

Structural Equilibrium Analysis of Political Advertising

Brett R. Gordon
Columbia University

Wesley R. Hartmann
Stanford University

October 2010

Preliminary and incomplete

Abstract

We present a structural model of political advertising in equilibrium. Candidates choose advertising across media markets in order to maximize the probability of winning the national election. The voter model takes the form of an aggregate random coefficients discrete choice model in which advertising affects a voter's incentive to vote for either candidate or not to vote at all. We estimate the model using detailed advertising and voting data from the 2000 and 2004 Presidential elections.

We use the model to conduct a counterfactual in which we eliminate the Electoral College, and consider a direct national vote. Changing the structure of the electoral process alters candidates' marginal incentives to advertise in a given market. This leads to a new equilibrium allocation of advertising and potentially a new voting outcome. Furthermore, our model could be used for other counterfactuals, such as considering the effects of 3rd-party candidates or certain campaign finance reforms, and could be applied or extended to races for other offices (e.g. house, senate or gubernatorial) or the primaries.

Please note our counterfactual results are in progress and are not yet in the paper.

Keywords: Political advertising, voter choice, electoral college, structural model, empirical game, endogeneity, moment inequality.

1 Introduction

Concerns about the structure of the Electoral College are twofold. First, that it induces biases in the election process that favor populous states (Nelson 1974, Bartels 1985, Edwards 2004). Second, that a candidate who captures a plurality of the popular vote may fail to receive a plurality of the Electoral College vote. Numerous proposals for comprehensive electoral reform have been made since the founding of the country.¹ Public interest in reform became especially strong following the 2000 Presidential election, when George W. Bush won with a majority in the Electoral College despite having a minority of the popular vote. This led many, particularly proponents of reform, to conclude that Al Gore would have won under a direct voting system that eliminated the Electoral College (Time Magazine 2000). Others were quick to note that the Electoral College and popular vote have disagreed only three times in history, suggesting that any such reform is unlikely to impact the final candidate selection.

However, such analyses are fundamentally misguided because they ignore the fact that changes in the electoral system may lead candidates to change their campaigning strategies. As a result the outcome of the vote may also change. Under the current winner-take-all allocation rule, the candidate who wins the largest share of the popular vote in a state wins all of that state's Electoral College votes. This means battleground states—such as Florida and Ohio, where candidates expect a narrow margin of victory—attract significantly more candidate resources than non-battleground states. In contrast, non-battlegrounds states such as New York and Texas barely attract any attention from the candidates; in recent elections, for example, neither major party candidate chose to advertise at all in several states.

An alternative vote allocation rule changes the marginal incentives to campaign in a given market and should result in candidates altering their equilibrium allocations. Thus, in the absence of an equilibrium model, it is difficult to determine whether a candidate would or would not have won an election under a different vote allocation rule. Despite the importance

¹The first reform proposal was brought before the Senate in 1816. More recently, between 1950 and 1979, proposed amendments for Electoral College reform were debated in the Senate on five occasions and in the House twice, once actually passing by a vote of 339 to 70.

of this subject, past empirical work on alternative Electoral College systems ignores the equilibrium implications of changing the allocation rule on candidates' strategies and voters' choices (Blair 1979, Gelman, King, and Boscardin 1998, Grofman and Feld 2005, Hopkins and Goux 2008, Strömberg 2008).

This paper defines a structural equilibrium model of advertising competition between U.S. Presidential candidates in the general election.² A structural model permits us to conduct counterfactual analyses to consider the implications of moving from the current Electoral College system to various alternatives, such as adopting a congressional district plan, a proportional allocation system, or a direct popular election. We specifically consider the last option since it has come closest to being passed. Our model allows us to examine whether candidates more equitably allocate their resources across states under a direct voting model. In addition, we can analyze how different electoral systems affect voter turnout and how the presence of third-party candidates may differentially affect voting outcomes.

We apply the model to data on advertising per candidate and voting outcomes in the 2000 and 2004 elections. First, we use county-level voting data to estimate an aggregate discrete choice model of voter candidate selection with unobserved heterogeneity. We specify voter preferences using the aggregate demand model of Berry, Levinsohn, and Pakes (1995), and estimate it as a mathematical program with equilibrium constraints (MPEC) as detailed in Dubé, Fox, and Su (2009). This rich model of voting behavior permits the returns to advertising to vary by market and provides us with a flexible basis to predict voters' responses to changes in candidate advertising.

Second, given the voter choice model, we estimate the candidate advertising model as a simultaneous move game. Candidates strategically choose advertising levels across markets while facing uncertainty over local 'demand' shocks that could alter voters' choices. The existence of candidate uncertainty in some form is crucial to the model: if voting outcomes are deterministic functions of advertising choices, then a losing candidate would never choose positive levels of advertising in equilibrium. Prior to Election Day, candidates form unbiased

²Over \$750 million was spent on media and advertising in the last Presidential election, and observers predict spending in 2012 to exceed \$1.5 billion. Driving this growth in spending is the increasing recognition that a candidate's marketing campaign plays a critical role in the ultimate election outcome.

beliefs about the nature of these shocks, and then set advertising to maximize the expected return from winning the election. On Election Day, voters perfectly observe the shocks and decide whether to vote for a candidate or not to vote at all.

The candidate model is computationally challenging to solve and estimate due to the large continuous action space, the existence of corner solutions, and the potential for multiple equilibria. We address these issues by estimating the model using the moment inequality approach in Pakes, Porter, Ho, and Ishii (2006) (hereafter PPHI). Two key benefits of following PPHI are that we can remain agnostic about the nature of agents’ beliefs over competitors’ private information and that we avoid explicitly solving for the game’s equilibrium. Estimation via moment inequalities circumvents these difficulties while still allowing us to find parameters that rationalize the observed outcomes.

To our knowledge, we are the first to investigate Electoral College reform using structural empirical methods.³ Brams and Davis (1974) and Owen (1975), among others, initiated a theoretical literature that examines the implications of different electoral allocation rules on election outcomes.⁴ More recently, Lizzeri and Persico (2001) show that a winner-take-all system provides an arbitrary public good less often than a proportional voting system, and then show that the Electoral College is subject to the same inefficiency. Strömberg (2008) incorporates a random-effects regression into a probabilistic voting-model and considers candidates’ resource allocation decisions in uncertain elections.⁵

Despite the strategic nature of political competition, little work formulates elections as empirical games.⁶ As in other discrete-choice models, our inclusion of unobserved heterogeneity

³Two recent structural papers in political economy focus on the voter side. Degan and Merlo (2009) model the decision of voters to participate in multiple elections. Kawai and Watanabe (2010) estimate a structural model of strategic voting using Japanese general election data.

⁴Our model shares some similarities with the theoretical literature on contests, especially recent work by Kaplan and Sela (2008) on political contests with private entry costs and by Siegel (2009) on all-pay contests. In the latter, each player chooses a costly “score,” representing a (possibly) sunk investment, and the player with the highest score wins the prize. Our model is similar in that candidates engage in a winner-take-all game where each must choose how much to “invest” in advertising, which becomes a sunk cost. Such contests arise naturally in settings where participants must expend resources no matter if they win or lose, such as elections, lobbying activities, and R&D races.

⁵Che, Iyer, and Shanmugam (2007) use a nested logit model to examine voter turnout and candidate ad type decisions in a non-strategic setting.

⁶Erikson and Palfrey (2000) investigate the simultaneity problem in estimating the effect of campaign spending on election outcomes. Abstracting away from the voter side entirely, the authors derive testable

should better capture substitution patterns across alternatives (i.e. candidates). The estimated joint model of voter decisions and the candidates' advertising game should provide the required structural basis to help make counterfactual predictions in substantially different election regimes. Our counterfactual results are relevant for two audiences. First, candidates and political consultants could use a model to help predict how voters and competitors would respond to a change in the own candidate's advertising strategy. Second, policy makers seek to understand how the structure of an election may affect voter participation rates and influence candidates' campaign fundraising and spending activities.

The rest of the paper is organized as follows. Section 2 discusses the data set. Section 3 presents the voter and candidate models. Section 4 explains our estimation strategy.

2 Data

This section details our data sources and provides some reduced-form evidence of a relationship between advertising and voting outcomes.

2.1 Advertising

The primary advertising data come from the Campaign Media Analysis Group (CMAG) for the 2000 and 2004 Presidential elections, and were made available through the University of Wisconsin Advertising Project. CMAG monitors political advertising activity on all national television and cable networks, and assigns each advertisement to support the proper candidate. The data provide a complete record of every advertisement broadcast in each of the country's top designated media markets (DMAs), representing 78% of the country's population. Television ads are the largest component of media spending for political campaigns according to AdWeek (2009). See Freedman and Goldstein (1999) for more details on the creation of the CMAG dataset.

The data contain a large number of individual presidential ads: 247,643 in 2000 and 807,296 in 2004. For each ad, we observe the precise date and time it aired, the candidate

implications from the equilibrium solution to a spending game between candidates, which they empirically test using a set of reduced-form regressions.

supported (e.g., Democrat, Republican, Independent, etc.), and the sponsoring group (e.g., the candidate, the national party, independent groups, or “hybrid/coordinated”). Another key variable we observe is an estimate of the ad’s cost calculated by CMAG, which will help serve as a basis of estimation in the candidate model. The data allow us to calculate the total length (in seconds) of ads supporting a particular candidate, which we sum over sponsoring groups to yield the observed (to the voter) advertising quantity.⁷ We restrict attention to advertisements appearing after Labor Day, when the primaries have concluded and competition in the general election begins in earnest.

Table 1 displays descriptive statistics. The third-party candidates (Nader and Badnarik) spent more on average per ad because much of their advertising was concentrated in larger, more expensive media markets.

Table 1: Descriptive Statistics: Full Sample

Election	Candidate	Party	Total Ads	Total Expenditure	Average Ad Cost	Popular Vote
2000	Bush	Republican	126814	\$89,202,830	\$703	47.9%
	Gore	Democrat	119300	\$76,902,197	\$645	48.4%
	Nader	Green	1256	\$1,227,463	\$977	2.7%
	Various	Other	269	\$373,241	\$2370	1.0%
	<i>Total</i>		247639	\$167,605,731	\$939	100.0%
2004	Bush	Republican	262293	\$209,595,807	\$799	50.7%
	Kerry	Democrat	544205	\$353,848,127	\$650	48.3%
	Badnarik	Libertarian	248	\$297,717	\$1201	0.3%
	Various	Other	550	\$197,719	\$360	0.7%
	<i>Total</i>		807296	\$563,939,370	\$752	100%
	<i>Total</i>		1054935	\$731,545,101	\$795	-

We also obtained separate advertising cost data from TNS Media Intelligence at the DMA level to use as instrumental variables. The instruments contain the aggregate average cost-per-thousand (CPM) impressions and cost-per-point (CPP) by DMA in each election year.⁸ These should be valid instruments for advertising because they must be correlated with

⁷Federal campaign finance law allows political parties to explicitly coordinate certain expenses, including advertising, on behalf of the general election candidates (Garrett and Whitaker 2007).

⁸Ideally, our data would contain a GRP-type advertising variable that helps measure exposure, but we only

candidates' actual advertising costs, but should be uncorrelated with local voting demand shocks.

2.2 Votes

The county-level vote data is available from www.polidata.org. For each of the 1,342 counties, we observe the number of votes cast for all possible candidates and the size of the voting-age population (VAP). The VAP estimates serve as our market size parameters, and allow us to calculate a measure of voter turnout at the county level.⁹

It is important to note that we observe advertising at the DMA level and voting outcomes at the county level. We assign the observed level of advertising at the market level to each of the counties contained in that market.¹⁰ We conduct our analysis for all counties for which we observe the DMA-level advertising. Voting behavior, and therefore advertising, in the counties representing the remaining 22% of the population is held fixed when estimating the candidate-side and analyzing the counterfactual candidate policies. We are currently exploring alternative solutions to this issue.

2.3 Reduced-Form Evidence and Discussion

To help explore our data, we obtained measures of the competitiveness of a state in a particular election from *The Cook Political Report*.¹¹ This periodical publishes an index of competitiveness based on factors such as polling, historical voting patterns, and expert opinion. The ratings are published irregularly throughout the election year, and we use the ratings closest to Labor Day.

Figure 1 plots the advertising per Electoral College vote for Democrats and Republicans

observe advertising quantity (in seconds) and expenditure. Although advertising quantity is more appropriate from the voter perspective, it does not account for variation in exposure rates across markets. Advertising costs should be positively correlated with exposure, but confounds the per unit price and overall quantity of advertising.

⁹Unfortunately, voting-eligible population (VEP), a more accurate measure to calculate turnout that removes non-citizens and criminals, is only available at the state level. See the web page maintained by Michael McDonald at http://elections.gmu.edu/voter_turnout.htm for more information on measures of voter turnout.

¹⁰Of the 1,342 counties, only five belong to multiple DMAs. We use zip code-level population data to weigh the advertising proportionally according to the share of the population in a given state.

¹¹The authors thank Mitchell Lovett for providing this data.

against the competitiveness index for each state. The size of each circle is proportional to the number of Electoral College votes in that state. The figure illustrates that candidates tend to spend more per vote in states where the outcome of the election is hardest to predict, such as the battleground states of Ohio, Pennsylvania, and Florida. In contrast California, which the Democrats won with a 10% margin in 2004, receives little advertising from either party.

Our model, described in the next section, estimates voter preference parameters by aggregating from the voter level to form county-specific vote shares. We run a series of regressions to test whether the data contain reduced-form evidence of a meaningful relationship between aggregate voting outcomes and candidate advertising. Table 2 contains the results of several regressions where the dependent variable is the county-specific, difference in logs between the candidate’s vote share and the “outside good’s” vote share, defined here as voters who either did not vote or voted for the third-party candidate.¹² The independent variables are a dummy indicating which election, a dummy indicating which party (e.g. Democrat or Republican), an interaction term between the election and party dummies, and the candidate’s advertising quantity observed in that county.

Table 2: Reduced-Form Evidence

DV: $\ln(\text{candidate vote share}) - \ln(\text{outside share})$

Variable	OLS	2SLS	2SLS
Intercept	-0.657	-0.668	-0.709
Election	0.298	0.299	0.303
Party	-0.314	-0.313	-0.307
Election*Party	-0.137	-0.141	-0.155
Advertising Quantity	0.109	0.118	0.153
DMA Fixed Effects	No	No	Yes

$N = 6,384$, standard errors clustered by DMA

All coefficients are significant with $p < 0.01$.

The first column of Table 2 reports OLS results. As expected, advertising appears to have a positive effect on the candidate’s vote share. The second column reports 2SLS results using our ad cost instruments and the third column includes fixed effects at the DMA level.

¹²We expect to formally model the third-party candidate’s presence in the near future, but chose to ignore them for now to keep the estimation simpler.

The advertising coefficient remains positive and strongly significant in each specification, suggesting that the data contain the appropriate variation to help estimate our structural model. Note that controlling for market-specific unobservable shocks through the fixed effects in (moving from the second to third column) leads to an increase in the advertising coefficient. The direction of the change suggests there is a negative correlation between the unobserved shocks and advertising. The results in Figure 1 support this observation: the data show that in stronghold states, where one candidate receives a large, positive unobserved net shock, we see little to no advertising. We find it encouraging to see such consistency across Table 2 and Figure 1.

3 Model

In this section we present a two-stage, static model of presidential advertising competition and voter behavior in the general election. We observe the advertising choices and voting outcomes from a collection of T elections. We consider voting outcomes at the county level, such that each voter lives in some county $c = 1, \dots, C$, inside a DMA $m = 1, \dots, M$, inside a given state $s = 1, \dots, S$. We use $c \in m$ to denote the set of counties in market m , and $m \in s$ to denote the set of markets in state s . The vector θ is the collection of parameters of interest.

Figure 2: Model Overview

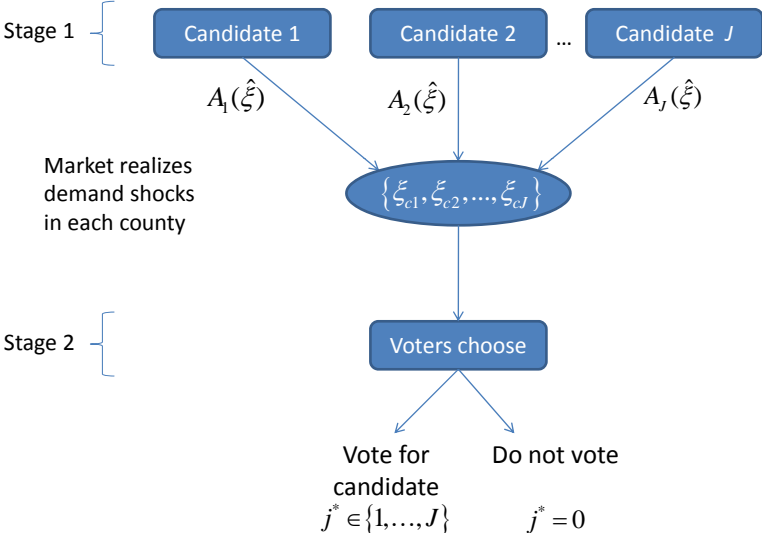


Figure 2 depicts the basic structure of the game. In the first stage, for a given election t , each candidate $j = 1, \dots, J$ sets advertising levels $\{A_{tmj}\}$ across markets. Advertising is the same across counties within a market, such that $A_{tcj} = A_{tmj}, \forall c \in m$. Candidates, however, must allocate advertising before votes are cast and are uncertain about future county-specific demand shocks $\{\xi_{tcj}\}$ that could influence voters' decisions. Candidates form rational expectations about these demand shocks and set advertising conditional on these beliefs. The second stage takes place on Election Day, and voters perfectly observe the demand shocks and advertising. If a voter decides to turnout for the election, she chooses for which candidate to vote. Voting outcomes across all counties are realized and one candidate is deemed the winner.

Formally, our model abstracts away from the campaign fundraising process, but still allows candidates' "budgets" to flexibly adjust under our counterfactual scenarios. We discuss this issue later at the end of Section 3.2 and consider it an interesting avenue for future research.

3.1 Voters

A voter's utility for candidate j given advertising quantity A_{tcj} in election t is:

$$u_{itcj} = \beta_{itj} + \alpha_i A_{tcj} + \gamma_{mj} + \xi_{tcj} + \varepsilon_{itcj} , \quad (1)$$

where β_{itj} is a voter-specific taste for a candidate from party j in election t , α_i is the marginal utility of advertising, γ_{mj} represents market-party fixed-effects, and ε_{itcj} captures idiosyncratic variation in utility, which is i.i.d. across voters, candidates, and periods. ξ_{tcj} is a time-county-candidate specific demand shock that is perfectly observed to voters when casting their votes, but unobserved to candidates (and the researcher) when making their advertising decisions. Candidates' beliefs about the demand shocks ξ_{tcj} induce endogeneity in candidates' advertising strategies. If a voter does not turnout for the election, she selects the outside good and receives a utility of

$$u_{itc0} = \varepsilon_{itc0} . \quad (2)$$

The γ_{mj} in our model serve the role of location-candidate specific dummies. As the only observed characteristic we include is advertising, this helps fit the mean utility level for a

candidate (or party) in a specific market. The dummies also help address the endogeneity of advertising by capturing any omitted and unobserved characteristics that vary by party and/or market. Thus, any correlation between advertising and market-specific party preferences is controlled for without the need for an instrument. We do require an instrument to address the remaining unexplained variation, which corresponds to time-specific deviations from the unobserved candidate-market mean utility.

To capture heterogeneity in voter preferences, we allow the candidate-election specific intercepts and the marginal utility of advertising to vary across voters. We assume that

$$\begin{bmatrix} \beta_{itj} \\ \alpha_i \end{bmatrix} \sim N\left(\begin{bmatrix} \bar{\beta}_{tj} \\ \bar{\alpha} \end{bmatrix}, \Sigma\right) \quad (3)$$

where Σ is the full covariance matrix of voter tastes. Allowing for off-diagonal terms in Σ is important to remove the property of independence from irrelevant alternatives (IIA) commonly found in logit demand models.

Each voter selects the candidate who gives her the highest utility, or decides not to vote.¹³ Assuming that $\{\varepsilon_{itcj}\}_j$ are multivariate extreme-valued and drawn independently from the taste distribution $F(\beta, \alpha; \theta)$, integrating over the idiosyncratic shocks we obtain the following vote shares:

$$s_{tcj}(A_{tc}, \xi_{tc}; \theta) = \int_{\beta, \alpha} \frac{\exp\{\beta_{itj} + \alpha_i A_{tcj} + \gamma_{mj} + \xi_{tcj}\}}{\sum_{k \in \{0, \dots, J\}} \exp\{\beta_{itk} + \alpha_i A_{tck} + \gamma_{mk} + \xi_{tck}\}} dF(\beta, \alpha). \quad (4)$$

The model of voter choice above does not consider whether voters act strategically based on whether they expect their vote to be pivotal in deciding the election outcome. Although voters' expectations of being pivotal can play a role in small elections (e.g., Coate, Conlin and Moro 2008), the effect vanishes in larger elections (Feddersen and Pesendorfer 1996, 1999).

3.2 Candidates

In an election at t , candidates simultaneously choose advertising $A_{tmj}, \forall m, j$ across DMA's given their beliefs about the demand shocks ξ_{tcj} . We introduce uncertainty over ξ_{tcj} because

¹³The model assumes that voters act sincerely in casting their votes.

candidates must set advertising before voters make their decisions. This is an important source of uncertainty in the outcome of the election, which is precisely what motivates candidate's to advertise.¹⁴

Prior to making their advertising decisions, candidates gather information through campaign research and other sources about the nature of potential demand shocks in each county. This information provides the candidate with an expectation $\bar{\xi}_{tcj}$ of each shock's realized value ξ_{tcj} , such that

$$\boldsymbol{\xi}_{tcj} = \bar{\xi}_{tcj} + \epsilon_{tmj}, \quad \epsilon_{tmj} \sim N(0, \sigma_\xi) \quad (5)$$

where ϵ_{tmj} is a random draw independent across (t, m, j) and $\boldsymbol{\xi}_{tcj}$ is bold faced to indicate that it is a random variable from the perspective of the candidates. We define the uncertainty in candidates beliefs as a DMA level shock common to all counties within because the DMA is the geographic level at which the candidate chooses advertising. In our implementation, we assume candidates' expectations are rational, and set the mean belief equal to the realized value, $\bar{\xi}_{tcj} = \xi_{tcj}$. Note that this implies the realized value of ϵ_{tmj} is always zero, but the candidates do not know this.¹⁵

What is the interpretation of the error term ϵ_{tmj} ? Consider, for example, that weather affects voter turnout on Election Day and might differentially favor one party over others (Gomez, Hansford, and Krause 2007). The candidate-DMA fixed effects in γ_{mj} capture the fact that, on average, it rains more in some DMAs than in others and that voters' responses may vary by candidate. The *realized* value of the demand shock ξ_{tcj} captures whether it *actually* rained on Election Day in the counties within the DMA, and ϵ_{tmj} represents a candidate's *ex-ante* uncertainty over this outcome. Candidates are also uncertain about a range of possible events between when they set advertising and Election Day; ϵ_{tmj} accounts for any potential source of uncertainty from the candidates' perspective that will be observable

¹⁴Including some form of uncertainty in the outcome of the voter model is critical. Alternative specifications exist for the precise nature of this uncertainty. For example, candidates could vary in their beliefs about the effectiveness of advertising across markets or about the share of undecided voters who will eventually turn out.

¹⁵One alternative that would make the realizations of ϵ_{tmj} non-zero would be to model candidate beliefs about the shocks nonparametrically. We could estimate a flexible function $\bar{\xi}_{tcj} = f(Z_{tcj}, A_{tmj}|\tilde{\theta})$ that allows a candidate to predict the shock given their information set.

to voters by the time they cast their votes on Election Day.

This specification implies two assumptions. First, all candidates have the same beliefs concerning the expected value of all ξ_{tcj} 's. This assumption is reasonable if candidates have access to roughly the same sources of information and are equally skilled at using this information. The second assumption is that candidates face identical degrees of uncertainty across all demand shocks, such that σ_ξ is constant across elections, DMAs, and candidates. This assumption is probably less realistic because some markets may have inherently higher levels of uncertainty in election outcomes and this uncertainty could vary over candidates. We make this assumption to reduce the number of parameters to estimate, and will consider relaxing it in the future.

Let $\xi_{ts} = \{\xi_{tcj} : \forall c \in s, j \in J\}$ be the set of demand shocks ξ_{tcj} across all counties and candidates within a state. We define A_{ts} in a similar manner. Let d_{tsj} indicate whether a candidate receives the majority of votes in a state, given by

$$d_{tsj}(A_{ts}, \xi_{ts}; \theta) = 1 \cdot \left\{ \sum_{m \in s} \sum_{c \in m} N_{tc} s_{tcj}(A_{tc}, \xi_{tc}; \theta) > \sum_{m \in s} \sum_{c \in m} N_{tc} s_{tck}(A_{tc}, \xi_{tc}; \theta), \forall k \neq j \right\} \quad (6)$$

where N_{tc} is the number of voters in a county. Under a winner-take-all rule, a candidate wins all of a state's Electoral College votes if he/she obtains a majority of the popular vote. If the state holds V_{ts} electoral votes, then the number of votes a candidate receives in the state is

$$D_{tsj}(A_{ts}, \xi_{ts}; \theta) = d_{tsj}(A_{ts}, \xi_{ts}; \theta) \cdot V_{ts}. \quad (7)$$

Let d_{tj} indicate whether a candidate wins the general election by obtaining a majority of the Electoral College votes

$$d_{tj}(A_t, \xi_t; \theta) = 1 \cdot \left\{ \sum_{s=1}^S D_{tsj}(A_{ts}, \xi_{ts}; \theta) > \bar{V}_t \right\} \quad (8)$$

where $\xi_t = (\xi_{t1}, \dots, \xi_{tS})'$, $A_t = (A_{t1}, \dots, A_{tS})'$, and $\bar{V}_t = \{269, 270\}$ for $t = 1, 2$ is the minimum number of votes required for a majority in each election.

Candidates choose advertising strategies $A_{tj} = (A_{tj1}, \dots, A_{tjm}, \dots, A_{tjM})'$ that maximize their expected return from winning the election. In the same vein as Downs (1957), Baron (1989), and others, the expected return represents the stream of benefits associated with the

candidate winning the office, which could include the perceived monetary value of winning the election, the ability to implement policies consistent with his or her preferences, or simply the candidate’s “hunger” for the office. Recall that candidates set advertising at the market level, and that advertising is constant for all counties within a market, such that $A_{tcj} = A_{tjm}, \forall c \in m$. The candidate’s expected return function is:

$$\pi_{tj}(A_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta) = R_{tj} \mathbb{E} [d_{tj}(A_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)] - \sum_{m=1}^M \omega_{tmj} A_{tjm} \quad (9)$$

such that $A_{tjm} \geq 0, \forall m, j, t$ and \mathbb{E} is the expectation with respect to the demand shocks ξ . The competing candidates’ advertising choices are bolded to indicate that candidate j potentially views their outcomes as random; in estimation we do not have to specify whether the game is of complete or incomplete information from the players’ perspectives. The marginal cost of advertising is

$$\omega_{tmj} = w_{tmj} + v_{tmj} \quad (10)$$

where w_{tmj} is an observable *estimate* of the marginal cost of advertising and v_{tmj} is measurement error observed by the candidate but unobserved to the econometrician.¹⁶

Our formulation of the candidate’s objective function deserves discussion on several points. First, we assume, as others do (e.g. Mitchell 1987, Stromberg 2008), that candidates maximize the probability of winning the election against the cost of advertising.¹⁷ An alternative candidate objective function commonly found in the literature assumes candidates maximize the expected number of electoral college votes (Brams and Davis 1974, Shachar 2009).¹⁸ In our notation, this corresponds to maximizing $\sum_s V_{ts} \mathbb{E} [d_{tsj}]$. The distinction arises from the fact that the market which yields the maximum return on a dollar of advertising could differ under either objective.

The second point concerns the model’s lack of a budget constraint. Common practice in the literature is to assume that a candidate’s advertising budget is exogenously specified and equal to the total amount spent in the election, $B_{tj} = \sum_m w_{tmj} A_{tjm}$. However, the true

¹⁶Our v_{tmj} error terms are akin to the ν_2 error terms in PPHL.

¹⁷To be precise, Stromberg (2008) models candidates as maximizing their approximate probability of winning using an argument based on the asymptotic distribution of votes.

¹⁸Snyder (1989) provides a theoretical comparison of these alternative candidate objectives.

budget is an endogenous outcome of the model arising from underlying primitives related to potential donors’ desires for the candidate to win and their assessment of the finances necessary to win. Imposing the budget constraint would create an exogenous restriction on an endogenous variable. While modeling the entire fundraising process of a candidate is beyond the scope and feasibility of this paper and data, the specification above instead models a policy invariant primitive, R_{tj} , that determines budgets.

To help interpret the parameter R_{tj} , consider the first-order condition (FOC) for advertising, defined by $h(A_{tmj}; \theta)$, such that

$$h(A_{tmj}; \theta) \equiv \frac{\partial \pi_{tj}(A_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)}{\partial A_{tmj}} = R_{tj} \frac{\partial \mathbb{E}[d_{tj}(A_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)]}{\partial A_{tmj}} - \omega_{tmj} \quad (11)$$

where the optimal advertising level sets $h(A_{tmj}^*; \theta) = 0$ (we later discuss the implications of corner solutions). The FOC shows that a candidate will choose advertising to set the marginal benefit of advertising in a given market equal to the expected marginal cost of buying advertising in that media market. Recall that the expected value of a binary random variable is equivalent to the probability the random variable is equal to one. Then the partial derivative adjacent to R_{tj} can be interpreted as the marginal increase in the probability of winning the election given a small change in the advertising level. Given that the cost term is in dollars, R_{tj} essentially converts a marginal increase in the odds of winning the election into dollar terms. Thus, we consider R_{tj} as a measure of the returns to winning the election.

4 Estimation

In this section we discuss the identification of our model’s parameters and detail our estimation strategy. We first estimate the voter model, and then take those parameters as given to estimate the candidate model.

4.1 Identification

In this section we discuss the intuition behind the identification of our model’s parameters and the counterfactuals. The voter-side parameters consist of $(\bar{\beta}_{tj}, \bar{\alpha}, \gamma_{mj}, \xi_{tcj}, \Sigma)$, and their identification follows from standard arguments when estimating aggregate market share

models. We observe variation in vote shares and advertising levels across time and many markets. The mean voter preference for a party j in election t , captured in $\bar{\beta}_{tj}$, is identified by variation over elections in the mean vote shares for candidates and differential turnout rates across each candidates' supporters. The market-candidate fixed effects, γ_{mj} , represent the mean vote shares for a candidate within a market over elections. The coefficient on advertising is identified through the variation in advertising over time, markets, and candidates. The ξ_{tej} control for unobserved factors at the election-candidate-county level that are common to all voters, and are identified as the residual of the vote shares predicted by the model. The covariance of voter tastes, Σ , is identified through variation in vote outcomes not already accounted for by the mean preference parameters.

The candidate-side parameters, (R_{tj}, σ_ξ) , require a more nuanced approach. The candidate's return to winning the election, R_{tj} , is identified based on the average advertising spending by each candidate, holding all else fixed. The identification of σ_ξ , the variance of candidates' beliefs about the demand shocks, is most easily seen by considering the extremes. When σ_ξ tends toward infinity, all outcomes approach randomness and it effectively scales down any advertising effects such that advertising is useless. Similarly, as it tends toward zero, candidates increasingly know the outcome ex-ante, in which case a laggard would have little incentive to advertise. Therefore, observing positive levels of advertising by both candidates suggests that the variance on the demand shock must be sufficiently large for the laggard to rationalize advertising at all. The expected value of the demand shocks varies over markets and candidates, leading the model to predict different margins of victory. Thus, σ_ξ is identified through variation in expected voting margins and advertising levels. If a candidate advertises in a market where they expect to lose by a wide margin, then σ_ξ must be sufficiently large to rationalize positive advertising.

4.2 Estimation of the Voter Model

To estimate the voter model, we formulate the estimation objective function as a Mathematical Program with Equilibrium Constraints (MPEC), following the work of Su and Judd (2008). We use the approach in Dubé, Fox, and Su (2009), who show how to estimate the aggregate

demand model in Berry, Levinsohn, and Pakes (1995) by formulating the GMM objective function as an MPEC problem. We extend their model to include county-party fixed effects and a full covariance matrix in taste heterogeneity.¹⁹ We briefly describe the method below and direct the reader to Dubé, Fox, and Su (2009) for more details.

The key insight to the approach is as follows. BLP (1995) use a nested optimization procedure that minimizes a GMM objective function in the outer loop while solving for the residuals ξ_{tcj} in an inner loop using the inversion in Berry (1994). One reason this procedure is computationally inefficient is that many inner loop evaluations are made when the structural parameters are still far from the optimal values.

Dubé, Fox, and Su (2009) show that formulating the problem as a MPEC substantially improves the numerical accuracy and speed of the estimator, both of which are particularly important for our application given the large number of election-county observations. Rather than explicitly solving for the residuals during each evaluation, the objective function optimizes over both the demand shocks ξ and the structural parameters θ . This eliminates a common source of error in the structural parameter estimates induced by using loose convergence tolerances for the inner loop market share inversion.

Assuming the standard orthogonality condition $\mathbb{E}[\xi_{tcj} \cdot h(z_{tcj})] = 0$ holds for some vector-valued function $h(\cdot)$ of our instruments, the empirical analog is

$$g(\xi) = \frac{1}{TCJ} \sum_{t=1}^T \sum_{c=1}^C \sum_{j=1}^J \xi_{tcj} \cdot h(z_{tcj})$$

S is the vector of observed market shares across all elections, counties, and candidates and let $s(\xi; \theta)$ be the corresponding vector of market shares implied by the model given particular values for the demand shocks and structural parameters. The MPEC objective function is

$$\begin{aligned} \min_{\{\theta, \xi, \nu\}} \quad & \nu'W\nu \\ \text{subject to} \quad & g(\xi) = \nu \\ & s(\xi; \theta) = S \end{aligned} \tag{12}$$

where W is an appropriate weighting matrix. The second constraint above enforces the

¹⁹We altered the Matlab coded posted at <http://faculty.chicagobooth.edu/jean-pierre.dube/vita/MPEC> to estimate the voter model.

normal market share inversion found in BLP (1995). We use Halton sequences (Bhat 2001) to reduce the computational burden of simulating the share integrals to compute $s(\xi; \theta)$.

4.3 Estimation of the Candidate Model

To estimate the parameters of the candidate’s return function, we use the moment inequality approach in PPHI (2006). The main benefit of following this approach is that we avoid explicitly solving for the equilibrium of the game, which could be difficult given it is a game with a large continuous action space, corner solutions, and potentially multiple equilibria. Explicitly solving for the equilibrium even once is computationally intensive and time consuming, as we note when discussing the counterfactuals. Furthermore, if the model was a game of incomplete information, we would need to make a parametric assumption about the beliefs, $\mu_j(v_{-j})$, that one candidate forms about the other candidate’s private cost shock, and then integrate over this belief when calculating each candidate’s expected response function. Using moment inequalities allows us to remain agnostic during estimation about whether the game is of complete or incomplete information and, if incomplete, the precise form of these beliefs.

Our primary source of randomness in constructing the moment inequalities comes from v_{tmj} , which is observed by candidate j , possibly observed by candidate $-j$, and unobserved by the econometrician. Recall that a complete strategy profile for a candidate is a vector of advertising levels A_{tj} across all markets. Let $A_{tj}^{m'} = \{A_{t1j}, \dots, A'_{tmj}, \dots, A_{tMj}\}$ be an alternative strategy profile with a different advertising level $A'_{tmj} \neq A_{tmj}$ substituted in market m . The key assumption in PPHI is a necessary condition for either a Nash Equilibrium or Bayesian Nash Equilibrium (BNE): candidate j ’s expected return from choosing the observed strategy A_{tj} must yield a higher expected return than choosing an alternative strategy $A_{tj}^{m'}$ holding all other candidates’ strategies fixed. Our formulation of the voter model ensures candidate’s returns are concave in A_{tmj} , implying the necessary condition is also sufficient for equilibrium.

More formally, the key necessary condition for an equilibrium can be stated as

$$\mathbb{E} [\pi_{tj}(A_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi) | z_{tmj}] \geq \mathbb{E} [\pi_{tj}(A_{tj}^{m'}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi) | z_{tmj}] \quad (13)$$

Define

$$\Delta d(A_{tj}, A_{tj}^{m'}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi) = \mathbb{E} [d_{tj}(A_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)] - \mathbb{E} [d_{tj}(A_{tj}^{m'}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)] \quad (14)$$

as the difference in the expected probability of winning under the observed and alternative advertising strategies. We can express the difference in overall returns as:

$$\Delta \pi(A_{tj}, A_{tj}^{m'}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, R_{tj}, \sigma_\xi) = R_{tj} \Delta d(A_{tj}, A_{tj}^{m'}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi) - \Delta A'_{tmj} (w_{tmj} + v_{tmj}) \quad (15)$$

where $\Delta A'_{tmj} = (A_{tmj} - A'_{tmj})$.

We now discuss how we form the moment inequalities. Provided that we have an instrument z_{tjm} in the sense that $\mathbb{E}_v[v_{tmj}|z_{tjm}] = 0$, then the difference $\Delta A'_{tmj}$ must also be uncorrelated with v_{tmj} . Starting with equation (15), dividing through by $\Delta A'_{tmj}$ ensures the unobservable is additively separable, which yields the inequality:

$$R_{tj} \frac{\Delta d(A_{tj}, A_{tj}^{m'}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)}{\Delta A'_{tmj}} - w_{tmj} - v_{tmj} \geq 0 \quad (16)$$

Thus, taking the expectation with respect to the error term leads to:

$$R_{tj} \frac{\Delta d(A_{tj}, A'_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)}{\Delta A'_{tmj}} - w_{tmj} - \mathbb{E}[v_{tmj}|z_{tjm}] \geq 0$$

This is now a useful point to discuss the implications of corner solutions, where we observe advertising levels of zero by a candidate in a market. Taken across the entire sample, $E[v_{tmj}|z_{tjm}] = 0$. However, behavior at corner solutions is substantially different such that, for example, we cannot consider a deviation in which advertising is reduced. This is problematic because the ability to take negative deviations, together with positive deviations, is what helps us obtain point identification of parameters relative to the limited identification of only analyzing whether candidates advertised or not. Ignoring the negative deviations for these observations could create a selection problem that biases the estimates. We therefore consider forming the moment for only those observations with positive advertising and adapt the above inequality as follows:

$$R_{tj} \frac{\Delta d(A_{tj}, A'_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)}{\Delta A'_{tmj}} - w_{tmj} - \mathbb{E}[v_{tmj}|z_{tjm}, A_{tjm} > 0] \geq 0$$

The next subsection on selection explicitly considers whether $\mathbb{E}[v_{tmj}|z_{tmj}, A_{tjm} > 0]$ is equal to zero or not and discusses some alternative approaches to dealing with the issue.

Given the voter parameters θ and a set of alternative advertising levels A' , the sample analogue of the condition in equation (16):

$$m(A', R_{tj}, \sigma_\xi; \theta) = \frac{1}{TMJ} \sum_{t,m,j} \left[\left(\frac{R_{tj} \Delta d(A_{tj}, A_{tj}^{m'}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)}{\Delta A'_{tjm}} - w_{tm} \right) \otimes h(z_{tmj}) \right] \geq 0$$

where $h(\cdot)$ is any positive-valued function and \otimes is the Kronecker product operator. We construct the set of alternative advertising strategies using percent deviations from the observed value. We search for the set of $\{R_{tj}, \sigma_\xi\}$ across elections, markets, and candidates that satisfy this system of inequalities, or values that minimize the extent to which these inequalities are violated:

$$\min_{\{R_{tj}, \sigma_\xi\}} \sum_{k=1}^K (\min \{0, m(A'_k, R_{tj}, \sigma_\xi; \theta)\})^2$$

where A'_k is a vector of alternative advertising strategies over candidates, markets, and elections. Any set of alternative advertising levels could be used to generate these moments. Currently, we consider the following range of alternative advertising levels: 5%, 10% and 20% deviations above and below the observed value.

4.4 Discussion of Potential Selection Issues

A necessary condition to implement the moment inequality estimation above is that $\mathbb{E}[v_{tmj}|z_{tmj}, A_{tjm} > 0] = 0$, such that we can avoid instances where we are unable to consider a downward deviation in advertising. Although the potential for such a selection problem exists in many applications, we believe the nature of our v_{tmj} term makes it safe to assume away this problem in our application.

Specifically, recall that our cost data are estimated costs produced by CMAG, the firm responsible for gathering the advertising data. The v_{tmj} represent minor measurement errors CMAG faced in estimating candidates' actual advertising costs. It seems unlikely that a candidate could receive a sufficiently large cost shock that would lead him or her not to advertise in a particular market. More likely is that a candidate would not even observe the

values of these shocks until *after* deciding to advertise some positive amount in a market. Our rationale for assuming $\mathbb{E}[v_{tmj}|z_{tmj}, A_{tjm} > 0] = 0$ is that candidates selected markets to advertise based on the w_{tmj} observed by us and their beliefs about ξ_t , which we know from the demand side. This assumption therefore allows us to estimate the supply side using only those observations where we observe positive advertising.

It is useful to compare this with the various ways PPHI suggest for dealing with selection. PPHI distinguish between two types of error terms, such that we might express $v_{tmj} = v_{2tmj} + v_{1tmj}$. The term v_{2tmj} is a structural error known by the agent when making a decision. v_{1tmj} is non-structural in the sense that it is realized after the decision is made. One approach PPHI suggest to resolving selection is assuming the structural error term is common to both agents, i.e. $v_{2tmj} = v_{2tm}$. In this case, v_{2tm} can be differenced out by taking a positive advertising deviation for one candidate ($A'_{tmj} > A_{tmj}$) and a negative deviation for the other candidate ($A'_{tmk} < A_{tmk}$). One problem with this approach is that it requires the existence of a v_{1tmj} term, which implies candidates did not know the price of advertising at the time they committed to a given quantity. If we were willing to include such a term, an alternative strategy would have been to assume away v_{2tmj} , such that $\mathbb{E}[v_{tmj}|z_{tmj}, A_{tjm} > 0] = 0$ by construction.

Given that we believe candidates knew advertising prices when their quantities were set, we prefer to assume away the existence of v_{1tmj} . We are therefore left only with a v_{2tmj} term, which we believe is unlikely to have affected a candidate's decision of whether or not to advertise in a market. Advertisers rely on publications such as SQAD's forecasts to develop their marketing plans, and then likely adjust the quantities in line with the forecast errors they realize during negotiations. The v_{2tmj} represent these errors in our model and we believe it is unreasonable to think they can be large enough to deter a candidate from advertising in a market. Furthermore, it is unlikely candidates ever knew the v_{2tmj} in a market in which they did not advertise.

4.5 Parameter Estimates (preliminary and incomplete)

Below we present parameter estimates for the voter and candidate models. Table 3 contains estimates for the voter model using fixed effects at the state, DMA, and DMA-Party level. We include a constant, an election dummy for the 2004 election, a party dummy for the Democrats, an Election-Party interaction term, and the length of advertising shown in the market. We have not calculated standard errors yet but given our sample size and first-stage estimates, we expect all the coefficients to be significant. The objective function improves substantially once we include DMA-Party fixed effects in the third column. The advertising coefficient is positive and consistently rises as we include more fixed effects. The direction of this change is consistent with the reduced-form estimates in Table 2, confirming the belief that there is a negative correlation between the unobservables and candidate advertising.

Table 3: Voter Parameter Estimates

Fixed Effects	State	DMA	DMA-Party
Constant	-0.6799	-0.2306	-0.4848
Election Dummy	0.2833	0.3055	0.2747
Party Dummy	-0.3263	-0.3138	-
Election-Party	-0.1255	-0.2228	-0.1710
Ads	0.0336	0.1064	0.3356
Obj Func	26.9403	25.2952	6.6301

To help interpret the demand estimates, we calculate the advertising elasticities using the model with DMA-Party fixed effects. The results appear in Table 4. The average elasticity of advertising with respect to a candidate’s own share of the state-wide vote is 0.102 in 2000 and is 0.160 in 2004. As expected advertising sensitivity is much higher in battleground states compared to non-battleground states.²⁰ Although these elasticities may appear small they are consistent with meta-analytic results from consumer-packaged goods in Hanssens, Parsons, and Schultz (2001, ch. 8), who report advertising elasticities of roughly 0.1 across

²⁰As defined by *The Washington Post*, the ten battleground states are: Colorado, Florida, Indiana, Iowa, Missouri, Nevada, North Carolina, Ohio, Pennsylvania, and Virginia.

Table 4: Advertising Elasticity

Election	Party	Battleground States	Non-battleground States
2000	Republicans	0.2011	0.0680
2000	Democrats	0.1936	0.0507
2004	Republicans	0.2817	0.0296
2004	Democrats	0.4264	0.0446

a variety of product categories.²¹ We must also note that our estimates may be biased due to the non-random sample of media markets in our data: the top 75 DMA's tend to lean slightly toward the Democrats.

Table 5 contains parameter estimates for the candidate model in the second column. The estimates are converted into dollar terms based on the observed spending levels. The third column of Table 5 displays the actual spending by each candidate, and shows that are estimates are roughly in the correct range. Presently we are investigating the precise reasons for the disparity in the estimates and observed spending levels. One important point to note is that our estimates of R_{tj} are implicitly normalized relative to the number of markets in the election. An election with fewer markets would presumably require less total advertising, even though spending per market could be higher or lower. Less overall advertising implies the value of winning to a candidate (our R_{tj}) must also be lower to rationalize less overall spending. If we were to estimate our model across elections with a varying number of markets, the relevant comparison would be the ratio of the return to the election-specific number of markets.

5 Counterfactual and Robustness Checks (preliminary and incomplete)

In this section we discuss our approach to conducting the counterfactual on Electoral College reform. We also describe several robustness checks on our model specification.

²¹Using a similar aggregate demand model but with an application to Canadian Parliamentary elections, Rekkas (2007) reports vote elasticities that range from 0.77 to 1.87 for overall *campaign spending*, significantly higher than our advertising elasticity estimates.

Table 5: Candidate Parameter Estimates

Parameters	In Millions of \$	Actual Spending
R_{tj} : 2000 Bush	65.9	66
R_{tj} : 2000 Gore	80.3	51
R_{tj} : 2004 Bush	69.3	84
R_{tj} : 2004 Kerry	65.2	101
σ_ξ	0.5	

5.1 Electoral College Reform

Given the structural estimates from the previous section, we are able to conduct a variety of counterfactuals. We focus on a counterfactual that substitutes the Electoral College system for a direct national election. Under this model the candidate who receives the most national votes wins the election. Compared to alternative electoral reforms, such as the proportional allocation of Electoral College votes or a congressional district-based voting allocation, voting via a direct national election has come the closest to being passed.

The primary reasons for conducting such a counterfactual are to understand how candidates reallocate their resources (e.g., advertising dollars) under a new electoral process and how voters subsequently respond to those changes.²² We expect the equilibrium allocation of advertising to change because candidates face significantly different marginal incentives to advertise in a national election. We currently observe zero advertising by a party in a number of large states, such as California and Texas, because the margin of victory is so large such that any advertising by lagging candidate is unlikely to yield any votes in the Electoral College. Thus, the return on those dollars of advertising would be close to zero. However, with a national election, advertising directly influences voters' decisions and each vote counts directly to a candidate's national tally. The incentive to direct at least some advertising funds to large, previously non-battleground states should be strong.

Computing the counterfactual requires us to explicitly solve for the equilibrium. We assume

²²For theories and evidence on how the Electoral College influences campaign resource allocation strategies and election outcomes, see Brams and Davis (1974), Colantoni, Levesque, and Ordeshook (1975), Bartels (1985), Edwards (2004)

candidates possess complete information about the structural error terms. We make this complete information assumption to avoid having to integrate over each candidate's beliefs about the competitor's private information. With a slight abuse of notation, we define

$$\tilde{d}_{tj}(A_{tj}, A_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi) = 1 \left\{ \sum_m \sum_{c \in m} N_{tc} s_{tcj}(A_{tc}, \bar{\xi}_{tc}, \epsilon_{tm}; \theta) > \sum_m \sum_{c \in m} N_{tc} s_{tck}(A_{tc}, \bar{\xi}_{tc}, \epsilon_{tm}; \theta), \forall k \neq j \right\}$$

to indicate whether candidate j receives a plurality (or majority in the case of $J = 2$) of the popular vote. Under a direct national election, each candidate chooses advertising levels to maximize the probability of winning the election given the total costs of advertising:

$$\tilde{\pi}_{tj}(A_{tj}, A_{t-j}, \boldsymbol{\xi}_t; \theta) = R_{tj} \mathbb{E} \left[\tilde{d}_{tj}(A_{tj}, A_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi) \right] - \sum_{m=1}^M \omega_{tmj} A_{tmj}.$$

We solve for the equilibrium that maximizes the total surplus, $(\tilde{\pi}_{tj} + \tilde{\pi}_{tk})$, across all the candidates in a given election.

To compute the equilibrium, we solve for the set of advertising levels $\{A_{tmj}^*, A_{tmk}^*\}$ across all markets to maximize the total surplus. This entails calculating the derivative of the probability of winning the election with respect to a candidate's advertising level in a market, given by:

$$\frac{\partial \mathbb{E}[\tilde{d}_{tj}(A_{tj}, A_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)]}{\partial A_{tmj}} = \frac{\partial}{\partial A_{tmj}} \int \dots \int \tilde{d}_{tj}(A_{tj}, A_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi) f(\epsilon_1) \dots f(\epsilon_M) \quad (17)$$

The challenge in computing this derivative is that the integrand is a non-differentiable indicator function. The indicator of winning only changes when there is a sufficient change in the number of votes received to change the outcome of the election. A shift in the advertising level A_{tmj} only changes the derivative when the candidate can win enough votes in market m to place him or her on the margin of winning the entire election. Letting $M_{-m} = \{m : M \setminus m\}$, then we can rewrite the case when candidate's vote shares are on the margin as:

$$\sum_{\tilde{m} \in M_{-m}} \sum_{c \in \tilde{m}} N_{tc} s_{tcj}(A_{t\tilde{m}}, \bar{\xi}_{tc}, \epsilon_{t\tilde{m}}; \theta) - \sum_{\tilde{m} \in M_{-m}} \sum_{c \in \tilde{m}} N_{tc} s_{tck}(A_{t\tilde{m}}, \bar{\xi}_{tc}, \epsilon_{t\tilde{m}}; \theta) = \sum_{c \in m} N_{tc} s_{tck}(A_{tm}, \bar{\xi}_{tc}, \epsilon_{tm}; \theta) - \sum_{c \in m} N_{tc} s_{tcj}(A_{tm}, \bar{\xi}_{tc}, \epsilon_{tm}; \theta)$$

In words, the equation above shows that the margin within the market (LHS) must equal the margin outside the market (RHS).

We can approximate the derivative in equation (17) using Monte Carlo methods for non-differentiable functions.²³ Define the following function:

$$h_j(A_{tj}, A_{tk}, \bar{\xi}_{tj}, \bar{\xi}_{tk}, \epsilon_{tj}, \epsilon_{tk}) = \sum_m \sum_{c \in m} N_{tc} s_{tcj}(A_{tm}, \bar{\xi}_{tc}, \epsilon_{tc}; \theta) - \sum_m \sum_{c \in m} N_{tc} s_{tck}(A_{tm}, \bar{\xi}_{tc}, \epsilon_{tc}; \theta)$$

When $h_j(\cdot) = 0$ each candidate has the same vote share. The key point is that $h_j(\cdot)$ is *weakly* increasing in all of candidate j 's arguments: a large increase in one element, say advertising A_{tmj} , will increase candidate j 's vote share until all the county vote shares within market m are (arbitrarily) close to one. Provided that the margin of votes outside the market is not too large relative to the margin within the market, we can find a value of the market shock $\epsilon_{tj}^{m*} = \{\epsilon_{t1j}, \dots, \epsilon_{tmj}^*, \dots, \epsilon_{tMj}\}$ such that $h_j(A_{tj}, A_{tk}, \bar{\xi}_{tj}, \bar{\xi}_{tk}, \epsilon_{tj}^{m*}, \epsilon_{tk}) = 0$. Then the FOC can be approximated using the following approach. We simulate NS draws of the vector ϵ . For each draw ϵ_r , we hold fixed all advertising levels $\{A_{tj}, A_{tk}\}$, all county-candidate shocks $\{\bar{\xi}_{tj}, \bar{\xi}_{tk}\}$, and the market-candidate shock $\epsilon_{tmk,r}$ for candidate k . Given these values, we can solve for the $\epsilon_{tmj,r}^*$ that sets $h_j(\cdot) = 0$, and averaging over draws yields:

$$\frac{\partial \mathbb{E}[\tilde{d}_{tj}(A_{tj}, A_{t-j}, \boldsymbol{\xi}_t; \theta, \sigma_\xi)]}{\partial A_{tmj}} \approx \frac{1}{NS} \sum_{r=1}^{NS} \delta \{h_j(A_{tj}, A_{tk}, \bar{\xi}_{tj}, \bar{\xi}_{tk}, \epsilon_{tj,r}^{m*}, \epsilon_{tk,r}) = 0\} \frac{\partial h_j(A_{tj}, A_{tk}, \bar{\xi}_{tj}, \bar{\xi}_{tk}, \epsilon_{tj,r}^{m*}, \epsilon_{tk,r})}{\partial A_{tmj}} f_{tmj}(\epsilon_{tmj,r}^*)$$

where $\delta \{\cdot\}$ is an indicator function that takes a value of one over values for ϵ_{tj}^{m*} that set $h_j(\cdot) = 0$. The approach provides a solution to our original computational difficulty. Although $\tilde{d}_{tj}(\cdot)$ is non-differentiable, the derivative of $h_j(\cdot)$ exists as closed form. It is possible that no value of $\epsilon_{tmj,r}^*$ exists such that $h_j(\cdot) = 0$, in which case $\delta \{\cdot\}$ is zero and the derivative at that r^{th} draw is zero. This will alternatively occur (a) if candidate j is the leading candidate and the margin of victory outside the market is large relative to the number of potential votes to lose inside the market or (b) if candidate j is the lagging candidate and the margin of loss outside the market is too large relative to the number of potential votes to gain inside the market.

Counterfactual results and discussion to go here

5.2 Campaign Finance Reform

Another counterfactual we consider alters the size of each candidate's budget...

²³See Glasserman (2004), chapter 7, for more information.

5.3 Robustness Checks

One goal of this paper to understand the implications of particular modeling choices in the context of political advertising. Below are two potential robustness checks we could consider. Please note that neither have been implemented yet.

5.3.1 Alternative Candidate Objective Function

Our original formulation for a candidate's objective function is to maximize the expected return to winning the election. This causes a candidate to balance the marginal increase in the odds of winning the election from an additional unit of advertising against the marginal cost of advertising.

However, some might argue that candidates seek to maximize their expected number of Electoral College votes, even if this sum is greater than the minimum required to win the election. A candidate might reasonably seek to maximize votes if a large margin of victory increases the candidate's future political capital, making it easier for them later to implement their desired policies, and thus raising the return to winning the election.

This alternative objective function is:

$$\max_{\{A_{tmj}\}_m} \pi_{tj}(A_{tj}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta) = R_{tj} \mathbb{E} \left[\sum_{s=1}^S V_{ts} \cdot d_{tsj}(A_{ts}, \mathbf{A}_{t-j}, \boldsymbol{\xi}_{ts}; \theta) \right] - \sum_{m=1}^M \omega_{tmj} A_{tmj} \quad (18)$$

such that $A_{tmj} \geq 0, \forall m, j, t$. Despite this subtle change, the model should predict different advertising levels because the new FOC for advertising is:

$$\frac{\partial \pi'_{tj}(A_{tj}; \theta)}{\partial A_{tmj}} = R_{tj} V_{ts} \frac{\partial \mathbb{E}_{\boldsymbol{\xi}} [d_{tsj}(A_{tsj}, \mathbf{A}_{ts-j}, \boldsymbol{\xi}_{ts}; \theta)]}{\partial A_{tmj}} - \omega_{tmj} \quad (19)$$

where the derivative is now only over the expected probability of winning in a particular state, as opposed to $\mathbb{E}_{\boldsymbol{\xi}} [d_{tj}(A_t, \mathbf{A}_{t-j}, \boldsymbol{\xi}_t; \theta)]$, the expected probability of winning the entire election. This simplifies the solution of the equilibrium considerably because candidates are effectively playing a set of M independent advertising games instead of one large game. This permits us to solve for the equilibrium advertising levels separately by market.

5.3.2 Decomposing Advertising by Sponsor

We observe the sponsor of each advertisement, whether it be from a candidate, a national political party, a “hybrid/coordinated” groups, or an independent group. From a voter’s perspective, the identity of the sponsor plays little to no role. Thus, we estimate the voter model including all observed advertisements A_{tcj} .

Our baseline model assumes that a candidate controls the allocation of all advertising. Although this assumption is reasonable for the first three groups (Garrett and Whitaker 2007), federal law prohibits candidates from coordinating with independent groups (such as the Swift Boat Veterans for Truth, a group formed during the 2004 presidential election). Advertising from independent groups varies as a percentage of the total advertising within a market, suggesting that the returns to advertising for local groups varies across markets.

We could estimate an alternative model that takes the advertising of independent groups A_{tcj}^I as fixed, such that candidates set A_{tcj}^C , and voters have utility:

$$u_{itcj} = \beta_{itj} + \alpha_i(A_{tcj}^I + A_{tcj}^C) + \gamma_{mj} + \xi_{tcj} + \varepsilon_{itcj} , \quad (20)$$

where $A_{tcj} = A_{tcj}^I + A_{tcj}^C$. Candidates now choose A_{tcj}^C in their objective function.

6 Conclusion

We have presented an equilibrium model of an election where candidates compete through advertising which affects voters’ decisions of whether and for whom to vote. We estimate the model using Presidential advertising data at the media market level from two recent elections and using county-level voter data.

Our model necessarily abstracts away from various realities of the election process, which creates several interesting avenues for future research. One potential extension could be to model forward-looking candidates who alter their advertising strategies as polling data shift their expectations about the ultimate outcome of the election. A second extension might endogenize the budget process by formally modeling campaign contributions. The simplest way to formulate such a model might be to consider a two-stage game consisting of a fundraising stage and an election stage. In the first stage, candidates maximize the size of their

campaign war chest by engaging in costly (e.g. either in money or time) fundraising activities. In the second stage, candidates allocate their resources (e.g., advertising, Get-Out-The-Vote campaigns, local visits) to win the election subject to the budget they obtained in the first stage.

Appendix: Computational Details

One challenging computational aspect of equation (15) is to calculate Δd . To evaluate this expression, we must reevaluate the demand-side outcomes under alternative levels of A and σ_ξ and integrate over the set of random shocks ξ_t . We use Monte Carlo simulations and importance sampling to perform this integration. First, we draw a set of NS^ξ simulated demand shocks $\{\xi_{tmj}^n\}_{n=1}^{NS^\xi}$, where $\xi_{tmj}^n \sim N(\bar{\xi}_{tmj}, \sigma_\xi^2)$ for each election, market, and candidate, for a total of $T \cdot M \cdot J \cdot NS^\xi$ draws. The importance sampling calculates the expected probability of candidate j winning election t as follows:

$$\hat{\mathbb{E}}_{t\xi} [d_{tj}(A_t, \mathbf{A}_{t-j}, \xi_t; \theta, \sigma_\xi)] = \sum_{n=1}^{NS^\xi} \left[1 \cdot \left\{ \sum_{s=1}^S D_{tsj}(A_{ts}, \mathbf{A}_{ts-j}, \xi_{0ts}^n; \theta) > \bar{V} \right\} g_t^n(\sigma_\xi | \sigma_{0\xi}) \right] \quad (21)$$

where ξ_{0ts}^n are a set of initial demand shock draws given the variance $\sigma_{0\xi}$. These draws are held fixed throughout the estimation as we solve for σ_ξ . Changing σ_ξ changes the weight of each set of draws, as follows:

$$g_t^n(\sigma_\xi | \sigma_{0\xi}) = \frac{\tilde{g}_t^n(\sigma_\xi | \sigma_{0\xi})}{\sum_{n=1}^{NS^\xi} \tilde{g}_t^n(\sigma_\xi | \sigma_{0\xi})} \quad (22)$$

$$\tilde{g}_t^n(\sigma_\xi | \sigma_{0\xi}) = \frac{\prod_{j=1}^J \left[\prod_{m=1}^M \phi \left(\frac{\xi_{0tmj}^n - \bar{\xi}_{tmj}}{\sigma_\xi} \right) \right]}{\prod_{j=1}^J \left[\prod_{m=1}^M \phi \left(\frac{\xi_{0tmj}^n - \bar{\xi}_{tmj}}{\sigma_{0\xi}} \right) \right]} \quad (23)$$

Estimation of the candidate side therefore proceeds in the following steps:

1. Calculate alternative vectors of advertising for each t, m, j deviation, e.g.

$$A_{tj}^{m'} = \{A_{t1j}, \dots, A'_{tmj}, \dots, A_{tMj}\}$$

2. Draw the initial demand shocks, $\{\xi_{0tmj}^n\}_{n=1}^{NS^\xi}$, $\xi_{0tmj}^n \sim N(\bar{\xi}_{tmj}, \sigma_{0\xi}^2)$
3. Calculate the outcome under every possible demand shock and advertising level:
 1. $\left\{ \sum_{s=1}^S D_{tsj}(A_{ts}, \mathbf{A}_{ts-j}, \xi_{0ts}^n; \theta) > \bar{V} \right\}$
4. Calculate the expected outcome for each set of simulated demand shocks and advertising levels following Equation 21, using σ_ξ^2 .

5. Calculate the moment inequalities and objective function following Equation ??, given the expected outcomes from Step 4 and the R_{tj} s.
6. Return to Step 4 as directed by the optimization routine with a new $\{R_{tj}, \sigma_\xi\}$ until convergence.
7. Return to Step 2, setting $\sigma_{0\xi}^2$ equal to the converged σ_ξ .
8. Repeat the above until $\sigma_{0\xi}$ has converged.

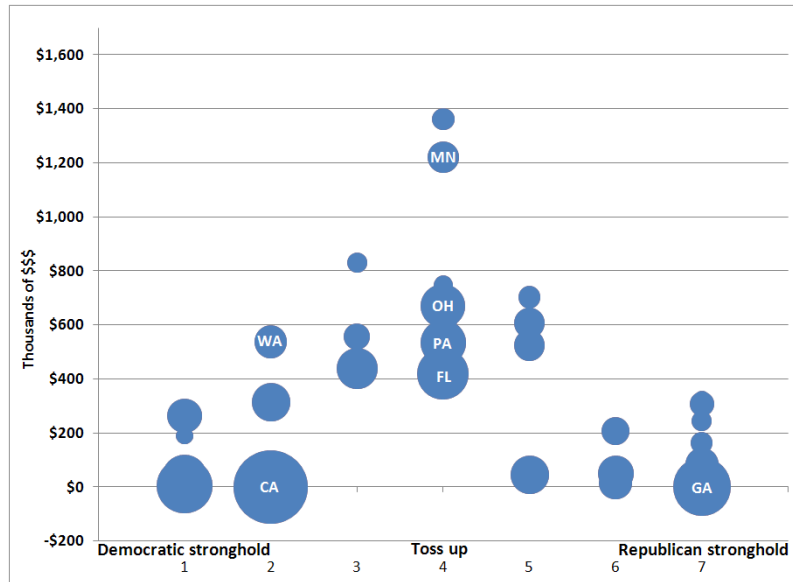
References

- [1] Ashworth, S. and J. D. Clinton (2006), “Does Advertising Exposure Affect Turnout?” *Quarterly Journal of Political Science*, 2, 27-41.
- [2] Bajari, P., L. Benkard, and J. Levin (2007), “Estimating Dynamic Models of Imperfect Competition,” *Econometrica*, 75(5), 1331-1370.
- [3] Baron, D. P. (1989), “Service-Induced Campaign Contributions and the Electoral Equilibrium,” *The Quarterly Journal of Economics*, 104(1), 45-72.
- [4] Bartels, L. M. (1985), “Expectations and Preferences in Presidential Nominating Campaigns,” *American Political Science Review*, 79(3), 804-815.
- [5] Berry, S., J. Levinsohn, and A. Pakes (1995), “Automobile Prices in Market Equilibrium,” *Econometrica*, 63(4), 841-890.
- [6] Bhat, C. (2001), “Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model,” *Transportation Research B*, 35,677-693.
- [7] Blair, D. H. (1979), “Electoral College Reform and the Distribution of Voting Power,” *Public Choice*, 34(2), 201-215.
- [8] Brams, S. J. and M. D. Davis (1974), “The 3/2’s Rule in Presidential Campaigning,” *American Political Science Review*, 68(1), 113-134.
- [9] Che, H., G. Iyer, and R. Shanmugam (2007), “Negative Advertising and Voter Choice,” Working paper, University of California, Berkeley.
- [10] Coate, S., M. Conlin, A. Moro (2008), “The Performance of Pivotal-Voter Models in Small-Scale Elections: Evidence from Texas Liquor Referenda,” *Journal of Public Economics*, 92(3-4), 582-596.
- [11] Colantoni, C. S., Levesque, T. J., and Ordeshook, P. C. (1975), “Campaign Resource Allocation Under the Electoral College,” *American Political Science Review*, 69(1), 141-154.
- [12] Degan, A. and A. Merlo (2009), “A Structural Model of Turnout and Voting in Multiple Elections,” Working paper, University of Pennsylvania.
- [13] Downs, A. (1957), *An Economic Theory of Democracy*, New York: Harper and Row.
- [14] Dubé, J. P., J. T. Fox, C. L. Su (2009), “Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation,” Working paper, University of Chicago.
- [15] Edwards, G. C. (2004), *Why the Electoral College is Bad for America*, New Haven, CT: Yale University Press.
- [16] Erikson, R. S. and T. R. Palfrey (2000), “Equilibria in Campaign Spending Games: Theory and Data,” *American Political Science Review*, 94(3), 595-609.

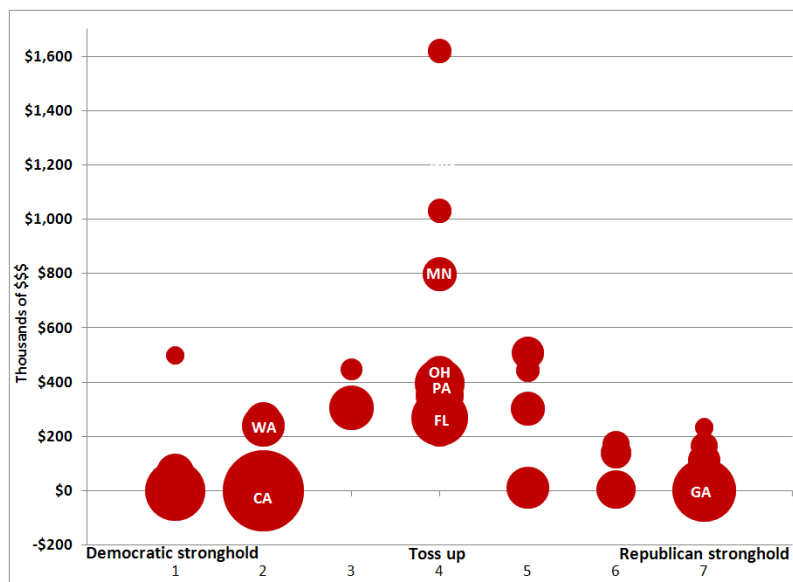
- [17] Feddersen, T. and W. Pesendorfer (1996), "The Swing Voter's Curse," *American Economic Review*, 86(3), 408-424.
- [18] Feddersen, T. and W. Pesendorfer (1999), "Abstention in Elections with Asymmetric Information and Diverse Preferences," *American Political Science Review*, 93(3), 381-398.
- [19] Freedman, P. and K. Goldstein (1999), "Measuring Media Exposure and the Effects of Negative Campaign Ads," *American Journal of Political Science*, 43(4), 1189-1208.
- [20] Garrett, R. S. and L. P. Whitaker (2007), "Coordinated Party Expenditures in Federal Elections: An Overview," *Congressional Research Service*, The Library of Congress, Order Code RS22644.
- [21] Gelman, A., G. King, and W. J. Boscardin (1998), "Estimating the Probability of Events that have Never Occurred: when is your vote decisive?" *Journal of the American Statistical Association*, 93, 1-9.
- [22] Gerber, A. (1998), "Estimating the Effect of Campaign Spending on Senate Election Outcomes Using Instrumental Variables," *American Political Science Review*, 92(2), 401.
- [20] Glasserman, P. (2004), *Monte Carlo Methods in Financial Engineering*, Springer Publishing.
- [23] Gomez, B. T., T. G. Hansford, and G. A. Krause (2007), "The Republicans Should Pray for Rain: Weather, Turnout, and Voting in U.S. Presidential Elections," *Journal of Politics*, 69(3), 649-663.
- [24] Green, D. P., and J. S. Krasno (1988), "Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections," *American Journal of Political Science*, 32(4), 884-907.
- [25] Grofman B. and S. L. Feld (2005), "Thinking About the Political Impacts of the Electoral College," *Public Choice*, 123, 1-18.
- [26] Hanssens, D., Parsons, L. J., and Schultz, R. L. (2001), *Market Response Models: Econometric and Time Series Analysis*, 2nd Edition, Kluwer Academic Publishers.
- [27] Hillygus, D. S. (2005), "The Dynamics of Turnout Intention in Election 2000," *Journal of Politics*, 67(1), 50-68.
- [28] Hopkins, D. and D. Goux (2008), "Repealing the Unit Rule: Electoral College Allocation and Candidate Strategy," Working paper, University of California, Berkeley.
- [29] Johnston, R., M. G. Hagen, and K. H. Jamieson, (2004), *The 2000 Presidential Election and the Foundations of Party Politics*, Cambridge, UK: Cambridge University Press.
- [30] Kaplan, T. R. and A. Sela (2008), "Effective Political Contests," Working paper, Department of Economics, Ben-Gurion University of the Negev.
- [31] Kawai, K. and Y. Watanabe (2010), "Inferring Strategic Voting," Working paper, Northwestern University.

- [32] Lizzeri, A. and N. Persico (2001), "The Provision of Public Goods under Alternative Electoral Incentives," *American Economic Review*, 91(1), 225-239.
- [33] Mitchell, D. W. (1987), "Candidate Behavior under Mixed Motives," *Social Choice and Welfare*, 4, 153-160.
- [34] Nelson, M. C. (1974), "Partisan Bias in the Electoral College," *Journal of Politics*, 36(4), 1033-1048.
- [35] Owen, G. (1975), "Evaluation of a Presidential Election Game," *American Political Science Review*, 69(3), 947-953.
- [36] Pakes, A., J. Porter, K. Ho, and J. Ishii (2006), "Moment Inequalities and Their Application," mimeo, Harvard University.
- [37] Powell, G. B. Jr. (1986), "American Voter Turnout in Comparative Perspective," *American Political Science Review*, 80(1), 17-43.
- [38] Rekkas, M. (2007), "The Impact of Campaign Spending on Votes in Multiparty Elections," *Review of Economics and Statistics*, 89(3), 573-585.
- [39] Shachar, R. (2009), "The Political Participation Puzzle and Marketing," *Journal of Marketing Research*, 46(6), 798-815.
- [40] Siegel, R. (2009), "All-Pay Contests," *Econometrica*, 77(1), 71-92.
- [41] Snyder, J. (1989), "Election Goals and the Allocation of Campaign Resources," *Econometrica*, 57(3), 637-660.
- [42] Stromberg, D. (2008), "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida," *American Economic Review*, 98(3), 769-807.
- [43] Su, C. L. and K. Judd (2008), "Constrained Optimization Approaches to Estimation of Structural Models," Working paper, University of Chicago.
- [44] Time Magazine (2000a), "Electoral College Debate: Election 2000: The Hidden Beauty Of The System," November 20, 2000. Accessed at <http://www.time.com/time/magazine/article/0,9171,998539-1,00.html>.
- [45] Time Magazine (2000b), "Electoral College Debate: Election 2000: It's A Mess, But We've Been Through It Before," November 20, 2000. Accessed at <http://www.time.com/time/magazine/article/0,9171,998527-2,00.html>

Figure 1: Ad Spending per Electoral College Vote in 2004 Presidential Election



(a) Democrat



(b) Republican