Journal of Marketing Research
Article Postprint
Volume XLV

The Political Participation Puzzle and Marketing

Ron Shachar
Tel Aviv University[1] and Duke University[2]

Acknowledgment:
My thanks go to Bharat Anand, Dmitri Byzalov, Ken Goldstein, Don Green, Puneet Manchanda, Carl Mela, Scott Neslin, Mitchell Lovett, Jason Roos, and Subrata Sen who were kind enough to share some of their expertise with me. I am especially grateful to the Brennan Center for Justice and Professor Ken Goldstein for allowing me to share their data and to Liraz Lasry, and Elizabeth Schneider for their gracious help as research assistants. This research was supported by the Israel Science Foundation (grant No. 1252/06)

[1]Tel Aviv University, Tel Aviv, Israel, 69978. Phone: +972-3-640-6311. E-mail: rroonn@post.tau.ac.il.
[2]Fuqua school of business, Duke University, Durham, North Carolina, USA 27708. Phone: (919) 660-8046. E-mail: shachar@duke.edu

ABSTRACT

This study shows that one of the most intriguing findings on political participation (that the participation rate is higher in close elections) is due to the omission of variables -- the marketing activities. The finding about the relationship between closeness and participation is intriguing because (1) it seems to imply that people participate in elections because their vote might be decisive, but (2) such an incentive to vote is unreasonable.

This study presents a theoretical model that suggests that closeness does not affect the turnout rate directly, but rather through the marketing activities of the parties. In other words, it is shown that in equilibrium, close elections attract higher marketing spending, which in turn increases turnout.

The author use data on the 1996-2004 presidential elections in the US to examine the model and its implications. Using structural (and non-structural) estimation the author finds that the data support the model and its implications. Furthermore, the effect of marketing on turnout is quite dramatic. For example, it turns out that if the marketing activity were cancelled in the 2004 elections, the number of voters would have decreased by 15 million.

Keywords: political marketing, marketing communications, structural estimation of an equilibrium model, advertising effectiveness, structural clustering

"Saying that closeness increases the probability of being pivotal [in elections]... is like saying that tall men are more likely than short men to bump their heads on the moon." (Schwartz 1987).

This study shows that one of the most intriguing findings on political participation is due to the omission of variables -- the marketing activities.

The intriguing empirical finding is that the participation rate is higher in close elections (see, for example, Shachar and Nalebuff 1999 who, as discussed below, did not only document this finding, but have also presented a theory to explain it). This finding seems to imply that people participate in elections because their vote might be decisive (and thus they are more likely to participate in close elections). However, such an incentive to vote is unreasonable because in a national election, the probability that someone's vote will change the outcome is essentially zero. So, why does the closeness of the race matter?

This study shows that closeness does not affect the turnout rate directly, but rather through the marketing activities of the parties. In other words, (a) the candidates devote more marketing effort and money to states with close races, (b) marketing increase turnout, and thus (c) when the marketing variables are not accounted for, it seems that closeness has a direct affect on turnout.

The model is structured to capture the setting of presidential elections in the USA. The sequence of events is the following. First, the candidates decide on (a) the allocation of their advertising budget across the 50 states, and (b) their (costly) grassroots effort in each state. Second, based on these marketing activities (and various attributes of each state), the participation rate and the vote share of each candidate are determined.

In equilibrium the marketing activity in each state is driven by the predicted closeness of the race, and by three additional state-level elements: (1) the number of electoral votes, (2) the size of the voting population, and (3) the effectiveness of marketing in stimulating turnout. Interestingly, the last factor was ignored by previous studies that implicitly assumed that marketing effectiveness does not vary across the states. However, it seems reasonable that the response to marketing does vary across the states (as suggested by this model) and that experienced political parties are aware of that and design their resource allocation accordingly. Thus, an interesting secondary research question would be: is there heterogeneity in marketing effectiveness across the states and if there is, what is its importance in explaining candidates' marketing decisions?

To test this model and its implication we have collected state-by-state data on the 1996, 2000, and 2004 presidential elections in the USA. The endogenous variables are the share of votes for each candidate, participation rates, advertising and the share of eligible voters contacted by the parties (i.e., grassroots campaign). The list of exogenous variables includes such variables as a state liberalism index and the percent of voters who moved to the state shortly before the elections.

The preliminary results, based on non-structural estimations, provide an initial support to the main hypotheses of the study -- when the marketing variables are accounted for, turnout does not depend on closeness. We also find that, as predicted by the model, the marketing variables depend on closeness and have a significant effect on turnout.

The structural estimation accounts for the endogeneity of the marketing variables and deals with additional estimation issues. It turns out that this prudent analysis also provides support to the main hypothesis—the effect of closeness on turnout is only through the marketing variables. In other words, the puzzling direct effect of closeness on turnout, found in earlier studies, was due to the exclusion of the marketing variables from the analysis.

Furthermore, our findings suggest that the effect of the marketing variables on turnout is quite dramatic. For example, using counterfactual experiments we show that if the marketing activity was cancelled in the 2004, the number of voters would have decreased by 15 million. This result indicates that marketing is an important factor in the functioning of the American democracy.

The structural estimation coupled with a unique segmentation (clustering) approach demonstrates that there is significant heterogeneity in the effectiveness of the marketing variables across the states. Furthermore, it also shows that this heterogeneity, ignored by previous studies, has an important impact on the allocation of the advertising budget. Specifically, allowing marketing effectiveness to vary across the states improves the R^2 of the turnout part of the model (from 70 percent to 96 percent) as well as the marketing part (e.g., 45 percent to 62 percent for ads).

The rest of this article is composed of six sections. In the next section we present the model. We then describe the data, followed by the preliminary results. In the following two sections, we discuss estimation issues and report the structural estimation. The final section concludes. The rest of this section presents the relevant literature. A reader who wishes to learn more about this paper can find the appendix on the journal's Web site.

Related Literature

There are three lines of work that are related to this study. First, studies that have speculated that the effect of closeness on turnout is indirect. Second, studies that have included a marketing variable in the turnout equation. Third, studies that have examined a resource allocation model. Clearly, none of the previous studies estimated a model in which closeness, turnout and the marketing variables are endogenous.

We describe the most relevant findings in the same order as above.

Indirect effect of closeness. Although they did not focus directly on the role of marketing activities in resolving the participation puzzle, Cox and Manger (1989) and Shachar and Nalebuff (1999) also conjectured that the effect of closeness on turnout is indirect.

Cox and Manger (1989) suggested that elite actors (candidates and their chief financial supporters, for example) respond to closeness with greater effort at mobilization. However, using data on the 1982 U.S. House elections, they found that *even* when campaign expenditures are accounted for, the closeness of the race *has* a direct effect on turnout. Note a couple of differences between their approach and ours: (1) unlike our data, theirs did not account for the marketing efforts of volunteers (and thus for a large portion of the grassroots campaign) and (2) they did not account for the potential endogeneity of the marketing variables and closeness.

Shachar and Nalebuff (1999) have raised a similar argument. They suggested that the population is divided into social groups, and that each group has a leader (who is not necessarily a political figure) and followers. They showed that, in equilibrium, leaders' effort is a function of the pivotalness of the state (which structurally depends not only on predicted closeness but also on the number of electoral votes and the size of the voting population), and that followers respond to such effort. Using state-by-state data on presidential elections in the U.S., they demonstrated that their model is consistent with the data. However, their structural estimation *did not* include data on effort. Thus, they could not examine whether the direct effect of closeness (and the other strategic variables) vanishes when marketing activities are accounted for.

Marketing variables in the turnout equation. The role of grassroots campaign in stimulating participation in elections was identified almost a century ago (see Gosnell 1927, Eldersveld and Dodge 1954, Eldersveld 1956, and Kramer 1970). Recently, using an extensive randomized field experiment, Gerber and Green (2000) demonstrated that personal (face-to-face) canvassing substantially increase turnout.

Lately, mostly due to the availability of data on advertising expenditures, scholars have studied the effect of advertising spending on the participation rate. Using detailed political advertising data (combined with survey data) Freedman, Franz and Goldstein (2004) demonstrated that exposure to ads has a positive effect on turnout.

There are various significant distinctions between this study and the previous work. The most important one is that previous studies (other than those using experimental data) have implicitly assumed that the marketing variables (advertising spending and grassroots campaign) are exogenous, while we account for the endogeneity of these variables.

A resource allocation model. Mostly due to the scarcity of data, the initial examinations of the allocation of campaign resources across the states were theoretical. Brams and

Davis (1973, 1974) showed that in equilibrium candidates will devote a disproportionate share of their resources to states with many electoral votes, while Snyder (1989) introduced the role of closeness in resource allocation. Following such theoretical studies, there was some evidence that ad spending (Nagler and Leighley 1992) and grassroots campaigning (Shachar and Nalebuff 1999) empirically depend on the closeness of the race.

Unlike previous studies, which either presented a theoretical model or estimated a non-structural model, we directly estimate an equilibrium-based model of resource allocation in elections. Furthermore, this study introduces a new element (ignored by previous theoretical and empirical studies) that is involved in the resource allocation decision -- the level of responsiveness of each state to marketing activities. Finally, unlike previous studies that try to endogenize either ads or grassroots efforts, this study simultaneously explains both.

Resource allocation is a central decision in marketing (consider, for example, the allocation of advertising budget across media outlets.) Recently, Manchanda, Rossi and Chintagunta (2004) study the distribution of sales force managers' efforts across physicians. They demonstrate that by accounting for the knowledge of sales force managers on the responsiveness of physicians, the econometrician can improve the precision of the estimates. Srinivasan, Raman and Naik (2005) presented a model of optimal resource allocation to corporate versus product branding.

Political Marketing

The scarcity of studies in marketing in the literature above is surprising. The volume and importance of political marketing is too significant to be ignored by marketing scholars. In the 2004 elections the two candidates spent together more than a billion dollars and the total spending on all the 2004 races were about 4 billion dollars (http://www.opensecrets.org/pressreleases/2004/04spending.asp). These numbers are very impressive since almost all the spending of political candidates is on marketing. In other words, in 2004 the spending on political marketing in the US was almost 4 billion dollars. In order to appreciate the volume of this industry one can compare it to an industry which receives a lot of attention by marketing scholars -- the movie industry. The total spending on media buys for the major studios in 2006 was estimated at a little more than $3.5 billion (http://www.hollywoodreporter.com/)

However, the importance of political marketing should not be based only on its monetary volume. Since the 1960 elections in the US it has become clear that political marketing might determine who wins the presidency.[1]

Intensifying the academic examination of political marketing can be beneficial for at least two reasons. First, some aspects of political campaigns—such as (a) the winner takes all, and (b) a common deadline for the campaign—create a semi-natural experimental setting. This setting (which is quite structured) makes it easier (than in the commercial arena) to examine some interesting research questions. Second, the data on political campaigns is rich. It is possible to get a lot of information on both voters (such as their knowledge, attitudes, perceptions, preferences, choice, etc) and candidates (such as their positions, declarations, ads, spending, etc). For example, in this study, we have data on the marketing spending of all the firms in the market. Such data is quite rare for commercial campaigns.

THE MODEL

The model is designed to capture the equilibrium relationship among turnout, marketing activities and the pivotal probability. In order to focus on turnout, we follow Shachar and Nalebuff (1999) and assume that there are no swing voters. Thus, the role of marketing activities is to increase turnout among supporters, not to change the voting tendencies of individuals. As discussed in more details below, in our model supporters of a *candidate* are not necessarily affiliated (or identify) with his *party*. Thus, our assumption does not imply that people do not vote outside of their party. This assumption is consistent with observations of scholars who have demonstrated that the campaign has a marginal effect on the choice of party and thus its main role is to stimulate turnout. Furthermore, it is important to realize (as formally demonstrated in the subsection titled as "Individuals' decisions") that even with this assumption the model still implies that an increase in the marketing spending of a candidate (ceteris paribus) leads to an increase in the number of votes that he gets. We return to this issue at the end of subsection titled as "Individuals' decisions" and show that, given our data, this assumption is not restrictive.

We start by a brief description of the setup of the model and proceed by characterizing the choices of voters and the equilibrium strategies of the parties.

The Setup

The model is structured to capture the setting of presidential elections in the USA. Two candidates, *r* and *d*, compete for the presidency. The elections are being held in *S* states on the same day. The candidate that gets more votes in state s wins all the electoral votes of that state, v_s. The candidate that gets more electoral votes wins the election.

The sequence of events is the following. In the first stage the political parties determine the marketing variables for each of the S states, and in the second stage each eligible voter in each state decides whether to participate in the elections.

Although our data consist of three election periods (1996–2004), to simplify the notations the theoretical model focuses on one election and thus ignores a time subscript (which will be added in the empirical section).

We start the analysis by characterizing the second stage decisions— the choices of the individuals (i.e., potential voters).

Individuals' Decisions

Individuals make two (related) decisions: (1) whom to vote for, and (2) whether to participate in the elections.

The first decision is exogenous to the *theoretical* model. (However, it is an endogenous variable in the estimation.) The population of eligible voters is exogenously divided between those who support d and those who support r. While we sometimes refer to the first group as "Democrats" and to the second as "Republicans," it *does not mean that these are registered voters*. It only means that each eligible voter prefers one party over the other and if she participates in the elections, she votes for the candidate of "her" party.

The share of Democrats, d_s, is a random variable with mean $z_s\theta$ and variance σ_d^2.[2] In other words, part of the variation in d_s is due to observable variables, z_s, and part is due to unobservables. Previous studies (e.g., Campbell 1992) have outlined a long list of variables that can be included in z_s (for example, a liberalism index of the state, and the composition of the state legislature party division). Such variables will be included in the empirical analysis.

The second decision, whether to participate in the election or not, is endogenous both in the theoretical model and the empirical one. This decision depends on various factors such as the individual's education, income and race. More importantly, the parties can affect the participation rate of their supporters with their marketing activities—advertising and grassroots campaign. Thus, the turnout rate among the supporters of j (i.e., the share of supporters of the candidate of party j who participate in the elections) is

$$\psi_{j,s} = \exp(\delta_s^a a_{j,s} + \delta_s^c c_{j,s} + \beta_{0,j} + \varepsilon_s) \tag{1}$$

where $a_{j,s}$ is the number of ads sent by party j in state s, $c_{j,s}$ is the share of supporters in state s who were contacted by a representative of party j, β_{0j} is the a-priori tendency to participate in the election (which is allowed to differ between the supporters of d and r), and ε is a random variable that represents the other factors influencing the participation decision. The mean of ε is $x_s\beta$ and its variance is σ_ε^2. In other words, the turnout rate is not only a function of the marketing variables, but also of other observable factors captured by x. While ε vary across states, it is common for all individuals in any specific state.

The parameters δ_s^a and δ_s^c represent the effectiveness of the marketing activities. These parameters are state-specific. This means that advertising (or grassroots campaign) might have a stronger effect in some states than in others. While we expect that on average an increase in the marketing variables would lead to an increase in the turnout rate, it is possible that δ_s^a and/or δ_s^c are negative for some of the states. In other words, while δ_s^a and δ_s^c are not necessarily non-negative, we expect that the marginal effect of marketing, across all states, is positive. The role of the heterogeneity in the δs will be discussed later.

Before presenting the next argument it is useful to restate the definition of the following symbols. The share of those who support d among the eligible voters of state s is denoted by d_s. The turnout rate among these supporters is $\psi_{d,s}$ (and $\psi_{r,s}$ is defined accordingly). Thus, the proportion of voters for d (out of all eligible voters in state s) is $d_s\psi_{d,s}$ and accordingly $(1-d_s)\,\psi_{r,s}$ is the proportion of voters for r.

Because we have assumed that the only role of marketing is to increase turnout among supporters, marketing does not change d_s. Although marketing does not change the unobserved d_s, it *does* affect the observed share of votes for d on election day, denoted by dv_s. Specifically, notice that:

$$dv_s = \frac{d_s\psi_d(a_{d,s}, c_{d,s}, \varepsilon_s)}{d_s\psi_d(a_{d,s}, c_{d,s}, \varepsilon_s) + (1-d_s)\psi_r(a_{r,s}, c_{r,s}, \varepsilon_s)} \tag{2}$$

This means that an increase in $a_{d,s}$ and/or $c_{d,s}$ leads to an increase in dv_s. Since d_s and $\psi_{j,s}$ are unobservables, the assumption above—that the marketing variables do not have a direct effect on the d_s— is not restrictive. In other words, while we observe the number of voters for d and r, we do not observe the number of supporters of each candidate and the turnout rates among supporters. Thus, our data cannot assist us in distinguishing between the effect of marketing on the choice of party and its effect on turnout rates. In this sense the assumption at hand can be considered as a normalization.

Parties' Decisions

The candidates have two marketing tools to stimulate turnout among their supporters—advertising and grassroots campaigning. Each candidate has a given national advertising budget and he needs to decide how to allocate it across the *S* states. The candidates also have local activists whose effort determines the share of supporters who are contacted. In other words, the candidates are facing a budget constraint with respect to ads and a cost function with respect to contacts.[3] Before formulating the production function of ads and the cost function of contact, we discuss the objective function of the candidates.

Candidates' objective function. Previous studies suggested two alternative candidates' goals: (1) candidates wish to maximize the expected number of electoral votes won (i.e., market share), and (2) candidates wish to maximize the probability that they will get the majority of electoral votes (i.e., probability of winning). While the second goal seems more reasonable, the first can also be justified. One such justification might be that the margin of victory can have a strong impact on the power the candidate has in office. In an early stage of the 1996 campaign it was fairly clear that the president, Bill Clinton, was likely to be reelected. Thus, it is sensible to assume that the main objective of both Clinton and his Republican competitor, Bob Dole, was to maximize their market share.

While the theoretical implications of these two alternative goals have been examined (Snyder 1989), the issue has not received any empirical attention. One of the secondary aims of this study is to take a first empirical look at this issue. To achieve this goal, the model will be estimated under two different scenarios—once assuming that the candidates' goal is only to win the election and once assuming that their goal is to maximize the expected number of electoral votes won. The comparison between the fit of both models will assist us in addressing the question at hand.

For simplicity of the presentation, the theoretical model is presented under the first assumption—the candidates maximize $\sum_s v_s p_{j,s}$, where $p_{j,s}$ is the probability that candidate j wins state s.

Advertising spending. The candidates decide about the allocation of their advertising budget, E_j, across the states. Formally,

$$\sum_s e_{j,s} \leq E_j \text{ for each } j \tag{3}$$

where $e_{j,s}$ represents the spending of candidate j on state s. When formulating the dependence of $a_{j,s}$ on $e_{j,s}$ we need to account for two real-world characteristics: (1) $a_{j,s}$ is concave in $e_{j,s}$, and (2) the cost of $a_{j,s}$ is an increasing function in the population size (For example, reaching the same proportion of voters in California is more expensive than in Rhode Island). Therefore the "production function" of ads is formulated as:

$$a_{j,s} = e_{j,s}^{\tau_a} n_s^{\gamma_a} \tag{4}$$

where n_s the size of the voting population, the parameter τ_a is bounded between 0 and 1 (i.e., $0 < \tau_a < 1$), and γ_a is a negative parameter. The assumption that $\gamma_a < 0$ implies that the cost is an increasing function of the population size (and thus $a_{j,s}$ is a decreasing function in n_s). The assumption that $0 < \tau_a < 1$ implies that the cost is convex in $a_{j,s}$ (and thus $a_{j,s}$ is concave in $e_{j,s}$).

Grassroots campaign. While the candidates face a national budget constraint when determining their advertising spending, they face a local cost function when deciding on their grassroots campaign (which is executed by local activists who invest costly effort in this activity).

When formulating the cost function of $c_{j,s}$ we need to account for two real-world characteristics—the cost is (1) a convex function, and (2) an increasing function in the population size. Formally, the cost function is:

$$Cost(c_{j,s}) = \frac{1}{\tau_c} c_{j,s}^{\tau_c} n_s^{\gamma_c} \exp(w_s \rho) \tag{5}$$

where γ_c is a positive parameter, and $\tau_c > 1$. The assumption that $\gamma_c > 0$ implies that the cost is an increasing function of the population size. The assumption that $\tau_c > 1$ implies that the cost is convex in $c_{j,s}$. Additional exogenous variables that affect the cost are denoted by w_s, and include, for example, the proportion of people living in a particular metro area. These variables are described and motivated in the data section.

Equilibrium

Next, we present the objective function of the candidates, and derive the equilibrium levels of $e_{j,s}$ and $c_{j,s}$.

We start by presenting $p_{j,s}$—the probability that candidate j wins state s. Notice that when candidates decide on their marketing activities, $e_{j,s}$ and $c_{j,s}$, they know the distribution of ε_s, and d_s, but not their realizations (which is determined on election day).

The Democratic candidate wins state s if

$$d_s \psi_d(a_{d,s}, c_{d,s}, \varepsilon_s) \geq (1-d_s)\psi_r(a_{r,s}, c_{r,s}, \varepsilon_s) \Rightarrow d_s \geq \alpha(e_{d,s}, e_{r,s}, c_{d,s}, c_{r,s}, n_s) \tag{6}$$

where $\alpha(e_{d,s}, e_{r,s}, c_{d,s}, c_{r,s}, n_s) \equiv \frac{1}{1+\exp\left[(\beta_{0,d}-\beta_{0,r})+\delta_s^a\left(e_{d,s}^{\tau_a}-e_{r,s}^{\tau_a}\right)n_s^{\gamma_a}+\delta_s^c\left(c_{d,s}-c_{r,s}\right)\right]}$

Thus $p_{d,s}$ can be written as:

$$p_{d,s}(e_{d,s}, e_{r,s}, c_{d,s}, c_{r,s}, n_s, z_s) = 1 - F_d(\alpha(e_{d,s}, e_{r,s}, c_{d,s}, c_{r,s}, n_s) \mid z_s) \quad (7)$$

where $F_d(\bullet \mid z_s)$ is the c.d.f. of d_s (and $p_{r,s}$ is obviously equal to $F_d(\bullet \mid z_s)$.)

The optimal budget-allocation and grassroots effort of candidate j (given the marketing activities of his competitor) is the solution of the following maximization problem:[4]

$$\max_{\{e_{j,s},c_{j,s}\}} \sum_s \left[v_s p_{j,s}(e_{d,s}, e_{r,s}, c_{d,s}, c_{r,s}, n_s, z_s) - \kappa \frac{1}{\tau_c} c_{j,s}^{\tau_c} n_s^{\gamma_c} \exp(w_s \rho) \right] \quad (8)$$

$$\text{s.t.} \sum_s e_{j,s} \le E_j$$

where the parameter κ "translates" the units of the cost function to electoral votes.

Using Lagrangians we get the following first order conditions for the Democratic candidate for every s (To simplify the exposition we do not present the condition $\lambda_d \left[\sum_s e_{j,s} - E_d\right] = 0$. It is easy to show that the condition is binding and thus $\lambda_d > 0$ and $\sum_s e_{d,s} = E_d$):

$$\tau_a v_s f_d\left(\alpha(\bullet) \mid z_s\right)\alpha(\bullet)\left(1-\alpha(\bullet)\right)\delta_s^a e_{d,s}^{\tau_a - 1} n_s^{\gamma_a} - \lambda_d = 0 \quad (9)$$

$$v_s f_d\left(\alpha(\bullet) \mid z_s\right)\alpha(\bullet)\left(1-\alpha(\bullet)\right)\delta_s^c - \kappa c_{d,s}^{\tau_c - 1} n_s^{\gamma_c} \exp(w_s \rho) = 0 \quad (10)$$

and for the Republican candidate (for every s):

$$\tau_a v_s f_d\left(\alpha(\bullet) \mid z_s\right)\alpha(\bullet)\left(1-\alpha(\bullet)\right)\delta_s^a e_{r,s}^{\tau_a - 1} n_s^{\gamma_a} - \lambda_r = 0 \quad (11)$$

$$v_s f_d(\alpha(\bullet) \mid z_s)\alpha(\bullet)(1-\alpha(\bullet))\delta_s^c - \kappa c_{r,s}^{\tau_c - 1} n_s^{\gamma_c} \exp(w_s \rho) = 0 \tag{12}$$

We can, now, solve for the budget allocation and grassroots effort in equilibrium. We first show that the allocation and effort of both candidates are identical, and then solve for the optimal *e* and *c* (denoted by *e** and *c**, respectively) as a function of the other variables of the model.[5]

Proposition 1 *In equilibrium,* $\frac{e^*_{d,s}}{E_d} = \frac{e^*_{r,s}}{E_r}$ *and* $c^*_{d,s} = c^*_{r,s}$ *for every s.*

Proof *From eq. 10 and eq. 12 we immediately get that* $c^*_{d,s} = c^*_{r,s}$. *From eq. 9 and eq. 11 we get:* $\lambda_d^{\frac{1}{1-\tau_a}} e^*_{d,s} = \lambda_r^{\frac{1}{1-\tau_a}} e^*_{r,s} \Rightarrow \lambda_d^{\frac{1}{1-\tau_a}} \sum_s e^*_{d,s} = \lambda_r^{\frac{1}{1-\tau_a}} \sum_s e^*_{r,s} \Rightarrow \lambda_d^{\frac{1}{1-\tau_a}} E_d = \lambda_r^{\frac{1}{1-\tau_a}} E_r$. *Since* $\lambda_d^{\frac{1}{1-\tau_a}} e^*_{d,s} = \lambda_r^{\frac{1}{1-\tau_a}} e^*_{r,s}$ *and* $\lambda_d^{\frac{1}{1-\tau_a}} E_d = \lambda_r^{\frac{1}{1-\tau_a}} E_r$ *we get that* $\Rightarrow \frac{e^*_{d,s}}{E_d} = \frac{e^*_{r,s}}{E_r}$. ■

This means that in equilibrium, the effort and the proportional allocation of the two candidates are identical.

It turns out that the complexity of the estimation decreases dramatically if in equilibrium $e^*_{d,s} = e^*_{r,s}$ (because in such a case there is an analytical solution for *e**). Of course, this would be the case if $E_d = E_r$. Such an assumption (i.e., $E_d = E_r$) is, actually, quite reasonable in general. For example, in the 2004 elections the total disbursements of George Bush and John Kerry were 359 and 333 million dollars respectively (Source: Federal Election Commission. Note, though, that this assumption might be less reasonable for the 2008 election, but the 2008 elections are not included in our data set). Furthermore, in the data section we show that the result $e^*_{d,s} = e^*_{r,s}$ is also quite reasonable empirically for our data. Thus, in order to (dramatically) simplify the estimation, we assume that $E_d = E_r = E$ and get that in equilibrium $e^*_{d,s} = e^*_{r,s}$, and as a result,

$\alpha(e_{d,s}, e_{r,s}, c_{d,s}, c_{r,s}, n_s) = \frac{1}{1+\exp(\beta_{0,d} - \beta_{0,r})}$. Notice also (from the proof of Proposition 1) that in this case $\lambda_d = \lambda_r$. We can now easily solve for the optimal marketing levels.

Proposition 2 *The equilibrium common values of the budget-allocation and the grassroots-effort are*

$$e^*_{\cdot,s} = \frac{\left(v_s f_d\left(\alpha^* \mid z_s\right)\delta^a_s n_s^{\gamma_a}\right)^{\frac{1}{1-\tau_a}}}{\sum_{\hat{s}\in S}\left(v_{\hat{s}} f_d\left(\alpha^* \mid z_{\hat{s}}\right)\delta^a_{\hat{s}} n_{\hat{s}}^{\gamma_a}\right)^{\frac{1}{1-\tau_a}}} E \tag{13}$$

where $\alpha^* \equiv \frac{1}{1+\exp(\beta_{0,d} - \beta_{0,r})}$ *and*

$$c^*_{\cdot,s} = \left[\frac{1}{\kappa} v_s f_d\left(\alpha^* \mid z_s\right)\alpha^*(1-\alpha^*)\delta^c_s n_s^{-\gamma_c}\exp(-w_s\rho)\right]^{\frac{1}{\tau_c - 1}} \tag{14}$$

Proof *To solve for* $e^*_{\cdot,s}$ *we first solve for λ (using either eq. 9 or eq. 11 and the result that* $e^*_{d,s} = e^*_{r,s}$ *and* $\lambda_d = \lambda_r$*):*

$$\lambda = \left(\frac{E}{\sum_s\left[\alpha^*(1-\alpha^*)\tau_a v_s f_d\left(\alpha^* \mid z_s\right)\delta^a_s n_s^{\gamma_a}\right]^{\frac{1}{1-\tau_a}}}\right)^{\tau_a - 1}$$

and then plug it back into 9 and 11. The solution of $c^*_{\cdot,s}$ *is immediate from 10 and 12.*

This result implies that the marketing variables depend on four elements. They increase with (a) the number of electoral votes, v_s, (b) the predicted closeness of the race, $f_d\left(\alpha^* \mid z_s\right)$, and (c) the responsiveness of voters to the marketing variables, δ^a_s and δ^c_s , and it decreases with the size of the voting population, n_s. The marketing variables also depend on the τ parameters, w_s (only for $c^*_{\cdot,s}$) and E (only for $e^*_{\cdot,s}$), but these elements are of lesser interest.

The role of three out of these four elements (electoral votes, predicted closeness and the size of the voting population) has already been highlighted by previous studies. These variables are usually referred to as the "strategic variables." However, previous studies ignored the potential heterogeneity in the effectiveness of the marketing variables. Here, we show theoretically that two states that are identical with respect to the three strategic variables might still see different marketing activity due to differences in their responsiveness level to the marketing variables. Although we do not have a theory behind the potential differences in the δs across states, it seems reasonable to assume that such variation exists and that experienced political parties are aware of it and design their resource allocation accordingly. This raises two empirical questions. First, is there heterogeneity in the δs? Second, what is the importance of such heterogeneity in explaining parties' strategies? The second question is especially interesting, since this element was ignored by previous studies.

The effect of closeness on the turnout rates. Since the marketing activities are a function of the strategic variables, the turnout rate also depends on these variables even if they do not have a direct effect. In other words, even if predicted closeness, for example, does not have a direct effect on turnout, it still has an indirect effect through the marketing variables.

Thus, without accounting for the marketing variables, it might seem that the turnout rate (conditional on x_s) is a function of the closeness of the race and the other strategic variables. However, according to this model, when the marketing variables are accounted for, the closeness of the race should not have any direct effect on turnout.

Summary of the research questions. The main research question is simple -- Is the turnout rate a function of the strategic variables when the marketing variables are accounted for? We have presented various secondary questions as well: (1) Is there heterogeneity in the δs and what is the role of such heterogeneity in explaining candidates' strategies? and (2) Is the objective function of the candidates to maximize their winning probability or the expected number of electoral votes? Finally, given the importance of turnout rates for the functioning of democracies, we would also be interested in the contribution of marketing to the turnout rate. We will elaborate on this secondary research question later.

DATA

To test the model's implications and structurally estimate it we have collected state-by-state data on the 1996, 2000 and 2004 presidential elections in the USA. These include information about the four endogenous variables of the model—election results (i.e., the share of votes for the Democratic candidate), turnout rate, ads and contact. They also contain exogenous variables such as the number of electoral votes, the percentage of eligible voters who moved into the state in the year before the elections, and a state liberalism index. The following four subsections (a) present the four endogenous variables and the exogenous variables associated with them and (b) provide a statistical description of these variables (i.e., summary statistics). For a more elaborating description of the data see the working paper version of this study (Shachar 2007).

Election Results

The endogenous variable is $dv_{s,t}$, the share of votes for the Democratic candidate in each state and election year. The election years are indexed by t. This subscript was ignored in the model section to simplify the notation. The mean and median of this variable (48.72 percent and 49.29 percent, respectively) are very close to 50 percent demonstrating that in general these elections were very close. Furthermore, in more than 71 percent of the races $dv_{s,t}$ was between 40 and 60 percent. Table 1a describes this variable and the relevant exogenous variables (such as a state liberalism index) formally and statistically.

Participation Rates

The turnout rate, denoted by $y_{s,t}$, is defined as the share of votes for both the Republican and Democratic candidates out of the voting age population. We use the voting age population rather than registered voters mainly because we consider not registering as another form of failure to participate.

The mean and median turnout rates are 50.97 percent and 51.51 percent, respectively. The simple relationship between the closeness of the race and turnout can be illustrated by the higher participation rate in close elections. For example, the average turnout rate for the 29 races in which the winner have had at most 52 percent of the votes (i.e., $.48 \leq dv_{s,t} \leq .52$) is 53.58 percent. This result is based on the ex-post closeness. The results with ex-ante closeness (i.e., $.48 \le \hat{dv}_{s,t} \le .52$, where $\hat{dv}_{s,t}$, the predicted $dv_{s,t}$, is described in the "preliminary results section") are similar.

Table 1b describes this variable and the relevant exogenous variables (such as the presence of a contemporaneous governor's race) formally and statistically.

Advertising

The raw data on advertising was created by Campaign Media Analysis Group (CMAG) for the 1996, 2000, and 2004 elections, and were made available to us by the University of Wisconsin Advertising Project (see Goldstein, Franz, and Ridout 2002 and Goldstein and Rivlin 2006). These data provide a comprehensive record of every ad broadcast on the national broadcast and cable television networks in each of the nation's top media markets.

Since the data is broken down by candidates, it allows us to examine the reasonability of the assumption $E_{d,t} = E_{r,t}$ and the result that $e^*_{d,s,t} = e^*_{r,s,t}$ (which implies, of course, $a_{d,s,t} = a_{r,s,t}$). What do we mean by "reasonability?" It is highly unlikely that (in every election year) the total budget of each of the candidates is exactly the same (i.e., $E_{d,t} = E_{r,t}$). However, this assumption is very helpful in dramatically simplifying the estimation. Thus, the empirical examination should not be whether $a_{d,s,t} = a_{r,s,t}$ is exactly right, but rather whether this is a reasonable approximation. Of course, there is no statistical test of whether an assumption is "a reasonable approximation." In other words, it is a subjective assessment. Fortunately the evidence reported below is positive enough and thus provides the support needed to determine that the assumption is indeed "a reasonable approximation."

The first piece of evidence to suggest that the assumption is reasonable is the very high correlation between $a_{d,s,t}$ and $a_{r,s,t}$. It is .84 in 1996, .93 in 2000 and .96 in 2004. These correlations imply that, as expected by our theory, the allocation of the two candidates across the various states is fairly similar.

The second, and even stronger, piece of evidence appears in Table 2. This table reports the estimates of b_0 and b_1 (for each election year) in the simple regression $a_{d,s,t} = b_0 + b_1 a_{r,s,t} + \varepsilon^b_{s,t}$ where $\varepsilon^b_{s,t}$ comes from a normal distribution with mean zero and variance σ^2_b. Strictly speaking, we expect to get $b_0 = 0$ and $b_1 = 1$. The estimates are pretty close to these values (e.g., b_1 is 1.007 in 1996, 1.008 in 2000 and 1.288 in 2004). These estimates satisfy the reasonability "requirement." Furthermore, it turns out that in most cases they even satisfy the strict statistical testing. Specifically, a simple Wald-test for each parameter separately cannot reject our null hypothesis, at the five percent level, in five out of six cases. When testing for both constraints together (i.e., $b_0 = 0$ and $b_1 = 1$) the F statistics are 2.99, .58, and 21.54 for 1996, 2000, and 2004, respectively. The critical F for testing these hypotheses (at the five percent significance level) is 3.19. This means that we can reject the null hypothesis for 2004 and the F statistics is quite close to the critical F for 1996, and thus this evidence is not the strongest vindication of the assumption. However, since we do not require a definite test, but rather just examine the reasonability of an assumption, the evidence is quite supportive. Furthermore, taken together, the collection of evidence above make a compelling argument that $a_{d,s,t} = a_{r,s,t}$ is indeed "a reasonable approximation."

Given that $a_{d,s,t} = a_{r,s,t}$, the variable used in the empirical work is the equilibrium common value presented in Proposition 2. Specifically, $a_{s,t}$, is the number of minutes of ads per day aired by the two candidates in each state between September 1st and election day.[6] The number of ad minutes is zero for 39 observations (12 in 1996, 20 in 2000, and 7 in 2004). The mean and median of this variable are 43.60 and 27.63, respectively. These numbers are not as high as they might seem at first. First, notice that they represent ad minutes *aired* by the candidates and *not exposure* to ads. Second, these minutes are spread over multiple channels, and thus the number of ad minutes per channel is much smaller.

The simple relationship between the closeness of the race and ads can be illustrated by the higher than average intensity of ads in close elections. The average ad minutes is 78.53 in races in which the winner had at most 52 percent of the votes, compared with 21.18 in the other races. Furthermore, the simple relationship between ads and turnout is demonstrated by their correlation, which is .23.

Finally, note that later we suggest that this variable might be subject to a measurement error (ME) problem.

Grassroots Campaign

This endogenous variable, $c_{s,t}$, is the share of eligible voters who were contacted by a representative of one of the parties in order to encourage turnout. This information comes from the American national elections studies conducted by the Center for Political Studies at the University of Michigan. Respondents were asked (in each election year) whether a person from one of the political parties called or visited to discuss the campaign. The share of respondents that were contacted in each state serves as our measure of grassroots campaign.

Although this measure was used in most of the previous studies, it is important to realize that it is a noisy measure of the contact rate, since it is (a) based on a sample, and (b) for several states the number of observations in the sample is small. For example, in 50 cases (out of the 150 of our data) there are fewer than 5 observations per state. Not surprisingly, many of these cases are for smaller states. For example, the average voting-age population for these 50 cases is 1.35 million compared with an average of 5.55 million for the other 100 cases.

We recognize that our variable is a noisy measure of the contact rate and thus depart from the treatment adopted by previous studies to this variable in two ways. First, we include in the data only the 100 cases in which there are at least 5 observations per state, and deal directly (and structurally) with the missing data issue. Second, as discussed below, we account for the ME in this variable.

The average contact rate in our sample (i.e., for the 100 observations) is .39. The median is .36 and the standard deviation of this variable is .17.

The simple relationship between the closeness of the race and the contact rate can be illustrated by the higher than average contact efforts in close elections. The average contact rate is 51.68 in races in which the winner has had at most 52 percent of the votes, compared with 35.61 in the other races.

Furthermore, the simple relationship between contact and turnout is demonstrated by their correlation, which is .67.

Finally, the variables that are allowed to affect the cost of the grassroots campaign (i.e., the w variables from eq. 5) are the proportion of people living in metropolitan areas, *Metro*, and a measure of the number of undergraduate students in the state, *Enrolled*.

Following Green and Gerber (2004) we assume that the cost of grassroots campaign increases with *Metro* and decreases with *Enrolled*.[7]

PRELIMINARY RESULTS

The research questions will be addressed via structural estimation in a later section. This section presents some preliminary results that offer initial support for the main hypotheses. These results can also be viewed as a description of the data. For a more elaborating description of the analysis and for some additional tests see the working paper version of this study (Shachar 2007).

The Share of Democrats

Table 3 presents the results of the regression in which the dependent variable is the share of votes for the Democratic candidates. The seven exogenous variables explain 83 percent of the variation of the dependent variable. Based on this regression we calculate $\hat{dv}_{s,t}$ (i.e., the predicted Democratic vote share) and two *rough proxies* of the predicted closeness of the race: (1) $-\left|\hat{dv}_{s,t} - .50\right|$, and (2) $I\left\{.48 \le \hat{dv}_{s,t} \le .52\right\}$ where $I\{\cdot\}$ is the indicator function. The first proxy is (minus) the margin of victory and the second is a binary variable identifying close races (i.e., the winner gets at most 52 percent of the votes).

Turnout Rate

Table 4 presents the results of a regression in which the dependent variable is the turnout rate. Each column of the table is using a different proxy for the predicted closeness. In both cases, the two strategic variables (the predicted closeness, and the ratio between the number of electoral votes and the voting age population) have the expected positive effect on the participation rate, but only the predicted closeness has a statistically significant effect (at the ten percent level for the first proxy and at the two percent for the second proxy).

The results in this table demonstrate that the effect, identified by previous studies, of the closeness of the race on participation rate exists also in the last three presidential elections in the U.S. Therefore, the question raised earlier (will the direct effect of closeness on the participation rate vanish when the marketing activities are accounted for) is relevant for this data set.

The Effect of the Strategic Variables on Marketing

In the working paper version of this study (Shachar 2007) we show that, as expected by the model, the predicted closeness of the race in each state is a strong driving force in the allocation decision. Specifically, the coefficient of the closeness variable implies that in a state in which the predicted election result is 52:48 the parties tend to air about 35 additional minutes of ads compared with a state in which the expected result is 62:38. We get a similar finding (about the role of the strategic variables in determining the marketing decisions) for the grassroots effort.

The Effect of Marketing on Turnout

The results above provided initial (and preliminary) support for some of the model's implications: the strategic variables (and especially the predicted closeness of the race) affect *both* the turnout rate, and the marketing variables (the number of ad minutes and the contact rate). We are now ready to examine the main hypothesis—that the direct effect of the strategic variables on turnout vanishes when the marketing variables are accounted for.

Table 5 addresses these questions using the binary variable proxy for the closeness of the race (i.e., $I\left\{.48 \leq \hat{dv}_{s,t} \leq .52\right\}$). The results with the other proxy are similar (see Shachar 2007).

The first column in Table 5 is the baseline model presented in Table 4 (i.e., the turnout regression without the marketing variables). The second column (marked as Model II) includes one more variable, the number of ad minutes aired in each state. The estimates provide strong initial and preliminary support to the model. First, ads significantly increase turnout. Specifically, sending ten additional minutes of ads per day increase turnout by a bit more than .5 percent. Second, the effect of predicted closeness is, now, not different from zero even at the ten percent significance level (its t-statistics decreased from 2.47 to 1.09). In other words, after accounting for the advertising spending, the direct effect of the predicted closeness vanishes. While the direct effect of the predicted closeness dies out, the effect of the other strategic variable—the ratio between the number of electoral votes and the voting population—does not. Actually, it becomes stronger and significant at the ten percent level. This phenomenon will reappear in the next columns of the table and it will be discussed in the "structural results section."

The next three columns present the estimation results for the 100 observations for which there is no missing data for the contact variable: Model III is the baseline regression, Model IV is the baseline regression with the advertising variable, and Model V is the baseline regression with both marketing variables.

Including the number of ads as an explanatory variable reduces the effect of closeness (from 4.12 to 1.11). Furthermore, the effect of predicted closeness is, now, not different from zero even at the ten percent significance level. When the contact rate variable is also included the effect of closeness drops further (to 1.06).

As expected, the two marketing variables have a *strong* positive and significant effect. As before, an increase in the number of ads by ten leads to an increase of a bit more than one half in the percent of eligible voters who participate in the elections. The effect of grassroots campaign is even more dramatic. Increasing the proportion of the population contacted by the parties by 10 percent, leads to an increase of about 2 percentage points in participation.

Finally, the missing data problem for the contact variable was treated here in a very nave fashion. As discussed in the next section, in the structural estimation we will be using all 150 observations and will account for the missing values in a structural fashion.

ESTIMATION ISSUES

This section presents the likelihood function of the structural model and discusses some additional estimation issues (such as endogeneity).

The Likelihood Function

Before presenting the likelihood, we suggest that there might be ME, hereafter in the marketing variables and formulate them. When constructing the likelihood we take into account both the issue of the measurement errors and the problem of missing observations for the contact variable. We offer a structural approach to the solution of these two problems.

Measurement Errors It is reasonable to assume that there are MEs in both marketing variables. There are several potential sources of ME in the advertising data. First, although the CMAG data cover almost all TV markets in the US, it does not cover all of them. Second, ME can be introduced also in the coding of the ads. The source of ME in the contact data is obvious—our variable is based on a sample of the population (and, in some cases, the number of observations in the sample can be small—e.g., there are 5 cases with fewer than 10 observations.)

Measurement errors introduce serious threats to the consistency and efficiency of the estimates. The most daunting risk is that of endogeneity. Recall that the marketing variables affect the turnout rate, and thus ME in these variables introduce the issue of endogeneity with respect to them.

Our solution to the ME problem is fully structural—formulating the ME function and accounting for it in the likelihood. The formulation is quite standard. Formally, (recall that the number of ad minutes cannot be negative and the share of supporters contacted is bounded between 0 and 1)

$$a^{o}_{s,t} = a^{*}_{s,t} + u_{s,t} \text{ if } a^{*}_{s,t} + u_{s,t} > 0$$

$$a^{o}_{s,t} = 0 \text{ otherwise} \tag{15}$$

and

$$c^o_{s,t} = c^*_{s,t} + \omega_{s,t} \text{ if } 0 < c^*_{s,t} + \omega_{s,t} < 1$$

$$c^o_{s,t} = 1 \text{ if } c^*_{s,t} + \omega_{s,t} \geq 1$$

$$c^o_{s,t} = 0 \text{ otherwise} \quad (16)$$

where (i) $a^*_{s,t} = \left(e^*_{s,t}\right)^{\tau_a} n^{\gamma_a}_{s,t}$, (ii) $e^*_{s,t}$ and $c^*_{s,t}$ are the equilibrium levels from equations (13) and (14), (iii) $a^o_{s,t}$ and $c^o_{s,t}$ are the ad minutes and contact rates observed (i.e., not necessarily the actual ones), and (iv) $u_{s,t}$ and $\omega_{s,t}$ are random variables with means $h_t\mu^u$ and $h_t\mu^\omega$ and variances σ^2_u and σ^2_ω, respectively. The $h_t\mu$ elements only allow differences across election years.

This formulation has important implications with respect to the estimated model. It implies that the turnout rate is a function of the (unobserved) equilibrium values of the marketing variables, $a^*_{s,t}$ and $c^*_{s,t}$, rather than the observed ones. Stated formally, it is easy to show that $\ln(y_{s,t}) = \delta^a_s a^*_{s,t} + \delta^c_s c^*_{s,t} + \beta_0 + \varepsilon_{s,t}$, where $y_{s,t}$ represents the turnout rate in state s in election year t, and it is assumed (for the simplicity of the presentation) that $\beta_{0,d} = \beta_{0,r} = \beta_0$.

Missing observations. The solution to the ME problem immediately resolves the missing data problem. Specifically, the turnout rate, $y_{s,t}$, is not a function of $c^o_{s,t}$, for which some of the observations are missing, but rather it depends on $c^*_{s,t}$, for which none of the observations are missing.

At the same time, $c^o_{s,t}$ is still a dependent variable in our model and thus the contact equation can be estimated only for a subset of the sample. Indeed, this is accounted for in eq. (18) below.

The likelihood function. We assume that all the random variables (d, ε, u and ω) come from normal distributions, and that the density function of the state-specific parameters, the δs, is discrete. We further discuss the estimation of the δs below.

To calculate the likelihood of observing $dv_{s,t}$, $y_{s,t}$, $a^o_{s,t}$, and $c^o_{s,t}$ conditional on all other observed variables, we derive the inferred values of the random variables ($d_{s,t}$, $\varepsilon_{s,t}$, $u_{s,t}$ and $\omega_{s,t}$). It is easy to show that these inferred values (denoted by*) are:[8]

$$d^*_{s,t} = \left[1 + \frac{1-dv_{s,t}}{dv_{s,t}}\exp(\beta_{0,d} - \beta_{0,r})\right]^{-1}$$

$$\varepsilon^*_{s,t} = \ln\left(\frac{y_{s,t}}{1 + d^*_{s,t}\left[\exp(\beta_{0,d} - \beta_{0,r}) - 1\right]}\right) - \beta_{0,r} - \delta^a_s a^*_{s,t}(z_t, v_t, n_t; \delta^a) - \delta^c_s c^*_{s,t}(z_{s,t}, v_{s,t}, n_{s,t}, w_{s,t}; \delta^c_s)$$

$$u^*_{s,t} = a^o_{s,t} - a^*_{s,t} = a^o_{s,t} - \left(e^*_{j,s,t}(z_t, v_t, n_t; \delta^a)\right)^{\gamma_a} n^{\gamma_a}_{s,t}$$

$$\omega^*_{s,t} = c^o_{s,t} - c^*_{s,t}(z_{s,t}, v_{s,t}, n_{s,t}, w_{s,t}; \delta^c_s) \qquad (17)$$

where z_t, v_t, n_t and δ^a are the S-element vectors whose s'th component are $z_{s,t}$, $v_{s,t}$, $n_{s,t}$ and δ^a_s, respectively. Indeed, from equations (13) and (14), we can see that while $c^*_{s,t}$ depends on $z_{s,t}$, $v_{s,t}$, $n_{s,t}$, $w_{s,t}$ and δ^c_s, the equilibrium ad minutes, $a^*_{s,t}$, is a function of z_t, v_t, n_t and δ^a. In other words, while the contact rate in state s depends only on the characteristics of that state, the number of ad minutes is a function of the attributes of all the states. These differences are due to the fact that when deciding on the grassroots effort the candidates are facing a local cost function, but when choosing ad spending they face a national budget constraint. The inter-state dependence will affect our ability to estimate δ_s using standard methods, as will be discussed later.

The relevant density functions are:[9]

$$f_1(dv_{s,t} \mid z_{s,t}) = \left[\frac{1}{\sigma_d} \phi\left(\frac{d^*_{s,t} - z_{s,t}\theta}{\sigma_d} \right) \right] \left(\frac{d^*_{s,t}}{dv_{s,t}} \right)^2 \exp(\beta_{0,d} - \beta_{0,r}) \tag{18}$$

$$f_2(y_{s,t} \mid x_{s,t}, z_t, v_t, n_t, w_{s,t}; \delta^a, \delta^c_s) = \left[\frac{1}{\sigma_\varepsilon y_{s,t}} \phi\left(\frac{\varepsilon^*_{s,t}(\delta^a, \delta^c_s) - x_{s,t}\beta}{\sigma_\varepsilon} \right) \right]$$

$$f_3(a^o_{s,t} \mid z_t, v_t, n_t; \delta^a) = \left[\frac{1}{\sigma_u} \phi\left(\frac{u^*_{s,t}(\delta^a) - h_t \mu^u}{\sigma_u} \right) \right]^{I\{a^o_{s,t} > 0\}} \left[\frac{1}{\sigma_u} \Phi\left(\frac{u^*_{s,t}(\delta^a) - h_t \mu^u}{\sigma_u} \right) \right]^{I\{a^o_{s,t} = 0\}}$$

$$f_4(c^o_{s,t} \mid z_{s,t}, v_{s,t}, n_{s,t}, w_{s,t}; \delta^c_s) = \left[\frac{1}{\sigma_\omega} \phi\left(\frac{\omega^*_{s,t}(\delta^c_s) - h_t \mu^\omega}{\sigma_\omega} \right) \right]^{I\{c^o_{s,t} < 1\}} \left[1 - \frac{1}{\sigma_\omega} \Phi\left(\frac{\omega^*_{s,t}(\delta^c_s) - h_t \mu^\omega}{\sigma_\omega} \right) \right]^{I\{c^o_{s,t} = 1\}}$$

where ϕ and Φ are the density function and cumulative density function of the normal distribution, respectively, and $I\{\}$ is the indicator function.

The likelihood function (conditional on the state specific parameters) is:

$$f_5(\Theta \mid Y, X; \delta^c, \delta^a) = \prod_{s=1}^{S} \prod_{t=1}^{3} f_1(dv_{s,t} \mid \cdot) f_2(y_{s,t} \mid \cdot; \delta^a, \delta^c_s) f_3(a^o_{s,t} \mid \cdot; \delta^a) f_4(c^o_{s,t} \mid \cdot; \delta^c_s) \tag{19}$$

where Y and X represent all the endogenous and exogenous variables (for all election years and states), respectively, and Θ stands for all the parameters of the model (other than the δs).

Clustering. The standard approach in the case of unobserved "individual" specific parameters is to integrate them out of the likelihood. This is not feasible in our case, because the ad spending in each state is a function of the unobserved δs in *each* of the other $S - 1$ states. As a result, even with only 2 segments, the number of combinations that one needs to integrate over is $1.1259E + 15$. One way to deal with such a challenge is via simulation. Another way, presented and adopted here, is via clustering. Next, we describe the clustering approach and demonstrate its effectiveness via a Monte Carlo experiment.

As in standard clustering, the basic idea is to allocate each "individual" to one of the clusters/segments. As just pointed out, it is practically impossible to search over all possible combination. Thus we suggest an algorithm that, while not necessarily identifying the clustering that leads to the highest likelihood, is most likely to do so. The rationale behind the algorithm is quite simple. In the spirit of the hierarchical clustering methods all 50 states start as one segment. Then, the state that is the most different from all the others, with respect to the δs, diverts to create a new segment. In the next step, a state that is more similar to this single state than to the average of the other 48 states joins the new segment, and so on.

The algorithm works as follows. Consider the case where the number of segments (clusters) is set at two. In the first stage, the model is fully estimated 50 times. In each of these cases there are two clusters: one that includes only one state and the other includes all other 49 states. The model is estimated 50 times, since there are 50 such combinations. In each of these estimations all the model's parameters are estimated (i.e., Θ, δ^c and δ^a).

By the end of the first stage there are 50 sets of estimates associated with 50 likelihood values. The combination that led to the highest likelihood is selected and serves as the starting point for the next stage. In the second stage, the model is estimated 49 times. In each one of these times, one of the 49 states that were grouped as one segment is moved to the other cluster. In other words, in each of the estimations there are two segments: one with two states and the other with 48 states. Once again, the combination that led to the highest likelihood is selected and serves as the starting point for the next stage. This process stops when moving states to the new cluster does not increase the likelihood any further. Of course, the model is then estimated under the assumption that the number of segments is higher than two, and the optimal number of clusters is determined via the usual information criteria.

This approach follows the tradition of the divisive methods (that are sometimes referred to as "top down") of the hierarchical clustering techniques (Everitt 1993) where in our case the likelihood function is used to "measure" the similarity of states.

To examine the effectiveness of the clustering approach, we have conducted a Monte Carlo experiment. One hundred models with two clusters were simulated and then estimated. On average, across all 100 runs, the percent of states that were identified correctly was 92. In other words, in only eight percent of the cases a state which was in one cluster was allocated by our method to another cluster. Furthermore, on average the estimates of the δ s were very precise. For example, the averages (standard deviations) of the $\hat{\delta}^{a}$ were .29988 (.0378) for the first segment and .10034 (.0368) for the second, while the true values were .3 and .1.

Testing The Main Hypothesis

The inferred values in eq. (17) represent the theoretical model. Since our model suggests that the effect of closeness on turnout is only through the marketing variables, the inferred values do not include a direct effect of closeness on turnout. However, in order to test the main hypothesis of this study, we need to allow for such effect, and thus rewrite the inferred value of ε accordingly (note: as before, we use the ratio of electoral votes to the size of the voting population instead of each one of them separately):

$$\varepsilon^*_{s,t} = \ln\left(\frac{y_{s,t}}{1 + d^*_{s,t}\left[\exp(\beta_{0,d} - \beta_{0,r}) - 1\right]}\right) - \beta_{0,r} - \delta^a_s a^*_{s,t}(z_t, v_t, n_t; \delta^a) - \delta^c_s c^*_{s,t}(z_{s,t}, v_{s,t}, n_{s,t}, w_{s,t}; \delta^c_s)$$
$$- \beta_{closeness} f_d\left(\alpha^* \mid z_{s,t}\right) - \beta_{\frac{v}{n}} \frac{v_{s,t}}{n_{s,t}} \qquad (20)$$

In other words, our model suggests that $\beta_{closeness}$ and $\beta_{\frac{v}{n}}$ are equal to zero. Indeed, $\beta_{closeness}$ and $\beta_{\frac{v}{n}}$ are freed (in the estimation) only when testing the main hypothesis. Otherwise, they are set at zero.

An alternative way to test the main hypothesis, which we also use, has two stages. In the first stage the model is estimated under the assumption $\beta_{closeness} = \beta_{\frac{v}{n}} = 0$ (i.e., using exactly the inferred values in eq. 17), and then the "residual" of the turnout equation is calculated—$\left[\varepsilon^*_{s,t}(\hat{\delta}^a, \hat{\delta}^c_s) - x_{s,t}\hat{\beta}\right]$. In the second stage, the "residual" is regressed against the strategic variables, $f_d\left(\alpha^* \mid z_{s,t}\right)$ and $\frac{v_{s,t}}{n_{s,t}}$. Our theory suggest that (a) when the model estimated in the first stage excludes the marketing variables (i.e., δ^a, δ^c_s are set at zero) the "residual" will depend on the strategic variables, but (b) when the marketing variables are included in the first stage, the strategic variables will be irrelevant (i.e., have a zero effect) in the regression of the second stage.

Identification and Endogeneity

Endogeneity. Unlike most previous studies that have examined the relationship among turnout, closeness and the marketing variables, we are not facing the endogeneity problem. (There are various reasons to believe that the inclusion of the marketing variables in the turnout equation violates the exogeneity assumption, and they are discussed in Shachar 2007.) The reason for this is quite simple—*none of these endogenous variables also serve as explanatory variables in our empirical model*. Specifically, while previous studies included the observed marketing variables and/or the ex-post closeness in the turnout equation, we use the equilibrium values of the marketing variables and the ex-ante structural measure of closeness in the turnout part of the model.

Identification. The identification of all the model parameters, with the exception of the δs, is straightforward. The identification of the δs is discussed in Shachar (2007).

STRUCTURAL ESTIMATION RESULTS

This section presents the results of the structural estimation. It starts with an examination of the main hypothesis of this study—that the effect of the strategic variables on turnout is only through the marketing variables. It then proceeds to analyze the contribution of the marketing activity to the turnout rate. Finally, we study the two secondary research questions: (1) do candidates maximize market share or winning probability? and more importantly (2) is there heterogeneity in the δs and if yes, what is its importance in the marketing decisions?

Instead of presenting only the results that directly examine the main hypothesis, Table 6 presents the parameters of interest from the structural estimation of six models. Such an approach enables an assessment of (1) the relative importance of various parts of the model and (2) the robustness of the results. Table 7 presents all the parameters (i.e., not only the parameters of interest) of the sixth model.

Table 6 specifies precisely the differences among the various models. These differences are related to the following issues: (a) The marketing variables in the turnout equation (included versus excluded), (b) the direct effect of the strategic variables in the turnout equation (included versus excluded), (c) the *w* variables in the contact rate equation (included versus excluded), (d) time effects—i.e., election years dummies—in all parts of the model (included versus excluded), and (e) the parameter τ^a (free or set at .5).

The first support for our main hypothesis comes by the comparison of models I and II. The only difference between these two models is the inclusion of the marketing variables in the second one. While the predicted closeness of the race has a positive significant effect in model I, it has a negative significant effect in model II. (The use of the term "significant" here and after implies that the relevant estimated parameter is different from zero at least at the five percent significance level.) In other words, as suggested by the main hypothesis, the positive effect of closeness is due to the exclusion of the marketing variables. While a couple of the results in model II are a bit strange (and are thus discussed shortly), they should not diminish the importance and significance implied by the comparison between models I and II. Furthermore, in the subsection titled as "testing the main hypothesis" we have suggested another way to test the main hypothesis. Using the alternative way leads to the same conclusion. Recall that the alternative way has two stages. In the first, the model is estimated without the strategic variables and the "residual" of the turnout equation is calculated. In the second, we let the strategic variables explain the "residual." We execute these two stages twice—once when the marketing variables are excluded in the first stage and once when they are included. We find that while in the first case (i.e., the marketing variables are excluded) the effect of the ex-ante closeness is .5807 (with standard error of .1925) in the second case the effect is –.2472 (with standard error of .1565).

Model I reinforces the findings of previous studies that closeness has a (significant) positive effect on turnout. Thus, the data from 1996–2004 is similar, in this sense, to the ones used by previous studies and the results reinforce the political participation puzzle. Model II demonstrates that the puzzling effect presented in model I is due to the exclusion of the marketing variables from the turnout equation. Furthermore, the effect of both marketing variables is positive and very significant, and including these variables in the turnout equation increases the R^2 of turnout from 48 percent to 61 percent.

This is not a trivial result for various reasons. First, unlike previous studies that ignored the potential endogeneity problem of the marketing variables, this study does not use the observed marketing variables in the turnout equation, but rather the unobserved equilibrium values of these variables. Thus, while previous evidence indicating that the marketing variables affect turnout might have been corrupted by the endogeneity problem, the results here are immune to such an issue. Second, model II allows the unobserved equilibrium values of the marketing variables to depend only on the strategic variables. Thus, for example, the *w* variables are excluded from the contact rate equation.

Still, there are a couple of surprising results in model II. First, the effect of predicted closeness on turnout is negative and significant. As will be seen below, this strange result vanishes in the richer models. Second, the direct effect of $\frac{v_{s,t}}{n_{s,t}}$ (electoral votes per eligible voter) is actually becoming significant (at the ten percent level) in model II. Furthermore, this strange result persists in the richer models. While this is indeed an interesting result, it is not important for the following two reasons. First, when previous studies discussed the political participation puzzle, they referred to the effect of predicted closeness on turnout not to the effect of the other two strategic variables (see, for example, the citation of Schwartz that starts this paper). In other words, what puzzled previous scholars was the effect of predicted closeness. Second, it is quite possible that the variable "electoral votes per voter" serves as a proxy for other characteristics of a state that are relevant for turnout. Such an explanation is not likely for the effect of predicted closeness.

The main hypothesis is re-examined in model V. Models III and IV are included to demonstrate the role of the *w* variables, the time effect and τ^a. In model III, the *w* variables are allowed to affect the cost of the grassroots effort and thus the equilibrium contact rate. The inclusion of these variables increases the R^2 of the contact equation from 15.6 percent to 22.5 percent, and as a result the R^2 of turnout from 61.3 percent to 63 percent.

In model IV, time effects (i.e., year-dummy variables) are included in the means of ε, u and ω. This leads to a significant improvement in the likelihood and in the fit of the turnout and contact equations. It also leads to a significant drop in δ^a, which is now not significant, and the direct effect of predicted closeness is now positive (but not significant).

In model V the parameter τ^a which was set so far at .5 is freed. Its estimate is .62 with s.e. of .06. This leads to an improvement in the R^2 of the ad equation.

Model V, the richest model in Table 6, enables us to take a final look at our main hypothesis. Once again we find that the direct effect of closeness is insignificant when the marketing variables are included. Furthermore, when model V is re-estimated without the marketing variables, the coefficient of the direct effect is 1.1679 with a standard error of .2212. In other words, the puzzling direct effect of closeness exists also in the richer version of the model.

Therefore, Table 6 demonstrates that closeness affects the turnout rate only through the marketing variables. In other words, the puzzling finding about the existence of the direct effect of closeness on turnout was due to the exclusion of the marketing variables from the analysis. We are now ready to examine the estimates of the structural estimation.

Model VI is the only one which is consistent with the theoretical model, since it is the only model that does not include the direct effect of the strategic variables. Table 7 presents all the estimated parameters of this model.[10]

The most interesting parameters in Table 7 are, obviously, δ^a and δ^c. Both of them are positive and both of them are different from zero even at the one percent level. As discussed above, this finding is especially impressive since unlike previous studies, there is no reason to expect that these δs are inconsistent.

This leads us to an assessment of the contribution of marketing to the functioning of democracies. In recent decades there is a growing concern in the US and in other democracies about the low turnout rate. Furthermore, during the last 50 years the turnout rate in many countries diminished. Our findings suggest that the marketing variables increase turnout and thus if the marketing efforts would have been curtailed, the participation rate would have been even lower. The structural estimate can be used to assess the magnitude of the decline in political participation as a result of cancellation of the marketing efforts. We find that in such a case the number of voters in the 2004 elections would have decreased by 15 million. For 2000 and 1996 the relevant numbers are 10 and 7 million. It turns out that advertising is responsible for a small fraction of this dramatic effect. For example, if the candidates would have not aired any ads in the 2004 elections the number of voters would have decreased by 3.5 million. On the other hand, canceling the grassroots effort would have led to a decrease of 12 million people. Interestingly, the chief strategists of both candidates in the 2004 elections mentioned the importance of grassroots campaigning in stimulating participation (Jamieson 2006)

The other parameters of interest in Table 7 usually have the expected sign and size and they are statistically significant. The cost of ads and contact increases in the size of the voting population (γ_a = –.43 and γ_c =.75). The cost of the grassroots campaign increases with the proportion of the population living in metropolitan areas and decreases with the proportion who enrolled into college. The cost function of producing ads is convex (τ_a =.62). The cost function of the grassroots campaign was assumed to be convex (i.e., τ_c was set at 2). While τ_c is identified, we have found that it is not estimable. In other words, in our data, the correlation between this parameter and others is too high and disables us from estimating it separately. Furthermore, using both a likelihood ratio test and a Wald test we could not reject the hypothesis that $\kappa = 1$. Since, in our data, this parameter is highly correlated with some of the other parameters it was set at 1.

The estimate of the advertising budget for 1996 is quite strange. It is smaller than the budget for the next two election years by a factor of 30. This is clearly an unreasonable estimate. Still, we need to realize that given the specific structure of the model a low estimate of E does not necessarily reflect a low budget. It is possible that the low estimate is only due to a low correlation between the ad minutes and the strategic variables. Specifically, notice that the optimal ad spending (equation 13) is a multiplication of an element which includes of all the strategic variables and E. Thus, if the empirical correlation between the ad minutes and the strategic variables is low, the model can fit the data well by setting E to be close to zero. This is exactly what happens for 1996: (a) there is a low correlation between the strategic variables and ad minutes and (b) the estimate of E is close to zero.

Finally, the other parameters in Table 7 (i.e., the θs, βs and σ s) usually have the expected sign. Although some of these estimates are not statistically different from zero, the fit of these parts of the model is quite high—the R^2 is 84 percent for the election results and 70 percent for the turnout rate. Interestingly, we find that the a-priori tendency to participate in elections is higher among Republicans than among Democrats (i.e., $\beta_{0,r} - \beta_{0,d} = .1884$ with a standard error of .1059). This finding is consistent with the results of Shachar and Nalebuff (1999).

Winning Or Market Share?

As discussed in the model section, the objective of presidential candidates is not completely clear. While the straightforward objective is winning the elections, one can argue that (at least in some cases) the aim of the candidates is to maximize their market share. The structural estimation, in which the marketing decisions of the candidates are endogenized, can shed some light on this issue. Specifically, we can assess how well the model explains the marketing choices under the assumption that the candidates maximize their winning probability versus the assumption that they maximize their market share.

The model estimated so far assumed that the candidates maximize their market share. If, instead, the candidates maximize their winning probability, the equilibrium values of the marketing variables depend on the pivotal probability of the state (rather than on its electoral votes). In other words, the element $v_{s,t}$ in equations 13 and 14 is replaced by $R_{s,t}$, where $R_{s,t}$ is the probability that state s will determine the winner in the t elections (i.e., it is the pivotal probability of state s in elections t).

Calculating the state-pivotal-probability is not trivial. Thus, we use simulation to assess it as part of our structural estimation. The resulting $R_{s,t}$ is, obviously, highly correlated with the state's number of electoral votes—.997.

The fit measures of the model under the assumption that the candidates' maximize their winning probability are worse than that of the alternative model. Specifically, replacing the number of electoral votes (in the structural estimation) with $R_{s,t}$ leads to a lower log-likelihood (115.77 versus 124.73), and lower R^2 for the marketing equations (43.25 versus 44.55 for ads and 26.67 versus 39.03 for contact).

This means that the model explains the candidates' behavior better under the assumption that they maximize their market share than under the assumption that they maximize their winning probability. While this is an interesting finding, it should be considered with caution, since it can be interpreted in other ways. For example, it is possible that the candidates wish to maximize their winning probability, but since calculating $R_{s,t}$ is so complex (recall that we were able to do it only through simulation) they use the number of electoral votes (which is highly correlated with $R_{s,t}$) as a proxy.

Clustering

Among the secondary research questions, the most intriguing and important one is about the heterogeneity of the δs: Is there heterogeneity in the δs and if there is, what is its importance in explaining candidates' marketing decisions? This question is important because this heterogeneity was ignored previously. Previous studies (theoretical and empirical) relied on the strategic variables in order to explain the allocation of the marketing budget in an election campaign. However, none of these studies examined the possibility that there is variation in the effectiveness of the marketing variables and that this variation can (at least partly) explain the allocation of the budget.

The fit measures will provide an answer to this research question for the following reason. The heterogeneity of the δs can improve the fit of the model in two ways—via an improvement in the prediction of the turnout rate and/or the prediction of the marketing variables. If the introduction of heterogeneity assists in explaining the turnout rate, it means that the effectiveness of marketing (in simulating turnout) varies across the states. If it assists in explaining the marketing variables, it means that when choosing their marketing activities across the states the candidates take into account the heterogeneity in the δs. Thus, only if the introduction of the heterogeneity in the δs will increase the fit of both the turnout rate equation and the marketing equation, we would conclude that the there is heterogeneity in the effectiveness of the marketing variables and that this heterogeneity plays a role in the marketing decisions.

Using the clustering approach described in the "estimation issues section" (and the two information criterion—BIC and CAIC) we find that the optimal number of clusters is seven, and that the data support the hypothesis. First, the fit of the turnout equation improves tremendously (R^2 increases from 69.94 with one cluster to 96.07 with seven clusters). This means that there is heterogeneity in the effectiveness of marketing across the states. Second, the fit of each marketing equation improves (the R^2 of the budget allocation part of the model increases from 44.73 with one cluster to 62.30 and the R^2 of the contact rate equation increases from 38.58 to 45.60). This means that the candidates take the heterogeneity in the δs into account when deciding on their marketing activities.

Table 8 presents the seven clusters and their δs. The largest cluster has 15 states (30 percent) and the smallest segment has only two states (4 percent). The differences in the δs across the clusters are large and significant. The δs of the largest cluster are very similar to those reported in Table 7. The cluster for which ads has the highest effectiveness (segment 2) is also the one for which the effect of contact is the lowest. Interestingly, the cluster for which grassroots campaign has the highest effectiveness (segment 7) is also the one for which the effect of ads is the lowest. Note, though, that these two segments are very small (with six states in both of them). Another interesting finding is that there is a segment for which the effectiveness of both ads and contact is negative. It is important to note that when a δ is negative, the equilibrium level of the marketing variable is zero.

The states that constitute each one of the clusters do not have any obvious connecting factors. (For a possible technical clarification of this point see the "Notes on Text Web Appendix.") While one can offer some rationale to some of these clusters (for example, it is nice to find California and New York in the same segment), the interpretation of these segments is quite challenging. Furthermore, recall that I did not offer any theoretical foundation for the heterogeneity of the δs. So, in a sense this heterogeneity is at least partly a "black box" and thus the results should be considered with some caution.

Still, these findings have two encouraging aspects. First, the significant increase in the likelihood and in the other fit measures by the introduction of these segments seems to suggest that the clustering approach presented here works well. In other words, from a methodological aspect this is quite reassuring. Second, our findings provide (at least) an initial support to the idea (which was ignored by previous, theoretical and empirical, studies) that the heterogeneity in the δs plays a role in candidates' strategic decisions.

CONCLUSION

This study shows that one of the most intriguing findings on political participation (that the participation rate is higher in close elections) is due to the omission of variables—the marketing activities. It is shown theoretically and empirically that the effect of closeness on turnout is only through the marketing variables. Furthermore, our findings suggest that the effect of the marketing variables on turnout is quite dramatic. For example, using counterfactual experiments I show that if the marketing activity was cancelled in the 2004, the number of voters would have decreased by 15 million. The structural estimation coupled with a unique segmentation (clustering) approach demonstrates that there is a significant heterogeneity in the effectiveness of the marketing variables across the states, and that this heterogeneity, ignored by previous studies, has an important impact on the allocation of the advertising budget.

The model was estimated using three elections since detailed data on political advertising does not exist for elections prior to 1996. In the near future the number of observations will increase enabling a more detailed examination of the model and its implications. In other words, with a larger sample one can estimate a richer model. Indeed, some of the assumptions and functional forms were selected in order to end-up with an analytical solution of the model because of the relatively small sample. (In general estimates of complex models using a small sample are not very reliable.)

Furthermore, the data used in future studies can be enriched in other dimensions. While this study focuses on two main elements of the marketing efforts (advertising and grassroots campaign) others can include another important resource allocated in political campaign—visits to each state of the presidential candidates, his running mate and family. Also, another factor that can create an indirect relationship between closeness and turnout is the coverage of the news media. Including such variables (campaign visits and media coverage) in the analysis can be especially useful when studying campaigns in which there are serious restrictions on the use of advertising (as is the case in some countries across the globe).

The model presented here and the rich data set used in this study can serve to address additional interesting questions. Here are three examples of such questions. First, by including swing voters in the model (and enriching the data with information such as their share in each state), one can analyze the impact of marketing on election results (i.e., on the share of votes for each of the candidates). In other words, the suggested model can be used to assess the role of marketing in determining the winner in the elections. Second (and related to the first), it seems that an important ingredient of political marketing and political advertising is negative advertising. It would be interesting to adopt the suggested model in order to examine the strategic use of negative advertising in political campaigns. Third, in this study I have treated each election year as one period. In practice, while the candidates cannot change their marketing strategy daily they can shift gears several times during the campaign. Extending the model to account for the dynamics of each election year can lead to some interesting insights on issues such as private information and signaling.

Finally, this study demonstrates, in various ways, the major role of marketing in political campaigns (for example, it has a dramatic impact on turnout). Hopefully, these results, coupled with the arguments raised in the introduction about the importance of political marketing (such as the volume of the industry) would encourage more interest in topics that exist on the border between marketing and political science.

REFERENCES

Brams, Steven and Morton Davis (1973), "Models of Resource Allocation in Presidential Campaigning: Implications for Democratic Representation," in *Democratic Representation and Apportionment: Quantitative Methods, Measures, and Criteria*, L. Papayanopoulos,ed. New York: Annals of the New York Academy of Sciences, 105–123.

Brams, Steven and Morton Davis (1974), "The 3/2's rule in Presidential Campaigning," *American Political Science Review*, 68 (1), 113-134.

Campbell, James (1992), "Forecasting the Presidential Vote in the States," *American Journal of Political Science*, 36 (2), 386–407.

Cox, Gary W. and Michael C. Munger (1989), "Closeness, Expenditures, and Turnout in the 1982 U.S. House Elections," *American Political Science Review*, 83 (1), 217–31.

Eldersveld, Samuel J. and Richard W. Dodge (1954), "Personal Contact or Mail Propaganda? An Experiment in Voting and Attitude Change," in *Public Opinion and Propaganda*, Daniel Katz, Dorwin Cartwright, Samuel Eldersveld, and Alfred M. Lee, eds. New York: Dryden Press.

Eldersveld, Samuel J. (1956), "Experimental Propaganda Techniques and Voting Behavior," *American Political Science Review*, 50 (March), 154–65.

Everitt, Brain (1993), *Cluster Analysis,* 3rd ed. London: Edward Arnold.

Freedman, P., M. Franz, and K. Goldstein (2004), "Campaign Advertising and Democratic Citizenship," *American Journal of Political Science*, 48 (4), 723–41.

Gerber, Alan S. and Donald P. Green (2000), "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment," *American Political Science Review*, 94 (3), 653–63.

Goldstein, Kenneth, Michael Franz, and Travis Ridout (2002) "Political Advertising in 2000," Combined File [dataset]. Final release. Madison, WI: The Department of Political Science at The University of Wisconsin-Madison and The Brennan Center for Justice at New York University.

Goldstein, Kenneth and Joel Rivlin (2006) "Political Advertising in 2004," Combined File [dataset]. Initial Release. Madison, WI: The Department of Political Science at The University of Wisconsin-Madison.

Gosnell, Harold (1927), *Getting-Out-The-Vote: An Experiment in the Stimulation of Voting*. Chicago: University of Chicago Press.

Green, Donald and Alan Gerber (2004), *Get Out The Vote! How to Increase Voter Turnout*. Washington, D.C.: Brookings Institution Press.

Jamieson, Kathleen (2006), *Electing the President 2004: The Insiders' View*. Philadelphia: University of Pennsylvania Press.

Kramer, Gerald H. (1970), "The Effects of Precinct-Level Canvassing on Voting Behavior," *Public Opinion Quarterly*, 34 (Winter), 560–72.

Manchanda, Puneet, Pradeep Chintagunta, and Peter Rossi (2004) "Response Modeling with Non-Random Marketing Mix Variables," *Journal of Marketing Research*, 41 (4), 467–78.

Nagler, Jonathan and Leighley, Jay (1992), "Presidential Campaign Expenditures: Evidence on Allocations and Effects," *Public Choice*, 73 (3), 319–33.

Niemi, Richard and Herbert Weisberg (2001), *Controversies in Voting Behavior*, 4th edition. Washington: CQ Press, 180–99.

Schwartz, Thomas (1987), "Your Vote Counts on Account of the Way It Is Counted: An Institutional Solution to the Paradox of Not Voting," *Public Choice*, 54 (2), 101–121.

Shachar, Ron (2007), "The Political Participation Puzzle and Marketing [Working Paper Version]," The Henry Crown Institute of Business Research in Israel, Working Paper No. 27/2007.

Shachar Ron and Barry Nalebuff (1999), "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review*, 89 (3), 525–49.

Snyder, James (1989), "Election Goals and the Allocation of Campaign Resources," *Econometrica*, 57 (3), 637–60.

Srinivasan, Shuba, Kalyan Raman and Prasad Naik (2005), "Modeling Corporate Advertising," Mimeo.

FOOTNOTES

1. The election results were practically a tie (50.0868% to Kennedy and 49.9132% to Nixon). Many pundits attributed the small advantage of Kennedy to his good performance in the first televised debate. Furthermore, many also argue that one of the main reasons that Kennedy preformed better was the fact that his advisors prepared him better to the debates. In other words, there is a perception that the better consultants were the reason to Kennedy's victory. The importance of marketing and consultants only grew since 1960. Consider, for example, the center role of Karl Rove in the Bush administration.

2. The share of democrats among the voters is obviously bounded between zero and one. However, instead of selecting a distribution that restricts d_s to be within the interval [0,1] we assume that σ_d^2 is small enough and thus the probability that d_s is smaller than zero or larger than one is practically zero.
We make a similar assumption with respect to σ_ε^2 which affects the turnout rate (and will be described shortly).
The estimates of these standard deviations (reported in the structural estimation section) are consistent with these assumptions.

3. Of course, the budget depends on the closeness of the race on the *national level*. In other words, the higher the uncertainty about the winner of the election, the higher the donations to the candidates. Notice that "the closeness of the race on the national level" varies across election years, but not across states. Thus, the most general way to account for the dependence of the budget on the closeness of the national race (in the estimation) is to consider the budget as a parameter and to estimate a specific budget for each election year. Indeed, this is the approach that we adopt.

4. To simplify the exposition we do not present the restrictions $e_{j,s} \geq 0$ and $1 \geq c_{j,s} \geq 0$ (for any j and s) in the optimization problem and, later, in the first order conditions associated with them. However, we apply these restrictions on the solution. For example, later we show that the optimal $c_{r,s}$ is equal to zero for states with negative δ_s^c.

5. Although the cost functions are convex (in the marketing variables) and ψ is log-concave, the second-order conditions are not necessarily satisfied, as demonstrated in the Appendix. The concavity of the objective function depends on the values of the model's parameters. Thus, it is tested once we have the estimated parameters in the "structural results section."

6. (i) The unit of analysis in the dataset is the broadcast of a single advertisement, with information on where and when it aired. We have aggregated the data along two dimensions—time and place. Specifically, since the proposed research is on the state level, we have aggregated the market level data to the state level using the population size in each market. Furthermore, since the model treats each election year (for each state) as one observation, we have averaged the number of ad minutes between September 1st and

elections day. This period is considered to be the most intense and critical in the campaign.

(ii) Another way to think about the approximation $a_{d,s,t} = a_{r,s,t}$ is that we are ignoring the variation due to the differences between the candidates and focus only on the variation across states. Such a strategy makes sense for two reasons. First, this is a study of the political participation puzzle. The puzzle is based on the variation in participation and closeness across states. Given that (a) for each state the closeness is the same for both parties, and (b) the data about participation is not party specific, it seems that allowing for variations in spending across candidates cannot assist us in addressing the main research question. Second, while the standard deviation of $a_{j,s,t}$ (i.e., across *candidates*, states and time) is 22.5, the standard deviation of $\frac{a_{s,t}}{2}$ (i.e., across states and time) is 21.95. In other words, by ignoring the differences between the two candidates we exclude a very small fraction of the variation in ads.

7. (a) It is harder to contact people in metro areas. On the other hand, undergraduate students are quite effective and cost efficient in contacting potential voters. We thank Don Green for suggesting these variables. (b) *Enrolled* is a binary variable which is equal to 1 for the top 25 percent of the observations in terms of the number of undergraduate students in the state, and to 0 for the other.

8. For example, we know that $dv_{s,t}$ depends on the unobserved $d_{s,t}$ as follows:

$$dv_{s,t} = \frac{d_{s,t}\psi_d(a_{d,s,t}, c_{d,s,t}, \varepsilon_{s,t})}{d_{s,t}\psi_d(a_{d,s,t}, c_{d,s,t}, \varepsilon_{s,t}) + (1-d_{s,t})\psi_r(a_{r,s,t}, c_{r,s,t}, \varepsilon_{s,t})}$$

Thus, we *infer* the true underlying percentage of Democrats, $d^*_{s,t}$:

$$d^*_{s,t} = \frac{dv_{s,t}\psi_r(a_{r,s,t}, c_{r,s,t}, \varepsilon_{s,t})}{dv_{s,t}\psi_r(a_{r,s,t}, c_{r,s,t}, \varepsilon_{s,t}) + (1-dv_{s,t})\psi_d(a_{d,s,t}, c_{d,s,t}, \varepsilon_{s,t})}$$

Inserting, $a^*_{d,s,t} = a^*_{r,s,t}$ and $c^*_{d,s,t} = c^*_{r,s,t}$ into the above, and using the functional form of ψ, we get $d^*_{s,t} = \left[1 + \frac{1-dv_{s,t}}{dv_{s,t}}\exp(\beta_{0,d} - \beta_{0,r})\right]^{-1}$.

9. (a) In these equations we account for the truncation of ads (as appear in eq. 15) but only for one type of contact's truncation. The reason for this is simple—there are no observations in our data with $c^o_{s,t} = 0$. Thus, although theoretically, the likelihood should account for the truncation of contact, we ignore it below to simplify the presentation.
(b) $f_4 = 1$ for the 50 observations for which $c^o_{s,t}$ is missing.
(c) In practice, we follow Shachar and Nalebuff (1999) and replace $f_2(y_{s,t} \mid \cdot)$ with $f_2(dp_{s,t} \mid \cdot)$ where $dp_{s,t} = d_{s,t}\psi_d$.
(d) $x_{s,t}$ includes a fixed effect for each election year.

10. As discussed in footnone 5 and in the Appendix, it is impossible, theoretically, to guarantee that the candidates' objective function is concave. Thus, concavity depends on the values of the model's parameters. We have examined this issue with the estimates which are reported in Table 7 and found that the objective function is concave. This means that the use of the first order conditions to solve for the equilibrium levels of ads and contact is justified. For more details see the Appendix .

The Political Participation Puzzle and Marketing

Ron Shachar

Web Appendix A

Technical Appendix

An interesting feature of this model is that it is impossible to guarantee that the candidates' objective function is concave.

In general: This limitation is not unique to the functional forms used in the theoretical model. To demonstrate this, we consider the objective function with general functions. To simplify the presentation we exclude ads from the model, and focus on the optimal contact level in only one state. In such a case the objective function of the Democratic candidate is

$$vp_d(c_d, c_r) - g(c_d)$$

where we denote the cost of contact by $g(\cdot)$ (and, of course, $g'' < 0$). Accordingly, the objective function of the Republican candidate is

$$v[1 - p_d(c_d, c_r)] - g(c_r)$$

If $p_d(c_d, c_r)$ is actually $p_d(c_d - c_r)$ (i.e., the probability of winning is a function of the difference in the grassroots effort), then it is impossible to guarantee that both objective functions are concave. Specifically, the second order conditions are:

$$vp_d'' - g'' \text{ for } d \text{ and } -vp_d'' - g'' \text{ for } r$$

Thus, both objective functions would be concave only if the convexity of g dominates the second derivatives. Otherwise, if the objective function is concave for one candidate, it is convex for the other.

Notice that the above is true for any functional form, as long as the probability of winning is a function of the difference in the grassroots effort.

In our case: In our case (i.e., with our functional forms), the second derivatives are as follows. For the grassroots effort:

$$\mp Q_s - \kappa n_s^{\gamma_c} \exp(w_s \rho)(\tau_c - 1) c_{j,s}^{\tau_c - 2}$$
$$\text{where } Q_s \equiv v_s \left(\delta_s^c\right)^2 f_d(\alpha \mid z_s) \alpha (1-\alpha) \left[(1 - 2\alpha) - \frac{(\alpha - z_s \theta)}{\sigma_d^2} \alpha (1 - \alpha) \right] \quad \text{(A1)}$$

and for ads:

$$v_s f_d(\alpha \mid z_s)\alpha(1-\alpha)\tau_a \delta_s^a n_s^{\gamma_a}\left[(\tau_a - 1)e_{j,s}^{\tau_a - 2} \mp \delta_s^a e_{j,s}^{2\tau_a - 2} n_s^{\gamma_a}\tau_a\left((1-2\alpha) - \frac{(\alpha - z_s\theta)}{\sigma_d^2}\alpha(1-\alpha)\right)\right] \quad \text{(A2)}$$

where in equations (A1 and (A2) the sign $\mp$ is a minus for d and a plus for r.

Concavity with the estimated parameters

Given that it is impossible to guarantee that the objective functions are concave, one needs to check for concavity with the estimated parameters. Such an analysis can include up to two steps. First, is the objective function strictly concave? In other words, is the objective function concave for every possible value of the decision variables? In our case we calculate the expression in equation (A1) for every value of α between 0 and 1 (recall that $\tau_c = 2$) and the expression in equation (A2) for every non-negative value of e_d and e_r that is smaller than the entire budget E. (Recall that α is a function of e_d and e_r).

If the objective function is strictly concave, then the values that solve the first order conditions maximize the objective functions. If the answer is negative, it is still possible that a^* and c^* are optimal, and this can be examined numerically.

Using the estimates reported in Table 7 we found that the objective function is strictly concave with respect to contact (ads) in 146 (94) out of 150 observations. For the other observations we have verified that the first order conditions are never equal to zero in areas in which the objective function is not concave. In other words, we have verified that a^* and c^* are indeed optimal.

Web Appendix B

Notes on the Text

This appendix consists of various notes on the main text. The appendix includes the titles of the sections and subsections of the paper in order to make it easy to relate the note to the relevant discussion.

INTRODUCTION

Related Literature

Marketing variables in the turnout equation. With regard to Freedman, Franz and Goldstein (2004): It is worth nothing that they have controlled for grassroots campaign and the closeness of the race. However, like most other studies, their measure of the closeness of the race was crude and ad-hoc. Specifically, it was a binary variable that is equal to one for nine states in which the race was expected to be close. On the other hand, in this study the pivotal probability is fully structural.

A resource allocation model. With regard to Nagler and Leighley 1992 and Shachar and Nalebuff 199: Nagler and Leighley (1992) used a data set, which was unique at that time, of state-by-state campaign expenditures on non-network television advertising in 1972. Using simple regression, they found that spending was higher, the greater the number of the electoral votes and the closer the race. Shachar and Nalebuff (1999) presented similar evidence with respect to grassroots campaigning. They showed that the conditional probability that a person was contacted by a party during the 1976 presidential campaign was a function of the predicted closeness of the race in her state.

THE MODEL

When discussing the impact of marketing on party choice and turnout the main text reads: “This assumption is consistent with observations of scholars who have demonstrated that the campaign has a marginal effect on the choice of party and thus its main role is to stimulate turnout.” Niemi and Weisberg (2001, p. 193) write “political parties view campaigns as ways to activate their supporters as much as ways of drawing in Independents and other nonsupporters. In other words, the campaign is necessary in order to translate the long-term forces, such as partisanship, into relevant considerations affecting the vote.”

Individuals’ Decisions

With regard to the two choices of individuals—(1) whom to vote for, and (2) whether to participate in the elections—the main text reads: “The first decision is exogenous to the *theoretical* model. (However, it is an endogenous variable in the estimation.)” In other words, I treat the share of votes for each candidate as an endogenous variable in the estimation, although I do not offer a behavioral theory for the choice of candidates.

With respect to equation (1) also notice that the turnout rate function can be made slightly more structural as follows: "a representative supporter of party j votes if his or her idiosyncratic cost μ is less than the benefit, $\psi_{j,s}$, where the distribution of μ in the population is uniform on $[0,1]$ (Shachar and Nalebuff 1999)."

Parties' Decisions

With respect to equation (3) notice that the term "production function" is used loosely. In other words, the "production function" does not account only for the cost of producing the ads but also for the cost of the media channels.

Equilibrium

With regard to Proposition 2, notice also that (a) when the objective of the candidate is to maximize the probability of winning, v_s, is replaced by the pivotal probability of state s (which is a non-linear function of all v_s and the predicted closeness of all state), and (b) the marketing variables depend on the predicted closeness not the actual closeness (which was used by some of the previous studies).

DATA

Election Results

With regard to Table 1a, notice that the exogenous variables are publicly available in early September, well prior to the November election. Thus, it is reasonable to assume that they are part of the candidates' information set when making the major marketing decisions.

Advertising

In table 2, there are various ways to examine whether $a_{d,s,t} = a_{r,s,t}$ and the exact results are sensitive to the method used (e.g., linear regression versus log-log regression). However, these differences are not critical, since I are not conducting a strict test but rather just examining the reasonability of an assumption.

Grassroots Campaign

With regard to the 100 cases in which there are at least 5 observations per state, note that the average number of observations (used to calculate $c_{s,t}$) in the ANES's survey is 41.03 and the median is 33.5.

PRELIMINARY RESULTS

The preliminary results should not be viewed as a test of any aspect of the model since the analysis does not account for various statistical concerns. For example, while the structural estimation includes all parts of the model simultaneously and accounts for the endogeneity of the marketing variables, here I estimate each equation separately and treat the marketing variables as exogenous in the turnout equation.

In Table 5, the effect of the predicted closeness is stronger in Model III than in Model I (4.12 versus 3.15). Recall that the only difference between these two models is that the first is estimated for all 150 observations and the second is only for the 100 observations that do not have missing data for the contact variable.

This interesting finding is actually consistent with the theory presented here for the following reason. The theory suggests that the effect of predicted closeness on turnout is indirect -- through the marketing variables. Thus, if the marketing variables are close to zero due to a third factor, I should expect a very small correlation between predicted closeness and turnout. This is exactly the case at hand. The 50 missing observations are for cases in which the marketing variables are close to zero because the relevant states are small and thus received very little attention from the parties. As mentioned above, the average voting population in these states is 1.35 million compared with 5.55 million for the other 100 observations. And, indeed, the average number of ads for these states is 24.40 (with a median of 0.87) compared with 36.20 for the other observations (with a median of 20.53). To test whether the correlation between predicted closeness and turnout is, indeed, low for these observations, I rerun the same regression (as in Model III) for the 50 observations with missing data on contact and found that the coefficient of predicted closeness is -.14 (with a t-statistic of -0.68). So, in a sense even this interesting finding can be considered as supporting the hypothesis that the relationship between the predicted closeness and turnout is indirect (through the marketing activities).

ESTIMATION ISSUES

The Likelihood Function

Measurement error.

(a) Recall that the turnout rate, $y_{s,t}$, is:

$$\psi_{d,s,t} d_{s,t} + \psi_{r,s,t}(1-d_{s,t}) =$$

$$\exp(\delta_s^a a_{d,s,t}^* + \delta_s^c c_{d,s,t}^* + \beta_{0,d} + \varepsilon_{s,t}) d_{s,t} + \exp(\delta_s^a a_{r,s,t}^* + \delta_s^c c_{r,s,t}^* + \beta_{0,r} + \varepsilon_{s,t})(1-d_{s,t}) =$$

$$\exp(\delta_s^a a_{s,t}^* + \delta_s^c c_{s,t}^* + \beta_0 + \varepsilon_{s,t})$$

where the last equality is due to the result $a_{d,s,t}^* = a_{r,s,t}^* = a_{s,t}^*$, $c_{d,s,t}^* = c_{r,s,t}^* = c_{s,t}^*$, and $\beta_{0,d} = \beta_{0,r} = \beta_0$.(b) One can still write the likelihood function with the observed marketing variables in the turnout equation as long as the likelihood directly accounts for the correlations (among the unobservables) implied by the ME issues. Such a likelihood is identical to the one discussed in the text.

Clustering. I have conducted a Monte Carlo experiment of another approach -- the standard simulation solution. Our preliminary examination indicated that the standard approach is inferior to the one presented here. A more through investigation of the advantages of each approach is left for future research.

Testing The Main Hypothesis

With regard to the equation presented in this subsection, recall that previous studies discussed the effect of three strategic variables on political participation—closeness, electoral votes and the size of the voting population. As before, I use the ratio of electoral votes to the size of the voting population instead of each one of them separately. The main reason for it is the very high correlation between the two.

Of course, the inclusion of these β s in the model for estimation could have been done also by offering a behavioral reason for the direct effect on individuals' participation function (eq. 1). In such a case, the inclusion of the β s above has had a structural interpretation. However, since I find this direct effect puzzling, I prefer to adopt the approach above—to include it directly in the model for estimation without a structural justification and only in order to test the main hypothesis.

STRUCTURAL ESTIMATION RESULTS

With regard to the results of model III, the jump, by a factor of about 10, in the magnitude of δ^c cannot be only attributed to its increased impact on turnout when measured more precisely. An important reason for this jump is the fact that the average $\exp(w_s\rho)$ is about 0.1, and as can be seen in equation (12), δ^c can balance this effect.

With regard to the results of model IV: What is happening here? Both the tendency to participate in the elections and the total number of ad-minutes increased between 1996 and 2004. When the turnout equation did not include time effects, the main explanation for the increase in the tendency to participate was the increase in the predicted ad-minutes across these elections years. (Notice that although the mean of u was not allowed to vary across election years, such variation was enabled via the estimate of E, as discussed later.) As a result, when the time effects were included, δ^a dropped, and this leads to the increase in the direct effect of predicted closeness. Specifically, when δ^a decreased the indirect effect of predicted closeness diminished and to balanced it the direct effect increased.

I estimate a separate $\beta_{0,d}$ for each election year and instead of estimating $\beta_{0,r}$, I estimate the difference between it and $\beta_{0,d}$.

Winning or Market Share?

With regard to the simulation of $R_{s,t}$ I adopt this approach from Shachar and Nalebuff (1999) who explain it as follows: "As part of our structural estimation, I assess the probability that the Democratic candidate would win each state in each election year. Using these probabilities, I then randomly draw the winner of each state race in each election year. Based on this and the number of electoral votes each state has in each election year, I determine for each state whether it would have changed the election results in this scenario. I repeat this 100,000 times and set $R_{s,t}$ to be equal to the proportion of times that state s was pivotal in period t."

Clustering

In the main text there is a sentence that reads, "The states that constitute each one of the clusters do not have any obvious connecting factors." Also, the results of this clustering procedure (i.e., the parameters and the states in each cluster) are somewhat sensitive to technical issues such as starting values. Thus, it is possible, for example, that a state that belongs to the first cluster here would belong to the sixth cluster in a different run (notice that the δ s of these segments are relatively similar). However, (a) the results reported here are the ones that lead to the highest likelihood in our analysis and (b) the main result of this section—that the fit of both the turnout equation and the marketing equations improve as a result of the introduction of heterogeneity in the δ s—is not sensitive to such technical issues.

CONCLUSION

With regard to the opportunity to use the model presented in this study in order to assess the role of marketing in determining the winner in the elections, there is an extensive literature in political science on this exact topic. However, it seems that the combination of (a) the marketing discipline, (b) the presented model, and the (c) data set might be able to shed a new light on this issue.