allow for affiliation among bidders' private costs. From a methodological perspective, with affiliation, estimation of the model becomes computationally intensive due to the multiple integrals in the likelihood function, and we propose a simulation-based method to resolve this issue. Also due to the affiliation, we need to modify the ways in which the hypotheses are constructed and also tested. We apply our testing procedure to the bridge construction procurement auctions in Japan. We find that during the period before mid-2005, not only these firms convicted by the JFTC colluded, the "competitive" firms that were not convicted also likely

8 We also re-conduct the test in Porter and Zona (1993) with the marginal bids. For the subsets of the "competitive" bids in the invited auctions and the cartel bids in the open auctions after mid-2005, the algorithm cannot converge, which may imply potential misspecification problems using the MNL model. For the rest of the subsets, the results are consistent with those in Table 3.4, as we cannot reject the null hypothesis of competitive behavior only for the "competitive" bids in the open auctions before mid-2005.

colluded in the invited auctions. On the other hand, our testing results find no evidence of collusion for all bidders after mid-2005, suggesting that the antitrust and anti-collusion actions by the government in mid-2005 were effective.

As discussed in Porter and Zona (1993), their test would no longer work should an antitrust authority publicly announce the adoption of their testing procedure, as all cartel firms could scale their competitive bids up by the same percentage. This is because their test statistic is based on the ratio between the likelihood of observing the low bidders from all auctions and the likelihood of observing all other rankings of bids from these auctions. While our test is based on a similar idea to Porter and Zona (1993) which is to detect differences in the ordering of higher bids, our testing procedure would still work when all cartel firms scale their competitive bids up by the same percentage, as a result of our hypotheses set-up, which is different from Porter and Zona (1993). On the other hand, as discussed in Porter and Zona (1993), their test would be less powerful with inclusive cartels as cost asymmetries between the cartel and competitive firms can give rise to the difference in the bidding behavior from the two groups of firms. Our test shares the same caveat in this case. We conclude our paper by quoting Porter and Zona (1993) "Attention would then have to focus on the determination of the identity of the lowest bidder or on rates of return."

References

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

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

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

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

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

Aradillas-Lopez, A. (2010). Semiparametric Estimation of a Simultaneous Game with Incomplete Information. *Journal of Econometrics*, *157*(2), 409–431.

Aradillas-López, A., Gandhi, A., and Quint, D. (2013). Identification and Inference in Ascending Auctions with Correlated Private Values. *Econometrica*, *81*(2), 489–534.

Aryal, G., and Gabrielli, M. F. (2013). Testing for Collusion in Asymmetric First-Price Auctions. *International Journal of Industrial Organization*, *31*(1), 26–35.

Aryal, G., and Gabrielli, M. F. (2020). An Empirical Analysis of Competitive Nonlinear Pricing. *International Journal of Industrial Organization*, *68*.

Aryal, G., Grundl, S., Kim, D. H., and Zhu, Y. (2018). Empirical Relevance of Ambiguity in First-Price Auctions. *Journal of Econometrics*, *204*(2), 189–206.

Aryal, G., and Zincenko, F. (2019). Empirical Framework for Cournot Oligopoly with Private Information. *Working Paper*, University of Virginia and University of Pittsburgh.

Asker, J. (2010). A Study of the Internal Organization of a Bidding Cartel. *American Economic Review*, *100*(3), 724–62.

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., and Haile, P. A. (2002). Identification of Standard Auction Models. *Econometrica*, *70*(6), 2107–2140.

Athey, S., and Haile, P. A. (2007). Nonparametric Approaches to Auctions. In *Handbook of Econometrics*, (pp. 3847–3965). Elsevier B.V.

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.

Bajari, P., Hong, H., Krainer, J., and Nekipelov, D. (2010). Estimating Static Models of Strategic Interactions. *Journal of Business and Economic Statistics*, *28*(4), 469–482.

Bajari, P., Hong, H., and Ryan, S. P. (2010). Identification and Estimation of a Discrete Game of Complete Information. *Econometrica*, *78*(5), 1529–1568.

Bajari, P., and Hortaçsu, A. (2005). Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data. *Journal of Political Economy*, *113*(4), 703–741.

Bajari, P., and Ye, L. (2003). Deciding Between Competition and Collusion. *Review of Economics and Statistics*, *85*(4), 971–989.

Baldwin, L. H., Marshall, R. C., and Richard, J. F. (1997). Bidder Collusion at Forest Service Timber Sales. *Journal of Political Economy*, *105*(4), 657–699.

Barelli, P., and Duggan, J. (2015). Purification of Bayes Nash Equilibrium with Correlated Types and Interdependent Payoffs. *Games and Economic Behavior*, *94*, 1–14.

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

Bartle, R. G. (1995). *The Elements of Integration and Lebesgue Measure*. John Wiley & Sons, Inc.

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

Beresteanu, A., Molchanov, I., and Molinari, F. (2011). Sharp Identification Regions in Models with Convex Moment Predictions. *Econometrica*, *79*(6), 1785–1821.

Bergstrom, T., Blume, L., and Varian, H. (1986). On the Private Provision of Public Goods. *Journal of Public Economics*, *29*, 25–49.

Berry, S. T., and Haile, P. A. (2014). Identification in Differentiated Products Markets Using Market Level Data. *Econometrica*, *82*(5), 1749–1797.

Bhattacharya, V., Roberts, J. W., and Sweeting, A. (2014). Regulating Bidder Participation in Auctions. *RAND Journal of Economics*, *45*(4), 675–704.

Börsch-Supan, A., and Hajivassiliou, V. A. (1993). Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models. *Journal of Econometrics*, *58*(3), 347–68.

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

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

Campo, S., Perrigne, I., and Vuong, Q. (2003). Asymmetry in First-Price Auctions with Affiliated Private Values. *Journal of Applied Econometrics*, *18*(2), 179–207.

Chassang, S., Kawai, K., Nakabayashi, J., and Ortner, J. M. (2019). Data Driven Regulation: Theory and Application to Missing Bids. *Working Paper*, New York University, University of California, Berkeley, Kindai University, and Boston University.

Chassang, S., and Ortner, J. (2019). Collusion in Auctions with Constrained Bids: Theory and Evidence from Public Procurement. *Journal of Political Economy*, *127*(5), 2269–2300.

Chen, X., Gentry, M. L., Li, T., and Lu, J. (2019). Identification and Estimation in First-Price Auctions with Risk-Averse Bidders and Selective Entry. *Working Paper*, Yale University, London School of Economics and Political Science, Vanderbilt University, and National University of Singapore.

Chernozhukov, V., Hong, H., and Tamer, E. (2007). Estimation and Confidence Regions for Parameter Sets in Econometric Models. *Econometrica*, *75*(5), 1243–1284.

Chesher, A., and Rosen, A. M. (2013). What Do Instrumental Variable Models Deliver with Discrete Dependent Variables? *American Economic Review: Papers & Proceedings*, *103*(3), 557–562.

Chesher, A., and Rosen, A. M. (2017). Generalized Instrumental Variable Models. *Econometrica*, *85*(3), 959–989.

Ciliberto, F., and Tamer, E. (2009). Market Structure and Multiple Equilibria in Airline Markets. *Econometrica*, *77*(6), 1791–1828.

Compiani, G., Haile, P. A., and Sant'Anna, M. (2019). Common Values, Unobserved Heterogeneity, and Endogenous Entry in U.S. Offshore Oil Lease Auctions. *Journal of Political Economy*, forthcoming.

Conley, T. G., and Decarolis, F. (2016). Detecting Bidders Groups in Collusive Auctions. *American Economic Journal: Microeconomics*, *8*(2), 1–38.

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

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

de Castro, L. I., and Paarsch, H. J. (2010). Testing Affiliation in Private-Values Models of First-Price Auctions using Grid Distributions. *Annals of Applied Statistics*, *4*(4), 2073–2098.

de Paula, Á., and Tang, X. (2012). Inference of Signs of Interaction Effects in Simultaneous Games with Incomplete Information. *Econometrica*, *80*(1), 143–172.

De Silva, D. G., Dunne, T., Kankanamge, A., and Kosmopoulou, G. (2008). The Impact of Public Information on Bidding in Highway Procurement Auctions. *European Economic Review*, *52*(1), 150–181.

De Silva, D. G., Li, T., and Zhao, J. (2020). Efficacy of a Bidder Training Program in Procurement Auctions with Selection and Entry. *Working Paper*, Lancaster University, Vanderbilt University, and Vanderbilt University.

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

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

D'Haultfœuille, X., and Février, P. (2020). The Provision of Wage Incentives: A Structural Estimation Using Contracts Variation. *Quantitative Economics*, *11*, 349–397.

Diamond, P. A. (1982). Aggregate Demand Management in Search Equilibrium. *Journal of Political Economy*, *90*(5), 881–894.

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.

Fey, M. (2008). Rent-Seeking Contests with Incomplete Information. *Public Choice*, *135*, 225–236.

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

Galichon, A., and Henry, M. (2011). Set Identification in Models with Multiple Equilibria. *Review of Economic Studies*, *78*(4), 1264–1298.

Gentry, M. L., Hubbard, T. P., Nekipelov, D., and Paarsch, H. J. (2018). Structural Econometrics of Auctions: A Review. *Foundations and Trends in Econometrics*, *9*(2-4), 79–302.

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.

Geweke, J. (1991). Efficient Simulation from the Multivariate Normal and Student-t Distributions Subject to

Linear Constraints and the Evaluation of Constraint Probabilities. *Computing Science and Statistics: the 23rd Symposium on the Interface*.

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.

Grundl, S., and Zhu, Y. (2019). Identification and Estimation of Risk Aversion in Frst-Price Auctions with Unobserved Auction Heterogeneity. *Journal of Econometrics*, *210*(2), 363–378.

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

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

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

Haile, P., Hong, H., and Shum, M. (2003). Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions. *NBER Working Paper Series*.

Haile, P. A., and Tamer, E. (2003). Inference with an Incomplete Model of English Auctions. *Journal of Political Economy*, *111*(1), 1–51.

Hajivassiliou, V., McFadden, D., and Ruud, P. (1996). Simulation of Multivariate Normal Rectangle Probabilities and Their Derivatives: Theoretical and Computational Results. *Journal of Econometrics*, *72*(1-2), 85–134.

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

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

He, W., and Sun, Y. (2019). Pure-Strategy Equilibria in Bayesian Games. *Journal of Economic Theory*, *180*, 11–49.

Hendricks, K., and Porter, R. H. (2007). An Empirical Perspective on Auctions. In *Handbook of Industrial Organization*, (pp. 2073–2143). Elsevier B.V.

Hickman, B. R., Hubbard, T. P., and Sağlam, Y. (2012). Structural Econometric Methods in Auctions: A Guide to the Literature. *Journal of Econometric Methods*, *1*(1), 67–106.

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

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

Hong, H., and Shum, M. (2003). Econometric Models of Asymmetric Ascending Auctions. *Journal of Econometrics*, *112*(2), 327–358.

Hortaçsu, A., and McAdams, D. (2010). Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market. *Journal of Political Economy*, *118*(5), 833–865.

Hu, Y., McAdams, D., and Shum, M. (2013). Identification of First-Price Auctions with Non-Separable Unobserved Heterogeneity. *Journal of Econometrics*, *174*(2), 186–193.

Hubbard, T. P., Li, T., and Paarsch, H. J. (2012). Semiparametric Estimation in Models of First-Price, Sealed-Bid Auctions with Affiliation. *Journal of Econometrics*, *168*(1), 4–16.

Ishii, R. (2008). Collusion in Repeated Procurement Auction: A Study of a Paving Market in Japan. *Working Paper*, Shiga University.

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

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

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

Jun, S. J., Pinkse, J., and Wan, Y. (2010). A Consistent Nonparametric Test of Affiliation in Auction Models. *Journal of Econometrics*, *159*(1), 46–54.

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 Nakabayashi, J. (2014). Detecting Large-Scale Collusion in Procurement Auctions. *Working Paper*, New York University and Kindai University.

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

Keane, M. P. (1994). A Computationally Practical Simulation Estimator for Panel Data. *Econometrica*, *62*(1), 95–116.

Khan, M. A., and Zhang, Y. (2014). On the Existence of Pure-Strategy Equilibria in Games with Private Information: A Complete Characterization. *Journal of Mathematical Economics*, *50*, 197–202.

Kline, B. (2015). Identification of Complete Information Games. *Journal of Econometrics*, *189*(1), 117–131.

Krasnokutskaya, E. (2011). Identification and Estimation of Auction Models with Unobserved Heterogeneity. *Review of Economic Studies*, *78*(1), 293–327.

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.

Levin, D., and Smith, J. (1994). Equilibrium in Auctions with Entry. *American Economic Review*, *84*(3), 585–599.

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.

Lewbel, A., and Tang, X. (2015). Identification and Estimation of Games with Incomplete Information Using Excluded Regressors. *Journal of Econometrics*, *189*(1), 229–244.

Li, T., Perrigne, I., and Vuong, Q. (2000). Conditionally Independent Private Information in OCS Wildcat Auctions. *Journal of Econometrics*, *98*(1), 129–61.

Li, T., Perrigne, I., and Vuong, Q. (2002). Structural Estimation of the Affiliated Private Value Auction Model. *RAND Journal of Economics*, *33*(2), 171–93.

Li, T., Perrigne, I., and Vuong, Q. (2003). Semiparametric Estimation of the Optimal Reserve Price in First-Price Auctions. *Journal of Business and Economic Statistics*, *21*(1), 53–64.

Li, T., and Zhang, B. (2010). Testing for Affiliation in First-Price Auctions Using Entry Behavior. *International Economic Review*, *51*(3), 837–50.

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.

Li, T., and Zheng, X. (2012). Information Acquisition and/or Bid Preparation: A Structural Analysis of Entry and Bidding in Timber Sale Auctions. *Journal of Econometrics*, *168*(1), 29–64.

Li, T., and Zheng, X. (2020). Econometric Analyses of Auctions : A Selective Review. *Handbook of Research Methods and Applications in Empirical Microeconomics*, forthcoming.

Li, T., and Zhu, Y. (2020). Identification of First-Price Auctions with Endogenous Entry and Possibly Biased Beliefs. *Working Paper*, Vanderbilt University and Bank of Canada.

List, J. A., Millimet, D. L., and Price, M. K. (2007). Inferring Treatment Status When Treatment Assignment is Unknown : Detecting Collusion in Timber Auctions. *Working Paper*, University of Chicago and NBER, Southern Methodist University, and University of Nevada-Reno.

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

Manski, C. F. (2003). *Partial Identification of Probability Distributions*. Springer-Verlag New York, Inc.

Manski, C. F., and Tamer, E. (2002). Inference on Regressions with Interval Data on a Regressor or Outcome. *Econometrica*, *70*(2), 519–546.

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

Marmer, V., Shneyerov, A. A., and Kaplan, U. (2017). Identifying Collusion in English Auctions. *Working Paper*, University of British Columbia, University of British Columbia, and Concordia University.

Marshall, R. C., and Marx, L. M. (2012). *The Economics of Collusion: Cartels and Bidding Rings*. The MIT Press.

Marshall, R. C., Richard, J.-F., and Shen, C. (2019). Bidder Collusion: Accounting for All Feasible Bidders. *Working Paper*, Penn State University, University of Pittsburgh, and East China Normal University.

Mason, R., and Valentinyi, A. (2010). The Existence and Uniqueness of Monotone Pure Strategy Equilibrium in Bayesian Games. *Working Paper*, University of Southampton and CEPR.

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

McAdams, D. (2003). Isotone Equilibrium in Games of Incomplete Information. *Econometrica*, *71*(4), 1191–1214.

McAfee, R. P., and McMillan, J. (1987). Auctions with Entry. *Economics Letters*, *23*(4), 343–347.

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

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

Milgrom, P. R., and Weber, R. J. (1982). A Theory of Auctions and Competitive Bidding. *Econometrica*, *50*(5), 1089–1122.

Milgrom, P. R., and Weber, R. J. (1985). Distributional Strategies for Games with Incomplete Information. *Mathematics of Operations Research*, *10*(4), 619–632.

Molinari, F. (2008). Partial Identification of Probability Distributions with Misclassified Data. *Journal of Econometrics*, *144*(1), 81–117.

Molinari, F. (2020). Microeconometrics with Partial Identification. *Working Paper*, Cornell University.

Paarsch, H. J. (1992). Deciding Between the Common and Private Value Paradigms in Empirical Models of

Auctions. *Journal of Econometrics*, *51*(1-2), 191–215.

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

Pesendorfer, M. (2000). A Study of Collusion in First-Price Auctions. *Review of Economic Studies*, *67*(3), 381–411.

Polydoro, A. (2011). A Note on Uniqueness of Bayesian Nash Equilibrium for Supermodular Incomplete Information Games. *Working Paper*, University of Rochester.

Porter, R. H. (2005). Detecting Collusion. *Review of Industrial Organization*, *26*(2), 147–167.

Porter, R. H., and Zona, J. D. (1993). Detection of Bid Rigging in Procurement Auctions. *Journal of Political Economy*, *101*(3), 518–538.

Porter, R. H., and Zona, J. D. (1999). Ohio School Milk Markets: An Analysis of Bidding. *RAND Journal of Economics*, *30*(2), 263–288.

Price, M. K. (2008). Using the Spatial Distribution of Didders to Detect Collusion in the Marketplace: Evidence from Timber Auctions. *Journal of Regional Science*, *48*(2), 399–417.

Raith, M. (1996). A General Model of Information Sharing in Oligopoly. *Journal of Economic Theory*, *71*(1), 260–288.

Reny, P. J. (2011). On the Existence of Monotone Pure-Strategy Equilibria in Bayesian Games. *Econometrica*, *79*(2), 499–553.

Roberts, J. W. (2013). Unobserved Heterogeneity and Reserve Prices in Auctions. *RAND Journal of Economics*, *44*(4), 712–732.

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

Rosa, B. V. (2019). Resident Bid Preference, Affiliation, and Procurement Competition: Evidence from New Mexico. *Journal of Industrial Economics*, *67*(2), 161–208.

Ryvkin, D. (2010). Contests with Private Costs: Beyond Two Players. *European Journal of Political Economy*, *26*(4), 558–567.

Samuelson, W. F. (1985). Competitive Bidding with Entry Costs. *Economics Letters*, *17*(1-2), 53–57.

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.

Schurter, K. (2020). Identification and Inference in First-Price Auctions with Collusion. *Working Paper*, Pennsylvania State University.

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

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

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

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

Sweeting, A. (2009). The Strategic Timing Incentives of Commercial Radio Stations: An Empirical Analysis Using Multiple Equilibria. *RAND Journal of Economics*, *40*(4), 710–742.

Tamer, E. (2003). Incomplete Simultaneous Discrete Response Model with Multiple Equilibria. *Review of Economic Studies*, *70*(1), 147–165.

Tamer, E. (2010). Partial Identification in Econometrics. *Annual Review of Economics*, *2*, 167–195.

Tang, X. (2010). Estimating Simultaneous Games with Incomplete Information under Median Restrictions. *Economics Letters*, *108*(3), 273–276.

Tang, X. (2011). Bounds on Revenue Distributions in Counterfactual Auctions with Reserve Prices. *RAND Journal of Economics*, *42*(1), 175–203.

The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel (2020). Scientific Background on the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2020: Improvements to Auction Theory and Inventions of New Auction Formats. Stockholm, Sweden.

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

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

Tullock, G. (1980). Efficient Rent Seeking. In *Toward a Theory of the Rent-Seeking Society*, (pp. 97–112). Texas A & M Univ. Press.

Van Zandt, T., and Vives, X. (2007). Monotone Equilibria in Bayesian Games of Strategic Complementarities. *Journal of Economic Theory*, *134*(1), 339–360.

Vives, X. (1990). Nash Equilibrium with Strategic Complementarities. *Journal of Mathematical Economics*, *19*(3), 305–321.

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

Wan, Y., and Xu, H. (2014). Semiparametric Identification of Binary Decision Games of Incomplete Information with Correlated Private Signals. *Journal of Econometrics*, *182*(2), 235–246.

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.

Zincenko, F. (2018). Nonparametric Estimation of First-Price Auctions with Risk-Averse Bidders. *Journal of Econometrics*, *205*(2), 303–305.