RON SHACHAR*

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, namely, the marketing activities. This relationship between closeness and participation is intriguing because (1) it implies 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, in equilibrium, close elections attract higher marketing spending, which in turn increases turnout. The author uses data on the 1996–2004 presidential elections in the United States to examine the model and its implications. Using structural (and nonstructural) estimation, the author finds that the data support the model and its implications. Furthermore, the effect of marketing on turnout is dramatic. For example, if the marketing activity were canceled 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

The Political Participation Puzzle and Marketing

> 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, p. 118)

This study shows that one of the most intriguing findings on political participation is due to the omission of variables, namely, the marketing activities. The empirical finding is that the participation rate is higher in close elections (see Shachar and Nalebuff [1999], who not only document this finding but also present a theory to explain it). This finding implies 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, (1) the candidates devote more marketing effort and money to states with close races, (2) marketing increases turnout, and thus (3) when the marketing variables are not accounted for, it seems that closeness has a direct affect on turnout.

The model herein is structured to capture the U.S. presidential election setting. The sequence of events is as follows: First, the candidates determine (1) the allocation of their advertising budget across the 50 states and (2) their (costly) grassroots effort in each state. Second, from 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

*Ron Shachar is Chairperson of the Marketing Department, Tel Aviv University, and Visiting Associate Professor, Fuqua School of Business, Duke University (e-mail: rroonn@post.tau.ac.il). The author thanks 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. The author is especially grateful to the Brennan Center for Justice and Professor Ken Goldstein for allowing him 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). Tülin Erdem served as guest editor and Pradeep Chintagunta served as associate editor for this article.

ISSN: 0022-2437 (print), 1547-7193 (electronic)

Journal of Marketing Research
Vol. XLVI (December 2009), 798–815

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. Notably, the last factor has been ignored in previous studies, which implicitly assume that marketing effectiveness does not vary across the states. However, it is reasonable that the response to marketing varies across the states (as suggested by this model) and that experienced political parties are aware of this and design their resource allocation accordingly. Thus, a 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, state-by-state data were collected for the 1996, 2000, and 2004 presidential elections in the United States. 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, for example, a state liberalism index and the percentage of voters who moved to the state shortly before the elections.

The preliminary results, based on nonstructural estimations, provide initial support for the main hypotheses of the study: When the marketing variables are accounted for, turnout does not depend on closeness. In addition, as the model predicts, 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. This prudent analysis also provides support for the main hypothesis that the effect of closeness on turnout is only through the marketing variables. In other words, the puzzling direct effect of closeness on turnout, as found in prior studies, is due to the exclusion of the marketing variables from the analysis.

Furthermore, the findings of the current study suggest that the effect of the marketing variables on turnout is dramatic. For example, counterfactual experiments show that if the marketing activity had been canceled in the 2004 election, the number of voters would have decreased by 15 million. This result indicates that marketing is an important factor in the functioning of democracy in the United States.

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, which previous studies have ignored, has an important impact on the allocation of the advertising budget. Specifically, allowing marketing effectiveness to vary across the states improves the R-square of both the turnout part of the model (from 70% to 96%) and the marketing part (e.g., from 45% to 62% for advertisements).

The rest of this article is organized as follows: The next section begins with a discussion of the relevant literature. Then, the model is presented. After this, the data are described, followed by the preliminary results. In the following sections, estimation issues are discussed, and the structural estimation is reported. The final section offers some conclusions. (To learn more about this article, see the Web Appendix at http://www.marketingpower.com/jmrdec09, which includes many of the footnotes that were removed from the main article because of space constraints.)

RELATED LITERATURE

Three lines of work are related to this study: (1) studies that have speculated that the effect of closeness on turnout is indirect, (2) studies that have included a marketing variable in the turnout equation, and (3) studies that have examined a resource allocation model. However, none of this prior research has estimated a model in which closeness, turnout, and the marketing variables are endogenous. The most relevant findings in each of these lines of work are discussed.

Indirect Effect of Closeness

Although they do not focus directly on the role of marketing activities in resolving the participation puzzle, Cox and Manger (1989) and Shachar and Nalebuff (1999) also conjecture that the effect of closeness on turnout is indirect. Cox and Manger (1989) suggest that elite actors (e.g., candidates and their chief financial supporters) respond to closeness with greater effort at mobilization. However, using data on the 1982 U.S. House elections, they find that even when campaign expenditures are accounted for, the closeness of the race has a direct effect on turnout. There are a couple of differences between their approach and that presented herein: (1) Unlike the current data, their data do not account for the marketing efforts of volunteers (and, thus, for a large portion of the grassroots campaign), and (2) they do not account for the potential endogeneity of the marketing variables and closeness.

Shachar and Nalebuff (1999) raise a similar argument. They suggest 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 show that, in equilibrium, leaders' efforts are 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 efforts. Using state-by-state data on presidential elections in the United States, they demonstrate that their model is consistent with the data. However, their structural estimation does 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 taken into account.

Marketing Variables in the Turnout Equation

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

Lately, mostly because of 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) demonstrate that exposure to advertisements has a positive effect on turnout.

There are various significant distinctions between the current study and prior 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 the current study accounts for the endogeneity of these variables.

A Resource Allocation Model

Mostly because of the scarcity of data, the initial examinations of the allocation of campaign resources across the states were theoretical. Brams and Davis (1973, 1974) show that, in equilibrium, candidates devote a disproportionate share of their resources to states with many electoral votes, while Snyder (1989) introduces 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 present a theoretical model or estimate a nonstructural model, the current study directly estimates 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—namely, the level of responsiveness of each state to marketing activities. Finally, unlike previous studies that try to endogenize either advertisements or grassroots efforts, this study simultaneously explains both.

Resource allocation is a central decision in marketing (e.g., consider the allocation of advertising budget across media outlets). Recently, Manchanda, Rossi, and Chintagunta (2004) studied 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, econometricians can improve the precision of the estimates. Srinivasan, Raman, and Naik (2005) present a model of optimal resource allocation to corporate versus product branding.

Political Marketing

The scarcity of studies in marketing in the previously cited literature 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 together spent more than $1 billion, and the total spending on all the 2004 races was approximately $4 billion (http://www.opensecrets.org/pressreleases/2004/04spending.asp). These numbers are impressive because almost all the spending of political candidates is on marketing. In other words, in 2004, the spending on political marketing in the United States was almost $4 billion. To appreciate the volume of this industry, compare it with an industry that receives a lot of attention by marketing scholars—for example, 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 United States, 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 (1) the winner takes all and (2) a common deadline for the campaign, create a seminatural experimental setting. This setting (which is structured) makes it easier (than in the commercial arena) to examine some noteworthy research questions. Second, the data on political campaigns are rich. It is possible to obtain a lot of information on both voters (e.g., their knowledge, attitudes, perceptions, preferences, and choice) and candidates (e.g., their positions, declarations, advertisements, and spending). For example, this study has data on the marketing spending of all the firms in the market. Such data are rare for commercial campaigns.

THE MODEL

The model is designed to capture the equilibrium relationship among turnout, marketing activities, and the pivotal probability. To focus on turnout, it is assumed that there are no swing voters (Shachar and Nalebuff 1999). Thus, the role of marketing activities is to increase turnout among supporters, not to change the voting tendencies of individuals. As discussed in greater detail subsequently, in the model, supporters of a *candidate* are not necessarily affiliated (or identify) with his *party*. Thus, the 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 (and is formally demonstrated in the subsection titled "Individuals' Decisions") that even with this assumption, the model still implies that an increase in the marketing spending of a candidate (all else being equal) leads to an increase in the number of votes the candidate receives. This issue is revisited subsequently, and it is shown that, given the data, this assumption is not restrictive.

The Setup

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

The sequence of events is as follows: 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 the data consist of three election periods (1996–

[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 argued that one of the main reasons Kennedy preformed better was that his advisors prepared him better for the debates. In other words, there is a perception that the better consultants were the reason for Kennedy's victory. The importance of marketing and consultants has only grown since 1960. For example, consider the center role of Karl Rove in the Bush administration.

2004), to simplify the notations, the theoretical model focuses on one election and thus ignores a time subscript (which is added in the empirical section). The analysis begins with a characterization of the second-stage decisions—the choices of the individuals (i.e., potential voters).

Individuals' Decisions

Individuals make two (related) decisions: (1) who 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. Although the first group is sometimes referred to as Democrats and the second as Republicans, it does not mean that the people who fall into these two categories are necessarily registered voters. It only means that each eligible voter prefers one party over the other, and if this person participates in the elections, he or she votes for the candidate of his or 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 (e.g., a state liberalism index, the composition of the state legislature party division). Such variables are included in the empirical analysis.

The second decision—whether to participate in the election—is endogenous both in the theoretical model and in the empirical one. This decision depends on various factors, such as the individual's education, income, and race. More important, the parties can affect the participation rate of their supporters with their marketing activities (i.e., advertising and grassroots campaigns). 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 as follows:

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

where $a_{j,s}$ is the number of advertisements 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, $\beta_{0,j}$ 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 a function not only of the marketing variables but also of other observable factors captured by x. Although ε varies 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. Although it is expected that, on average, an increase in the marketing variables will 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, although δ_s^a and δ_s^c are not necessarily nonnegative, the marginal effect of marketing, across all states, is expected to be positive. The role of the heterogeneity in the δs is discussed subsequently.

Before the next argument is presented, it is useful to restate the definition of the following symbols: The share of voters 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 (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 it is assumed that the only role of marketing is to increase turnout among supporters, marketing does not change d_s. However, it does affect the observed share of votes for d on election day, denoted by dv_s. Specifically,

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

This means that an increase in $a_{d,s}$ and/or $c_{d,s}$ leads to an increase in dv_s. Because d_s and $\psi_{j,s}$ are unobservables, the assumption that the marketing variables do not have a direct effect on d_s is not restrictive. In other words, although the number of voters for d and r is observed, the number of supporters of each candidate and the turnout rates among supporters are not observed. Thus, the data cannot help distinguish 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 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 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 advertising and a cost function with respect to contacts.[3] Before the production function of advertisements and the cost function of contact are formulated, the objective function of the candidates is discussed.

Candidates' objective function. Previous studies have suggested two alternative candidate goals: Either the candidates want to maximize the expected number of electoral votes won (i.e., market share), or they want to maximize the probability that they will receive the majority of electoral votes (i.e., probability of winning). Although the second goal is 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 Bill Clinton was likely to be reelected. Thus, it is

[2]The share of Democrats among the voters is bounded between zero and one. However, instead of selecting a distribution that restricts d_s to be within the interval [0, 1], it is assumed 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. A similar assumption is made with respect to σ_ε^2, which affects the turnout rate. The estimates of these standard deviations (reported in the structural estimation section) are consistent with these assumptions.

[3]The budget depends on the closeness of the race on the national level. In other words, the greater the uncertainty about the winner of the election, the higher are the donations to the candidates. Note 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 a parameter and to estimate a specific budget for each election year. Indeed, this is the approach adopted here.

sensible to assume that the main objective of both Clinton and his Republican competitor, Bob Dole, was to maximize their market share.

Although 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 empirically examine this issue. To achieve this goal, the model is estimated under two scenarios—(1) assuming that the candidates' goal is only to win the election and (2) assuming that their goal is to maximize the expected number of electoral votes won. A comparison between the fit of both models will help address the question at hand.

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

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

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

where $e_{j,s}$ represents the spending of candidate j on state s. When formulating the dependence of $a_{j,s}$ on $e_{j,s}$, it is necessary 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 advertisements is formulated as follows:

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

where n_s is 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. The candidates face a national budget constraint when determining their advertising spending, but 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 $c_{j,s}$, it is necessary 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 as follows:

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

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 in the "Data" section.

Equilibrium

Next, the objective function of the candidates is presented, and the equilibrium levels of $e_{j,s}$ and $c_{j,s}$ are derived. The probability that candidate j wins state s is represented as $p_{j,s}$. Note that when candidates determine 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

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

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 follows:

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

where $F_d(\bullet|z_s)$ is the cumulative density function of d_s (and $p_{r,s}$ is equal to $F_d[\bullet|z_s]$).

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]

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

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

Through the use of Lagrangians, the following first-order conditions are obtained for the Democratic candidate for every s (to simplify the exposition, the condition $\lambda_d[\Sigma_s e_{j,s} - E_d] = 0$ is not presented, though it is easy to show that the condition is binding, and thus $\lambda_d > 0$ and $\Sigma_s e_{d,s} = E_d$):

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

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

For the Republican candidate (for every s),

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

(12) $$v_s f_d\left[\alpha(\bullet)|z_s\right]\alpha(\bullet)\left[1 - \alpha(\bullet)\right]\delta_s^c - \kappa c_{r,s}^{\tau_c - 1} n_s^{\gamma_c} \exp(w_s \rho) = 0.$$

[4]To simplify the exposition, the restrictions $e_{j,s} \geq 0$ and $1 \geq c_{j,s} \geq 0$ (for any j and s) are not presented in the optimization problem and, subsequently, in the first-order conditions associated with them. However, these restrictions are applied to the solution. For example, it is shown that the optimal $c_{r,s}$ is equal to zero for states with negative δ_s^c.

It is now possible to solve for the budget allocation and grassroots effort in equilibrium. The allocation and effort of both candidates are first shown to be identical, and then the optimal e and c (denoted by e* and c*, respectively) are solved as a function of the other variables of the model.[5]

P_1: In equilibrium, $e^*_{d,s}/E_d = e^*_{r,s}/E_r$ and $c^*_{d,s} = c^*_{r,s}$ for every s.

Proof: As derived from Equation 10 and Equation 12, $c^*_{d,s} = c^*_{r,s}$. As derived from Equation 9 and Equation 11,

$$\lambda_d^{\frac{1}{1-\tau_a}} e^*_{d,s} = \lambda_r^{\frac{1}{1-\tau_a}} e^*_{r,s} \Rightarrow \lambda_d^{\frac{1}{1-\tau_a}} \sum_s e^*_{d,s} = \lambda_r^{\frac{1}{1-\tau_a}} \sum_s e^*_{r,s}$$

$$\Rightarrow \lambda_d^{\frac{1}{1-\tau_a}} E_d = \lambda_r^{\frac{1}{1-\tau_a}} E_r.$$

Because

$$\lambda_d^{\frac{1}{1-\tau_a}} e^*_{d,s} = \lambda_r^{\frac{1}{1-\tau_a}} e^*_{r,s} \text{ and } \lambda_d^{\frac{1}{1-\tau_a}} E_d = \lambda_r^{\frac{1}{1-\tau_a}} E_r, \text{ then}$$

$$\Rightarrow \frac{e^*_{d,s}}{E_d} = \frac{e^*_{r,s}}{E_r}. \text{ Q.E.D.}$$

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*). This would be the case if $E_d = E_r$. In general, such an assumption (i.e., $E_d = E_r$) is reasonable. For example, in the 2004 elections, the total disbursements of George Bush and John Kerry were $359 million and $333 million, respectively (see http://www.fec.gov/disclosure.shtml). However, note that this assumption might be less reasonable for the 2008 election, but the 2008 elections are not included in the data set. Furthermore, in the "Data" section, it is shown that the result $e^*_{d,s} = e^*_{r,s}$ is also empirically reasonable for the data. Thus, to (dramatically) simplify the estimation, it is assumed that $E_d = E_r = E$, and it is found that, in equilibrium, $e^*_{d,s} = e^*_{r,s}$. As a result,

$$\alpha(e_{d,s}, e_{r,s}, c_{d,s}, c_{r,s}, n_s) = \frac{1}{1+\exp(\beta_{0,d} - \beta_{0,r})}.$$

Note also (from the proof of P_1) that in this case, $\lambda_d = \lambda_r$. It is now possible to solve for the optimal marketing levels.

P_2: The equilibrium common values of the budget allocation and the grassroots effort are as follows:

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

where

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

and

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

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

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

Then, it is plugged back into Equations 9 and 11. The solution of $c^*_{.,s}$ is immediate from Equations 10 and 12.

This result implies that the marketing variables depend on four elements. They increase with (1) the number of electoral votes, v_s, (2) the predicted closeness of the race, $f_d(\alpha^*|z_s)$, and (3) the responsiveness of voters to the marketing variables, δ^a_s and δ^c_s, but decrease with the size of the voting population, n_s. The marketing variables also depend on the τ parameters, w_s (only for $c^*_{.,s}$) and E (only for $e^*_{.,s}$), but these elements are of lesser interest.

The role of three of these four elements (electoral votes, predicted closeness, and the size of the voting population) has already been emphasized by previous studies. These variables are usually referred to as the "strategic variables." However, previous studies have ignored the potential heterogeneity in the effectiveness of the marketing variables. Here, it is theoretically shown that two states that are identical with respect to the three strategic variables might still have different marketing activity because of differences in their responsiveness level to the marketing variables. Although there is no theory behind the potential differences in the δs across states, it is 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 noteworthy because previous studies have ignored this element.

The effect of closeness on the turnout rates. Because 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.

[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 Web Appendix (at http://www.marketingpower.com/jmrdec09). The concavity of the objective function depends on the values of the model's parameters. Thus, it is tested after obtaining the estimated parameters in the "Structural Results" section.

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 taken into account? Various secondary questions have been presented 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, it would be worthwhile to understand the contribution of marketing to the turnout rate. This secondary research issue is elaborated on subsequently.

DATA

To test the model's implications and structurally estimate it, state-by-state data on the 1996, 2000, and 2004 presidential elections in the United States were collected. 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, advertising, and grassroots campaigns. They also contain exogenous variables, such as the number of electoral votes, the percentage of eligible voters who moved to the state in the year before the elections, and a state liberalism index. The following four subsections present the four endogenous variables and the exogenous variables associated with them and provide a statistical description of these variables (i.e., summary statistics). For a more elaborating description of the data, see 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% and 49.29%, respectively) are close to 50%, demonstrating that, in general, these elections were very close. Furthermore, in more than 71% of the races, $dv_{s,t}$ was between 40% and 60%. Table 1, Panel A, describes this variable and the relevant exogenous variables (e.g., a state liberalism index) formally and statistically.

Turnout Rates

The turnout rate, denoted by $y_{s,t}$, is defined as the share of votes for both the Republican and the Democratic candidates from the voting age population. Voting age population, rather than registered voters, is used mainly because not registering can be considered another form of failure to participate.

The mean and median turnout rates are 50.97% and 51.51%, 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 has had at most 52% of the votes (i.e., $.48 \leq dv_{s,t} \leq .52$) is 53.58%. This result is based on the *ex post* closeness. The results with *ex ante* closeness (i.e., $.48 \leq \hat{d}v_{s,t} \leq .52$, where $\hat{d}v_{s,t}$, the predicted $dv_{s,t}$, is described in the "Preliminary Results" section) are similar. Table 1, Panel B, describes this variable and the relevant exogenous variables (e.g., the presence of a contemporaneous governor's race) formally and statistically.

Advertising

The raw data on advertising were created by Campaign Media Analysis Group for the 1996, 2000, and 2004 elections and were made available by the University of Wisconsin Advertising Project (see Goldstein, Franz, and Ridout

Table 1

SUMMARY STATISTICS

A: Election Results

Variable	M	SD	Mdn	Minimum	Maximum
Democratic vote share	48.73	9.09	49.28	26.60	69.00
Democratic share in the national Gallup Poll in early September	49.67	4.12	49	45.00	55.00
State liberalism index (ADA and ACU)[a]	0	1.00	–.14	–1.68	1.90
Prior state vote deviation from previous national vote	–2.34	7.19	–1.63	–22.04	12.61
Prior state vote deviation from twice previous (eight years prior) national vote	–1.72	6.13	–1.64	–18.36	12.24
Vice presidential candidate's home state	0	.20	0	–1.00	1.00
Presidential candidate's home state	0	.20	0	–1.00	1.00
Standardized first-quarter state economic growth × incumbent party	.003	1.49	0	–4.89	5.86

B: Participation Rate

Variable	M	SD	Mdn	Minimum	Maximum
Percent of eligible voters who participated in the elections	51.51	7.83	50.97	33.27	73.25
Proportion of African Americans	.10	.10	.07	.00	.37
Personal income (per capita) in 2000: $10,000	2.76	.43	2.75	1.95	4.15
Percentage of people with at least high school degree	84.94	4.44	85.65	73.80	92.30
Percentage of people who moved to the state in the year before the elections	.12	.76	.06	–2.10	3.53
Concurrent governor's race	.23	.42	0	.00	1.00

[a]The Americans for Democratic Action (ADA) tracks how members of Congress vote on key issues and gives each member a rating from 0, meaning complete disagreement with ADA policies, to 100, meaning complete agreement with ADA policies. A score of 0 is considered conservative, and a score 100 is considered liberal. The American Conservative Union (ACU) builds a similar score that measures the degree of conservatism. The state liberalism index is based on the average score (by both ADA and ACU) of all members of Congress in that state. It is then standardized.

2002; Goldstein and Rivlin 2006). These data provide a comprehensive record of every advertisement broadcast on the national broadcast and cable television networks in each of the country's top media markets.

Because the data are broken down by candidates, it was possible 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 that $a_{d,s,t} = a_{r,s,t}$). What does "reasonability" mean here? 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 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. 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 subsequently 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 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 the theory presented herein predicts, 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, it is expected that $b_0 = 0$ and $b_1 = 1$. The estimates are 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, in most cases, they even satisfy the strict statistical testing. Specifically, a simple Wald test for each parameter separately cannot reject the null hypothesis, at the 5% level, in five 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 5% significance level) is 3.19. This means that the null hypothesis for 2004 can be rejected and the F statistics is close to the critical F for 1996; thus, this evidence is not the strongest vindication of the assumption. However, because a definite test is not required here, but rather just an examination of the reasonability of an assumption, the evidence is supportive. Furthermore, taken together, the collection of evidence makes a compelling argument that $a_{d,s,t} = a_{r,s,t}$ is indeed a reasonable approximation.

Table 2

EXAMINATION OF THE REASONABILITY OF THE EQUALITY ASSUMPTION

	$a_{d,s,t} = b_0 + b_1 a_{r,s,t} + \varepsilon^b_{s,t}$		
	1996	*2000*	*2004*
b_0	2.878	–1.326	.435
	(1.662)	(1.445)	(1.668)
b_1	1.007	1.008	1.288
	(.093)	(.059)	(.055)

Notes: Standard errors are in parentheses.

Given that $a_{d,s,t} = a_{r,s,t}$, the variable used in the empirical work is the equilibrium common value presented in P_2. Specifically, $a_{s,t}$ is the number of advertising minutes per day aired by the two candidates in each state between September 1 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, note that they represent ad minutes aired by the candidates and not exposure to advertisements. 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 advertising can be illustrated by the higher-than-average intensity of advertisements in close elections. The average ad minutes is 78.53 in races in which the winner had, at most, 52% of the votes, compared with 21.18 in the other races. Furthermore, the simple relationship between advertising and turnout is demonstrated by their correlation, which is .23.

(Finally, note that this variable might be subject to a measurement error [ME] problem, which is suggested subsequently.)

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 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 contacted in each state serves as the measure of grassroots campaigning.

Although this measure was used in most of the previous studies, note that it is a noisy measure of the contact rate because (1) it is based on a sample and (2) the number of observations in the sample is small for several states. For example, in 50 cases (of the 150 in the data), there are fewer than five observations per state. Not surprisingly, many of these cases are for smaller states. For example, the average

[6]First, the unit of analysis in the data set is the broadcast of a single advertisement, with information on where and when it aired. The data have been aggregated along two dimensions—time and place. Specifically, because this study is on the state level, the market-level data have been aggregated to the state level using the population size in each market. Furthermore, because the model treats each election year (for each state) as one observation, the number of ad minutes between September 1 and elections day has been averaged. This period is considered the most intense and critical in the campaign. Second, another way to think about the approximation $a_{d,s,t} = a_{r,s,t}$ is that the variation due to the differences between the candidates is being ignored, and only the variation across states is considered. 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 for each state, the closeness is the same for both parties and the data about participation is not party specific, it seems that allowing for variations in spending across candidates cannot help address 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 $a_{s,t}/2$ (i.e., across states and time) is 21.95. In other words, ignoring the differences between the two candidates excludes a small fraction of the variation in advertising.

voting age population for these 50 cases is 1.35 million, compared with an average of 5.55 million for the other 100 cases.

This variable is a noisy measure of the contact rate, and thus it is necessary to depart from its treatment in previous studies in two ways. First, this study includes in the data only the 100 cases in which there are at least five observations per state, and it deals directly (and structurally) with the missing data issue. Second, the ME in this variable is taken into account.

The average contact rate in the sample (i.e., for the 100 observations) is .39. The median is .36, and the standard deviation 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% 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 Equation 5) are the proportion of people living in metropolitan areas, Metro, and a measure of the number of undergraduate students in the state, Enrolled. In line with Green and Gerber (2004), it is assumed that the cost of the grassroots campaign increases with Metro and decreases with Enrolled.[7]

PRELIMINARY RESULTS

The research questions are addressed through structural estimation in a subsequent 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 elaborate description of the analysis and for some additional tests, see 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% of the variation of the dependent variable. On the basis of this regression, $\hat{d}v_{s,t}$ (i.e., the predicted Democratic vote share) and two rough proxies of the predicted closeness of the race—that is, $-|\hat{d}v_{s,t} - .50|$ and $I\{.48 \le \hat{d}v_{s,t} \le .52\}$, where $I\{\cdot\}$ is the indicator function—are calculated. The first proxy is (minus) the margin of victory, and the second is a binary variable identifying close races (i.e., the winner receives at most 52% 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 uses 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 10% level for the first proxy and at the 2% level for the second proxy). The results demonstrate that the effect identified in previous studies of the closeness of the race on participation rate also exists in the last three U.S. presidential elections. Therefore, the question raised previously (Will the direct effect of closeness on the participation rate vanish when the marketing activities are taken into account?) is relevant for this data set.

Table 3
THE DEPENDENT VARIABLE: DEMOCRATIC VOTE SHARE

Variable	*Coefficient*	*SE*	*t-Statistic*
Democratic share in the national Gallup Poll in early September	.600	.135	4.43
State liberalism index (ADA and ACU)	1.601	.410	3.90
Prior state vote deviation from previous national vote	.793	.086	9.20
Prior state vote deviation from twice previous (eight years prior) national vote	.183	.099	1.83
Vice presidential candidate's home state	.430	1.565	.27
Presidential candidate's home state	.856	1.569	.55
Standardized first-quarter state economic growth × incumbent party	–.078	.373	–.21
Constant	21.084	6.749	3.12
R^2		83.82	

Notes: ADA = The Americans for Democratic Action, and ACU = The American Conservative Union.

Table 4
THE DEPENDENT VARIABLE: TURNOUT RATE

Variable	*Coefficient (SE)*	*Coefficient (SE)*
Predicted closeness is $-\lvert\hat{d}v_{s,t} - .50\rvert$	.18 (.11)	
Predicted closeness is $I\{.48 \le \hat{d}v_{s,t} \le .52\}$		3.15 (1.27)
The ratio between the number of electoral votes and the voting population	.28 (.43)	.31 (.42)
Proportion of African Americans	7.15 (6.71)	8.48 (6.65)
Concurrent governor's race	1.04 (1.28)	1.37 (1.27)
Percentage of residents who moved to the state a year before the election	–1.27 (.71)	–1.25 (.69)
Income per capita	.40 (1.44)	.88 (1.43)
Percentage with at least four years of high school education	1.12 (.16)	1.08 (.15)
Constant	–45.06 (12.14)	–45.05 (11.98)
R^2	38.95	40.28

[7]First, it is more difficult to contact people in metro areas. Conversely, undergraduate students are effective and cost efficient in contacting potential voters. (Don Green is acknowledged for suggesting these variables.) Second, the enrolled variable is binary and equal to 1 for the top 25% of the observations in terms of the number of undergraduate students in the state and to 0 for the other.

The Effect of the Strategic Variables on Marketing

Shachar (2007) shows that as the model hypothesizes, 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 approximately 35 additional minutes of advertising compared with a state in which the expected result is 62:38. A similar finding (about the role of the strategic variables in determining the marketing decisions) for the grassroots effort is obtained.

The Effect of Marketing on Turnout

The foregoing results provide 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). Next, the main hypothesis is examined—namely, that the direct effect of the strategic variables on turnout vanishes when the marketing variables are taken into account.

Table 5 addresses these questions using the binary variable proxy for the closeness of the race (i.e., $I\{.48 \leq \hat{d}v_{s,t} \leq .52\}$). 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 2) 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, advertising significantly increases turnout. Specifically, airing ten additional minutes of advertising per day increases turnout by more than .5%. Second, now, the effect of predicted closeness is not different from zero, even at the 10% significance level (its t-statistics decreased from 2.47 to 1.09). In other words, after the advertising spending is accounted for, the direct effect of the predicted closeness vanishes. Although 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. On the contrary, it becomes stronger and significant at the 10% level. This phenomenon reappears in the next columns of the table and is 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 3 is the baseline regression, Model 4 is the baseline regression with the advertising variable, and Model 5 is the baseline regression with both marketing variables.

Including the number of ad minutes as an explanatory variable reduces the effect of closeness (from 4.12 to 1.11). Furthermore, now, the effect of predicted closeness is not different from zero even at the 10% significance level. When the contact rate variable is also included, the effect of closeness drops even further (to 1.06).

As expected, the two marketing variables have a strong positive and significant effect. Again, an increase in the number of ad minutes by ten leads to an increase of more than .5% of eligible voters who participate in the elections. The effect of grassroots campaigning is even more dramatic. Increasing the proportion of the population contacted by the parties by 10% leads to an increase of approximately 2 percentage points in participation.

Finally, the missing data problem for the contact variable was treated here in a naive way. As discussed in the next section, the structural estimation uses all 150 observations and accounts for the missing values in a structural way.

Table 5

THE ROLE OF THE MARKETING VARIABLES (DEPENDENT VARIABLE: TURNOUT RATE)

Variable	*Model 1*	*Model 2*	*Model 3*	*Model 4*	*Model 5*
Predicted closeness is $I\{.48 \leq \hat{d}v_{s,t} \leq .52\}$	3.15**	1.41	4.12**	1.11	1.06
	(1.27)	(1.30)	(1.44)	(1.47)	(1.22)
The number of electoral votes divided by the voting population	.31	.75*	.41	1.15*	1.33**
	(.42)	(.42)	(.68)	(.64)	(.53)
Number of advertisements		.05**		.07**	.05**
		(.01)		(.02)	(.01)
The share of people who were contacted					20.07**
					(3.07)
Proportion of African Americans	8.48	14.15**	12.53	19.08**	22.17**
	(6.65)	(6.51)	(8.44)	(7.82)	(6.49)
Concurrent governor's race	1.37	1.91	.67	1.91	.62
	(1.27)	(1.22)	(1.51)	(1.40)	(1.18)
Percentage of residents who moved to the state a year before the election	–1.25*	–1.48**	–1.42	–.92	–.39
	(.69)	(.66)	(.95)	(.87)	(.72)
Income per capita	.88	1.01	1.22	2.49	2.31*
	(1.43)	(1.37)	(1.64)	(1.52)	(1.26)
Percentage with at least four years of high school education	1.08**	1.06**	1.09**	1.00**	.76**
	(.15)	(.15)	(.19)	(.17)	(.15)
Constant	–45.05	–47.60	–48.11	–49.58	–36.07
	(11.98)	(11.45)	(14.55)	(13.24)	(11.16)
R^2	38.95	45.99	45.41	55.27	69.67
Number of observations	150			118	

*$p < .10$.
**$p < .05$.
Notes: Standard errors are in parentheses.

ESTIMATION ISSUES

The Likelihood Function

Before the likelihood is presented, it is suggested that there might be ME hereafter in the marketing variables, and thus they are formulated. In the likelihood construction, both the issue of the MEs and the problem of missing observations for the contact variable are taken into account. A structural approach to the solution of these two problems is offered.

ME. 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 data from the Campaign Media Analysis Group cover almost all television markets in the United States, they do not cover all of them. Second, ME can be introduced in the coding of the advertisements. The source of ME in the contact data is obvious—the 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 five cases with fewer than ten 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 introduces the issue of endogeneity with respect to them.

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

$$\text{(15)} \qquad a^o_{s,t} = a^*_{s,t} + u_{s,t} \text{ if } a^*_{s,t} + u_{s,t} > 0, \text{ and}$$

$$a^o_{s,t} = 0 \text{ if otherwise;}$$

$$\text{(16)} \qquad c^o_{s,t} = c^*_{s,t} + \omega_{s,t} \text{ if } 0 < c^*_{s,t} + \omega_{s,t} < 1,$$

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

$$c^o_{s,t} = 0 \text{ if otherwise,}$$

where (1) $a^*_{s,t} = (e^*_{s,t})^{\tau_a} n^{\gamma_a}_{s,t}$, (2) $e^*_{s,t}$ and $c^*_{s,t}$ are the equilibrium levels from Equations 13 and 14, (3) $a^o_{s,t}$ and $c^o_{s,t}$ are the ad minutes and contact rates observed (i.e., not necessarily the actual ones), and (4) $u_{s,t}$ 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. 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; 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 the model, and thus the contact equation can be estimated only for a subset of the sample. Indeed, this is accounted for in Equation 18.

The likelihood function. It is assumed 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. The estimation of the δs is discussed in greater detail subsequently.

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, the inferred values of the random variables ($d_{s,t}$, $\varepsilon_{s,t}$, $u_{s,t}$, and $\omega_{s,t}$) are derived. These inferred values (denoted by the asterisk) are as follows:[8]

$$\text{(17)} \qquad d^*_{s,t} = \left[1 + \frac{1 - dv_{s,t}}{dv_{s,t}} \exp(\beta_{0,d} - \beta_{0,r})\right]^{-1},$$

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

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

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

where z_t, v_t, n_t, and δ^a are the S-element vectors whose sth component are $z_{s,t}$, $v_{s,t}$, $n_{s,t}$, and δ^a_s, respectively. Indeed, Equations 13 and 14 show 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 settings faced by the candidates: When deciding on the grassroots effort, the candidates face a local cost function, but when choosing ad spending, they face a national budget constraint. The interstate dependence affects the ability to estimate δ_s using standard methods, as discussed subsequently.

The relevant density functions are as follows:[9]

---

[8]For example, $dv_{s,t}$ 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, the true underlying percentage of Democrats, $d^*_{s,t}$, is inferred:

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

Inserting, $a^*_{d,s,t} = a^*_{r,s,t}$ and $c^*_{d,s,t} = c^*_{r,s,t}$ into the preceding equation and using the functional form of ψ, it follows that

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

[9]First, these equations account for the truncation of advertisements (see Equation 15) but only for one type of contact's truncation. The reason for this is simple—there are no observations in the data with $c^o_{s,t} = 0$. Thus, although the likelihood should theoretically account for the truncation of contact, it is ignored here to simplify the presentation. Second, $f_4 = 1$ for the 50 observations for which $c^o_{s,t}$ is missing. Third, in practice, Shachar and Nalebuff's (1999) approach is followed, replacing $f_2(y_{s,t}|\bullet)$ with $f_2(dp_{s,t}|\bullet)$, where $dp_{s,t} = d_{s,t}\psi_d$. Fourth, $x_{s,t}$ includes a fixed effect for each election year.

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

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

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

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

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 as follows:

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

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 this 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 two segments, the number of combinations needed to integrate over is 1.1259E + 15. One way to deal with such a challenge is through simulation. Another way, presented and adopted here, is through clustering. Next, the clustering approach is described, and its effectiveness is demonstrated with a Monte Carlo experiment.

As in standard clustering, the basic idea is to allocate each "individual" to one of the clusters/segments. As pointed out previously, it is almost impossible to search over all possible combinations. Thus, an algorithm is suggested that, though not necessarily identifying the clustering that leads to the highest likelihood, is most likely to do so. The rationale behind the algorithm is 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 in which 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 includes only one state, and the other includes all other 49 states. The model is estimated 50 times because 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 leads 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, 1 of the 49 states grouped as one segment is moved to the other cluster. In other words, in each of the estimations, there are two segments: one with 2 states and the other with 48 states. Again, the combination that leads 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. 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 with the usual information criteria.

This approach follows the tradition of the divisive methods (sometimes referred to as "top-down") of the hierarchical clustering techniques (Everitt 1993). In this case, the likelihood function is used to "measure" the similarity of states.

To examine the effectiveness of the clustering approach, a Monte Carlo experiment was conducted. One hundred models with two clusters were simulated and then estimated. On average, across all 100 runs, the percentage of states identified correctly was 92. In other words, a state that was in one cluster was allocated to another cluster in only 8% of the cases. Furthermore, on average, the estimates of the δs were 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, and the true values were .3 and .1.

Testing the Main Hypothesis

The inferred values in Equation 17 represent the theoretical model. Because the 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, to test the main hypothesis of this study, it is necessary to allow for such effect and, thus, to rewrite the inferred value of ε accordingly (note that, as before, the ratio of electoral votes to the size of the voting population is used instead of each one of them separately):

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

In other words, the model suggests that $\beta_{closeness}$ and $\beta_{(v/n)}$ are equal to zero. Indeed, $\beta_{closeness}$ and $\beta_{(v/n)}$ 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 is also used, has two stages. In the first stage, the model is estimated under the assumption that $\beta_{closeness} = \beta_{(v/n)} = 0$ (i.e., using exactly the inferred values in Equation 17), and then the "residual" of the turnout equation is calculated: $[\varepsilon^*_{s,t}(\hat{\delta}^a,\ \hat{\delta}^c_s) - x_{s,t}\hat{\beta}]$. In the second stage, the "residual" is regressed against the strategic variables: $f_d(\alpha^* | z_{s,t})$ and $v_{s,t}/n_{s,t}$. The theory suggests that (1) 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 depends on the strategic variables, but (2) when the marketing variables are included in the first stage, the strategic variables are 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, the endogeneity problem is not an issue here. (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 simple—none of these endogenous variables also serve as explanatory variables in the empirical model. Specifically, while previous studies have included the observed marketing variables and/or the *ex post* closeness in the turnout equation, here, the equilibrium values of the marketing variables and the *ex ante* structural measure of closeness in the turnout part of the model are used.

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 begins 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 analyzes the contribution of the marketing activity to the turnout rate. Finally, two secondary research questions are considered: (1) Do candidates maximize market share or winning probability? and (2) More important, is there heterogeneity in the δs, and if so, 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: (1) the marketing variables in the turnout equation (included versus excluded), (2) the direct effect of the

Table 6

STRUCTURAL ESTIMATION RESULTS: THE PARAMETERS OF INTEREST OF SIX MODELS

		1	*2*	*3*	*4*	*5*	*6*
Marketing variables in the turnout equation		Excluded	Included	Included	Included	Included	Included
Strategic variables in the turnout equation		Included	Included	Included	Included	Included	Excluded
The w variables in the contact rate equation		Excluded	Excluded	Included	Included	Included	Included
Time effects		Excluded	Excluded	Excluded	Included	Included	Included
τ_a		Set	Set	Set	Set	Estimated	Estimated
$\beta_{Closeness}$ [.01]		.9998***	–1.0657**	–1.421***	.3035	.3680	0
		(.2346)	(.544)	(.4993)	(.4167)	(.3987)	
$\beta_{v/n}$ [.1]		.088	.1262*	.0728	.1291**	.1264*	0
		(.080)	(.0727)	(.0734)	(.0651)	(.0653)	
δ^a [.01]		1	.322***	.2904***	.0617	.0545	.0715***
			(.0716)	(.0619)	(.0433)	(.0387)	(.0247)
δ^c [.1]		.33***	.2972***	2.5622***	1.62***	1.6494***	2.1444***
		(.094)	(.0837)	(.7159)	(.6265)	(.6387)	(.5487)
Log-likelihood		59.38	80.47	88.80	125.06	126.68	124.73
	Democrats	83.98	84.26	84.12	84.05	83.97	84.13
R^2	Turnout	47.67	61.26	63.04	70.85	71.22	70.55
	Ad minutes	40.62	39.78	41.14	41.29	44.54	44.55
	Contact	16.60	15.59	22.48	39.12	39.72	39.03

*$p < .10$.
**$p < .05$.
***$p < .01$.

Notes: In Model 1, the δs cannot be identified with the turnout equation, because the marketing variables are excluded from this equation. The parameter δ^c is identified because it affects the choice of grassroots effort. The parameter δ^a influences the choice of ad minutes. However, when there is no heterogeneity in δ^a, this effect does not exist (check Equation 13). Thus, δ^a is normalized at 1 in Model 1. Some of the parameters were scaled in the estimation. The scales appear in brackets (e.g., the scale of δ^a is .01). The unscaled value of each parameter is the multiplication of the parameter and the scale (e.g., the unscaled value of δ^a in Model 2 is .0032). Standard errors are in parentheses.

Table 7
STRUCTURAL ESTIMATION RESULTS

	Estimate	*SE*	*t-Statistic*
θ: State liberalism index (ADA and ACU) [.1]	.1356	.0258	5.264
θ: Vice presidential candidate's home state [.01]	–.2151	1.2404	–.173
θ: Presidential candidate's home state [.01]	–.2381	1.1142	–.214
θ: Prior state vote deviation from previous national vote [.01]	.9753	.063	15.471
θ: Prior state vote deviation from the twice previous (eight years prior) national vote [.001]	.2876	.677	.425
θ: Standardized first-quarter state economic growth × incumbent party [.001]	.9456	2.1751	.435
σ_d [.1]	.3549	.0196	18.125
$\beta_{0,r} - \beta_{0,d}$	.1884	.1059	1.778
β: Proportion of African Americans	.0713	.0892	.799
β: Concurrent governor's race [.1]	.2701	.1686	1.602
β: Percentage of residents who moved to the state a year before the election [.1]	–.3365	.0931	–3.613
β: Income per capita [.1]	.1756	.2583	.68
β: Percentage with at least four years of high school education [10]	.1445	.0229	6.313
σ_e	.0824	.0048	17.169
γ_a	–.4286	.0585	–7.327
E_{1996} [100]	.3235	1.7549	.184
E_{2000} [100]	10.1998	2.5775	3.957
E_{2004} [100]	11.2041	2.8589	3.919
σ_u [100]	.3908	.0275	14.217
τ_a	.6192	.0606	10.215
δ^a [.01]	.0715	.0247	2.889
γ_c	.748	.2325	3.217
σ_w	.1302	.0095	13.672
δ^c [.1]	2.1444	.5487	3.908
ρ: Metro [10]	.3915	.0897	4.363
ρ: Enrolled	–.6997	.3022	–2.316

Notes: ADA = The Americans for Democratic Action, and ACU = The American Conservative Union.

strategic variables in the turnout equation (included versus excluded), (3) the w variables in the contact rate equation (included versus excluded), (4) time effects (i.e., election years dummies) in all parts of the model (included versus excluded), and (5) the parameter τ^a (free or set at .5).

The first support for the main hypothesis comes from the comparison of Models 1 and 2. The only difference between these two models is the inclusion of the marketing variables in Model 2. While the predicted closeness of the race has a significant, positive effect in Model 1, it has a significant, negative effect in Model 2. (Hereinafter, the use of the term "significant" implies that the relevant estimated parameter is different from zero at least at the 5% significance level.) In other words, as the main hypothesis suggests, the positive effect of closeness is due to the exclusion of the marketing variables. While a couple of the results in Model 2 are surprising (these are discussed subsequently), they should not diminish the importance and significance implied by the comparison between Models 1 and 2. Furthermore, the subsection "Testing the Main Hypothesis" 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, the strategic variables explain the "residual." These two stages are executed twice—once when the marketing variables are excluded in the first stage and once when they are included. In the first stage, the effect of the *ex ante* closeness is .5807 (SE = .1925), and in the second stage, the effect is –.2472 (SE = .1565). These results also support the main hypothesis.

Model 1 reinforces the findings of previous studies that closeness has a positive, significant effect on turnout. Thus, the data from 1996 to 2004 are similar in this sense to those used in previous studies, and the results reinforce the political participation puzzle. Model 2 demonstrates that the puzzling effect presented in Model 1 is due to the exclusion of the marketing variables from the turnout equation. Furthermore, the effect of both marketing variables is positive and significant, and including these variables in the turnout equation increases the R-square of turnout from 48% to 61%.

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 2 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 2. First, the effect of predicted closeness on turnout is negative and significant. As is shown subsequently, this unexpected result vanishes in the richer models. Second, the direct effect of $v_{s,t}/n_{s,t}$ (electoral votes per eligible voter) becomes significant (at the 10% level) in Model 2. Furthermore, this unexpected result persists in the richer models. Although this is indeed an unexpected 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 Schwartz 1987). In other words, what puzzled previous scholars was the effect of predicted closeness. Second, it is 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 reexamined in Model 5. Models 3 and 4 are included to demonstrate the role of the w variables, the time effect, and τ^a. In Model 3, 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-square of the contact equation from 15.6% to 22.5% and, as a result, the R-square of turnout from 61.3% to 63%.

In Model 4, 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 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 5, the parameter τ^a, which thus far has been set at .5, is freed. Its estimate is .62 (SE = .06). This leads to an improvement in the R-square of the ad equation. Model 5, the richest model in Table 6, offers a final proof of the main hypothesis. Again, the direct effect of closeness is insignificant when the marketing variables are included. Furthermore, when Model 5 is reestimated without the marketing variables, the coefficient of the direct effect is 1.1679 (SE = .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. Now, after the main hypothesis has been addressed, the estimates of the structural estimation can be examined in detail.

Model 6 is the only one that is consistent with the theoretical model because 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 noteworthy parameters in Table 7 are δ^a and δ^c. Both are positive, and both are different from zero even at the 1% level. As discussed previously, this finding is especially impressive because unlike previous studies, there is no reason to expect that these δs are inconsistent.

This leads to an assessment of the contribution of marketing to the functioning of democracies. In recent decades, there has been a growing concern in the United States and in other democracies about the low turnout rate. Furthermore, during the past 50 years, the turnout rate in many countries has diminished. The findings in this study 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. 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 million and 7 million. It turns out that advertising is responsible for a small fraction of this dramatic effect. For example, if the candidates had not aired any advertisements in the 2004 elections, the number of voters would have decreased by 3.5 million. Conversely, canceling the grassroots effort would have led to a decrease of 12 million people. Notably, 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 advertisements 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 enrolled in college. The cost function of producing advertisements 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). Although τ_c is identified, it is not estimable. In other words, in the data, the correlation between this parameter and others is too high, rendering it impossible to estimate separately. Furthermore, using both a likelihood ratio test and a Wald test, it was not possible to reject the hypothesis that κ = 1. Because in the 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 unusual. It is smaller than the budget for the next two election years by a factor of 30. This is an unreasonable estimate. Still, 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, note that the optimal ad spending (Equation 13) is a multiplication of an element that includes 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: (1) There is a low correlation between the strategic variables and ad minutes, and (2) the estimate of E is close to zero.

Finally, the other parameters in Table 7 (i.e., the θ, β, and σ parameters) 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 high (R^2 = 84% for the election results and 70% for the turnout rate). Notably, the a priori tendency to participate in elections is higher among Republicans than among Democrats (i.e., $\beta_{0,r} - \beta_{0,d}$ = .1884; SE = .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. Although the straightforward objective is winning the election, it can be argued 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,

[10]As discussed in n. 5 and in the Web Appendix (at http://www.marketingpower.com/jmrdec09), it is theoretically impossible to guarantee that the candidates’ objective function is concave. Thus, concavity depends on the values of the model’s parameters. This issue is examined with the estimates reported in Table 7, and the objective function is concave. This means that the use of the first-order conditions to solve for the equilibrium levels of advertisements and contact is justified. For more details, see the Web Appendix (at http://www.marketingpower.com/jmrdec09).

it is possible to 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.

Thus far, the model estimated has assumed that the candidates maximize their market share. However, if 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, simulation is used to assess it as part of the structural estimation. The resultant $R_{s,t}$ is 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-square for the marketing equations (43.25 versus 44.55 for advertisements 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. Although this is a noteworthy finding, it should be considered with caution because it can be interpreted in other ways. For example, it is possible that the candidates want to maximize their winning probability, but because calculating $R_{s,t}$ is so complex (recall that it was only possible to do so 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 has been ignored previously. Previous studies (theoretical and empirical) relied on the strategic variables 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 can be used to 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—by improving 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 increases the fit of both the turnout rate equation and the marketing equations will the conclusion be reached that there is heterogeneity in the effectiveness of the marketing variables and that this heterogeneity plays a role in the marketing decisions.

The clustering approach described in the "Estimation Issues" section (as well as the Bayesian information criterion and consistent Akaike information criterion) reveals that the optimal number of clusters is seven. The fit measures provide support for the hypothesis. First, the fit of the turnout equation improves tremendously (R-square 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-square of the budget allocation part of the model increases from 44.73 with one cluster to 62.30, and the R-square 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 determining their marketing activities.

Table 8 presents the seven clusters and their δs. The largest cluster has 15 states (30%), and the smallest segment

Table 8

THE SEVEN CLUSTERS

Cluster	*1*	*2*	*3*	*4*	*5*	*6*	*7*
δ^a [.01]	.0588	.1349	.0455	–.4123	.0733	.0601	–2.838
δ^c [.1]	1.7488	–6.7146	.1238	–.9999	–2.7265	2.9669	7.2014
Number of states	15	2	11	5	5	8	4
States	Alabama	Hawaii	Arkansas	California	Arizona	Alaska	Delaware
	Colorado	Nevada	Florida	Georgia	New Mexico	Idaho	Kansas
	Connecticut		Illinois	Indiana	Texas	Maine	Maryland
	Iowa		Mississippi	New York	Utah	Minnesota	New Hampshire
	Kentucky		New Jersey	South Carolina	West Virginia	Oregon	
	Louisiana		North Carolina			South Dakota	
	Massachusetts		Oklahoma			Wisconsin	
	Michigan		Pennsylvania			Wyoming	
	Missouri		Tennessee				
	Montana		Virginia				
	Nebraska		Washington				
	North Dakota						
	Ohio						
	Rhode Island						
	Vermont						

Notes: When δ is negative, the equilibrium level of the marketing variable is zero.

has only 2 states (4%). The differences in the δs across the clusters are large and significant. The δs of the largest cluster are similar to those reported in Table 7. The cluster for which advertising has the highest effectiveness (Segment 2) is also the one for which the effect of contact is the lowest. Notably, the cluster for which grassroots campaign has the highest effectiveness (Segment 7) is also the one for which the effect of advertising is the lowest. However, note that these two segments are small (with 6 states in both). Another notable finding is that there is a segment for which the effectiveness of both advertising and contact is negative. Note also that when a δ is negative, the equilibrium level of the marketing variable is zero.

The states that constitute each of the clusters do not have any obvious connecting factors. (For a possible technical clarification of this point, see the Web Appendix, "Notes on Text," at http://www.marketingpower.com/jmrdec09). Although some rationale can be offered for some of these clusters (e.g., California and New York being in the same segment), the interpretation of these segments is challenging. Furthermore, recall that no theoretical foundation for the heterogeneity of the δs has been offered. Thus, in a sense, this heterogeneity is at least partly a "black box," and 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 suggests that the clustering approach presented here works well. In other words, from a methodological aspect, this is reassuring. Second, the findings provide (at least) initial support for 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 regarding political participation (that the participation rate is higher in close elections) is due to the omission of variables—namely, marketing activities. It is shown theoretically and empirically that the effect of closeness on turnout is only through the marketing variables. Furthermore, the findings suggest that the effect of the marketing variables on turnout is dramatic. For example, this study uses counterfactual experiments to show that if the marketing activity had been canceled in the 2004 election, 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, which has been ignored by previous studies, has an important impact on the allocation of the advertising budget.

The model was estimated using three elections because detailed data on political advertising do not exist for elections before 1996. In the 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, it will be possible to 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 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), further research could include another important resource allocated in political campaign—namely, visits to each state of the presidential candidates, their running mates, and their families. In addition, 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).

The model presented here and the rich data set used can serve to address additional issues. Three examples of such issues are as follows: First, by including swing voters in the model (and enriching the data with information such as their share in each state), it is possible to 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 could be used to assess the role of marketing in determining the winner in the elections. Second (and related to the first), an important ingredient of political marketing and political advertising seems to be negative advertising. The suggested model could be adopted to examine the strategic use of negative advertising in political campaigns. Third, the current study treated each election year as one period. In practice, although 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 might provide insights into issues such as private information and signaling.

Finally, in various ways, this study demonstrates the major role of marketing in political campaigns (e.g., it has a dramatic impact on turnout). It is hoped that these results, coupled with the arguments raised at the outset of this article about the importance of political marketing (e.g., the volume of the industry), will 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.

——— and ——— (1974), "The 3/2's Rule in Presidential Campaigning," *American Political Science Review*, 68 (1), 113–34.

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

——— 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, 532–42.

Everitt, Brain (1993), *Cluster Analysis*, 3d 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 [data set], final release, Department of Political Science, University of Wisconsin–Madison, and the Brennan Center for Justice, New York University.

——— and Joel Rivlin (2006), "Political Advertising in 2004," combined file [data set], initial release, Department of Political Science, 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, DC: 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, Peter Rossi, and Pradeep Chintagunta (2004), "Response Modeling with Nonrandom Marketing-Mix Variables," *Journal of Marketing Research*, 41 (November), 467–78.

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

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 No. 27/2007, The Henry Crown Institute of Business Research in Israel, Faculty of Management, Tel Aviv University.

——— 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," working paper, Anderson School of Management, University of California at Riverside.