The Dissertation Committee for Kyle James Kretschman
certifies that this is the approved version of the following dissertation:

Designing Structural Election Models for New Policy Analysis

Committee:

Jason Abrevaya, Supervisor

Richard Dusansky, Supervisor

Kenneth Hendricks

Eugenio Miravete

Brian Roberts

Designing Structural Election Models for New Policy Analysis

by

Kyle James Kretschman, B.S.; M.S.Eco.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2011

Acknowledgments

I thank my supervisors Jason Abrevaya and Richard Dusansky for all of their advice and encouragement. I am also grateful to Kenneth Hendricks, Eugenio Miravete, and Brian Roberts for their comments and suggestions.

Designing Structural Election Models for New Policy Analysis

Publication No. ____________________

Kyle James Kretschman, Ph.D.
The University of Texas at Austin, 2011

Supervisors: Jason Abrevaya
Richard Dusansky

This dissertation focuses on designing new structural election models and applying modern estimation techniques to quantify policy reform questions. All three chapters use models that are based on individual decision-making and estimate the parameters using a novel data set of U.S. House of Representative elections. These models provide new opportunities to analyze and quantify election policy reforms.

The first chapter utilizes a unique compilation of primary election expenditures to see if general election voters value the primary nomination signal. While producing new results on the relationships between primary elections and general elections and between candidate characteristics and vote shares, this model allows me to show that campaign finance reform can have an unintended consequence. A limit on expenditures would have little effect on the

competitiveness of elections and substantially decrease voter turnout in the U.S. House elections. In contrast, it is shown that a mandatory public funding policy is predicted to increase competitiveness and increase voter turnout.

The second chapter examines why unopposed candidates spend massive amounts on their campaign. The postulated answer is that U.S. House of Representative candidates are creating a barrier to entry to discourage candidates from opposing them in the next election. This barrier reduces competition in the election and limits the voters choices. An unbalanced panel of congressional districts is used to quantify how an incumbents expenditure in previous elections impacts the probability of running unopposed in a later election.

The third chapter estimates the value of a congressional seat based on the observed campaign expenditures. Campaign expenditures are modeled as bids in an asymmetric all-pay auction. The model produces predictions on how much a candidate should spend based on the partisanship leaning of each district. The predictions and observed expenditures are then used to estimate the value of a congressional seat. Along with analyzing how expenditures would change with new campaign finance reforms, this model has the capability of quantifying the effect of redistricting. After 2010 Census results become available, the majority of states will redraw their congressional districts changing the distribution of partisan votes. This model can be used to quantify the effect that the change in voter distribution has on campaign expenditures.

Table of Contents

Acknowledgments **iv**

Abstract **v**

List of Tables **x**

List of Figures **xi**

Chapter 1. Do Primary Elections Matter? Empirical Implications for Campaign Finance Reform Analysis **1**

1.1 Introduction 1
1.2 Two Stage Discrete Choice Model 7
1.2.1 First Stage: Primary Elections 8
1.2.1.1 A Political Party's Discrete Choice 8
1.2.1.2 Endogeneity of Candidate Expenditures in the Primary Election 9
1.2.1.3 Probability of Nomination 10
1.2.2 Second Stage: General Elections 11
1.2.2.1 Utility of a Registered Voter 11
1.2.2.2 Correlation of Unobserved Valuations 13
1.2.2.3 Predicted Vote Share Function 14
1.2.2.4 Endogeneity of Candidate Expenditures in the General Election 16
1.3 Estimation 17
1.3.1 First Stage Estimation 17
1.3.1.1 Control Function for Primary Expenditures . . 17
1.3.1.2 Primary Parameter Estimates 18
1.3.2 Second Stage Estimation 18
1.3.3 Standard Errors 19

1.4 Data 21
1.4.1 Overview 21
1.4.2 Control Variables 23
1.4.3 Instruments 27
1.5 Empirical Results 29
1.5.1 First Stage Results 29
1.5.1.1 Estimated Coefficients 29
1.5.1.2 Predicted $\hat{\psi}$ Values 32
1.5.2 Second Stage Results 33
1.5.2.1 Evidence Against Valuing the Primary Election Information and Selection Bias 34
1.5.3 Marginal Effects of Control Variables 37
1.5.3.1 Expenditure Elasticities 38
1.6 Campaign Finance Reform Analysis 40
1.6.1 Expenditure Cap 41
1.6.2 Public Funding 43
1.7 Conclusion 46

Chapter 2. Barriers to Entry in U.S. Congressional Elections 48
2.1 Introduction 48
2.2 Challenger Entry Decision Model 51
2.2.1 Theoretical Model 52
2.2.2 Single Challenger Pure Nash Equilibria 53
2.3 Econometric Model and Estimation Method 58
2.3.1 Estimation Equation 59
2.3.2 Estimation Method 60
2.4 Data 61
2.4.1 Control Variables 63
2.5 Results 65
2.5.1 Average Partial Effects 68
2.6 Conclusion 71

Chapter 3. What is a U.S. House Seat Worth? Estimating the Value Using Auction Theory **72**
3.1 Introduction 72
3.2 Finance Augmented Median Voter Model (FAMVM) 75
3.2.1 Campaign Expenditure Equilibria 77
3.2.2 Electoral Outcome Probabilities 81
3.2.2.1 Prediction Intuition 82
3.3 Testable Predictions and Estimation 83
3.4 Data 84
3.5 Estimation Results 88
3.6 Conclusion 90

Appendix **93**

Appendix 1. **94**
1.1 Reduced Form Regression on Expenditures 94

Bibliography **96**

List of Tables

1.1 Summary Statistics 26
1.2 First Stage Parameters 31
1.3 Second Stage Parameters 35
1.4 Average Elasticities: Specification 4 39

2.1 Conditional Means of the Incumbent Expenditures ($) 64
2.2 Summary Statistics 64
2.3 Estimated Coefficients 66
2.4 Average Partial Effects 69

3.1 Notation Summary 77
3.2 Model Predictions in Region *ii* 85
3.3 Summary Statistics 88
3.4 Estimated *v* Values 89

1.1 Reduced Form Regressions 95

List of Figures

1.1 Histogram of Predicted ψ Values 33

1.2 Predicted Average Winner's Margin of Victory Under Different Expenditure Caps 41

1.3 Predicted Average Change in Number of Voters Under Different Expenditure Caps 43

1.4 Predicted Average Winner's Margin of Victory Under Different Public Funding Levels 44

1.5 Predicted Average Change in Number of Voters Under Different Public Funding Levels 45

2.1 State Space 58

3.1 The District State Space 78

3.2 Strategy CDFs 81

3.3 Changing Variables 83

3.4 Congressional Districts in the (b,μ) State Space 91

3.5 Valuations of Congressional Districts 92

Chapter 1

Do Primary Elections Matter? Empirical Implications for Campaign Finance Reform Analysis

1.1 Introduction

How do campaign expenditures influence votes? This is a question that has many factors and very important implications when considering campaign finance reform policies. Primary elections are an integral part of the election process in the United States. The primary elections whittle down all the potential candidates to only the most electable and, hopefully, qualified candidates. The primary election nomination process provides information to the general election voters that can be accounted for when determining how campaign expenditures turn into general election votes.

The majority of the empirical election literature has ignored primary elections when estimating functions that predict voting outcomes. For a long time, the empirical literature focused on reduced form estimates to quantify the impact of candidate characteristics such as expenditures and incumbency on the general election vote share.[1] A selection of these papers are Jacobson

[1] For clarification, the terms "percentage of the vote" and "vote share" have been used

(1978) & (1990), Green & Krasno (1988), Levitt (1994), Gerber (1998) and Erikson & Palfrey (2000).[2] This work has focused on the general election results. They have not utilized any information on primary election expenditures or modeled a full election cycle. The prior literature that models primary and general elections together has been theoretical work that focuses on the individual candidates choosing their campaign platform. For recent work, see Owen & Grofman (2006), Callander & Wilkie (2007), Adams & Merrill (2008) and Agranov (2010). These models are generally based on candidates picking optimal platforms based on the median voter. So far, these models have not produced estimable equations that link the theory to observed data since many elements of the models are econometrically unobservable.

Recently, the empirical election literature has used discrete choice methods to provide empirical results that are based on individual optimization; see Rekkas (2007) and Kretschman & Mastronardi (2010). This work is based on adapting the empirical discrete choice methods to model elections as the result of utility maximizing behavior from voters. Individual voting decisions are not observed but the models use aggregation methods to directly link the theoretical predictions to observed elections results. This type of model estimates the voter preference parameters instead of estimating reduced form

interchangeably in the past. However, in this chapter, vote share is defined as the number of votes a candidate receives divided by the number of registered voters. This is not equal to the percentage of vote, which equals the number of votes received divided by the total number of votes cast. The difference arises because this model allows for a registered voter to choose not to vote. The choice to register is beyond the scope of this chapter.

[2] See the introduction of Gerber (2004) for a thorough summary of the reduced form literature.

correlations. Both papers use the seminal work of Berry (1994) and Berry, Levinsohn & Pakes (1995) to estimate the impact of candidate expenditures and other characteristics on vote shares.

The purpose of this chapter is to develop an election cycle model that uses primary election data to determine if general election voters value the information from the primary election and provide new insights on the effects of campaign policy reforms. In almost all U.S. House elections, the candidates in the general election are explicitly selected for the election by primary elections that occur earlier in the year. Therefore, the population of general election candidates is a selected sample. If general election voters incorporate the information from the primary nomination process when deciding how to vote in the general election, the estimates of the impact of campaign expenditures and incumbency would be biased. The voters would be basing their voting decision partially on the information about the candidate that was observed in the primaries, but the estimates would be attributing the decision as a reaction to expenditures, incumbency, and any other general election candidate characteristic.

This chapter develops a model to deal with this issue. An election cycle is modeled in two stages. In the first stage, each "political party"[3] chooses to

[3] "Political party" is included in quotes because the political parties do not actually choose the candidates to nominate. Each candidate is nominated through a primary process that normally involves individual voters. However, this chapter will abstract away from the primary process and therefore "political party" is used as shorthand for "the primary process by which each political party nominates a candidate."

nominate their most preferred candidate for the general election. Then, in the second stage, registered voters vote for the candidate that maximizes their own utility, or choose not to vote if neither candidate is preferred to abstaining. The two stages are linked by the fact that general election voters observe the number of opponents each candidate defeated in his primary election and that the valuation of unobserved candidate characteristics is correlated in each stage.[4] The correlation of the two valuations is a parameter estimated in the model that tests if, on average, general election voters value the primary election unobserved characteristics.

This model also allows for new insights into the effects of potential campaign finance reform policies. Since the model is based on voters making optimal decisions, it can be used to predict the effects of policies that change the observed candidate characteristics. Specifically, the model and estimation results will be used to predict the effects of a policy that places a strict limit on campaign expenditures and a policy that provides equal public funding to all candidates.

This chapter contributes to the election literature in many ways. Similar to Rekkas (2007) and Kretschman & Mastronardi (2010), instead of postulating reduced form relationships between vote shares and candidate characteristics, the estimable equations in this chapter are derived from individual agents maximizing an objective function. Therefore, the estimates from this

[4] In some election literature such as Adams & Merrill (2008), these unobserved characteristics have been called "valence" factors.

model are preference parameters for the decision making agents rather than correlation coefficients. This chapter advances beyond those two papers by developing a method that shows a direct link between a theory of how candidates are selected in the primary and how this affects the general election vote shares. It is the first research to use empirical discrete choice methods to address how candidates are selected for the general election.

This method is an advancement beyond the two stage theoretical work because it provides a theoretical model that can directly be estimated with the available data. The only reliable and comparable outcome in the primary elections is who is chosen as the winner of the primary. There exist a large number of primaries that do not have reported vote results because a single candidate ran unopposed. Some states does not require a primary election to be held in this case, so no vote total is reported in the data for these elections. For primaries that do report election vote results, the participation rates are incredibly low in primary elections and information on the size of the electorate for each primary is either unavailable or unreliable. Therefore, this model is designed to use the basic information of who wins the primary election and the opponents in the primary since more detailed information is unavailable. Instead of postulating how candidates behave, the main assumption is that the "political parties" pick the candidate whom they prefer most.[5] While this method does abstract from explaining why the particular candidate wins a

[5] An alternative assumption would be that the "political parties" are strategic and pick the candidate with the highest chance of winning the general election. This model could be adapted for that assumption in future work.

primary, the method uses the available information to identify the influences on primary elections and incorporate the primary election information when estimating general election results.

The model is estimated using 2004 and 2006 U.S. House of Representative election data. The estimated primary election results show that primary elections are almost completely determined by campaign expenditures. Therefore, the main contribution of this work is to show that, on average, primary elections are uninformative to general election voters about the candidates beyond observing the amount of campaign expenditures. The main reason for this result is that in two thirds of the primary elections in the sample, a candidate runs unopposed or greatly outspends his next closest competitor. There is little additional information to infer about the candidate in either of these situations. The general election results confirm this result by showing that general elections voters do not incorporate the information from the primary elections when deciding for whom to vote.

The estimated results are used to predict the effect of two different campaign finance reform policies on the average competitiveness of the elections, as measured by the average difference in vote share by the winner and loser, and on voter turnout. A strict expenditure cap is predicted to have very little effect on the average competitiveness and to reduce the voter turnout in these elections. These effects are generally undesirable and provide an argument against enacting this sort of policy. Alternatively, a mandatory public funding policy is predicted to increase the competitiveness of these elections and vastly increase voter turnout. These are generally desirable properties. Further research will determine if this type of policy is socially desirable.

The organization for the remainder of this chapter is as follows. Section 1.2 presents the two stage discrete choice model of an election cycle. Section 1.3 describes the multistep estimation method for recovering the preference parameters of the model while controlling for the endogeneity of expenditures. Section 1.4 explains the unique data set gathered for this analysis. Section 1.5 provides and interprets the empirical results. Section 1.6 shows the effects of two potential campaign finance reforms: a strict expenditure cap and a public funding policy. Section 1.7 concludes.

1.2 Two Stage Discrete Choice Model

This chapter models a U.S House of Representative election cycle as a two stage discrete choice model. The first stage represents the primary election and the decision makers are the "political parties." As mentioned in the introduction, the leaders of the political parties do not actually select the candidates. The candidates are nominated by the primary process. However, this model will abstract away from how the primary process operates and use "political party" as shorthand for the "the primary process that leads each political party to nominate a candidate for the general election."[6] In this stage, each political party nominates a candidate to run in the general election.

[6]The quotations will be omitted from here on.

The party has to choose from a discrete set of candidates. It is assumed that the political party nominates the candidate that they prefer most.

The second stage represents the general election. In this stage, the decision makers are registered voters. Each voter in the congressional district makes the discrete choice between the candidates in the general election. Even though the voting decision is not observed at the individual level, a predicted vote share share function for each candidate is derived from a representative agent model similar to Berry, Levinsohn & Pakes (1995).

The two stages of this model are connected in two ways. The the number of primary election opponents is included as a characteristic of the general election candidates. Also, the valuation of the candidate's unobserved characteristics in each stage is allowed to be correlated. The primary parties' valuation of the unobserved characteristics provides information about the primary candidate. This valuation will not be the same as the valuation by a registered voter in a general election. However, the unobservable valuations could be correlated. The correlation links the two stages econometrically and tests to see if the general election voters value the primary nomination unobserved characteristics.

1.2.1 First Stage: Primary Elections

1.2.1.1 A Political Party's Discrete Choice

In every congressional district, for each primary election, it is assumed that the political party, p, chooses the candidate that they prefer the most.

The political party's decision is:

$$\max_{c \in \{1,\ldots,C_p,NoC\}} U_c = \check{\beta}\check{X}_c + \psi_c \tag{1.2.1}$$

The subscript c denotes the different candidates in the primary election and C_p is the total number of candidates in each election which varies by party and congressional district. The political party also has the outside option of not nominating any candidate, denoted by NoC.[7] $\check{X}_c$ is the vector of observable primary candidate characteristics at the time of the primary election and ψ_c is the party valuation of the unobservable candidate characteristics. It is assumed that utility is a linear function of the observed candidate characteristics.

1.2.1.2 Endogeneity of Candidate Expenditures in the Primary Election

One of the observed characteristics is the candidate's primary election expenditures. It is important to address the endogeneity of the primary expenditures at this point since this stage does not produce a linear model. The candidate's unobserved characteristics, ψ_c, not only affect the chance of being nominated in the primary election, but also affect the amount of contributions a candidate raises. A candidate with greater contributions can expend more on the congressional race. Therefore, the unobserved characteristics are expected

[7]This is observed in the data when a primary election does not occur for a one of the parties in the district.

to be correlated with the candidate's expenditures.

To account for the endogeneity, decompose the unobserved party valuation into the part that is correlated with primary expenditures and the part that is not, $\psi_c = \rho v_c + \check{\psi}_c$, where v_c is the residual of the control function regression of the primary expenditures on a set of exogenous instruments.[8] ρ is the correlation between the primary expenditure and ψ, and $\check{\psi}_c$ is the remaining political party's unobserved value that is uncorrelated with the primary expenditures. This decomposition allows for the control function approach from Petrin & Train (2010) to control for endogeneity in this stage:

$$\max_{c \in \{1,\dots,C_p,NoC\}} \check{\beta}\check{X}_c + \rho v_c + \check{\psi}_c \tag{1.2.2}$$

1.2.1.3 Probability of Nomination

The solution to this optimization problem is to pick the candidate with the highest value of the objective function. This solution allows for the derivation of the probability that each candidate is nominated for the general election, P_c. For conciseness, define $\tilde{\beta}\tilde{X}_c \equiv \check{\beta}\check{X}_c + \rho v_c$. Since a discrete choice model can only identify relative utility levels, normalize the the mean utility of the outside option to 0, $\tilde{\beta}\tilde{X}_{NoC} = 0$. Then, define $\tilde{\psi}_c$ as the difference between the valuation for candidate c and the outside option, $\tilde{\psi}_c = \check{\psi}_c - \check{\psi}_{NoC}$. With these assumptions, the probability that candidate c is nominated is writ-

[8]If there are multiple endogenous variables, then v_c is the vector of residuals from the reduced form regressions of each of the endogenous variables on the set of exogenous instruments.

ten as:

$$\begin{aligned} P_c &= Prob(\tilde{\beta}\tilde{X}_c + \tilde{\psi}_c > 0, \tilde{\beta}\tilde{X}_c + \tilde{\psi}_c > \tilde{\beta}\tilde{X}_d + \tilde{\psi}_d \quad \forall d \neq c) \quad (1.2.3) \\ &= Prob(\tilde{\psi}_c > -\tilde{\beta}\tilde{X}_c, \tilde{\psi}_c > -\tilde{\beta}\tilde{X}_c + \tilde{\beta}\tilde{X}_d + \tilde{\psi}_d \quad \forall d \neq c) \quad (1.2.4) \end{aligned}$$

To estimate the model, assume that $\tilde{\psi}_c$ are independently distributed type 1 extreme value. Therefore, the probability that each candidate is nominated has the logit form.[9]

$$P_c = \frac{e^{\tilde{\beta}\tilde{X}_c}}{1 + \sum_{c' \in \{1,...C_p\}} e^{\tilde{\beta}\tilde{X}_{c'}}} \quad \forall c \in \{1, ..., C_p\} \quad (1.2.5)$$

$$\text{with } \tilde{\beta}\tilde{X}_c \equiv \check{\beta}\check{X}_c + \rho v$$

Once this model is estimated, the primary nomination signal can be derived as the predicted $\tilde{\psi}_c$ for the primary winner. This derivation will be shown in the estimation section.

1.2.2 Second Stage: General Elections

1.2.2.1 Utility of a Registered Voter

Assume that there exists a representative utility-maximizing agent in each congressional district.[10] This agent is a registered voter and is facing a discrete voting decision: vote for the Democratic Party candidate, vote for the Republican Party candidate, or not vote for one of the two candidates from

[9] See Train (2009) for the derivation.

[10] There is obviously many dimensions of voter heterogeneity such as the distribution of income within each congressional district. I ignore voter heterogeneity, i.e random coefficients, at this time to keep the model simple.

the dominant parties. The third option is the outside option that includes the choice of voting for a third party candidate or abstaining from the election. I define the outside option this way because the overwhelming majority of congressmen have been members of the Democratic or Republican parties during the past decade. Assume each voter wants to maximize his utility from voting.

$$\max_{j \in \{D,R,NotVote\}} U_{nj} = V(X_j) + e_{nj} \tag{1.2.6}$$

The Democratic Party is denoted by the subscript D and the Republican Party is denoted by the subscript R. For this decision problem, the n subscript denotes the representative voter that varies by congressional district and the j subscript denotes the different voting options: Democratic, Republican, and Not Vote. For example, U_{nR} is the utility that a registered voter receives from voting for the Republican candidate. It is important to note that this utility function is defined as the utility of voting for a specific candidate and not the utility from having that candidate be your representative. $V(\cdot)$ is the indirect utility function from the vector of general election observable characteristics, X_j, and e_{nj} is the unobserved utility that equates $V(\cdot)$ and the actual utility of each individual.

In this stage, there are two sources of unobserved utility. One is the utility from the candidate characteristics that are observed by the agent but unobserved by the econometrician, denoted by ξ_j. For example, some of the characteristics that may be included in ξ_j are beauty, charisma, and public speaking ability. These characteristics are all observed by the voters and affect the voting decision, but are unobserved by the econometrician. All of these characteristics combined can be thought of as the unobserved candidate quality that differs for each candidate in each district. The second source of unobserved utility is the idiosyncratic utility shock, denoted by ϵ_{nj}. This shock represents the unobserved "taste" for each candidate. Each voter is going to have a certain predisposition or affinity for each of the candidates. This is captured by ϵ_{nj} and varies by each candidate and each voter. With this decomposition of the unobserved utility, $e_{nj} \equiv \xi_j + \epsilon_{nj}$ and assuming $V(\cdot)$ is linear in its arguments, equation (1.2.6) is rewritten as:

$$\max_{j \in \{D,R,NotVote\}} U_{nj} = \beta X_j + (\xi_j + \epsilon_{nj}) \tag{1.2.7}$$

1.2.2.2 Correlation of Unobserved Valuations

In addition to the observed candidate characteristics, the representative voter knows the the general election candidate won a primary nomination. In a small abuse of notation, let the subscript c^* represent the nominated candidate in the first stage, so that subscript $c^* = j$ in the second stage. This means that $\tilde{\psi}_j$ references the primary valuation of the candidate's unobserved characteristics by the political party for candidate j. To incorporate this into the voter's utility function, decompose ξ_j into the linear projection on the party's valuation of the unobserved characteristics, $\xi_j = \gamma\tilde{\psi}_j + \tilde{\xi}_j$, where $E(\tilde{\xi}_j|\tilde{\psi}_j) = 0$ by the definition of a linear projection. The voters maximization problem is now:

$$\max_{j \in \{D,R,NotVote\}} U_{nj} = \beta X_j + \gamma \tilde{\psi}_j + \tilde{\xi}_j + \epsilon_{nj} \tag{1.2.8}$$

The parameter γ is the correlation between primary and general election unobserved characteristics and measures the representative voter's valuation of the primary election unobserved characteristics.

1.2.2.3 Predicted Vote Share Function

While U_{nj} is not observable, the derivation of the probability of voting for each of the candidates is done using the fact that the voter is a utility maximizer and assuming a distribution on ϵ_{nj}. For notational and expositional convenience, define the mean utility level for each candidate in each district as:

$$\delta_j \equiv \beta X_j + \gamma \tilde{\psi}_j + \tilde{\xi}_j \tag{1.2.9}$$

Therefore, equation 1.2.8 is simplified to $U_{nj} = \delta_j + \epsilon_{nj}$. The mean utility level of each candidate, δ_j, collects all the candidate characteristics and coefficients within a district. Simplifying the utility this way shows that the mean utility level of each candidate and the idiosyncratic taste valuation completely determine the voting choice for each agent, and thereby, determine the probability of each agent voting for each candidate.

Again, since a discrete choice model can only identify relative utility levels, define $\tilde{\epsilon}_{nj}$ as the utility shock difference with the outside option shock, $\tilde{\epsilon}_{nj} \equiv \epsilon_{nj} - \epsilon_{nNotVote}$, and normalize the mean utility of the outside option to zero, $\delta_{nNotVote} \equiv 0$. Now, the probability of voting for candidate j is:

$$\begin{aligned} P_j &= Prob(\delta_j + \tilde{\epsilon}_{nj} > 0, \delta_j + \tilde{\epsilon}_{nj} > \delta_k + \tilde{\epsilon}_{nk}) \\ &= Prob(\tilde{\epsilon}_{nj} > -\delta_j, \tilde{\epsilon}_{nj} > -\delta_j + \delta_k + \tilde{\epsilon}_{nk}) \end{aligned} \tag{1.2.10}$$

Assume that all $\tilde{\epsilon}_{nj}$ are independently distributed type 1 extreme value. Similar to the first stage, this assumption leads to the result that the predicted probability of voting for candidate j has the logit form.

$$P_j = \frac{e^{\delta_j}}{1 + \sum_{j' \in \{D,R\}} e^{\delta_{j'}}} \quad \forall j \in \{D, R\} \tag{1.2.11}$$

By the law of large numbers, the predicted vote share for each candidate is equal to the probability that the representative agent votes for candidate j.

$$\hat{s}_j(\beta, \xi; X) = \frac{e^{\delta_j}}{1 + \sum_{j' \in \{D,R\}} e^{\delta_j}} \quad \forall j \in \{D, R\} \tag{1.2.12}$$

Different from the first stage, there is much more information available in the observed general election results. This allows for a closed form solution for δ_j if we equate the predicted vote share function with the observed vote share for each candidate. Define s_j as the observed vote share of candidate j. s_j equals the total number of votes received divided by the total number of registered voters. Define s_N as the vote share of the outside option. s_N equals the total number of voters who abstained or voted for a third party candidate divided by the total number of registered voters. Some straightforward algebra produces:

$$ln(s_j) - ln(s_N) = \delta_j \tag{1.2.13}$$

$$\therefore ln(s_j) - ln(s_N) = \beta X_j + \gamma\tilde{\psi}_j + \tilde{\xi}_j \tag{1.2.14}$$

The model assumptions and the ability to accurately measure the general election vote shares allow for a linear model to estimates the impact of candidate characteristics on the voting outcomes.

1.2.2.4 Endogeneity of Candidate Expenditures in the General Election

$\tilde{\xi}_j$ is defined as all the characteristics of the candidate that are unobserved by the econometrician that affect the mean utility level of the candidate. Similar to the primaries, these characteristics not only affect the number of votes a candidate receives, but also affect the amount of contributions a candidate raises. A candidate with greater contributions can expend more on the congressional race. Again, the unobserved characteristics are expected to be correlated with the candidate's expenditures in this stage also. The estimation procedure will account for the endogeneity of general election campaign expenditures by using linear instrumental variables methods.

1.3 Estimation

1.3.1 First Stage Estimation

1.3.1.1 Control Function for Primary Expenditures

The first stage of the model is estimated using the control function method of Petrin & Train (2010) to control for endogeneity of expenditures. The method is a two step procedure where the first step estimates a control function that regresses the endogenous variable on a set of instruments and the exogenous regressors. Then, the residual for each observation, $\hat{v}_c$, is used as an additional regressor in the maximum likelihood estimation.

In this environment, the endogeneity in this stage arises through the fact that candidate quality affects expenditures and outcomes simultaneously. A control function for this environment is one that predicts the average expenditure for a candidate. The control function is the linear regression of the candidate's primary expenditure on an instrument that measures average candidate expenditure in the rest of the state and the rest of the exogenous variables.[11] Then, the predicted residual, $\hat{v}_c$, from the control function is an indication if the candidate spent more the predicted average, which would suggest he is of higher quality, or if he spent less, which would suggest he is of lower quality.

[11] Section 1.4 will explain the instruments in more detail and the actual control function estimates are provided in the Appendix.

1.3.1.2 Primary Parameter Estimates

The assumptions in the first stage yields a standard multinomial logit model for the probability that each candidate will be nominated. Substituting in $\hat{v}_c$ for all candidates, the likelihood function for this stage is defined as:

$$\begin{aligned} \ell &= \prod_{m=1}^{M} \prod_{p \in \{D,R\}} \prod_{c=1}^{C_p} P_c^{i(c)} \qquad (1.3.1) \\ \text{with } P_c &= \frac{e^{\tilde{\beta}\tilde{X}_c}}{1 + \sum_{c' \in \{1,\ldots C_p\}} e^{\tilde{\beta}\tilde{X}_{c'}}} \\ \text{and } \tilde{\beta}\tilde{X}_c &\equiv \check{\beta}\check{X}_c + \rho\hat{v}_c \end{aligned}$$

where i(c) is an indicator function defined as

$$i(c) = \begin{cases} 1 & : \text{ Party } p \text{ in district } m \text{ nominates candidate } c \\ 0 & : \text{ Otherwise} \end{cases} \qquad (1.3.2)$$

Estimates for the $\tilde{\beta}$ vector are obtained by Maximum Likelihood Estimation of $ln(\ell)$.

1.3.2 Second Stage Estimation

The estimation equation for the second stage results from equation (1.2.14). Estimation of the parameter vector (β, γ) is done using linear methods. $\tilde{\psi}_j$ is unobservable, so it must be estimated by $\hat{\psi}_j = E(\tilde{\psi}_j | \tilde{\psi}_j > -\tilde{\beta}\tilde{X}_j, \tilde{\psi}_j + \tilde{\beta}\tilde{X}_j - \tilde{\beta}\tilde{X}_d > \tilde{\psi}_d \quad \forall d \neq j)$. $\hat{\psi}_j$ is the expected value of the unobserved candidate characteristics in the first stage for candidate j given the selection rule that he won the primary election. This expected value can be found for each candidate based on the results of the first stage estimation and the distributional assumption.

$$\hat{\psi}_j = \frac{1}{1 - F(-\tilde{\beta}\tilde{X}_j)} \int_{-\tilde{\beta}\tilde{X}_j}^{\infty} \tilde{\psi}_j \prod_{d \neq j} F(\tilde{\psi}_j + \tilde{\beta}\tilde{X}_j - \tilde{\beta}\tilde{X}_d) f(\tilde{\psi}_j) d\tilde{\psi}_j \tag{1.3.3}$$

For conciseness, redefine: $s \equiv \tilde{\psi}_j, V_j \equiv \tilde{\beta}\tilde{X}_j,$ and $V_d \equiv \tilde{\beta}\tilde{X}_d$

$$\hat{\psi}_j = \frac{1}{1 - \frac{e^{-V_j}}{1+e^{-V_j}}} \int_{-V_j}^{\infty} s \prod_{d \neq j} \frac{e^{s+V_j-V_d}}{1 + e^{s+V_j-V_d}} \frac{e^s}{(1+e^s)^2} ds \tag{1.3.4}$$

$$\hat{\psi}_j = (1 + e^{-V_j}) \int_{-V_j}^{\infty} \frac{s}{1+e^s} \prod_{d} \frac{e^{s+V_j-V_d}}{1 + e^{s+V_j-V_d}} ds \tag{1.3.5}$$

If there is only one candidate in the primary, the closed form solution is:

$$\hat{\psi}_j = V_j e^{-V_j} + (1 + e^{-V_j}) ln(1 + e^{-V_j}) \tag{1.3.6}$$

Otherwise, it is straightforward to approximate the single dimensional integral in equation (1.3.5). With $\hat{\psi}_j$ for every general election candidate, the estimation of the second stage model can be done using optimally weighted IV estimation on:[12]

$$ln(s_j) - ln(s_N) = \beta X_j + \gamma \hat{\psi}_j + \tilde{\xi}_j \tag{1.3.7}$$

1.3.3 Standard Errors

The estimation of the preference parameters in this model is done in four distinct steps. Theoretically it would be possible to derive the standard

[12] The instruments used for this analysis will be discussed in the data section.

error formula after each step. However, this generally becomes difficult for a two step estimation process and would be incredibly difficult in this model because of the four steps and the inclusion of the approximated regressor $\hat{\psi}_j$. Therefore, the standard errors are found through a bootstrap process with 1000 repetitions. The resampling for the bootstrap must be done at the general election level to maintain consistency of the sample.

Additionally, the estimation includes state dummy variables. As discussed in Cameron, Gelbach & Miller (2007), bootstrapping the original sample at the general election level can produce pseudo samples for which it is impossible to estimate the coefficient vector. This occurs when the bootstrapped sample does not draw a single general election from a certain state. For any state excluded from the pseudo sample, the state dummy variable would be zero for all observations and the pseudo sample would not satisfy the rank condition. To prevent this failure, the bootstrap process is augmented to ensure all states are represented in the pseudo samples. The pseudo samples are created by selecting general election clusters in each state with replacement. This ensures that each pseudo sample has the same number of general elections in each state as the original sample and no state is excluded from a pseudo sample.

1.4 Data

1.4.1 Overview

For this analysis, a unique data set has been compiled from multiple sources. All of the election results and candidate expenditure data is provided by the Federal Election Commission (FEC). Biographical information on the candidates was obtained from Natalie Adona at VoteSmart.org. Voter registration statistics and the date for the primary elections were gathered from each state's board of elections. This chapter uses data from the 2004 and 2006 election cycles.

The FEC provides summary files that document each candidate's expenditures during the two year election cycle. The amounts reported in this file provide the total amounts spent for the whole election but they do not provide the timing of the expenditures. Most important to this analysis, these files do not provide the total amount spent before the primary election. While the summary files are able to provide the expenditure information for primary losers, the preprimary expenditure for primary winners must be calculated by looking at the individual candidate's expenditure report filings.

Each candidate is required to file quarterly reports in addition to prepirmary reports and pre-general election reports. Every individual expenditure is dated in these reports and is also supposed to indicate if it was for the primary election or the general election. However, the election indicator is missing for the majority of candidates. Therefore, it is necessary to make a few assumptions to determine the primary expenditure variable for each candidate.

If the candidate won the primary election, primary expenditures are defined as the total amount spent during the election cycle until and including the day of the primary. This value can be found by looking at the pre-primary expenditure report and the next quarterly expenditure report. The pre-primary report is filed 20 days before the primary election and provides the total amount spent up until the day of the report. This number is then added with all expenditures from the next quarterly report that are dated on or before the day of the primary. If the candidate did not win the primary, it is assumed that all the candidates' expenditures during the election cycle were used for the primary election, and the primary losers' primary expenditures are the amount reported in the election summary file.[13]

Biographical information for the candidates is provided by VoteSmart.org. Every election, this organization sends out a survey to every candidate requesting biographical information. The survey includes the standard questions of birthdate, gender, and marital status. The survey also asks for information on educational and occupational history. Due to the limited response rates for most of the questions, this chapter will only use the gender information at this time. If a candidate did not report his or her gender on the survey or did not return the survey, the gender data was supplemented by finding the information for the candidate through internet searches.

[13]If a primary election resulted in a runoff election, the runoff election and information is used in place of the primary election for that congressional district.

Estimation of the model requires that some states be excluded from the sample. To estimate the second stage of the model, it is necessary to have the total number of registered voters in each congressional district at the time of the general election. The majority of states either provide this information on their website or the numbers can be obtained by contacting their board of elections. However, without this information, the state must be excluded from the sample. There are two reasons that states do not have the total number of registered voters by congressional district. Either the state does not have the data aggregated at all or the data is aggregated at the county level and the congressional districts do not follow county boundaries. I also exclude the congressional districts in Connecticut, Louisiana, and Utah because they do not have the standard procedural process for primary and general elections.[14] Given these exclusion, I am able to use the congressional districts from 35 states.[15]

1.4.2 Control Variables

The candidate and election characteristics that are used as control variables are similar in the first and second stages. Both stages use a dummy

[14]Connecticut nominates its candidates through party conventions. Utah has a two step primary of a party convention followed by a runoff if one candidate did not gain enough of the convention delegate support. Only the nominated candidates are reported in the data if the candidate was elected by a convention. Louisiana does not have primary elections. Instead, all candidates enter the general election and the top two proceed to a runoff if no candidate earns a majority of the vote.

[15]The states included in the sample are AK, AR, AZ, CA, CO, DE, FL, ID, IN, IA, KY, ME, MD, MN, MT, NE, NV, NH, NJ, NM, NY, ND, OH, OK, OR, PA, RI, SC, SD, TX, VT, VA, WA, WV, WY.

variable for incumbency, denoted *Incumbent*, party specific dummy variables, *dDem* and *dRep*, and a gender dummy variable, *dFemale*. These variables vary by candidate and do not change in the sample from the primary election to the general election. Both stages also have dummy variables that are equal to one if a Senate election, *SenateRace*, or a Presidential election, *PresidentialRace*, is happening concurrently. The main function of these dummy variables is to control for the differences in overall voter turnout when the voter has a more publicly visible election to vote for at the same time as the U.S. House election.

The second stage includes additional covariates that are not included in the primary election estimation. The number of candidates in the primary election is included as a covariate for each general election candidate. This variable is a proxy for the competitiveness of the primary election for each candidate. The state level unemployment rate is also included as a proxy for the economic conditions at the time of the election.[16] Lastly, state dummy variables are also included in the second stage, but these are of limited interest individually and will be excluded from the summary statistics and estimation results.

Both stages include candidate campaign expenditures. This covariate changes greatly between the two stages. The expenditure variation between

[16] This information comes from the Bureau of Labor Statistics. While it would be preferable to include the district level unemployment, the BLS does not provide district level unemployment rates and the American Community Survey does not provide congressional district level data for all districts until the 2006 survey.

primaries and general elections occurs for many reasons. Primary elections occur months before the general election, so typically the primary expenditures are much less than the general election expenditures. Also, the different level of competition in a primary and general election cause large variations in expenditure levels across candidates. There are many cases in which a candidate spends large amounts of money on the primary because he is facing stiff competition for the party nomination, but then spends little additional money on the general election because he is running unopposed. The are also many cases of a candidate spending very little on the primary election and then spending large amounts on the general election. The actual control variables used in the estimation are the natural log of one plus the candidate expenditures to allow for the fact that some candidates spend zero dollars on the election. These values are also interacted with the incumbency dummies to allow for different marginal effects of expenditures for incumbents and challengers.

Table 1.1 shows the summary statistics for the data set.[17] In the primaries, this table shows that the vast majority of candidates are male and the average amount spent per candidate is close to a quarter of a million dollars. Also, there are an average of 1.55 candidates per primary with the maximum number of candidates in a single primary being 8 candidates. In the general elections, the average amount spent per candidate is almost $850,000 and almost half of the candidates are incumbents. Additionally, the majority of general elections in this sample have one Democrat and one Republican run-

[17] All expenditures are converted into 2006 dollars

Table 1.1: Summary Statistics

	Mean	St. Dev.	Min	Max
Primary Elections				
Campaign Expenditures ($)	232,235	336,639	0	3,612,543
Incumbent	.3202	.4667	-	-
Female	.1565	.3634	-	-
Number of Candidates		1674		
Number of Primary Elections		1077		
General Elections				
Campaign Expenditures ($)	845,283	906,561	0	8,112,752
Incumbent	.4958	.5002	-	-
Female	.1681	.3741	-	-
State Unemployment Rate	4.892	.8311	2.8	6.9
Vote Share	.2961	.1227	.0083	.6395
Number of Candidates		1077		
Number of General Elections		575		

Note: This table presents the summary statistics for the primary election sample of candidates and the general election sample of candidates. The primary elections have between 1 and 8 candidates with an average 1.55 candidates per primary election. All general elections have either 1 or 2 candidates and an average of 1.87 candidates per general election.

ning against each other. A candidates ran unopposed in the general election in only 73 of the 575 general elections. The table also shows that the average candidate has a .30 vote share but there is significant variation in this number.

1.4.3 Instruments

As mentioned earlier, the endogeneity problem in this model arises from the fact that there are candidate characteristics that a voter observes but the econometrician does not. These characteristics are referred to as unobserved candidate quality or "valence" factors. To deal with the endogeneity, instruments are constructed that are similar to Hausman (1996) type instruments that were popularized by Nevo (2000). These instruments are the characteristics of the same products in other regions. In this analysis, the congressional districts provide a distinct boundary for the elections within each state. An instrument for candidate quality is created using these boundaries. In the first stage, the instrument is the average candidate expenditure in all other primaries within the same state as the district that is being instrumented. In the second stage, the instrument is the average candidate expenditure in all other general elections within the same state.[18] The instruments exclude the district that is being instrumented when calculating the averages to help ensure that these instruments are not correlated with the election outcome of the district that is being instrumented.

These instruments provide a proxy for candidate quality by exploiting

[18] These instruments make it necessary to exclude any state that only has one congressional district.

the variation in each candidates' expenditure level from the average expenditure. The instruments provide a baseline average expenditure for each candidate in the state. Each candidates' variation from this average is a proxy of unobserved candidate quality. A candidate who spends less than the average is probably not as high quality as a candidate who spends more than the average for the rest of the state.

In the first stage, the actual instrument used is the natural log of the average candidate primary expenditures in the rest of the state. This instrument is interacted with the incumbency dummy variable to create two instruments.[19] In the second stage, a similar set of four instruments is constructed. The average is now calculated over general election candidates only and the average is calculated for candidates in the same party and candidates in the opposing party. The natural log of the own party average and the natural log of the opposing party average are both interacted with the incumbency dummy variable to create four instruments. The Appendix provides the reduced form regressions of the endogenous variables on the instruments and the exogenous variables to show the high correlation between the instruments and the endogenous variables.

[19]The interaction is included because expenditure is interacted with the incumbency dummy variable in the estimation.

1.5 Empirical Results

1.5.1 First Stage Results

1.5.1.1 Estimated Coefficients

The estimated coefficient vector from the first stage is presented in Table 1.2 along with the marginal effects.[20] The coefficient on $ln(PrimaryExpend)$ is positive and significant, while the coefficient is for $ln(PrimaryExpend) * Incumbent$ is negative and significant. The magnitudes indicate an expenditure increase leads to a higher probability of being nominated for the general election for both incumbents and challengers. The marginal effects show that challengers are able to increase the probability of being nominated much more from an increase in expenditures than incumbents. The difference in marginal effects is driven by the fact that, on average, incumbents vastly outspend challengers. The average incumbent spent $440,220 before the primary and the average challenger spent $132,390.

Along with the expenditure effects, the results estimate that an inherent incumbent advantage does not exist in the primary elections. It would be expected that an incumbent has an inherent advantage in the primary elections since all previous empirical election work estimates an inherent advantage for incumbents in the general election and because only 2 of the 536

[20] The marginal effects are only reported for the first four variables. The party dummy variables and concurrent elections variables do not vary within each primary so the marginal effects are omitted. $\hat{v}_1$ and $\hat{v}_2$ are the control function variables and so marginal effects are not suitable because they model the correlation between campaign expenditures and the unobserved characteristics.

incumbents in the sample lost in the primary election. However, the coefficient on *Incumbent*, which measure's an incumbent's inherent advantage in the election, is positive but relatively small and insignificant. The relatively small coefficients for $ln(PrimaryExpend) * Incumbent$ and *Incumbent* indicate that an incumbent's primary advantage is explained by his much higher spending level rather than an inherent advantage.

All of the other coefficients are estimated significantly. The female dummy has a negative coefficient which implies that a woman is less likely to be nominated by the political party than a man. Recall that the political party's have the the outside option of not nominating a candidate. Since the value of the *SenateRace* dummy variable does not vary by the different candidates within a primary, the positive significant value implies that all candidates are more likely to be nominated for the general election when a senate race is happening concurrently. The same interpretation holds true for when a presidential race is happening concurrently.

$\hat{v}_1$ and $\hat{v}_2$ are the additional regressors used to control for the endogeneity of expenditures. These variables are the residuals from the control function regressions of the endogenous variables on the instruments and all the exogenous variables. The coefficients are individually significant. This provides evidence that the estimates would suffer from endogeneity bias if the control function residuals were not included in this estimation.

Table 1.2: First Stage Parameters

VARIABLES		Marginal Effects
$ln(PrimaryExpend)$	2.1777***	1.1164%***
	(0.1340)	(0.1922)
$ln(PrimaryExpend) * Incumbent$	-0.2255*	0.0116%***
	(0.1303)	(0.0011)
$Incumbent$	0.1663	0.0069
	(1.4116)	(0.0282)
$dFemale$	-1.3157***	-0.0569***
	(0.1391)	(0.0036)
$dDem$	0.6107***	
	(0.0706)	
$dRep$	0.6724***	
	(0.1550)	
$PresidentialRace$	0.6316***	
	(0.1354)	
$SenateRace$	1.0162***	
	(0.1210)	
$\hat{v}_1$	-1.8621***	
	(0.1339)	
$\hat{v}_2$	-3.0774***	
	(0.2328)	
Observations	1674	
LLR Index	0.8350	
LL at Max	-156.5797	

Note: This table presents the coefficient results from the maximum likelihood estimation for the primary process. $\hat{v}_1$ and $\hat{v}_2$ are the estimated residuals from the control function regressions of $ln(PrimaryExpend)$ and $ln(PrimaryExpend)*Incumbent$ on the instrument set. Bootstrapped standard errors are reported in parentheses and the asterisks indicate significance levels of the coefficients, *** p<0.01, ** p<0.05, * p<0.1. The third column reports the marginal effects, measured in change in probability of nomination for the discrete variables and % change in probability of nomination for the continuous variables. The marginal effect for $ln(PrimaryExpend) * Incumbent$ is the change in probability from a marginal increase in expenditure by an incumbent candidate.

1.5.1.2 Predicted $\hat{\psi}$ Values

The first stage coefficients are used to estimate the value of $\hat{\psi}$ for each primary winner. Recall that $\hat{\psi}$ is the expected primary valuation of the unobserved candidate characteristics, and the formula for this value is given by equation (1.3.5). Figure 1.1 provides a histogram of these values for all the primary winners.[21] The histogram shows that the majority of primary elections result in $\hat{\psi}$ of about 0. The unrestricted expected value of ψ is 0, so this means that the primary election does not provide any additional information for these candidates. This occurs because the valuation of the *observed candidate characteristics* is high enough that it does not place any restriction on the values of the unobserved characteristics. For example, if a candidate greatly outspends the other candidates in the primary election, the observed expenditure explains why the candidate was nominated for the primary and it is impossible to infer any additional information about his unobserved characteristics. While the majority of candidates have a $\hat{\psi}$ close to zero, 289 of the 1077 general election candidates have an estimated $\hat{\psi} > .01$. Therefore, the empirical gain from incorporating the primary process is to have a measure of the candidate's unobserved characteristics to include in the second stage estimation for candidates whose observed characteristics do not completely explain why the candidate was nominated by the primary process.

[21] It is impossible to calculate $\hat{\psi}$ for the losers of the primary election. Since they were not chosen in the primary, no restrictions can be imposed on the valuation of their unobserved characteristics.

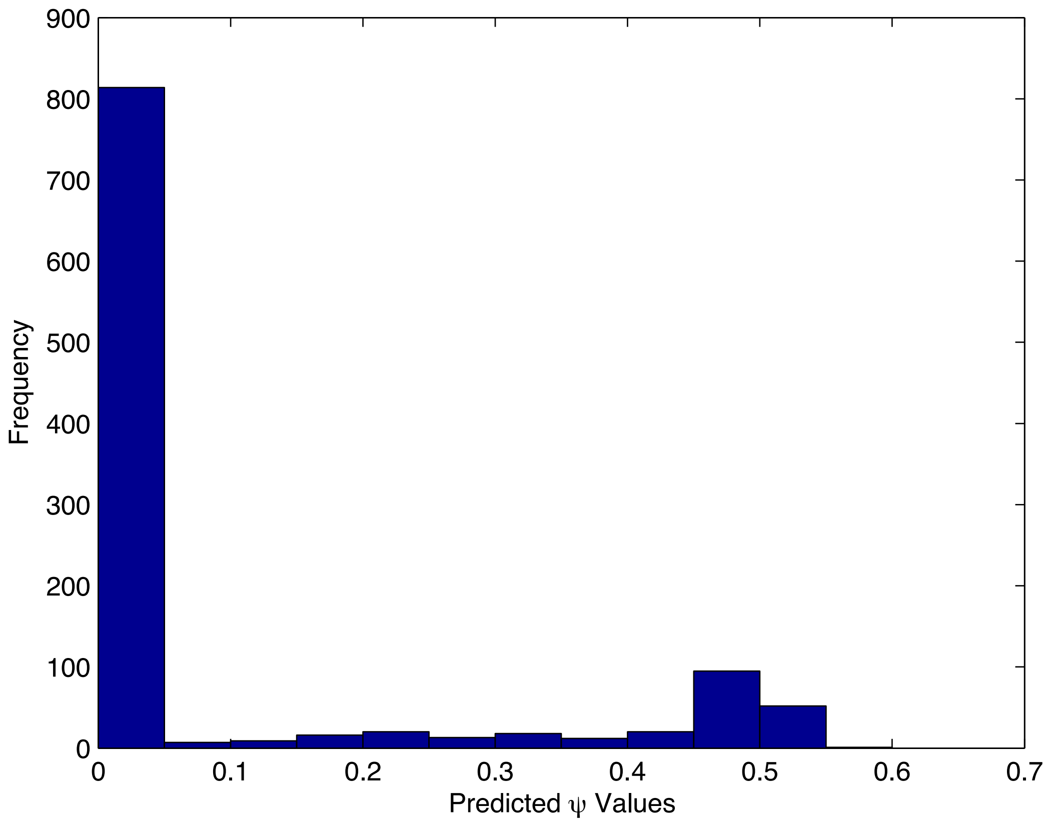


Figure 1.1: Histogram of Predicted ψ Values

1.5.2 Second Stage Results

Table 1.3 reports the estimated coefficients from four different specifications of the second stage model to show the impact of incorporating the primary information. The first two columns estimate the second stage, equation (1.3.7), without controlling for the endogeneity of expenditure. The third column and fourth columns estimate the model controlling for endogeneity.

The second and fourth columns include the $\hat{\psi}$ term to control for potential selection from the primary elections and test if voters value this information. The fifth column reports the estimated marginal effects from column (4). The marginal effects for the dummy variables ($Incumbent$, $dFemale$, $PresidentialRace$, and $SenateRace$) show the change in predicted vote share for an average candidate when the dummy variable switches from 0 to 1. The marginal effect for the number of primary candidates shows the change in vote share from going from 1 primary candidate, i.e. running unopposed in the primary, to 2 primary candidates. Lastly, the elasticities for the continuous variables ($UnemploymentRate$ and $\hat{\psi}$) are reported instead of the marginal effect to show the change in vote share. These values show the percentage change in vote share from a 1% increase in the variable.

1.5.2.1 Evidence Against Valuing the Primary Election Information and Selection Bias

Column (1) estimates the model without controlling for endogeneity or testing if the voter's value the estimated valuation of the primary election unobservables. This column shows that even without controlling for endogeneity, this method produces results that are qualitatively similar to the accepted results from prior literature. A challenger has a higher marginal return to expenditure and an incumbent has a significant inherent advantage. Column (2) includes $\hat{\psi}$ to test for the voter's valuation of primary signal and control for potential selection bias. The coefficient on $\hat{\psi}$ is relatively small and is positive but insignificant. This is evidence that the general elections voter's do not

Table 1.3: Second Stage Parameters

VARIABLES	No Selection, No Endogeneity (1)	Selection, No Endogeneity (2)	Endogeneity, No Selection (3)	Selection and Endogeneity (4)	Marginal Effect or Elastiticy (5)
$ln(GeneralExpend)$	0.0672***	0.0675***	0.1496***	0.1472	
	(0.0064)	(0.0058)	(0.0571)	(0.0907)	
$ln(GeneralExpend) * Incumbent$	-0.0462**	-0.0423**	-0.1230	-0.1245	
	(0.0200)	(0.0193)	(0.3353)	(0.5019)	
$Incumbent$	0.9623***	0.9986***	1.77578	1.7987	0.0555
	(0.2714)	(0.2645)	(4.4390)	(6.5863)	(0.0920)
$dFemale$	-0.03346	-0.0338	-0.0838*	-0.0815	-0.0166
	(0.0281)	(0.0275)	(0.0429)	(0.0591)	(0.0121)
$NumPrimaryCandidates$	0.0498***	0.0369***	-0.0131	0.0040	0.0008
	(0.0108)	(0.0143)	(0.0458)	(0.0551)	(0.0112)
$dDem$	-1.4704***	-1.4425***	-2.2013***	-2.2041**	
	(0.1942)	(0.2258)	(0.5641)	(0.9203)	
$dRep$	-1.5905***	-1.5624***	-2.2469***	-2.2521***	
	(0.1958)	(0.2260)	(0.4996)	(0.8116)	
$UnemploymentRate$	-0.0774*	-0.0806*	-0.1071*	-0.1034	-0.3590%
	(0.0436)	(0.0487)	(0.0565)	(0.0714)	(0.2481)
$PresidentialRace$	0.8073***	0.8084***	0.8393***	0.8373***	0.1012***
	(0.0380)	(0.0430)	(0.0544)	(0.0760)	(0.0090)
$SenateRace$	-.1120***	-0.1120***	-0.1077**	-0.1091***	-0.0133*
	(0.0304)	(0.0328)	(0.0448)	(0.0593)	(0.0071)
$\hat{\psi}$		0.1114		-0.1273	-0.0093%
		(0.0939)		(0.2516)	(0.0175)
Observations	1077	1077	1077	1077	

Note: This table presents the coefficient results from four specifications of estimation the general election process, omitting the coefficients for the state fixed effects. The dependent variable in each specification is $ln(s_j) - ln(s_N)$ ($\equiv \delta_j$). $\hat{\psi}$ is the estimated unobservable valuation from the first stage that controls for selection in columns (2) and (4). Columns (3) and (4) are estimated using the natural log of average candidate expenditure in the rest of the state and the interaction with the incumbency dummy variable as instruments. Robust standard errors are reported for columns (1) and (3) and bootstrapped standard errors are reported in columns (2) and (4). The asterisks indicate significance levels of the coefficients, *** p<0.01, ** p<0.05, * p<0.1. Column (5) reports the marginal effects, measured in vote shares for the discrete variables and % change in vote share for the continuous variables from column (4). The magnitude of effect for the candidate expenditures are presented in section 1.5.3.1.

value the signal. Additionally, the insignificant coefficient cannot reject the null hypothesis of no selection bias. More evidence against selection bias is that the coefficients on the other covariates barely change. The coefficient on $NumPrimaryCandidates$ does decrease but is not significantly different from the coefficient in the first specification. These facts imply that voter's do not value the primary valuation and are not voting based on this unobservable characteristic.

Columns (3) and (4) show the effect of controlling for the primary unobservable characteristics when the estimates also control for the endogenous expenditures. As is standard with instrumental variable estimation, these results are much less precise. Again, the coefficients barely change with and without the $\hat{\psi}$ term. In fact, the main effect of including this regressor is to increase the standard errors for most of the coefficients. This reinforces the fact that there is no additional information to be gained from including the estimated unobserved primary valuation, $\hat{\psi}$.

The estimates of the impact on the number of primary elections candidates provide more evidence that the instrumental variables adequately control for the candidate quality and voters do not incorporate the primary election information when deciding for whom to vote. In the first two columns, the coefficients are both estimated positive and significant. This indicates that, after controlling for the other factors, the winner of a more competitive primary will receive a higher vote share in the general election. However, this effects becomes small and insignificant after controlling for endogeneity. This should be expected because the number of primary candidates is a proxy for candidate quality. The instruments are chosen to control for candidate quality in the general election and therefore the number of primary election candidates defeated does not provide additional information.

1.5.3 Marginal Effects of Control Variables

The average marginal effects from column (4) are reported in column (5) to show the impact of the candidate characteristics on their vote shares. The expenditure effects on the candidate vote shares are discussed in the following subsection. The incumbent advantage is similar to the estimated advantage from Levitt (1994) and Lee (2008). With an average of 382,852 registered voters in each district, the .0555 incumbent advantage is equal to about a 21,200 vote advantage. The marginal effect for $dFemale$ shows that female candidates have a sizable disadvantage. The marginal effect estimates that a female candidate is inherently disadvantaged by about 6,400 votes.

While incumbency and gender vary by candidate, the state unemployment rate, presidential race, and senate race variables do not and therefore are included to control for the large variation in turnout rates. The marginal effect for $PresidentialRace$ implies that voter turnout is .10 higher during presidential election years. Surprisingly, the effect of a concurrent Senate race is estimated to be negative, and therefore voter turnout is slightly lower when a senate race is occurring. However, this effect is much smaller than the presidential race effect on turnout. Lastly, the effect from unemployment rate shows signs of voter discouragement. As the overall state unemployment rate increases, the turnout is predicted to decrease. The average unemployment rate in the sample is 4.892%. An increase to 4.941% is predicted to reduce turnout by 1400 votes.

1.5.3.1 Expenditure Elasticities

The best way to interpret the magnitude of the expenditure coefficients is through the elasticities. The expenditure elasticities show the change in predicted vote shares for each of the voting options from an expenditure increase. The elasticities calculated using the estimated coefficients from specification (4) are presented in Table 1.4. This table reports the average own- and cross-expenditure elasticities. These elasticities are calculated for every congressional district in the sample and the average elasticity across all the districts is reported. The values show the average percentage change in predicted vote share for each option if a candidate were to increase his expenditure by 1%.

These elasticities quantify a very important fact. A candidate can increase his vote share by increasing expenditures and the vote share increase has two different sources. The increase comes not only from stealing votes from the other candidate, but also from convincing voters to vote who would have otherwise abstained. This can be seen from the negative percentage change in the share of registered voters who would choose to not vote, $\Delta\hat{s}_{NotVote}$. Greater expenditures in these elections leads to higher voter turnout. In these estimates, the outside share percentage decrease is larger than the opposing candidates' share percentage decrease for all types of candidates.[22]

[22]However, these substitution patterns are susceptible to the IIA property of standard

Table 1.4: Average Elasticities: Specification 4

	Challenger		Incumbent	
	Democrat	Republican	Democrat	Republican
$\Delta\hat{s}_D$	0.1133%	-0.0551%	0.0144%	-0.0045%
	(0.0088)	(0.0153)	(0.0025)	(0.0019)
$\Delta\hat{s}_R$	-0.0510%	0.1189%	-0.0033%	0.0147%
	(0.0147)	(0.0138)	(0.0022)	(0.0021)
$\Delta\hat{s}_{NotVote}$	-0.0623%	-0.0638%	-0.0111%	-0.0101%
	(0.0177)	(0.0205)	(0.0041)	(0.0031)

Note: Values are the average elasticities from all the districts. These values show the percentage change in the predicted vote share for each option from a 1% increase in expenditure by the column candidate. The standard deviation of each elasticity is in parentheses.

These elasticities also provide a calculation of the cost of an additional vote. The average vote share for a challenger in this sample is .2135 and the average challenger spent $528,791. The model predicts that by spending 1% more, a Democratic challenger can increase his vote share by 0.1133%. Therefore, an additional $5,288 expenditure increases his vote share to .2138. With an average of 382,852 registered voters in each district, the vote share difference is equal to 94 new votes. This implies that it cost about $56 per new vote for a Democratic challenger and the number is very similar for a Republican challenger. An incumbent faces a much higher cost of garnering new votes. The average incumbent spent $1,170,727 and has a .3804 vote share. A 1% increase in expenditure only gains a Democratic incumbent 20 new votes, which implies a cost of $585 per new vote. This large difference

logit models. Further work will modify the model to relax the IIA property.

is driven by the much larger average expenditure and average vote share for incumbents combined with the decreasing returns to campaign expenditures.

1.6 Campaign Finance Reform Analysis

The second stage of this model is based on individual voters choosing the option that maximizes their utility based on the observed and unobserved candidate characteristics. The preference parameters of the utility function are then estimated. This allows the model to analyze the effects of policy reforms that change the the candidate's campaign expenditures. In addition to quantifying the effect on candidate vote shares, this model can quantify the effect on voter participation because the model incorporates the outside option of abstaining from voting.

The model is used to analyze the effects of two different campaign expenditure reforms. The first reform an expenditure cap policy. This experiment puts a strict limit on the amount of campaign expenditure allowed by any candidate, . The second reform simulates mandatory public funding of the campaign expenditures. This experiment puts a limit on the amount of expenditures by any candidate but then provides that amount of funding for every candidate. These experiments will only change the expenditure levels for the general election candidates. All other variables remain the same, including the values of the unobserved characteristics, ξ. These values are kept constant because the estimation method explicitly estimates these values to be uncorrelated with the candidate expenditures. The experiments only change the candidate expenditures so the ξ values should not change.

1.6.1 Expenditure Cap

In this experiment, the model is estimated and then the expenditures of some candidates are changed. Any candidate's expenditure that is more than the cap is replaced by the cap value. Then, using the new expenditure values, the predicted vote shares for each candidate and vote share of the outside option are calculated. This is done for 100 different expenditure cap levels between $10,000 to $1,000,000 and the predicted results are presented in Figures 1.2 and 1.3.

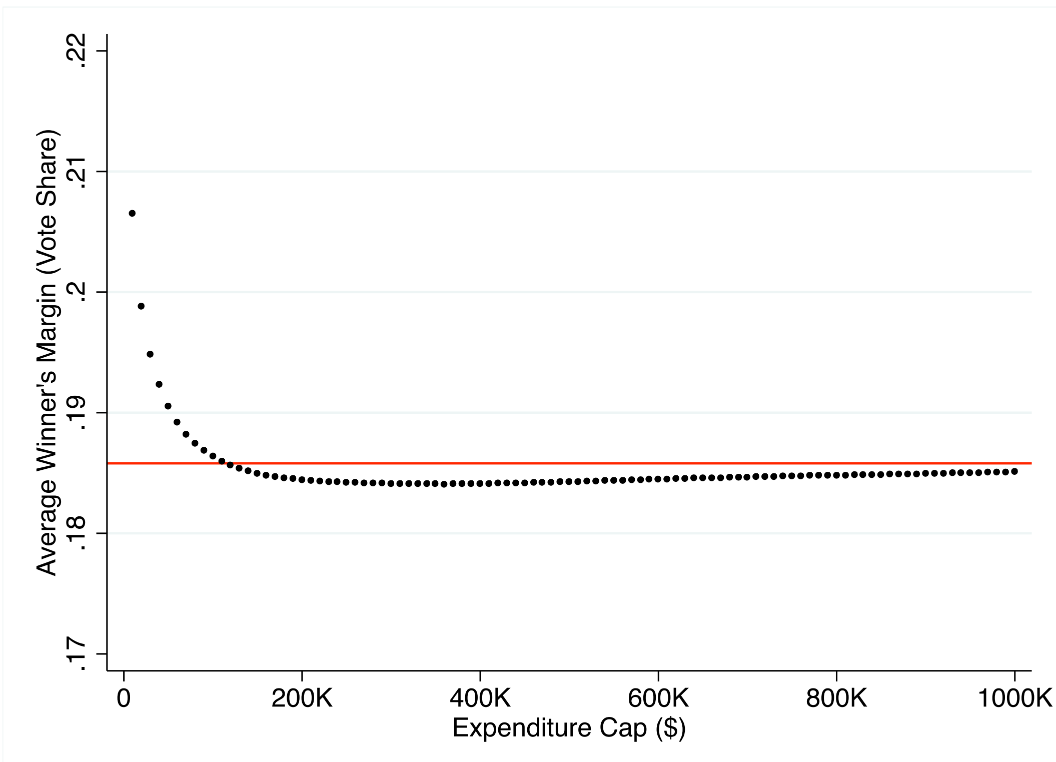


Figure 1.2: Predicted Average Winner's Margin of Victory Under Different Expenditure Caps

Figure 1.2 measures the average absolute difference in vote shares, the average winner's margin, between the two candidates for each of the 100 expenditure cap levels. This is a measure of the overall competitiveness of the elections. The average winner's margin without any policy reform is .1858 and is included in the graph as a reference line. This shows that for expenditure caps less than about $120,000, this policy makes the election less competitive on average. Based on the estimated results, the reason this happens is that incumbents have a very large intrinsic advantage. Imposing spending limits reduces the opportunities for challengers to convince voters to vote for them through campaign expenditures. As the cap increases, the predicted average winner's margin does drop slightly below the baseline model. However, the difference is almost negligible.

The effects of an expenditure cap on voter turnout are presented in Figure 1.3. The vertical axis measures the average change in number of voters as compared to the baseline model. The figure shows that an expenditure cap would reduce voter turnout for all levels on the cap. This result is driven by the fact that all voters have a positive response to campaign expenditures, as seen by the parameters in Table 1.3. Any restriction on the expenditures will decrease the relative utility level of the candidates as compared to the outside option. The predicted decrease in voter turnout also makes intuitive sense in these elections because a large portion of campaign expenditures are focused on getting people out to vote. Any restrictions on expenditures will lessen the effort to convince people to vote. Overall, neither of the predicted effects of an expenditure cap are generally desirable and so an expenditure cap does not seem to be a reasonable policy for U.S. House elections.

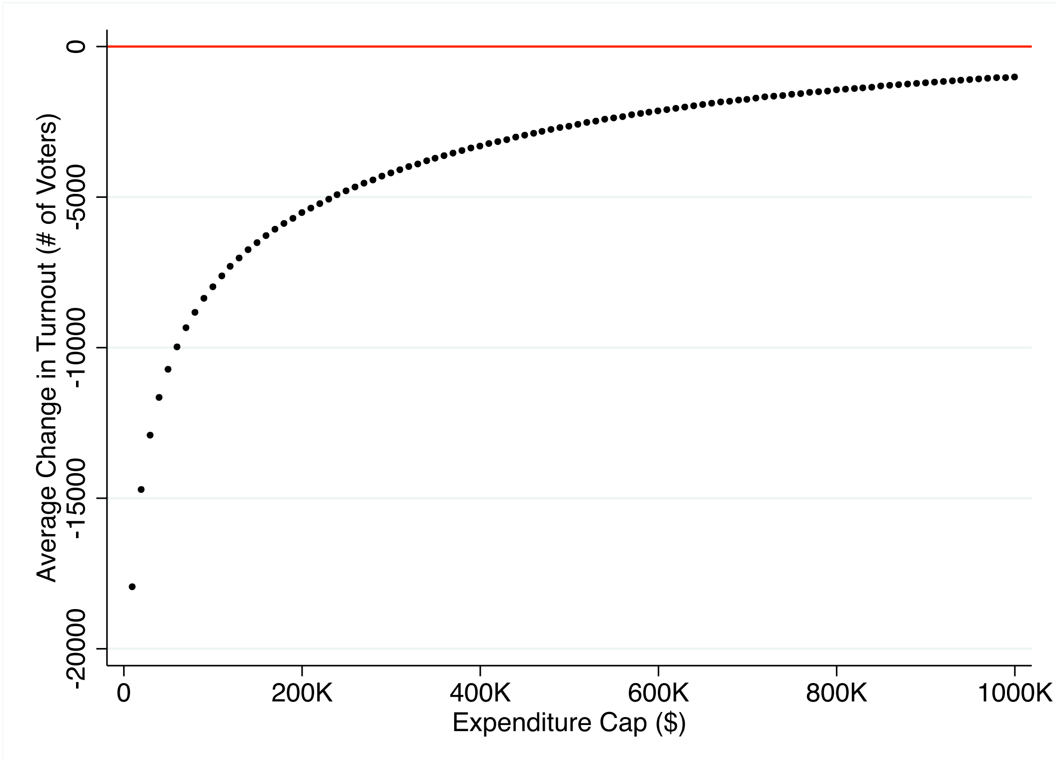


Figure 1.3: Predicted Average Change in Number of Voters Under Different Expenditure Caps

1.6.2 Public Funding

This policy experiment is meant to show the effects on competitiveness and voter turnout from implementing a policy analogous to mandatory public funding of campaign expenditures. Under this policy, a funding level is set for all candidates. All candidates are restricted to spending the publicly funded money and it is assumed that every candidate will spend the full allotment of public funds on the election. Therefore, the funding level replaces the actual expenditure level for all the candidates. This works as a cap against a candidate who spent more than the funding level and as a subsidy for any candidate who spent below the funding level. Again, this experiment is done for 100 different funding levels from $10,000 to $1,000,000 and the predicted results are presented in Figures 1.4 and 1.5.

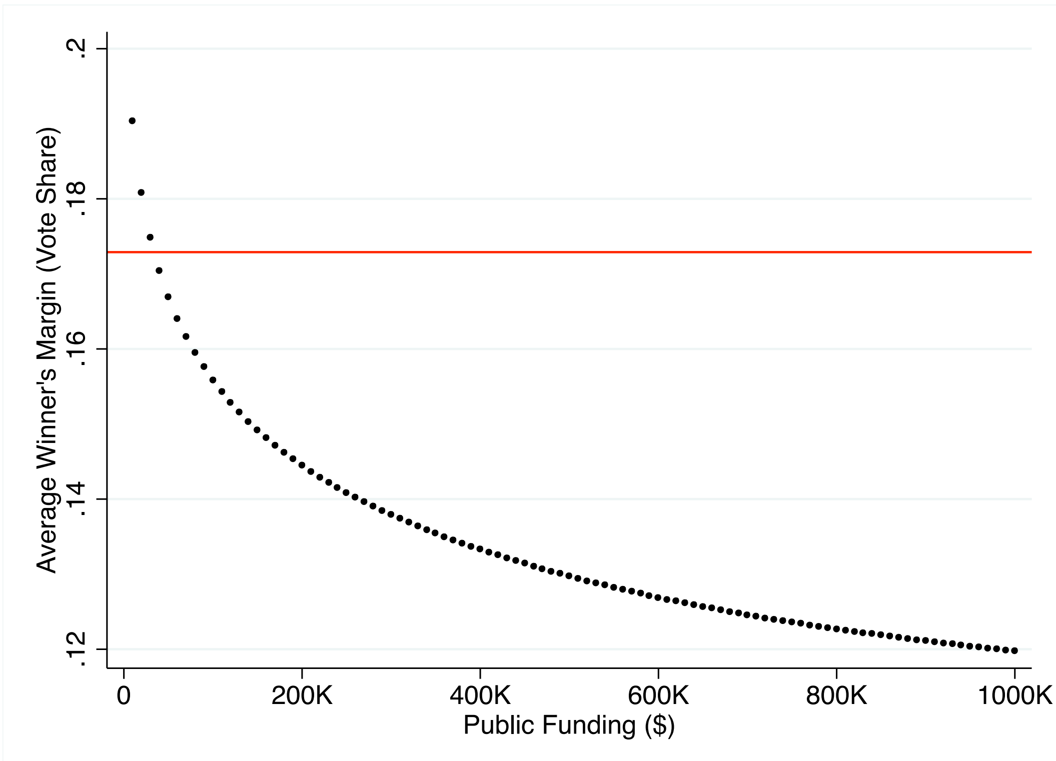


Figure 1.4: Predicted Average Winner's Margin of Victory Under Different Public Funding Levels

A major difference in this experiment from the expenditure cap is that this reform only analyzes the effects in general elections with both a Republican and Democratic candidate. This restriction is imposed because a public funding policy such as this would induce more candidates to run for the office

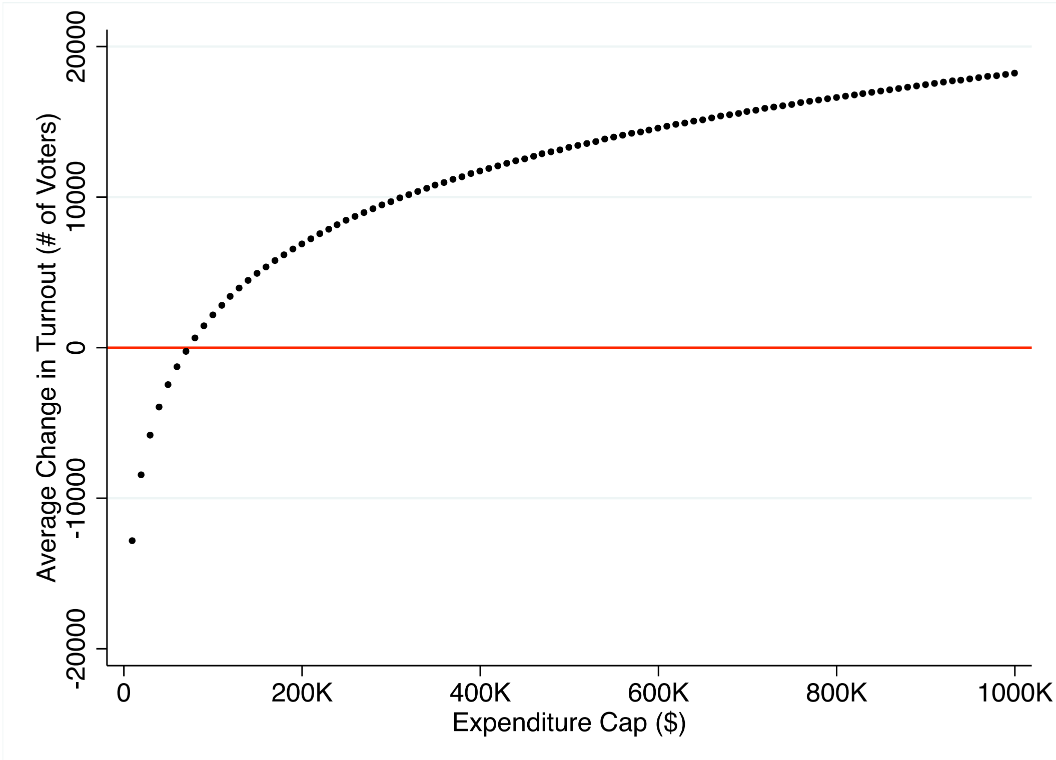


Figure 1.5: Predicted Average Change in Number of Voters Under Different Public Funding Levels

even if the policy included an entry stipulation. The model does not account for the endogenous decision whether to run for office or not, so this experiment makes the comparison only for districts with two candidates in the general election. Modeling the decision to run would provide more sophisticated results for this experiment, however, that is left for future work.

Figure 1.4 shows that this policy has a drastic effect on the average winner's margin of the elections. For very low levels of public funding, the average winning margin is higher than the baseline margin of .1729.[23] However, the average winner's margin quickly drops below the baseline and falls below .12 at a public funding level of $1,000,000. This is about a 30% decrease in the average winning margin from the baseline model.

The effects on voter turnout are just as striking. Figure 1.5 measures the increase in the number of voters as compared to the baseline model on the vertical axis. Considering that an average candidate spends over $800,000 in the baseline model, this figure shows that even relatively low levels of public funding can have a substantial increase in voter turnout. A public funding level of $200,000 is predicted to induce about 7,000 additional voters per district, which is close to a 2% increase in overall voter turnout. A public funding level near the average candidate expenditure of $800,000 is predicted to increase overall voter turnout by about 4%.

Both effects of the public funding policy are generally desirable qualities for elections. However, the cost of implementing this type of policy would be very high. Further research will explore whether the welfare benefits of this type of policy outweigh the costs of providing the public funds.

1.7 Conclusion

This chapter develops a model aimed at providing better policy analysis of potential campaign finance reforms by incorporating primary elections in a

[23]The baseline margin is lower in this experiment because the policy analysis is only done for districts with both major party candidates running.

structural election model. The method links the data and theory in a direct way that is based on individual agents making utility maximizing choices. A theoretical election cycle model and an unique data set of primary election expenditures provide a test that shows that voters do not value the primary election information. This chapter also produces very interesting predictions on the effects from two separate campaign finance reforms. An expenditure cap is predicted to have generally undesirable results while a public funding policy would reduce the average margin of victory and greatly increases voter turnout. The public funding policy would be quite expensive to implement so future research will determine if this policy is worth the costs. In addition to analyzing campaign finance reform policies, future research can adapt this model to study new primary structures and their effects on general elections.

Chapter 2

Barriers to Entry in U.S. Congressional Elections

2.1 Introduction

During the 2004, 2006, and 2008 U.S. House of Representative elections, the winning candidate spent an average of about $1,300,000.[1] In 13% of these elections, a candidate was not facing any competition in the general election. Even with no competition for the house seat, these candidates still spent an average of over $800,000. Why would a candidate spend this much money when running unopposed in the general election?[2]

The purpose of this chapter is to show that a candidate's campaign expenditures while running unopposed increases the probability of him running unopposed in the subsequent election. The intertemporal effect of unopposed expenditures create a barrier to entry for potential challengers by increasing the cost of running for the election. As in any market, barriers to entry in

[1] All values are reported in 2008 dollars.

[2] Almost all of these candidates who ran unopposed in the general election did not face any significant competition in the primary election either.

elections are generally undesirable. In this case, a barrier that prevents challengers to enter removes the threat of not being reelected for the incumbent. This chapter adapts the citizen-candidate model of Osborne & Slivinski (1996) to develop a single threshold model that predicts an incumbent to run unopposed on one side of the threshold and to face a challenger on the other side. This threshold is based on the relationship between the costs of running for election and the benefit of holding office. The relative costs are assumed to depend on previous election expenditures and current district characteristics. The influences of these characteristics on the probability of an incumbent running unopposed is estimated using the dynamic panel data method introduced by Wooldridge (2005).

Previous work examining deterrence to political campaign entry has focused on the incumbents' war chest. See Box-Steffensmeier (1996) and Goodliffe (2001). Both of these studies estimate the effect of the amount of cash on hand on the likelihood of running unopposed. In contrast to this work, this chapter quantifies the effects of prior expenditure patterns instead of the cash on hand. Prior expenditures are a more appropriate measure to use when looking for deterrence in political campaigns because the actual expenditure is observed behavior while the war chest is a measure of potential expenditure. Goodliffe (2001) explains all the deficiencies in using the war chest measure and finds no effect on the entry decision by challengers.

As mentioned previously, the theoretical model for this chapter is adapted from Osborne & Slivinski (1996). In their model, all the citizens simultaneously make the decision of whether or not to enter the political race. This chapter adapts their model by assuming that an incumbent candidate exists and is seeking reelection. The incumbent's policy position is known to all other citizens who decide whether or not to challenge the incumbent. One of the main results of their model is, for a certain range of parameter values, both a single candidate running unopposed and two candidates running in the election are equilibrium outcomes. This chapter adapts their model so there is only a single threshold, not a range of parameter values, that delineates where the model predicts more than one candidate running for election versus predicting the incumbent runs unopposed.

The single threshold allows the model to be estimated in a similar fashion to Bresnahan and Reiss (1991). Their paper recognizes that it is very difficult to identify which particular firms enter and exit a market. Instead, they use a theoretical model to predict the total number of firms in a market. This idea to applied political campaign entry. It is very difficult to identify all of the candidates' political positions. However, it is easy observe the number of candidates running for the office. This chapter estimates the effects of prior campaign expenditures on the likelihood of seeing only one candidate in the election versus two candidates.

The contribution of this work is that it quantifies an intertemporal effect of prior campaign expenditures on challenger entry using a simple theoretical model. The dynamic panel data method from Wooldridge (2005) is utilized since the empirical model is nonlinear and the lagged dependent variable is a regressor of interest. The model and results provide empirical evidence that unopposed campaign expenditures act as a deterrent to challenger entry in the next election.

The estimation results show that an increase in unopposed campaign expenditures increases the probability that the incumbent candidate will run unopposed in the next election. On average, an unopposed candidate that spends $10,000 more increases the probability of running unopposed in the subsequent election by about 0.0045. These results are estimated controlling for both observed and unobserved district characteristics. The results also show that there is a high level of state dependence across elections in the same district. This means that an election is much more likely to be unopposed if the prior election in the district was unopposed.

The organization for the remainder of this chapter is as follows. Section 2.2 presents the challenger entry decision model. Section 2.3 derives the estimation equation and explains the estimation method used for this analysis. Section 2.4 explains the unique data set gathered for this analysis. Section 2.5 provides and interprets the empirical results. Section 2.6 concludes.

2.2 Challenger Entry Decision Model

The citizen-candidate model from Osborne & Slivinski (1996) is adapted to examine the decision making process of potential challengers to an incumbent running for reelection. Different from the Osborne & Slivinski (1996) model, it is assumed that an incumbent is present with an observed policy position. All other citizens in the district must then decide whether to enter the congressional race. This decision is based on their own policy position, the cost of running for election and the benefit of winning the election and holding office.

The model predicts a single threshold that separates the pure Nash equilibria into two distinct types. On one side of the threshold, the model predicts an incumbent would face a challenger whose policy position is constrained by the district parameters. On the other side of the threshold, the model predicts that no challenger will enter the race and therefore the incumbent runs unopposed. The model presented in this chapter is a stylized model that makes assumptions to keep the model simple and maintain the focus on the threshold between opposed and unopposed elections. This chapter will only identify pure Nash equilibrium of a single challenger entering the race and will not characterize all the mixed Nash equilibrium in this model.

2.2.1 Theoretical Model

Assume that there is a continuum of citizens in each congressional district and each citizen is denoted by their ideal policy point, x_j. Let $F(X) \sim U[0,1]$ be the distribution of the citizens' ideal points and m be the median of this distribution. Each citizen has the utility preference for representation of $u(x_j; a) = -|x_j - a|$ where a is the policy position of the elected representative for that district. Each citizen makes the decision of whether to run for election or not. If a citizen runs for election, he incurs a utility cost of running, $c > 0$. The winner of the election then receives a utility benefit, $b > 0$, of holding the public office.

Assume that there exists an incumbent candidate and, without loss of generality, his policy is given by $x_I = m - \epsilon$, assuming $\epsilon < \frac{1}{3}$. This position is assumed to be fixed and known by all other citizens in the district. Given the presence of the incumbent candidate, all citizens simultaneously make the entry decision and the winner is determined by the the candidate who receives the the most votes. It is assumed that all citizens vote and they vote "sincerely". This means that each citizen votes for the candidate with the policy position that is closest to their own. In the event that an incumbent and a challenger receive an equal amount of votes, assume that the challenger wins the election.

2.2.2 Single Challenger Pure Nash Equilibria

This first proposition shows the main result of the theoretical model. At or above the threshold, $(b - c) \geq -2\epsilon$, there exists a pure Nash equilibrium where a citizen with ideology opposite of the incumbent chooses to run for the election. This shows that the model predicts the incumbent to be challenged at all parameter values above the threshold. Proposition 2 then characterizes all the equilibria near the threshold to show that the challenger can have a range of policy positions as the parameters get close to the threshold.

Proposition 1. When $(b - c) \geq -2\epsilon$, a single citizen, $x_j = m + \epsilon$ enters and all other citizens do not enter is a pure strategy Nash equilibrium.

Proof. Assume that all other citizens do not enter. $x_j = m + \epsilon$ enters the race iff:

$$\begin{aligned} \text{Utility from entering the race} &\geq \text{Utility from not entering} \\ (b - c) - |x_j - x_j| &\geq -|x_j - x_I| \quad \Rightarrow \\ (b - c) - |0| &\geq -|m + \epsilon - (m - \epsilon)| \\ (b - c) &\geq -2\epsilon \end{aligned} \tag{2.2.1}$$

This shows the threshold for when x_j has an equilibrium strategy of entering the race.

Now assume that x_j chooses to enter the race. I will now show that $x_k = [0, 1]$ will choose not to enter the race.

Case 1: $x_k \in [0, x_I] \cup [x_j, 1]$

Any x_k in this range will lose the election for certain and the citizen would receive $-c - |x_k - a|$ which is always less than $-|x_k - a|$.

Case 2: $x_k \in (x_I, x_j)$

The citizen in this range that would receive the maximum number of votes is $x_k = m$ and his vote total is:

$$Votes(x_k) = \frac{1}{2}(m - (m - \epsilon)) + \frac{1}{2}((m + \epsilon) - m) = \epsilon$$

The vote totals for x_j and x_I would both be the same.

$$Votes(x_I) = Votes(x_j) = m - \epsilon + \frac{1}{2}(m - (m - \epsilon)) = m - \frac{1}{2}\epsilon$$

Given the distribution assumption $F(X) \sim U[0,1]$, $m = \frac{1}{2}$. For x_k to win the election or tie with the other challenger:

$$\epsilon \geq m - \frac{1}{2}\epsilon \quad \Rightarrow \quad \frac{3}{2}\epsilon \geq \frac{1}{2} \quad \Rightarrow \quad \epsilon \geq \frac{1}{3}$$

However, it is assumed that $\epsilon < \frac{1}{3}$, so x_k would always lose the election for certain and will choose not to enter.

These two cases show that no other citizen has an incentive to run for election and completes the proof. □

Proposition 1 proves that a Nash equilibrium of challenger entry exists with no upper bound. Proposition 2 completely characterizes all the pure Nash equilibria above this threshold up to $(b - c) < -\epsilon$.

Proposition 2. When $-\epsilon > (b - c) \geq -2\epsilon$, a single citizen, x_j, chooses to enter the race and x_j can be any value in $[m - \epsilon - (b - c), m + \epsilon]$.

Proof. Any citizen with $x_j > m + \epsilon$ will not enter because they will lose the election to incumbent with certainty. By the same logic, $x_j < m - \epsilon$ will never enter the race.

Let us find the citizen, $x_j = m - \epsilon + \Delta$ such that 1) x_j is indifferent between entering and not entering the race, and 2) $x_j \leq m + \epsilon$.

Condition 1) requires that the utility from entering and winning for certain is greater than or equal to the utility form not entering and having the incumbent as a representative.

$$\begin{aligned}(b-c)-|x_j-x_j| \geq -|x_j-x_I| &\Rightarrow (b-c) \geq -|m-\epsilon+\Delta-(m-\epsilon)| \\ &\Rightarrow b-c \geq -\Delta \\ &\Rightarrow \Delta \geq -(b-c) \quad (\text{ Condition 1) })\end{aligned}$$

This condition represents the fact that a challenger only wants to enter the election if he is sufficiently far enough away from the incumbent. The utility increase from representing himself must be greater that relative cost of running for election.

Condition 2) requires that the challenger candidate enters in a position close enough to the median so that he always wins the election.

$$\begin{aligned}x_j \leq m+\epsilon &\Rightarrow m-\epsilon+\Delta \leq m+\epsilon \\ &\Rightarrow \Delta \leq 2\epsilon \quad (\text{ Condition 2) })\end{aligned}$$

Combining Conditions 1) and 2) provides the lower bound, in terms of the parameters, for when a challenger will choose to enter.

$$-(b-c) \leq 2\epsilon \quad \Rightarrow \quad (b-c) \geq -2\epsilon$$

To find the upper bound of the equilibria , x_j will only choose to enter such that his position does not allow another citizen to enter and win the election. As shown from Condition 1), a citizen will only want to enter if they are at least Δ away from the potential representative. Therefore, Condition 3) is that given $x_j = m - \epsilon + \Delta$ enters the race, citizen $x_k = m - \epsilon + 2\Delta$ will not be able to win the election. If x_k enters, he loses if his vote total is less than the incumbent's vote total.

$$\begin{aligned}
Votes(x_I) > Votes(x_k) \quad &\Rightarrow \quad m - \epsilon + \frac{1}{2}\Delta > 1 - (m - \epsilon + \frac{3}{2}\Delta) \\
&\Rightarrow \quad -\epsilon + \frac{1}{2}\Delta > \epsilon - \frac{3}{2}\Delta \\
&\Rightarrow \quad 2\Delta > 2\epsilon \\
&\Rightarrow \quad \Delta > \epsilon \quad (\text{ Condition 3) })
\end{aligned}$$

Condition 3) says that the single candidate must enter at a position to the right of the median to ensure that no other citizen would want to run. To complete the proof, the upper bound of the equilibria, in terms of the model parameters, is the combination of Conditions 1) and 3).

$$-(b - c) > \epsilon \quad \Rightarrow \quad (b - c) < -\epsilon \tag{2.2.2}$$

□

Proposition 2 provides very intuitive results for when we would expect to see an incumbent run unopposed. When the relative utility benefits to costs, $(b - c)$, of holding office is negative, a citizen will only run if the change in having a representative with closer ideology outweighs the relative costs. Equation 2.2.1 provides the lower bound for when a challenger would choose to run. Below this threshold, it does not benefit any citizen to run and oppose the incumbent.

Figure 3.4 summarizes these results. Whenever $(b-c)$ is less than -2ϵ, the model predicts that no challenger will enter the election and the incumbent will run unopposed. Whenever $(b-c)$ is greater than -2ϵ, a pure Nash strategy of a challenger entering the election to oppose the incumbent exists. The complete characterization of equilibria when $(b-c) > -\epsilon$ is left for future research because it involves the existence of a continuum of mixed Nash equilibria.

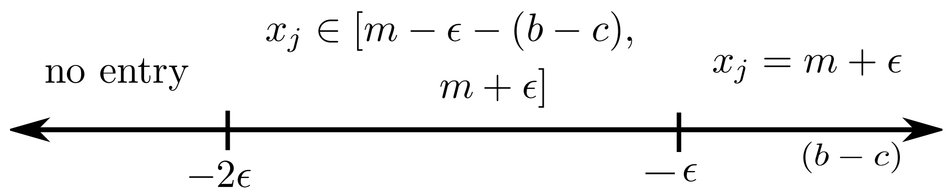


Figure 2.1: State Space

Note: This figure shows the model predictions for pure Nash equilibria. The relative benefits to cost, $(b-c)$, varies from $-\infty$ to $-\infty$ and the important threshold points are denoted at -2ϵ and $-\epsilon$. Below -2ϵ, the model predicts no challenger will enter the race. Above -2ϵ, a pure Nash equilibrium of a single citizen entering the race exists. Additionally, Proposition 2 showed that the single entrant can have any policy position in the closed range $[m - \epsilon - (b-c), m + \epsilon]$ when $(b-c) < -\epsilon$.

2.3 Econometric Model and Estimation Method

The theoretical model provides a single threshold for when the model predicts a challenger to the incumbent candidate. This threshold is based on the parameters that vary by congressional district. This section will use the threshold to derive the probability that an incumbent runs unopposed. This probability is assumed to be based on the incumbent expenditure patterns in the previous election, observed district demographic characteristics and unobserved district fixed effects. The proper estimation of the derived empirical model necessitates the use of dynamic panel data methods to deal with the fact that the current probability of being unopposed depends on the whether the previous election was unopposed.

2.3.1 Estimation Equation

The probability that an incumbent runs unopposed is the probability that the relative benefits to costs is less than negative two times the distance between the median policy position and the incumbent's policy position. All of these parameters vary by district and election year so let subscript dt denote district d at time t.

$$\begin{aligned} Prob(Unopposed_{dt}) &= Prob((b_{dt} - c_{dt}) \leq -2\epsilon_{dt}) \\ &= Prob((b_{dt} - c_{dt}) + 2\epsilon_{dt} \leq 0) \\ &= Prob((c_{dt} - b_{dt}) - 2\epsilon_{dt} > 0) \end{aligned}$$

Since c_{dt} and b_{dt} are both unobservable utility values, assume that the difference in these values is linear with respect to the prior election characteristics, $X_{d,t-1}$, current district demographics, Z_{dt}, a district specific fixed effect, ω_d, and a time varying residual term, e_{dt}.

$$Prob(Unopposed_{dt}) = Prob(\beta_0 X_{d,t-1} + \beta_1 Z_{dt} + \omega_d + e_{dt} - 2\epsilon_{dt} > 0) \quad (2.3.1)$$

$$\text{with } X_{d,t-1} \equiv \begin{bmatrix} ln(Incumbent\ Expend)_{d,t-1} \\ ln(Incumbent\ Expend)_{d,t-1} * Unopposed_{d,t-1} \\ Unopposed_{d,t-1} \end{bmatrix}$$

Lastly, assume that the combination of unobservables, $(e_{dt} - 2\epsilon_{dt})$, is distributed standard normal. With these assumptions, the estimation equation for the theoretical model is a dynamic panel probit model with a district specific fixed effect.

$$Prob(Unopposed_{dt}) = \Phi(\beta_0 X_{d,t-1} + \beta_1 Z_{dt} + \omega_d) \quad (2.3.2)$$

2.3.2 Estimation Method

The estimation of this model is complicated by the fact that $X_{d,t-1}$ includes the lagged dependent variable, $Unopposed_{d,t-1}$. To consistently estimate the parameters, this analysis will use the method introduced by Wooldridge (2005) that assumes the district specific constant can be estimated as a function of the initial condition and all the time varying coefficients. One of the main advantages of using this method is that it allows for the consistent estimation of a dynamic panel data model with a fixed effect using standard software packages.

To implement the Wooldridge (2005) method, assume that

$$\omega_d = \alpha_0 + \alpha_1 \bar{X}_d + \alpha_2 \bar{Z}_d + a_d \tag{2.3.3}$$

$$\text{with } \bar{X}_d \equiv \begin{bmatrix} \overline{ln(Incumbent\ Expend)_d} \\ \overline{ln(Incumbent\ Expend)_d * Unopposed_{d,0}} \\ Unopposed_{d,0} \end{bmatrix}$$

$$\text{and} \quad a_d | (\bar{X}_d, \bar{Z}_d) \sim N(0, \sigma_a^2)$$

This assumes the district fixed effect is a linear function of the initial election dependent variable, $Unopposed_{d,0}$, and the mean of all the time varying covariates. Substituting Equation 2.3.3 into 2.3.2 provides the final estimation equation. It is a function of the district characteristics, the average over time of the district characteristics, and a random district component.

$$Prob(Unopposed_{dt}) = \Phi(\beta_0 X_{d,t-1} + \beta_1 Z_{dt} + \alpha_0 + \alpha_1 \bar{X}_d + \alpha_2 \bar{Z}_d + a_d) \tag{2.3.4}$$

The parameters of Equation 2.3.4 can be estimated using standard random effects probit methods.

2.4 Data

This chapter utilizes the data from the 2004, 2006, and 2008 U.S. House of Representative elections. The data is collected from two publicly available sources, the Federal Election Commission (FEC) and the American Community Survey (ACS) that is administered by the U.S. Census Bureau. The FEC provides the candidate campaign expenditures and the elections information.

The campaign expenditures used in this chapter are the values reported in the "Candidate Financial Summaries (All Candidates)" files.[3] These expenditures are the total amount spent throughout the full election cycle for every candidate. The election results summary files[4] provide the information on whether the incumbent candidate ran unopposed or not.

The ACS is a annual survey that provides demographic information. The novel feature of the ACS is that it provides aggregate demographic information for congressional districts. This is the first research to take advantage of the congressional district characteristics from this survey and allows the analysis to control for socioeconomic and demographic differences between districts. The ACS began providing congressional district level information in 2006. Therefore, this analysis uses 2004 as the initial year of observation since the demographic data is not needed in the initial year. The 2004 elections provide the lagged election characteristics for the 2006 elections and the 2006 elections provide the lagged characteristics for the 2008 elections.

This chapter restricts the number of congressional districts used in the empirical analysis. Congressional districts from the state of Louisiana are omitted because Louisiana does not have a standard election process. The Louisiana election cycle does not include primary elections and instead has a runoff election if a single general election candidate does not gain a majority of the vote. This is a very different environment from the rest of the U.S. and so these districts are not included. Also excluded are the five congressional

[3] http://www.fec.gov/finance/disclosure/ftpsum.shtml

[4] http://www.fec.gov/pubrec/electionresults.shtml

districts in Texas[5] that were affected by the mid-decade redistricting. Lastly, the analysis excludes open elections without an incumbent candidate. The model is based on challenger entry given the presence of an incumbent, so open elections are not represented by the model.

2.4.1 Control Variables

Since the model represents the decision by a challenger of whether to enter or not, the control variables are a combination of the lagged election characteristics and current congressional district demographics. The main variables of interest for this analysis are the incumbent's expenditure in the prior election and whether the previous election was unopposed. Table 2.1 shows the average incumbent expenditures conditional on whether the previous election was unopposed and also whether the current election is unopposed. This table shows that the incumbent candidate spends less money on average if he is running unopposed. However, the unopposed candidates still spend over $800,000. The table also shows that the majority of elections are opposed but there is substantial variation between opposed and unopposed elections.

Table 2.2 presents the summary statistics for the control variables that capture the demographics of the congressional district. The variables used in this analysis are the median household income, the gini index, total population, the area in square miles, and a dummy variable that represents whether a senate election is happening concurrently. The median income level and gini index are used to control for the income distribution in each district. The

[5]Districts 15, 21, 23, 25, and 28.

Table 2.1: Conditional Means of the Incumbent Expenditures ($)

	Currently Opposed	Currently Unopposed
Prior Opposed Election	1,397,178 (n = 604)	804,787 (n = 64)
Prior Unopposed Election	1,121,080 (n = 69)	912,740 (n = 39)

Note: This table presents a matrix of the mean expenditure by the incumbent candidate conditional on whether the prior election was unopposed and whether the current election is unopposed. The number of elections in each cell is reported in parentheses.

Table 2.2: Summary Statistics

	Mean	St. Dev.	Min	Max
Median Household Income ($)	52,123	13,907	21,088	99,811
Gini Index	.4465	.0335	.3650	.6120
Total Population	692,765	65,987	515,004	991,439
Area (sq. miles)	10,818	46,120	10	571,951
Senate Race	.6482	.4778	-	-
Number of Elections		776		
Number of Unique Congressional Districts		421		

Note: This table presents the summary statistics for the 2006 and 2008 congressional districts excluding Louisiana, the redistricted Texas districts, and any open seat elections.

median income controls for differences in overall levels of income and the gini coefficient controls for the spread of the income distribution. The total population and area of the district control for differences in population density that could affect the decision of whether to enter a race. Lastly, the senate race variable controls for any effects of having a higher seat being contested at the same time. This table shows that congressional district demographics vary along many dimensions and the ACS data allows this analysis to control for these factors when estimating the effect of prior expenditure patterns on the challenger's entry decision.

2.5 Results

The estimated coefficients from three different specifications of Equation 2.3.4 are presented in Table 2.3 to show how the Wooldridge (2005) method changes the coefficients. The first specification estimates the coefficients using pooled probit estimation. This estimates the parameters assuming that the data is cross sectional. The second specification estimates the model using random effects probit estimation. This specification utilizes the panel nature of the data but does not make any additional assumptions about the district specific heterogeneity, ω_d. Column (3) presents the coefficient results using the Wooldridge (2005) method. Again, this method assumes that the district specific heterogeneity can be estimated as a function of the mean of the district characteristics and the estimation method is done using a random effects probit routine. All three specifications omit the coefficients for the year and region dummy variables.

Table 2.3: Estimated Coefficients

VARIABLES	Pooled Probit (1)	RE Probit (2)	Wooldridge (2005) (3)
$ln(Incumbent\ Expend)_{d,t-1}$	-0.365***	-0.385**	0.353
	(0.135)	(0.150)	(0.252)
$ln(Incumbent\ Expend)_{d,t-1} * Unopposed_{d,t-1}$	0.699**	0.725**	0.792*
	(0.298)	(0.318)	(0.416)
$Unopposed_{d,t-1}$	-8.732**	-9.128**	-10.16*
	(4.027)	(4.330)	(5.632)
$ln(Median\ Household\ Income)_{dt}$	-0.700**	-0.743**	-0.477
	(0.300)	(0.337)	(3.403)
$Gini\ Index_{dt}$	0.340	0.291	-23.17*
	(1.994)	(2.110)	(12.72)
$ln(Total\ Population)_{dt}$	-2.608***	-2.693***	3.671
	(0.933)	(1.007)	(4.951)
$ln(Area)_{dt}$	-0.0891**	-0.0939**	-0.109**
	(0.0352)	(0.0393)	(0.0480)
$Senate\ Race_{dt}$	0.317**	0.327**	0.370**
	(0.150)	(0.157)	(0.177)
$Constant$	46.66***	48.55***	59.99***
	(12.33)	(13.92)	(18.74)
$ln(Incumb\bar{e}nt\ Expend)_d$			-1.156***
			(0.340)
$ln(Incumb\bar{e}nt\ Expend)_d * Unopposed_{d,0}$			0.126
			(0.464)
$Unopposed_{d,0}$			-1.372
			(6.272)
$ln(Median\ Hou\bar{s}ehold\ Income)_d$			-0.205
			(3.446)
$ln(Total\ \bar{P}opulation)_d$			-6.899
			(5.180)
$Gini\ \bar{I}ndex_d$			25.10*
			(12.99)
Observations	776	776	776
Number of Districts		421	421

Note: This table presents the coefficient results, with standard errors in parentheses, from three specifications to show how the Wooldridge (2005) method affects the coefficients. The asterisks indicate significance levels of the coefficients, *** p<0.01, ** p<0.05, * p<0.1 and all specifications are reported omitting the coefficients for the region and year fixed effects. Column (1) estimates the model using pooled probit estimation. Column (2) estimates the model using standard random effects probit estimation. Column (3) estimates the model using the Wooldridge (2005) method.

The first three rows of coefficients are the main coefficients of interest for this analysis. The first coefficient, on $ln(Incumbent\ Expend_{d,t-1})$, is the effect of incumbent expenditure from the previous election when the prior election was opposed. This coefficient is significantly negative in the first two specifications and insignificantly positive in the third specification. This implies that, once district specific unobserved characteristics are adequately controlled for, opposed incumbent expenditure has an ambiguous effect on the probability that the subsequent election is unopposed. The second coefficient, on $ln(Incumbent\ Expend_{d,t-1}) * Unopposed_{d,t-1}$, represents the intertemporal effect of incumbent expenditure in the previous period when the prior election was unopposed. This coefficient is significantly positive in all three specifications. This provides evidence that unopposed incumbent expenditure increases the probability of running unopposed in the next election and subsection 2.5.1 shows the magnitude of effect by estimating the incumbent expenditure elasticities.

The third coefficient, on $Unopposed_{d,t-1}$, represents the effect of having the previous election be unopposed. The coefficient is estimated significantly negative in all specifications. The significance of the coefficient implies that there is state dependence for elections within a given district. State dependence means that the probability of an unopposed election significantly changes based on whether the prior election was opposed or unopposed. However, the negative coefficient does not mean that there is negative state dependence because $Unopposed_{d,t-1}$ is interacted with the incumbent expenditure. The next subsection shows that there is positive state dependence between elections in the same district.

The rest of the coefficients in the table are estimates of how district characteristics influence the probability of an unopposed election. Most of these are not statistically significant in the third specification because it is difficult to differentiate the observed effect of the regressor from the unobserved effect, ω_d. Since ω_d is estimated as a combination of the means of the observed characteristics, there is a close relationship between the observed and unobserved district characteristics that requires more variation to significantly estimate all the coefficients. However, analyzing all three specifications leads to some general conclusions about how the district characteristics affect the number of candidates in the election. The median household income has a generally negative relationship with unopposed elections which means we would expect to see more unopposed elections in less affluent districts. The significant negative coefficient on the gini index in the third specification implies that as the income inequality in a district increases, there is a smaller chance of an unopposed election in that district. Lastly, the negative coefficient on the district area implies that unopposed elections occur more frequently in smaller districts.

2.5.1 Average Partial Effects

The average partial effects are estimated to show the magnitude of state dependence and the magnitude of effect from incumbent expenditures. The average effect is calculated for six different example districts, previously opposed and unopposed elections when prior incumbent expenditure is set to the 25th, 50th and 75th percentile. These percentiles equal \$687,307, \$981,576,

Table 2.4: Average Partial Effects

Expenditure Percentile	25th	50th	75th
Predicted Probability of Unopposed Election			
Prior Opposed Election	.1420	.1563	.1764
Prior Unopposed Election	.2147	.2802	.3793
Estimated Effect of Prior Unopposed Election			
	.0665***	.1193***	.1945***
	(.0025)	(.0034)	(.0049)
Estimated Incumbent Expenditure Elasticity (%)			
Prior Opposed Election	.7407***	.6920***	.6336***
	(.0292)	(.0267)	(.0238)
Prior Unopposed Election	1.9486***	1.6079***	1.2342***
	(.0394)	(.0290)	(.0192)

Note: This table presents the predicted probability of an opposed or unopposed election, estimated effect of a prior unopposed election and the estimated incumbent expenditure elasticity. The partial effects are calculated for the 25th, 50th, and 75th expenditure percentiles. Bootstraped standard errors are in parentheses and the asterisks indicate significance levels of the coefficients, *** p<0.01.

and $1,522,870 respectively. All of the other control variables are set to their sample means and these estimates average out the unobserved district characteristic. Table 2.4 summarizes the predicted probability of an unopposed election for each example district and presents the estimated average partial effects.

The average partial effect from a prior unopposed election is the magnitude of state dependence. This equals the estimated difference in predicted probability from an opposed to an unopposed election and is calculated for each expenditure percentile. The change from opposed prior election to unopposed is estimated to be positive and significant for all three expenditure percentiles. This shows that there exists positive state dependence between elections in a district. A prior unopposed election increases the probability of a subsequent unopposed election by .05 to .20 depending on the amount of money spent by the unopposed candidate.

The incumbent expenditure average partial effect is easily interpreted as an elasticity. The average expenditure elasticity is estimated for each of the six example congressional districts and are all positive and significant. The elasticity shows the effect of a 1% incumbent expenditure increase on the probability of running unopposed in the subsequent election. For example, the estimated expenditure elasticity of a unopposed incumbent who spends in the 50th percentile, $981,576, is 1.6079%. The predicted probability of running unopposed in the next election is .2802. Therefore, a 1%, or about an additional $10,000, expenditure increase is predicted to increase the probability of running unopposed in the next election to .2847. The elasticity predicts a 1% expenditure increase leads to a .0045 increase in the probability of running unopposed. Even though the elasticity decreases by percentile, the predicted increase in probability stays relatively constant at .0045. This shows that incumbent expenditures while running unopposed increase the probability of running unopposed in the subsequent election.

2.6 Conclusion

This chapter empirically quantifies an intertemporal effect of U.S. House of Representative campaign expenditures. It shows one reason unopposed candidates spend a large amount of money is to increase the probability of running unopposed in the next election. The candidates' expenditure effectively creates a barrier to entry against potential challengers in the next election by increasing the relative cost of running for election. Future research will incorporate these results when designing campaign finance reform proposals. This chapter's results suggest that campaign finance reform would not only effect the outcome of current elections, it would also change the composition of future elections.

Chapter 3

What is a U.S. House Seat Worth? Estimating the Value Using Auction Theory

Preface

This chapter was coauthored with Nick Mastronardi and Scott Macdonell. The author modified the model originally developed by Nick Mastornardi, collected and managed the data, performed all the empirical analysis, and wrote this version.

3.1 Introduction

What is the the optimal amount to spend on a U.S. House of Representatives election? While one common answer is "As much as possible", campaign expenditure filings shows that candidates generally do not expend their full budget on an election. Another common answer would be "More than the other candidate." However, the candidates do not know the amount that their opponent is going to spend. The optimal campaign expenditure will obviously depend upon the value of a House seat. A formal model of expenditure strategies based upon valuations will allow us to answer the question of what a House seat is worth.

We model the candidate expenditure decision as a bid in a first-price asymmetric all-pay auction, solve for the mixed Nash expenditure strategies, and estimate the value of a House of Representative seat based on the observed campaign expenditure behavior. We model the campaign expenditure game as an asymmetric all-pay auction for three reasons. First, it is an auction because electoral outcomes are a discrete set. Second, since losing candidates do not recover their bids (campaign expenditures), the auction is all-pay. Finally, it is asymmetric as district demographics may favor one candidate. The model predicts mixed Nash expenditure strategies that are highly intuitive and yield realistic electoral outcome probabilities. In the limit, as the fraction of voters who respond to expenditures diminishes, the model returns the classic Median Voter Model.

Other work that models the optimal amount of campaign expenditure include Prat (2002b) and Erikson & Palfrey (2000). Prat (2002b) develops a structural model of campaign spending in which advertising sways pivotal voters. Like previous models, he assumes that candidates always exhaust their entire budget. As observed in the data, candidates often spend much less or much more than their official budget. In this chapter, we seek to explain expenditure behavior without assuming that candidates always expend exactly their budget. To do so, we model expenditures as bids in an all-pay auction. Erikson & Palfrey (2000) develop a simple model of optimal campaign expenditures to theoretically explain the bias when estimating the impact of campaign spending. There model assumes away the existence of mixed Nash equilibrium and therefore cannot capture the large variation in observed spending when the district characteristics are similar. We follow both of these papers in assuming that campaign expenditure decisions are made simultaneously[1]

This chapter utilizes previously derived auction results to model the campaign expenditures as bids. Amann & Leininger (1996) solve generic mixed Nash equilibrium strategies for asymmetric all-pay auctions. Che & Gale (1998) apply the equilibrium analysis to a political lobbying setting and show how the lobbying expenditures would change under different levels of a campaign contributions cap. Neither of these papers use the model predictions to estimate the parameters of the bid distributions. Hendricks & Porter (1988) empirically analyze asymmetric mixed Nash bid strategies by oil companies on federal off-shore tracts. Their paper uses the ex-post realization of the object value to test whether bid behavior was competitive or collusive. In this chapter, we use the observed expenditure behavior to estimate the object valuation, the per partisan valuation of a House seat.

The contribution of this chapter is to produce an empirically testable model in which candidates do not necessarily expend their entire budgets and

[1]Allowing for a dynamic expenditure decision would be much more realistic as candidates do observe their opponents expenditures throughout the race. However, we leave that model to future research.

which accounts for large variation in expenditure behavior between congressional districts with similar demographics. As far as the authors are aware, this is the first research to empirically estimate an auction model when the underlying data is not produced from an actual auction. We differ from previous empirical auction work by estimating the parameter of the model using the moments of the predicted distributions instead of using Maximum Likelihood techniques. We use the first two predicted moments of the optimal expenditure distributions to estimate candidates' valuation of a House seat. We estimate a House seat to be worth about \$4.5 million.[2]

The organization for the remainder of this chapter is as follows: Section 3.2 explains the model, derives the equilibrium, and explains its intuition. Section 3.3 explicitly outlines the testable predictions of the model. Section 3.4 describes the unique data set and the sample restrictions used to estimate the value of the House seat. Section 3.5 presents the results and Section 3.6 concludes.

3.2 Finance Augmented Median Voter Model (FAMVM)

In this model, campaign expenditures earn candidates a fraction of the advertising-responsive voters. Modeling expenditures in this fashion implies that candidates' expenditures are bids in an asymmetric all-pay auction where the object of valuation is the elected office. This is appropriate because in

[2] All dollar amounts are in 2008 terms.

an election, the set of outcomes is a discrete set. Further, we use the all-pay auction framework since the losing candidates' expenditures are never recovered. Last, we allow asymmetry in the model to accommodate districts where one candidate has an advantage. We assume that the candidates' moves are simultaneous and information is complete.

The players are a congressional district's two campaigning candidates, one Democratic candidate denoted with subscript D and one Republican candidate denoted with subscript R. Each candidate simultaneously chooses a level of campaign expenditure, e_D and e_R. In our model, candidates are assured of the support of their base of voters, b_p. Candidates behave as partisan social planners, attempting to maximize the payoffs to their base. There is a common per partisan value of the House seat, v. Therefore if she wins, a candidate's payoff consists of the total value of the seat to their base, $b_p v$, minus their expenditures. If they lose the election their payoff is equal to the negative of their expenditures.

A candidate gets elected to office if they accumulate more votes than their opponent. Auction asymmetries enter in that each candidate may have a different mass of resolute voters behind them, $b_D \neq b_R$. Swing voters, denoted μ, respond to the relative amounts of candidate expenditures (e.g. advertising). We assume the simplest case of a linear proportional advertising technology. So, for example, candidate P_D's total number of votes received is $b_D + \mu \frac{e_D}{e_D + e_R}$. Table 3.1 summarizes the set up of this auction.

Given a district environment (b_D, b_R, μ) and a v value, the model pre-

Table 3.1: Notation Summary

Candidates	P_D	P_R
Resolute Voters	b_D	b_R
Expenditures	e_D	e_R
Total Votes	$b_D + \mu\frac{e_D}{e_D+e_R}$	$b_R + \mu\frac{e_R}{e_D+e_R}$
Objective Function	$\max[b_D v * Prob(win_D \vert e_D) - e_D]$	$\max[b_R v * Prob(win_R \vert e_R) - e_R]$

dicts candidate expenditures. The optimal expenditure strategy is a probability distribution over a continuum of expenditure levels. From this point on, we assume without loss of generality that the Republican candidate has a base advantage, $b_R > b_D$, and is referred to as the advantaged candidate. For notational simplicity, define $b \equiv b_R - b_D$ as her net base advantage. The Democratic candidate is the disadvantaged candidate.

3.2.1 Campaign Expenditure Equilibria

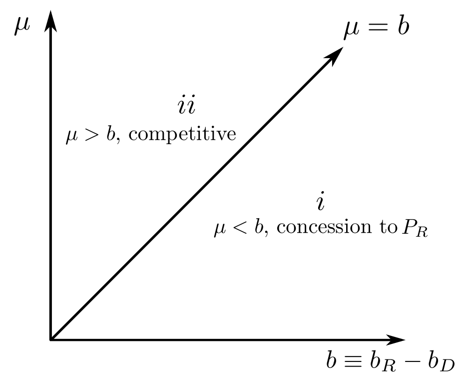
First, consider the simplest case, when the base advantage is larger than the mass of advertising responsive voters. Call this case i. Even if the disadvantaged candidate wins all the swing mass, he would still lose.

Theorem 3.2.1. *If $\mu < b$, in the unique Nash equilibrium neither candidate expends, and therefore the advantaged candidate will win by concession.*

Proof. There is no benefit to expenditures by either candidate, as P_R will always win.

$$b_R - b_D > \mu \Rightarrow b_R > b_D + \mu$$

Figure 3.1: The District State Space



Therefore, any expenditures by either player only decreases that player's expected payoff. □

Now, we consider the slightly more complicated, but more interesting case: when the mass of advertising responsive voters is potentially pivotal. Now $\mu > b$. This is region ii on the state space graph, Figure 3.1. Define $s \equiv \frac{\mu - b}{\mu + b}$ and $\gamma \equiv \frac{b_D}{b_R}$. Before continuing, we prove the following two propositions:

Proposition 3.2.2. *The disadvantaged candidate expends no more than their valuation $b_D v$.*

Proof. Spending more than her valuation is strictly dominated by spending zero. The former always results in a negative payoff, whereas the later must result in a non-negative payoff. □

Proposition 3.2.3. *The advantaged candidate never spends more than $b_D vs$.*

Proof. Given Proposition 3.2.2, $e_R = b_D vs$ will result in P_R tying with P_D. Thus, for any $\epsilon > 0$, spending $b_D vs + \epsilon$ or $b_D vs + \frac{\epsilon}{2}$ both provide a guaranteed electoral victory, but the later does so at a lower cost. In other words, all expenditures strictly greater than $b_D vs$ are strictly dominated. □

Theorem 3.2.4. *If $\mu > b$, the following pair of CDFs is a mixed Nash equilibrium to the game:*

$$G_D(e_D) = \begin{cases} 0 & \textit{if } e_D < 0 \\ 1 - \gamma s + \frac{e_D s}{b_R v} & \textit{if } e_D \in [0, b_D v] \\ 1 & \textit{if } e_D > b_D v \end{cases} \tag{3.2.1}$$

$$G_R(e_R) = \begin{cases} 0 & \textit{if } e_R < 0 \\ \frac{e_R}{b_D vs} & \textit{if } e_R \in [0, b_D vs] \\ 1 & \textit{if } e_R > b_D vs \end{cases} \tag{3.2.2}$$

Proof. Let each party's's expenditure CDF be denoted by G_p. The candidate who earns more total votes wins the election. Simple algebra shows that P_D wins if $e_D s > e_R$, and similarly for P_R. Thus the maximization problems faced by the players are:

$$P_D : \max_{e_D} \quad b_D v G_R(e_D s) - e_D \tag{3.2.3}$$

$$P_R : \max_{e_R} \quad b_R v G_D(\frac{e_R}{s}) - e_R \tag{3.2.4}$$

In equilibrium each player must receive a constant expected payoff (denoted k_p) from any level of expenditure in their support. Combining this with a change of variables yields:

$$G_D(e_D) = \frac{k_R + e_D s}{b_R v} \tag{3.2.5}$$

$$G_R(e_R) = \frac{k_D + \frac{e_R}{s}}{b_D v} \tag{3.2.6}$$

Proposition 3.2.2 allows us to pin down k_R by setting $G_D(b_D v) = 1$. Plugging this into Equation 3.2.5 and solving gives us

$$k_R = (b_R - b_D s)v \tag{3.2.7}$$

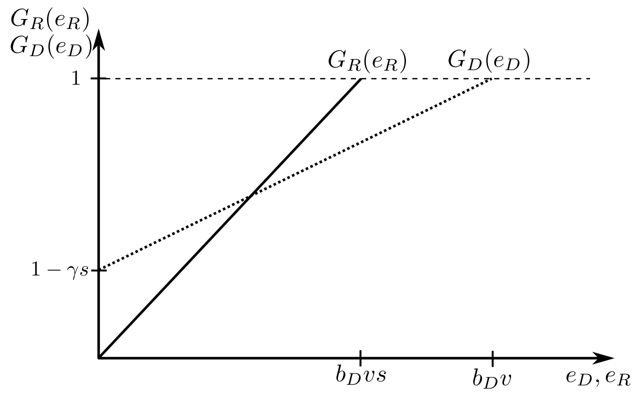
Similarly, Proposition 3.2.3 and Equation 3.2.6 allow us to do the same for k_D:

$$k_D = 0 \tag{3.2.8}$$

Plugging the above constants in to Equations 3.2.5 and 3.2.6 proves the theorem. □

By modeling campaign expenditures as bids in a first-price asymmetric all-pay auction, we have solved candidates' optimal campaign expenditure strategies for both candidates as mixed Nash equilibrium best-responses to each other. In region i, the optimal strategy by the candidates is trivial (both candidates always expend 0). In region ii, the optimal expenditure strategy is to play a distribution over possible expenditures. The distribution depends on the district's demographics, specifically, the particular location in the (b, μ)

Figure 3.2: Strategy CDFs



state space. The disadvantaged candidate spends 0 (concedes) with positive probability, and mixes uniformly between expenditures up to the office valuation the remainder of the time. The advantaged candidate's optimal strategy is to expend according to a uniform distribution over all expenditure levels up to a fraction of her valuation.

3.2.2 Electoral Outcome Probabilities

The model's predicted expenditure distributions imply electoral outcome predictions. We use the derived strategies to calculate the probabilities of an advantaged and a disadvantaged win, respectively (in region *ii*).

$$\int_0^{b_D vs} \int_0^{\frac{e_R}{s}} dG_D(e_D) dG_R(e_R) = 1 - \frac{1}{2}\gamma s \quad > \frac{1}{2} \tag{3.2.9}$$

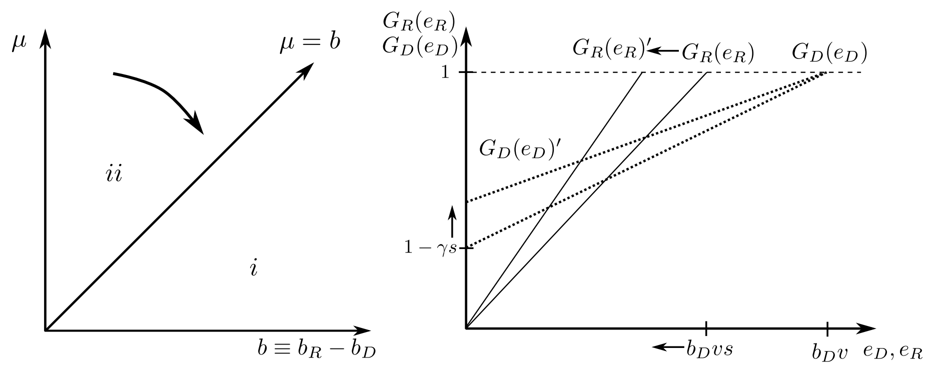
$$\int_0^{b_D v} \int_0^{e_D s} dG_R(e_R) dG_D(e_D) = \frac{1}{2}\gamma s \quad < \frac{1}{2} \tag{3.2.10}$$

When $\mu = b$, P_R's base advantage equals the swing mass and her probability of winning, $1 - \frac{1}{2}\gamma s$, equals one. The advantaged candidate wins with certainty. When b=0, the candidates have the same size base. Now, the fractions $s = \frac{\mu - b}{\mu + b}$ and $\gamma = \frac{b_D}{b_R}$ are both equal to 1, and the district is perfectly competitive. Both candidates have the same expenditure strategy, and both candidates have a $\frac{1}{2}$ chance of winning.

3.2.2.1 Prediction Intuition

Refer to Figure 3.3. Consider moving in an arc clockwise from the μ-axis ($b = 0$) toward the $\mu = b$ region-boundary. Along the vertical axis, district demographics favor neither candidate; the candidates have the same size base. The mixed Nash expenditure strategies are identical (neither candidate has a mass point at 0) and both mix uniformly up to their seat valuation. Also, both candidates have probability $= \frac{1}{2}$ of winning. As we move along the arc clockwise, meaning that the advantaged candidate's base advantage grows and the advertising-responsive swing mass decreases, the probability the disadvantaged candidate expends 0 (or concedes) increases and his probability of winning diminishes. The probability that the advantaged candidate wins increases and the upper bound of the advantaged candidate's support diminishes. In fact, as the arc approaches the $\mu = b$ region-boundary (the mass of advertising-responsive voters diminishes), the model expenditure predictions converge pointwise to the Median Voter Model predictions; no expenditures and the advantaged candidate, who must have the median voter (otherwise could not be in region i), wins with probability 1.

Figure 3.3: Changing Variables



3.3 Testable Predictions and Estimation

Previous research by Hendricks & Porter (1988) and Baldwin et al. (1997) used ex-post realizations of object values to compare observed bid behavior with Nash behavior and test for collusive bid behavior. Alternatively, we use observed expenditure behavior to infer object valuation. Table 3.2 explicitly states some of the model predictions.[3] While previous work has relied on maximum likelihood (ML) techniques to estimate auction models, in this chapter we estimate the value of a House seat using optimal General Methods

[3]There are many more moment predictions but these are the moments that are pertinent to estimating the value of a House seat.

of Moments (GMM), matching the empirical moments to the moment predictions. We do this because the model does not have an idiosyncratic error term. Instead, we use the model's predicted strategic uncertainty to estimate the valuation. The lack of error term has the benefit of not making any distributional assumptions about unobservables. The cost of excluding an error term is that ML estimates of the valuation would be driven by the outlier expenditures. ML estimation matches every single distributional assumption of the model to estimate v, including the order statistics. As the Table 3.2 shows, v determines the support of the expenditure distributions and is therefore a parameter in the order statistics. Any extreme candidate expenditure would drive the ML estimate of v up to unreasonable levels.

To prevent this, we estimate the value of the House seat based on the first two predicted moments, the mean and the variance of expenditures. These moment predictions are used to estimate the valuation of a House seat that would have given rise to the observed candidate expenditure behaviors. Again, this allows us to estimate the per partisan seat value v without introducing an error term. Since there is no strategic uncertainty in region i , we estimate v only with data from districts in region ii.

3.4 Data

For this analysis, a unique data set of voter registration statistics and candidate expenditures was compiled. We compiled data on voter registration

Table 3.2: Model Predictions in Region *ii*

Candidate	$p = D$	$p = R$
Moment Predictions		
$E(e_p \mid \mu, b_D, b_R)$	$\frac{1}{2}\gamma(b_D v s)$	$\frac{1}{2} b_D v s$
$Var(e_p \mid \mu, b_D, b_R)$	$(b_D v)^2(\frac{1}{3}\gamma s - \frac{1}{4}(\gamma s)^2)$	$\frac{1}{12}(b_D v s)^2$
Order Statistics		
$Min(e_p \mid \mu, b_D, b_R)$	0	0
$Max(e_p \mid \mu, b_D, b_R)$	$b_D v s$	$b_D v$

statistics by congressional district from the years 2002, 2004, 2006, and 2008. We were forced to exclude some Congressional districts for two main reasons. Not all states require voters to register for a political party when they register to vote. Generally, only the states with closed primary elections require and record the registered party for each voter. Also, some states keep voter registration statistics by county and their counties do not have the same borders as their districts. We use the voter registration statistics by congressional district as the measure of each candidate's base voters, b_p, and therefore we do not include any districts where this measure is unavailable. We then define μ as the number of registered voters who are not registered for either of the major parties.

Voters register with a party mainly so that they may participate in their party's primary. As was touted during the 2008 presidential primaries, it is possible that voters register with their less favored party so they can do damage to that party resulting in greater benefit for their favored party. In such a case, a strategic registrant would help elect the candidate from their less favored party that they feel has the greatest chance of losing to their favored party's candidate. Registrants employing such tactics are predicted to be a rare occurrence. Nonetheless, to mitigate any such strategic effect, we used registration statistics that were reported as close to the general election as possible.

The candidate campaign expenditure data comes from the United States Federal Election Commission (FEC) and all expenditures are converted to 2008 dollars. The FEC provides summary files of total election cycle expenditures for all candidates. Overall, we were able to collect voter registration statistics by party and link the candidate expenditures from 814 congressional district-election pairs. Our model does not incorporate any potential third party candidates and additional candidates have the potential to change the equilibrium in an all-pay auction. Therefore, we exclude any district where a third party candidate ran and spent more than $10,000. We only include districts that have a Democrat facing a Republican, or a single Democrat or Republican running unopposed. This restriction excludes 65 districts. The remaining 749 districts are plotted in the (b, μ) state space in Figure 3.4. Recall that the theory was derived assuming that the Republican candidate was advantaged. This state space graph now shows all the districts. The Republican candidate is advantaged in regions *i* and *ii* and the Democrat is advantaged in regions *iii* and *iv*. Region *iii* is analogous to region *ii* in that the model predicts mixed Nash expenditure strategies and region *iv* is analogous to region *i* in that the model predicts that the Republican candidate should concede to the

Democratic candidate.

Figure 3.4 shows that there is very large variation in the strategic environment of these congressional districts.[4] Since we do not introduce an error term into our model, and estimate the value of a House seat solely off of the strategic uncertainty, we dismiss districts for which our model predicts concession.[5] While this exclusion is necessary to estimate v from the model, the removal of the extreme regions i and iv should intuitively help provide a better estimate of the value of a House seat. The regions ii and iii are much less skewed to either party. The elections in these regions should be much more competitive and expenditures will matter more in the determination of the winner. Therefore, the expenditures should provide a better indication of the valuation of the House seat.

The data used to estimate v consists of partisan voter registration statistics from the 438 congressional districts located in the competitive regions of the district state-space, regions ii and iii. The summary statistics are presented in Table 3.3. Even though Figure 3.4 shows large variation in the strategic state space, these statistics show that the districts are very similar on average. There are only about 8,000 less registered Democrats on average and there is a sizable amount of swing voters. Additionally, these statistics

[4]One striking feature about this graph is that there are numerous districts in region iv, where the Democrat is measured to be heavily advantaged. There is a surprising lack of balance in region i, where a Republican is heavily advantaged. This fact does not affect the results of the this chapter, but could have implications about the partisanship of congressional districts in state's with closed primaries.

[5]Districts where $b \geq \mu$.

Table 3.3: Summary Statistics

	Mean	Std. Dev	Min	Max
Registered Democrats (b_D)	140,724	33,532	33,953	280,575
Registered Republicans (b_R)	148,856	42,545	18,338	252,511
Swing Voters (μ)	100,052	45,347	5,097	296,248
Democratic Expenditure (e_D) (\$)	911,372	990,843	0	5,591,222
Republican Expenditure (e_R) (\$)	1,101,913	1,096,256	0	8,680,645
Districts		438		

Note: This table presents the summary statistics for all the districts in regions *ii* and *iii*.

show that the Republican candidate outspends the Democrat by about 10% on average.

3.5 Estimation Results

This section presents the estimated v values, $\hat{v}$ and uses this estimate to calculate the distribution of House seat values. Table 3.4 presents three different estimates of v. The first column estimates v using only one moment, the expected value of the advantaged candidate's expenditure. The second column uses both expected values for the advantaged and the disadvantaged candidate. The third specification estimates v with both expected values and also both predicted variance moments. All estimates are done using optimally weighted GMM.

All three estimates are positive and significant at the 1% level. The value becomes smaller and more precisely identified once the variance moments are included. This occurs because the variance moments incorporate much

Table 3.4: Estimated v Values

	(1)	(2)	(3)
$\hat{v}$	44.797***	40.198***	29.996***
	(2.2552)	(2.0228)	(0.6678)
Observations	438	438	438
Objective Value	0	0.0757	0.27979
OverID Stat		33.135	122.55
$p =$		0	0

Note: This table presents the estimated v values for three separate specifications. Column (1) estimates $\hat{v}$ using only the mean advantaged candidate expenditure moment. Column (2) uses both candidates' mean expenditure moments. Column (3) uses both candidates' mean and variance expenditure moments. Optimally weighted standard errors are reported in parentheses and the asterisks indicate significance levels of the coefficients, *** p<0.0.

stronger model predictions than the mean predictions. The tradeoff is that the estimates are statistically over identified. However, this is expected given that we only have one parameter in the model.[6] Additionally, the estimates become statistically over identified as soon as one additional moment is included.

Therefore, we use the third estimate, $\hat{v} = 29.996$, to calculate the valuation of all the House seats. $\hat{v}$ is interpreted as the 2008 dollar value of the House seat to a partisan voter. The estimates are multiplied by the number of partisan voters to get the valuation to each candidate and these valuations are shown in Figure 3.5. These valuation vary greatly over the two candidates but the overall average valuation is about $4.5 million.

[6] Future work will parameterize the valuation to allow for more variation based on other district characteristics, such as the mean income level in the district.

3.6 Conclusion

This chapter models U.S. House of Representative campaign expenditures as bids in an asymmetric all-pay auction. The model predicts a wide variety of expenditure patterns and outcomes, even between districts that have similar demographics. Using the observed expenditure patterns and predictions from the model, we estimate that a Congressional seat is worth about \$30 per registered partisan and that on average the seat is worth about \$4.5 dollars. In future research, the model and estimates can be used to analyze the effects of campaign finance reforms and redistricting policies. The model can predict show candidates' campaign expenditures will react to the new policies and the effect that these change in expenditure patterns will have on election outcomes.

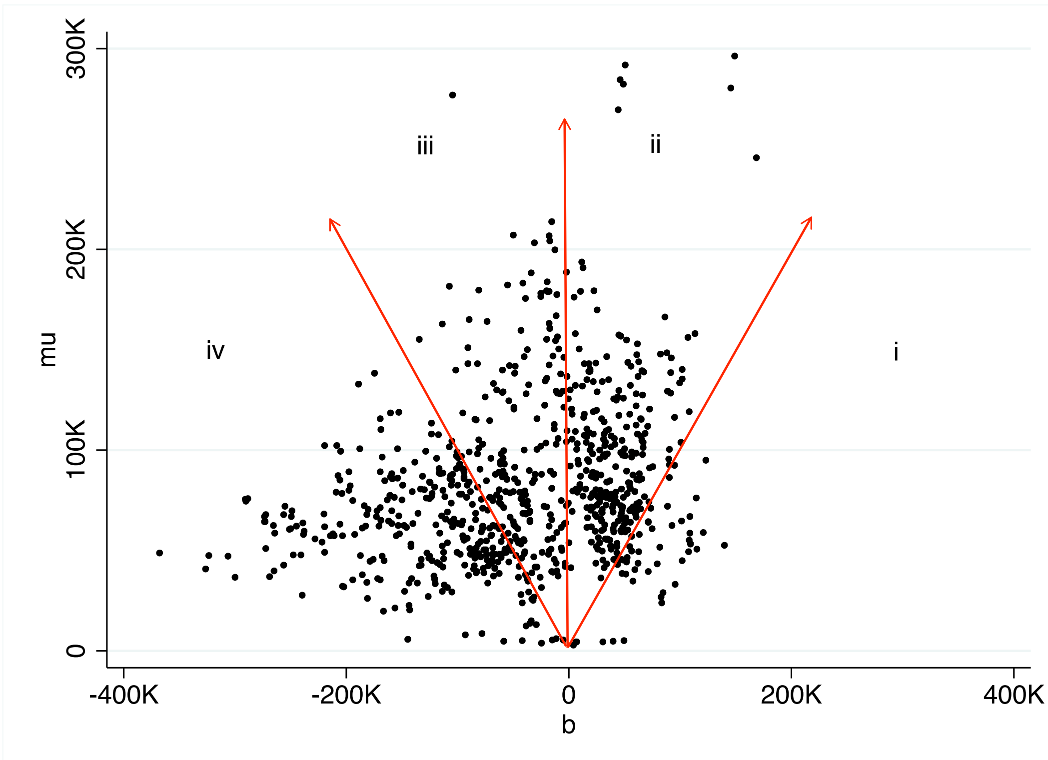


Figure 3.4: Congressional Districts in the (b,μ) State Space

Note: This figure plots all the congressional districts in the strategic state space. The difference in the Republican base and the Democratic base of voters is measured on the horizontal axis. The vertical axis measures the number of swing voters. The Republican candidate is advantaged in regions *i* and *ii* and the Democrat is advantaged in regions *iii* and *iv*.

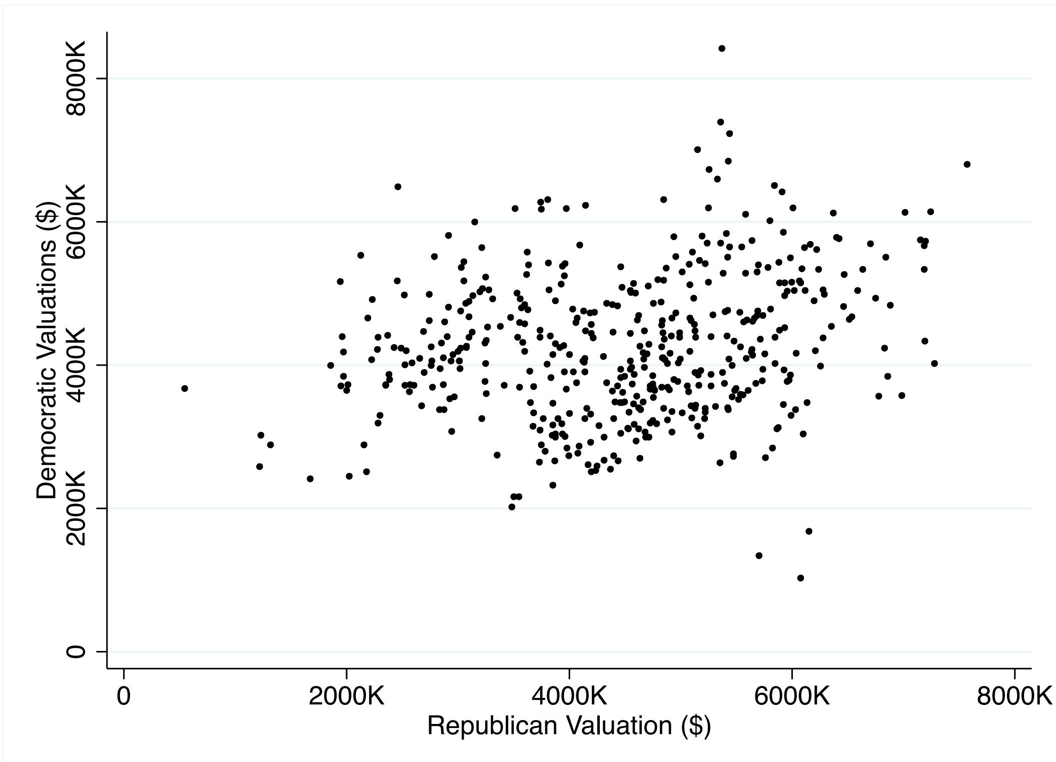


Figure 3.5: Valuations of Congressional Districts

Appendix

Appendix 1

1.1 Reduced Form Regression on Expenditures

Table 1.1 shows that the instruments are highly correlated with the endogenous expenditures, the first condition that must be satisfied for a valid instrument set. The first two columns regress the first stage endogenous variables, $ln(PrimaryExpend)$ and $ln(PrimaryExpend) * Incumbent$ on the instrument set and the exogenous variables. These regressions are the control functions for the primary elections. Columns (3) and (4) show the results of the reduced form regression for $ln(GeneralExpend)$ and $ln(GeneralExpend) * Incumbent$. The F Test at the bottom of each column test the joint hypothesis that all the instruments are insignificant. This hypothesis is rejected at the .001 level for all four regressions.

Table 1.1: Reduced Form Regressions

VARIABLES	(1)	(2)	(3)	(4)
$ln(RestPrimAvg)$	2.474***	-0.000312		
	(0.317)	(0.00322)		
$ln(RestPrimAvg) * Incumb$	-2.216***	0.301***		
	(0.324)	(0.0682)		
$ln(RestGenOwnPartyAvg)$			-0.447	-0.157***
			(0.312)	(0.0298)
$ln(RestGenOwnPartyAvg) * Incumb$			-0.438	0.0706
			(0.350)	(0.0719)
$ln(RestGenOppPartyAvg)$			-0.709*	-0.128***
			(0.414)	(0.0311)
$ln(RestGenOppPartyAvg) * Incumb$			0.170	0.0316
			(0.440)	(0.0508)
$Incumbent$	32.27***	9.080***	7.426	12.37***
	(3.980)	(0.837)	(5.918)	(0.975)
$dFemale$	1.071***	-0.00740	0.554**	0.0102
	(0.261)	(0.0261)	(0.215)	(0.0370)
$NumPrimaryCandidates$			0.727***	-0.00379
			(0.0974)	(0.0110)
$dDem$	-22.23***	-0.00368	24.65***	3.745***
	(3.953)	(0.0597)	(6.498)	(0.729)
$dRep$	-22.47***	0.0291	23.88***	3.846***
	(3.950)	(0.0572)	(6.476)	(0.731)
$UnemploymentRate$			0.447	0.0614
			(0.396)	(0.0628)
$PresidentialRace$	-0.119	-0.0240	-0.793**	-0.178***
	(0.203)	(0.0182)	(0.396)	(0.0614)
$SenateRace$	-0.499**	0.00644	-0.351	-0.139***
	(0.250)	(0.0252)	(0.266)	(0.0455)
Observations	1674	1674	1077	1077
R^2	0.849	0.997	0.947	0.998
F Test: F =	32.77	9.756	5.650	7.817
p =	0	6.13e-05	0.000169	3.31e-06

Note: This table presents the coefficient results from the reduced form regressions of the endogenous variables on the instruments and all other exogenous variables, omitting the coefficients for the state fixed effects. The dependent variables are: Column (1) $ln(PrimaryExpend)$, Column (2) $ln(PrimaryExpend)*Incumbent$, Column (3) $ln(GeneralExpend)$, and Column (4) $ln(GeneralExpend)*Incumbent$. Robust standard errors are reported in parentheses and the asterisks indicate significance levels of the coefficients, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The F Test in each column tests the joint hypothesis that all the instruments are insignificant.

Bibliography

[1] J. Adams and S. Merrill III. "Candidate and Party Strategies in Two-Stage Elections Beginning With a Primary". *American Journal of Political Science*, 52(2):344–359, April 2008.

[2] M. Agranov. "Flip-Flopping: Contrasting Effects of Intense Primaries and General Elections on Selection of Candidates". *Working Paper.*

[3] E. Amann and W. Leininger. "Asymmetric All-Pay Auctions with Incomplete Information: The Two Player Case". *Games and Economic Behavior*, 14(1):1–18, May 1996.

[4] S. Ansolabehere and A. Gerber. "The Mismeasure of Campaign Spending: Evidence from the 1990 U.S. House Elections". *The Journal of Politics*, 56(4):1106–1118, November 1994.

[5] S. Athey and P. A. Haile. "Identification of Standard Auction Models". *Econometrica*, 70(6):2107–2140, November 2002.

[6] P. Bajari, J. T. Fox, and S. P. Ryan. "Linear Regression Estimation of Discrete Choice Models with Nonparametric Distributions of Random Coefficients". *American Economic Review*, 97(2):459–463, May 2007.

[7] L. Baldwin, R. Marshall, and J.-F. Richard. "Bidder Collusion at Forest Service Timber Sales". *Journal of Political Economy*, 105(4):657–699, 1997.

[8] M. R. Baye, D. Kovenock, and C. G. de Vries. "Rigging the Lobbying Process: An Application of the All-Pay Auction". *The American Economic Review*, 83(1):289–294, March 1993.

[9] G. S. Becker. "A Theory of Competition Among Pressure Groups for Political Influence". *The Quarterly Journal of Economics*, 98(3):371–400, August 1983.

[10] G. S. Becker. "Public Policies, Pressure Groups, and Dead Weight Costs". *Journal of Public Economics*, 28(3):329–347, 1985.

[11] S. Berry. "Estimating Discrete-Choice Models of Product Differentiation". *RAND Journal of Economics*, 25(2):242–262, Summer 1994.

[12] S. Berry, M. Carnall, and P. Spiller. *"Airline Hubbing, Costs and Demand" in Advances in Airline Economics, Vol. 1: Competition Policy and Anti-Trust*. Elsevier Press, 2006.

[13] S. Berry, J. Levinsohn, and A. Pakes. "Automobile Prices in Market Equilibrium". *Econometrica*, 63(4):841–890, 1995.

[14] S. Berry, J. Levinsohn, and A. Pakes. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market". *Journal of Political Economy*, 112(1):68–105, 2004.

[15] S. Callander and S. Wilkie. "Lies, Damned Lies, and Political Campaigns". *Games and Economic Behavior*, 60:262–286, 2007.

[16] A. C. Cameron, J. B. Gelbach, and D. L. Miller. "Bootstrap-Based Improvements For Inference With Clustered Errors". *The Review of Economics and Statistics*, 90(3):414–427, August 2008.

[17] Y.-K. Che and I. L. Gale. "Caps on Political Lobbying". *The American Economic Review*, 88(3):643–651, June 1998.

[18] S. Coate. "Parto-Improving Campaign Finance Policy". *The American Economic Review*, 94(3):628–655, June 2004.

[19] G. B. Dahl. "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets". *Econometrica*, 70(6):2367–2420, November 2002.

[20] S. G. Donald and W. K. Newey. "Choosing the Number of Instruments". *Econometrica*, 69(5):1161–1191, September 2001.

[21] J.-P. H. Dubé, J. T. Fox, and C.-L. Su. "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation". *Chicago Booth School of Business Research Paper No. 09-07*, Available at SSRN: http://ssrn.com/abstract=1338152, 2009.

[22] R. S. Erikson and T. R. Palfrey. “Campaign Spending and Incumbency: An Alternative Simultaneous Equations Approach”. *The Journal of Politics*, 60(2):355–373, May 1998.

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

[24] A. Gerber. “Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables”. *The American Political Science Review*, 92(2):401–411, June 1998.

[25] A. Gerber. “Does Campaign Spending Work?: Field Experiments Provide Evidence and Suggest New Theory”. *American Behavioral Scientist*, 47(5):541–574, January 2004.

[26] A. Goolsbee and A. Petrin. “The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV”. *Econometrica*, 72(2):351–381, March 2004.

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

[28] J. Hausman. *“Valuation of New Goods Under Perfect and Imperfect Competition”*. In T.F Bresnahan and R. Gordon, eds., *The Economics of New Goods.* Chicago: National Bureau of Economic Research, 1996.

[29] J. J. Heckman. “Sample Selection Bias as a Specification Error”. *Econometrica*, 47(1):153–161, January 1979.

[30] K. Hendricks and R. H. Porter. “An Empirical Study of an Auction with Asymetric Information”. *The American Economic Review*, 78(5):865–883, Dec 1988.

[31] G. Jacobson. “The Effects of Campaign Spending in Congressional Elections”. *American Political Science Review*, 72(2):469–491, June 1978.

[32] G. Jacobson. “The Effects of Campaign Spending in House Elections: New Evidence for Old Arguments”. *American Journal of Political Science*, 34(2):334–362, May 1990.

[33] K. Kretschman and N. Mastronardi. “Partisanship and the Determination of Election Outcomes”. *Working Paper*, 2010.

[34] D. S. Lee. “Randomized Experiments From Non-Random Selection in U.S. House Elections”. *Journal of Econometrics*, 142:675–697, 2008.

[35] S. Levitt. “Policy Watch: Congressional Campaign Finance Reform”. *The Journal of Economic Perspectives*, 9(1):183–193, Winter 1995.

[36] S. D. Levitt. “Using Repeat Challengers to Estimate the Effect of Campaign Spending on Election Outcomes”. *Journal of Political Economy*, 102(4):777–798, August 1994.

[37] S. D. Levitt and C. D. Wolfram. “Decomposing the Sources of Incumbency Advantage in the U.S. House”. *Legislative Studies Quarterly*, 22(1):45–60, February 1997.

[38] P. R. Milgrom and R. J. Weber. “A Theory of Auctions and Competitive Bidding”. *Econometrica*, 50(5):1089–1122, September 1982.

[39] P. R. Milgrom and R. J. Weber. “Distributional Strategies for Games with Incomplete Information”. *Mathematics of Operations Research*, 10(4):619–632, November 1985.

[40] R. B. Myerson. “Analysis of Democratic Institutions: Structure, Conduct and Performance”. *The Journal of Economic Perspectives*, 9(1):77–89, Winter 1995.

[41] A. Nevo. “Mergers with Differentiated Products: The Case of the Ready-to-Eat cereal Industry. *The RAND Journal of Economics*, 31(3):395–421, Autumn 2000.

[42] G. Owen and B. Grofman. “Two-Stage Electoral Competition in Two-Party Contests: Persistent Divergence of Party Positions”. *Social Choice and Welfare*, 26(3):547–569, June 2006.

[43] F. Palda and K. Palda. "The Impact of Campaign Expenditures on Political Competition in the French Legislative Elections of 1993". *Public Choice*, 94(1):157–174, January 1998.

[44] A. Petrin and K. Train. "A Control Function Approach to Endogeneity in Consumer Choice Models". *Journal of Marketing Research*, 47:1–11, February 2010.

[45] A. Prat. "Campaign Advertising and Voter Welfare". *The Review of Economic Studies*, 69(4):999–1017, Oct 2002.

[46] A. Prat. "Campaign Spending with Office-Seeking Politicians, Rational Voters, and Multiple Lobbies". *Journal of Economic Theory*, 103:162–189, 2002.

[47] M. Rekkas. "The Impact of Campaign Spending on Votes in Multiparty Elections". *The Review of Economics and Statistics*, 89(3):573–585, August 2007.

[48] K. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd edition, 2009.

[49] J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, 2002.