Campaign Allocations Under Probabilistic Voting

Deborah Fletcher
Miami University

Steven Slutsky
University of Florida

September 2008

Working Paper # - 2008-09
Abstract: We develop a non-Downsian probabilistic voting model where candidates compete by running campaign ads in different media markets. Ads are viewed by everyone within a media market and cannot be targeted to subgroups such as undecided voters or partisans of one candidate. Based on observable factors, candidates can estimate the distribution of voter preference intensities in a media market, and campaign ads then shift this distribution. Due to unobservable factors, individuals with any intensity vote with some probability for each candidate. We derive comparative static implications of changes in such factors as the mix and intensities of partisans in a market on the advertising decisions of a candidate. Using campaign advertising data from governor and senate races in the US in 2002, we find these results to be consistent with actual campaign allocation behavior. In addition, the empirical results also shed insight into whether union members have become swing voters and whether voting behavior in the South continues to differ significantly from the rest of the country.

JEL keywords: D72, campaign allocation, probabilistic voting
(I) Introduction

Advertising expenditures by political candidates have been large in recent decades and continue to grow. For example, $600 million were spent on broadcast television and radio advertising in the 2004 race for President, nearly triple the spending in 2000 (Associated Press (2004)). Similarly, the Christian Science Monitor (2006) estimates that it costs at least $1 million to win a House seat and several million dollars to win a Senate seat. Despite this trend, factors affecting advertising allocations across media markets have not been widely studied.

Snyder (1989) and Stromberg (2008) consider allocation decisions in an Electoral College setting, where parties or candidates decide how much of their resources to allocate to the separate contests in different states. There exists an important distinction between allocations in an Electoral College (Presidential) race and races in a single state or district containing multiple media markets, such as races for governor, U.S. Senate or the House of Representatives. In the former, the magnitude of victory in any state is irrelevant, whereas in the latter, candidates care about the vote margin in each media market, as only the vote totals in the entire contest is important. Our purpose in this paper is to develop an empirically testable model of campaign allocations for single-state or single-district contests where there are multiple media markets in the state or district. To do this, we build upon, but must modify, the existing analyses of two-party competition.

Two-party competition has typically been modeled in a Downsian context, where the candidates compete by choosing platform positions in some issue space. If the space is multidimensional and individuals vote in a deterministic manner for the candidate
whose platform they prefer, then pure strategy Nash equilibria rarely exist. An alternative approach is to assume probabilistic voting, where candidates are uncertain about how an individual will vote.\(^1\) Even in multidimensional contexts this can lead to pure strategy equilibria if the probability that an individual votes for a candidate is a concave function of that candidate’s strategy and is a convex function of the opponent’s strategy.\(^2\)

For many political campaigns, the standard Downsian approach does not seem appropriate for analyzing advertising allocation decisions. Candidates rarely announce platform changes in ads for fear of being damaged by charges of flip-flopping. Instead, they compete by trying to change voter attitudes about the salience of issues or about non-policy attributes such as character and competence.\(^3\)

In this paper, we develop a non-Downsian probabilistic voting model in which individuals have preference intensities for one candidate over the other, based upon a variety of factors including the candidates’ issue positions. The candidates can estimate these intensities for voters of particular types based on observable factors. These estimates differ from individuals’ true intensities due to the existence of unobservable traits of each individual. Individuals with any estimated intensity will vote for each candidate with some probability. The candidates engage in campaign activities to seek to shift the intensities and hence the expected vote they receive.\(^4\) One important aspect of advertising is that although candidates attempt to do so, advertising cannot be perfectly

\(^1\) See Mueller (2003) for a survey of the literature on probabilistic voting.
\(^2\) Coughlin (1992) proves that if candidate uncertainty about voters is based upon the presence of unobservable traits which follow a binomial logit distribution, then the concave-convex assumption will be satisfied. Kirchgassner (2000) argues that such concavity-convexity cannot hold if the strategy spaces are unbounded.
\(^3\) See Enelow and Hinich (1984) for one approach to introducing such factors in a Downssian framework.
\(^4\) Other non-Downsian models similar to the one here are Skaperdas and Grofman (1995) and Stromberg (2008). The relations between their approaches and the model here are discussed below.
targeted to specific, narrow groups of viewers. Instead, an ad in any media market may be seen by individuals with very different political preferences. It is important that models include this kind of imperfect targeting when testing their empirical relevance.

The specific formulation of probabilistic voting that we use is most closely related to the analysis of Lindbeck and Weibull (1987), in which candidates compete by proposing alternative redistributions of income.\(^5\) The basic structure by which candidates’ incomplete information on voter preferences generates probabilistic voting is similar, and the sufficient conditions we develop for the existence of pure strategy Nash equilibria are similar to theirs. However, while there is overlap in the analyses of redistributive policies and advertising allocations, there are some significant differences between them. One difference is that in the redistribution contest, candidates have the same budget constraints, where candidates often differ significantly in their advertising budgets. Another difference is that individual voters react identically on the margin to promised redistributions by the different candidates, whereas a marginal advertising dollar from each candidate can have a very different effect on voters.\(^6\)

These two differences have important consequences in the analyses. In the redistribution context, Lindbeck and Weibull (1987) show that candidates choose identical policies, while in the advertising context, the equilibrium advertising strategies

---

\(^5\) Coughlin (1986, 1991) considers candidates competing by proposing redistribution policies in a probabilistic voting framework based on a logit distribution. The analyses of Snyder (1989) and Stromberg (2008) also utilize probabilistic voting. Snyder (1989) assumes a particular form of the probabilistic voting function, where a candidate who spends nothing in a district has a zero probability of winning even if the opponent spends only one dollar in that district. Stromberg (2008) includes significant complicating factors such as a type of uncertainty about voter preferences that leads to random outcomes with candidates maximizing the probability of winning instead of expected vote. For tractability, he then must make some strong structural assumptions which are not needed in our framework.

\(^6\) Another difference is in the underlying structure of the contest: advertising is analogous to an all-pay auction, whereas in redistributive contests, only the winner pays. This might affect welfare analysis but does not seem significant for a positive analysis of allocation decisions.
of the candidates may not be identical. Thus, certain interesting comparative static questions arise in the advertising setting that do not arise in the redistributive context. Although Lindbeck and Weibull (1987) assume a form of imperfect targeting, in which individuals are combined into groups with members of a group receiving the same redistribution, they do not focus on how changing the composition of the groups would change the candidates’ equilibrium policies, except in some polar cases. Our main focus is comparative static analysis of the effects that changes in group compositions have on equilibrium group advertising allocations.

Political competition based on advertising decisions instead of policy proposals is not only realistic, but also has advantages in specifying an empirical test of the probabilistic voting model. Direct tests of probabilistic voting in the Downsian context have been relatively rare. In part, this is due to the fact that if changes in candidate platforms during the course of the campaign occur at all, they are difficult to observe or quantify except in very broad measures.

While there is a substantial empirical literature addressing campaign expenditures, it differs fundamentally from our approach. The existing literature looks at how total campaign expenditures affect election outcomes, and does not consider how these expenditures are allocated across media markets. The main issue that has arisen in this literature is the potential simultaneity of campaign contributions, and thus expenditures, with respect to vote share. Because we explore campaign allocations

---

7 This literature spans more than 35 years, going at least as far back as Dawson and Zinser (1971), who estimate the winner’s vote share in House and Senate elections as a function of expenditures, incumbency, and party affiliation of the candidates.

8 This problem was first recognized by Giertz and Sullivan (1977), and a variety of authors including Jacobson (1985), Grier (1989), Levitt (1996), Gerber (1998), and Erikson and Palfrey (1998, 2000) have attempted to address the issue in a number of different ways. These include making assumptions about the relation between the expected and actual vote, using simultaneous equation systems or instruments for
across media markets, we are able to escape this simultaneity problem. First, how the
level of the advertising a candidate undertakes in different media markets affects the
candidate’s overall budget is likely to be a second order effect. Second, vote outcomes
for the candidates are not included in our reduced form equations, so we can look at the
candidates’ activities directly without having to estimate how these activities affect votes
in the different markets.

We derive empirically testable comparative static implications of the model and
then use campaign advertising data from U.S. Senate and gubernatorial races in 2002 to
test the predictions of the model. We also use the model to shed insight into some
debates about voter behavior. In particular, we consider whether union members have
become swing voters or remain, at least in the candidates’ views, intensely partisan
Democrats, and whether voting behavior differs significantly in the South from the rest of
the country. In general, the evidence seems consistent with the probabilistic voting
model, especially in the non-Southern states. We find that candidates seem to treat union
members as partisan Democrats, and that in some crucial respects, the South continues to
differ from the rest of the country.

The model is presented in Section II, including discussions of the error
distributions, the campaign effectiveness function, and the assumption we make that
candidates can only imperfectly target voters with their advertisements. Section III looks
in more depth at the role of imperfect voter characterization by candidates and considers
when the equilibrium exists, is unique, and is in the interior of the strategy spaces under
the restrictions giving rise to probabilistic voting. The comparative static implications

campaign contributions and expected vote, and only considering repeated contests between the same pairs
of candidates.
are derived in Section IV, with some specific testable implications drawn from them given in Section V. The empirical tests are presented in Section VI, and conclusions are given in Section VII. All proofs are contained in an Appendix.

(II) The Model

Consider an election campaign between two candidates L and R, whose policy positions have been set prior to the start of the campaign. The candidates engage in campaign activities to alter voter preferences in their favor. Since no voter is likely to prefer one candidate’s position over the other’s on every issue, candidates may affect voter preferences by trying to change the saliency of different issues in the voter’s mind. Independent of policies, candidates may also try to affect preferences by focusing on non-issue factors such as character and competence. Each voter is characterized by a preference intensity whose post-campaign value is denoted as I. The initial pre-campaign value is I₀. Positive values of I indicate a preference toward R, negative values a preference toward L, and 0 indifference.

To further specify the model in the context of probabilistic voting, three questions must be considered: First, what information do candidates have about voter preferences, and what do they then believe about voter behavior? Second, how do campaign activities affect voter preferences? Finally, what restrictions exist on candidate efforts to target voter types?

a. Candidate information and beliefs

We assume voters have traits that are only imperfectly observed by candidates. These errors in observation make voters appear to the candidates to have a random element in their behavior, even though individuals vote with certainty for their preferred
candidate. We call this general approach one of imperfect voter characterization.\(^9\) In this model, candidates will typically characterize individuals by observable traits, and may partition individuals into a finite set of types based on these observables with each type having a representative value of \(I\). A typical voter of that type will differ from the representative value. Denote the individual’s actual intensity by \(I^a\), where \(I^a = I + \varepsilon\), for \(\varepsilon\), an unobserved random variable whose pdf is \(\varphi(\varepsilon)\). We make the following general assumptions about \(\varphi(\varepsilon)\) which are comparable to those of Lindbeck and Weibull (1987):

\[
\begin{align*}
\varphi(\varepsilon) &= \varphi(-\varepsilon) \quad (1a) \\
\varphi'(\varepsilon) &\geq 0 \text{ if } \varepsilon < 0 \text{ and } \varphi'(\varepsilon) \leq 0 \text{ if } \varepsilon > 0 \quad (1b) \\
\exists c, \varphi(\varepsilon) &> 0 \text{ if } -c < \varepsilon < c \text{ and } \varphi(\varepsilon) = 0 \text{ if } \varepsilon < -c \text{ or } \varepsilon > c \quad (1c)
\end{align*}
\]

That is, \(\varphi(\varepsilon)\) is symmetric, weakly single peaked, continuously differentiable and strictly positive on its support \((-c, c)\).

For any observed value of \(I\), the candidates believe that the probability an individual votes for candidate R is given by a function \(V(I)\), defined by:

\[
V(I) \equiv \Pr\{\varepsilon > -I\} = \int_{-I}^{\infty} \varphi(\varepsilon) d\varepsilon. \quad (2)
\]

Then \(V\) is differentiable with \(\partial V/\partial I = \varphi(-I)\) so that \(V\) is nondecreasing in \(I\) and strictly increasing if \(-c < -I < c\). Further properties of \(V\) depend on the shape of \(\varphi(\varepsilon)\), as discussed below in section III.

This model of voter behavior is the same as the one underlying Stromberg (2008): voters have intensities of preferences toward candidates that are swayed by the campaign

---

\(^9\) An alternative justification for probabilistic voting would be that individuals make mistakes in voting given their true preferences. The possibility of an error might decline with intensity since more intense voters might take more care when casting their ballots.
activities of the candidates. These intensities are affected by random factors. Stromberg uses a reduced form approach to describing the distribution of voter intensities. Here, instead of making assumptions about the reduced form, separate assumptions are made on the distribution of voter types and on the errors with which a type is observed. Stromberg makes the strong simplifying assumption that the overall distribution of voter intensities is normal with a constant variance and a mean that may shift over time. In our context, a reduced form distribution can be calculated by combining the distribution of types with \( \varphi(\varepsilon) \). However, this distribution will generally not be normal or unimodal even when nice properties are assumed about \( \varphi(\varepsilon) \). The two approaches complement each other. His approach makes tractable analysis of the complications of the Electoral College and the additional random elements in his model. Our approach allows for consideration of important questions such as how increased polarization of the electorate in a market affects candidate resource allocation. This in effect would add weight in the tails of the reduced form distribution and take it from the middle, creating a new distribution that may not be unimodal. This type of change cannot be considered if the reduced form distribution is restricted to being normal.

b. Effectiveness of candidate activities

The candidates engage in activities designed to alter the sign or magnitude of individual preference intensities. Candidate R seeks to increase intensity for an individual of pre-campaign intensity \( I_0 \) with expenditures \( x(I_0) \) while L seeks to lower \( I_0 \) with expenditures \( y(I_0) \). Let \( h(I_0, x(I_0), y(I_0)) \) denote the effectiveness of spending by the candidates in changing the preference intensity of an individual with pre-campaign
intensity $I^0$, where $h$ may be positive or negative depending on the magnitudes of $x(I^0)$ and $y(I^0)$. Hence, $I = I^0 + h(I^0, x(I^0), y(I^0))$.

Certain properties of $h$ are significant for the results. One issue is what types of individuals are most susceptible to being swayed by campaign advertising: those with low intensities, that is, ones with $I^0$ near 0, those with strong intensities favorable to the candidate, or those with strong opposing intensities.\(^{10}\) Another issue involves the shape of the advertising response function: do ads have diminishing effectiveness, or is there at least some initial increasing effectiveness at low levels of advertising?\(^{11}\) The literature on the effectiveness of advertising in general and of political advertising in particular is contradictory or silent about these properties.

Finally, there are also questions about clutter; that is, whether ad effectiveness depends only on the level of the candidate’s own ads or also depends on the level of the opponent’s efforts, or on the level of advertising in other political races as well. For empirical analyses, clutter across races may be significant. For example, clutter may create differences between Presidential and non-Presidential election years in the U.S. For the theoretical model here with a single race, the crucial issue is whether there is cross-candidate clutter.

\(^{10}\) Most of the campaign advertising literature has concluded that ads reinforce partisanship (see Iyengar and Simon (2000) for a discussion of this literature). However, Freedman, Franz and Goldstein (2004) present evidence that political ads have the strongest effect on those with the lowest levels of political information, who would presumably be those voters with weak intensities.

\(^{11}\) This question has been a source of debate in the advertising literature for some time. In an extensive survey article, Simon and Arndt (1980) conclude that the advertising response function is concave rather than S-shaped, but this question has not been settled in the literature. Bronnenburg (1998) suggests that the empirical evidence supports a concave function because it is not optimal for a firm to advertise in the increasing part of the response function. He also finds evidence that long-term and short-term response functions may have a different shape, and that the functions may even differ over different consumer types. In the campaign advertising literature, Mueller (2003) discusses the S-shaped curve, but most of the literature incorporates either linear or concave response functions; for examples, see Grier (1989) and Levitt (1994). See Morton (2006) for a general discussion of how political advertising affects voters.
One specific example of $h(I^0, x, y)$ is $I = I^0 + g(x(I^0)) - \hat{g}(y(I^0))$ which assumes that advertising is equally effective at all values of $I^0$ and no clutter exists. The functions $g$ and $\hat{g}$ could be strictly concave, could have initial increasing returns so that they are S-shaped, or could be linear with constant marginal effectiveness so that $g = x(I^0)$ and $\hat{g} = y(I^0)$. One possible alternative specification that incorporates clutter between the ads of the opposing candidates is $I = I^0 + h[(x(I^0) - y(I^0))/(x(I^0) + y(I^0))]$. For $h$ linear or near linear, this function is strictly concave in $x(I^0)$ and strictly convex in $y(I^0)$. If, however, $h$ is nonlinear with significant curvature, then these may be violated for $x$ if $h'' > 0$ and for $y$ if $h'' < 0$. The specific condition for concavity in $x(I^0)$ and for convexity in $y(I^0)$ are that $\frac{1 - y(I^0)}{x(I^0)} < h''/h' < \frac{1 + x(I^0)}{y(I^0)}$.

In general, we assume diminishing effectiveness, which is independent of voter intensity, but initially make no assumption about clutter.\textsuperscript{12} Assuming $h$ is twice continuously differentiable, these imply:

\begin{equation}
    h_x > 0, \quad h_{xx} < 0, \quad h_y < 0, \quad h_{yy} > 0, \quad \text{and} \quad h_I \equiv 0 \tag{3}
\end{equation}

In addition, a useful benchmark case adds no cross-candidate clutter and has the same effectiveness function for both candidates:

\begin{equation}
    I = I^0 + g(x(I^0)) - g(y(I^0)), \text{ with } g' > 0, \quad g'' < 0. \tag{4}
\end{equation}

\textbf{c. Targeting}

In an ideal world for candidates, they would be able to precisely target their expenditures with different messages and expenditure amounts for each voter type. Candidates certainly attempt to accomplish this by running different ads on television.

\textsuperscript{12} The assumption that $h$ is independent of voter intensity is not used in the existence or uniqueness results of Theorem 1 but helps simplify the comparative statics results.
programs whose viewers have different demographics, or by giving different speeches depending on the nature of the audience.

In reality, perfect targeting is not possible, and spillovers occur where a message intended for one type is received by another. To model this formally, we assume that candidates face two media markets, denoted n and m. Within each market, there are three types of voters: partisans for each of the two candidates and neutrals. The pre-campaign intensities of preference for voters preferring candidate L in the two communities are $-\theta_n$ and $-\theta_m$, while those for R are $\gamma_n$ and $\gamma_m$. The neutral types have intensities of 0 in both markets. Let N and M be the total number of voters in communities n and m, respectively. In community n, there are $n_1$ partisans of L, $n_2$ neutrals, and $n_3$ partisans of R. In community m, these numbers are $m_1$, $m_2$, and $m_3$. We will call those individuals initially preferring L submarkets 1, those initially neutral submarkets 2, and those initially preferring R submarkets 3. Candidates can target the markets differently but cannot separately target submarkets within a market. Thus, candidate R can allocate resources with $x_n$ going to market n and $x_m$ to market m, while L can set $y_n$ and $y_m$ for the two communities, respectively. The post campaign situation in the two communities under (3) is shown in Figure 1.

The two candidates face resource constraints

$$x_n + x_m = E_R \text{ and } y_n + y_m = E_L$$

(5)

---

Skaperdas and Grofman (1995) similarly assume three types of voters, partisans of each candidate and undecideds. They have one market with the candidates deciding the mix of positive and negative campaign activity. Intensities of partisanship and the effectiveness of campaign activities are combined in reduced form functions specifying the extent to which voters move from being partisans to undecided or the reverse.
where $E_R$ and $E_L$ are their available resources, which we assume are fixed.\textsuperscript{14} Without loss of generality, we assume that $R$ has at least as much resources as $L$, so that $E_L \leq E_R$. We also assume that the resource difference is not too great, relative to voter intensities:

$$h\left(\frac{1}{2} E_R, \frac{1}{2} E_L\right) < \min[h_n, h_m]$$

(6)

The candidates simultaneously make their allocation decisions with the goal of maximizing their expected net votes.\textsuperscript{15} Thus candidate $R$ chooses $x_n$ and $x_m$ to maximize

$$n_1 V[-\theta_n + h(-\theta_n, x_n, y_n)] + n_2 V[h(0, x_n, y_n)] + n_3 V[\gamma_n + h(\gamma_n, x_n, y_n)] +$$

$$m_1 V[-\theta_m + h(-\theta_m, x_m, y_m)] + m_2 V[h(0, x_m, y_m)] + m_3 V[\gamma_m + h(\gamma_m, x_m, y_m)]$$

(7)

while $L$ chooses $y_n$ and $y_m$ to minimize this. Denote the function in (7) as

$$\hat{V}(x_n, x_m, y_n, y_m).$$

(III) Existence and uniqueness of equilibria

Consider further the general framework of imperfect voter characterization by the candidates. Then $V(I)$ will have a shape like that given in Figure 2. Note that for $c$ finite, regardless of the shape of $\phi$ between $-c$ and $c$, $V(I)$ cannot be globally concave or globally convex because of its upper and lower bounds of 1 and 0. This nonconcavity can lead to nonexistence of pure strategy equilibria. As an example, consider the limiting case as $c$ goes to 0. $V(I)$ converges to the correspondence in Figure 3, which is essentially the case of deterministic voting, where candidates believe that individuals vote

\textsuperscript{14} Initially we assume that advertising prices are the same in both markets but later we allow for different prices related to differences in market size.

\textsuperscript{15} Maximizing expected plurality is commonly assumed as the objectives of the candidates in probabilistic voting models in the spatial context. This objective is easier to analyze than the standard alternative of maximizing the probability of winning, as done in Stromberg (2008). A number of papers including Aranson, Hinich and Ordeshook (1974), Ledyard (1984), Snyder (1989), Duggan (2000) and Patty (2005) have considered the relation between outcomes under the two objectives. In some, perhaps special, circumstances, they are equivalent. In the redistribution context, Lindbeck and Weibull (1987) develop some differences in outcomes when candidates maximize their probability of winning the contest rather than expected plurality. It should be noted that expected plurality maximization is not only more convenient but in some circumstances can be justified as more realistic. Candidates with a low probability of winning may desire to lose by as small a margin as possible.
with certainty for their preferred candidate. This leads to a Colonel Blotto game modified by imperfect targeting. Fletcher and Slutsky (2008a) show that pure strategy equilibria generally do not exist in this case. Similar results will hold for cases near deterministic voting when $c$ is positive but sufficiently small.

For other parameter values, the vote function $V(I)$ will be concave over the relevant range of intensities $I$. First, if $c$ is large relative to the possible values of $I$, the upper and lower bounds of 1 and 0 for $V$ do not come into play. For $\lambda \equiv \max\{\theta_n - h(0, E_L), \theta_m - h(0, E_L), \gamma_n + h(E_R, 0), \gamma_m + h(E_R, 0)\}$, assume:

$$\lambda < c$$

(8)

In effect, (8) asserts that no candidate has enough resources to win a submarket with probability 1 even if the other candidate does not compete in that market. Under (8), for all $x_i$ and $y_i$, $-\theta_i + h(x_i, y_i) + \varepsilon > 0$ and $\gamma_i + h(x_i, y_i) + \hat{\varepsilon} < 0$ both hold for some $\varepsilon$ or $\hat{\varepsilon}$ arising with positive probability. Clearly (8) holds if $\varphi$ has unbounded support with $c = \infty$.

In addition to ruling out the boundary nonconcavity/ nonconvexity, restrictions are also needed on $\varphi$ and $h$ to ensure the appropriate local concavity and convexity in the relevant range of intensities. The following condition is sufficient for this:

---

16 Deterministic voting is usually specified with $V(0) = \frac{1}{2}$ so that an indifferent voter has equal probability of supporting either candidate, whereas in Figure 3 any probability of voting for either candidate is allowed when $I = 0$. This distinction would make little difference in the analysis of deterministic voting.

17 Lindbeck and Weibull (1987) do not explicitly address the global convexity issue, but implicitly assume a condition analogous to (8) with $c = \infty$, by assuming there is positive density everywhere.

18 This condition is directly analogous to condition C1 in Lindbeck and Weibull (1987). The right hand side of their condition depends on curvature conditions of the individual voter’s utility function instead of the advertising effectiveness function. Note that their weaker condition C2 does not directly apply in this context, since their condition is based on the candidates having identical equilibrium strategies, which is not true here.
\[
\max_{\varepsilon \in [-\lambda, \lambda]} \left| \frac{\phi'(<)}{\phi(<)} \right| < \min_{x \in [0,E_x], \varepsilon \in [0,E_i]} \left[ -\frac{h_{xx}}{(h_x)^2}, \frac{h_{yy}}{(h_y)^2} \right] \tag{9}
\]

Note that \(\phi'(<)/\phi(<) = (\partial^2 V/\partial I^2)/ (\partial V/\partial I)\) is a curvature measure of \(V(I)\) and \(-h_{xx}/(h_x)\) and \((h_{yy}/(h_y)\) are partial curvature measures of \(h\). These are mathematically analogous to measures of absolute risk aversion, in that both measure the degree of concavity.\(^{19}\) Thus (9) in a sense asserts that the curvature of \(h\) exceeds the variation in \(\phi\).

For (9) to hold, the right hand side expression related to \(h\) must be strictly positive. For the benchmark case in (4), consider the following three examples of the function \(g\), where \(k\) is a positive constant. For \(g(z) = k \ln z\), \(-h_{xx}/(h_x)^2 = (h_{yy}/(h_y)^2 = -g''/ (g')^2 = 1/k\) for all \(x\) or \(y\). Clearly, in this case, the right hand side of (9) is strictly positive. For \(g(z) = k z^{1/2}\), \(-h_{xx}/(h_x)^2 = (1/k)x^{-1/2}\) and \((h_{yy}/(h_y)^2 = (1/k)y^{-1/2}\). The right hand side of (9) is then \((1/k)E_R^{-1/2} > 0\). On the other hand, the constant marginal effectiveness case of \(g(z) = kz\) is ruled out since the right hand side of (9) would be 0 and the condition could not be satisfied.

Next consider three particular cases of \(\phi(<):\) uniform, normal, and quadratic distributions of errors. For the uniform distribution, \(\phi(<) = 1/(2c)\), and condition (9) always holds since \(\phi' = 0\). Condition (8) then imposes a sufficiently large value of \(c\) given the values of the other parameters of the model. For the normal distribution,

\[
\phi(<) = \left(1/((2\pi)^{1/2} \sigma)\right)e^{-<^2/(2\sigma^2)}, \text{ where } \sigma \text{ is the standard deviation of } <. \text{ Since this has}
\]

\(^{19}\) The terms for \(h\) here as compared to the coefficient of absolute risk aversion have an additional first partial derivative in the denominator, and thus are not invariant to linear transformations. Since \(h\) is being compared to another function \(\phi\), such transformations are not admissible. Condition (9) must be evaluated for each specific transformation. Most directly, the right hand side of (9) relates to the concavity index of Debreu and Koopmans (1982), which for a twice continuously differentiable function \(f(x)\) equals \(\inf_x [f'(x)/(f'(x))^2]\). They show that the sum of two functions is quasiconvex if the sum of their concavity indices is nonnegative. Here, condition (9) relates to a composite of two functions instead of their sum.
unbounded support, (8) holds automatically. For this distribution, \( \varphi'/\varphi = -\varepsilon/\sigma^2 \) for which the absolute value on the range \(-\lambda \leq \varepsilon \leq \lambda\) has a maximized value of \( \lambda/\sigma^2 \). Given any allowable function \( h \), (9) will hold for sufficiently large \( \sigma \). For the quadratic distribution, 
\[
\varphi(\varepsilon) = 3(1-(\varepsilon/c)^2)/4c. \]
Then \( \varphi'/\varphi = -2\varepsilon/(c^2-\varepsilon^2) \). If (8) holds, then the left hand side of (9) equals \( 2\lambda/(c^2-\lambda^2) \). For given \( \lambda \), both (8) and (9) will then hold for sufficiently large \( c \). In all three cases, the requirements for these conditions to hold are that the error distribution be sufficiently diffuse either by having a large support in the uniform and quadratic cases or a large variance for the normal distribution.

**Lemma 1**: Under conditions (8) and (9), the objective function \( \hat{V}(x_n, x_m, y_n, y_m) \) in (7) is strictly concave in \( x_n \) and \( x_m \) and strictly convex in \( y_n \) and \( y_m \).

In the Downsian context, to get existence of an equilibrium, the probability that an individual votes for a candidate must be a concave function of that candidate’s platform and is a convex function of the opponent’s platform (see Coughlin (1992) for a discussion of such models). Under conditions (8) and (9), imperfect voter characterization yields similar concave/convex vote functions in the non-Downsian context considered here. Probabilistic voting, then, is imperfect voter characterization under the restrictions of (8) and (9).

Note that deterministic voting arises when the error term is sufficiently concentrated, while probabilistic voting occurs when the error term is sufficiently diffuse.

---

\(^{20}\) For this function, \( \varphi(c) = \varphi(-c) = 0 \) and \( \int_{-c}^{c} \varphi(\varepsilon) d\varepsilon = 1 \).
The conditions for each case given above are sufficient but certainly not necessary. Deterministic-like voting behavior may arise even when $c$ is not near zero, and probabilistic voting behavior may arise even with some nonconcavity in the objective function, if the nonconcavity is for strategies not near the equilibrium ones. Determining better boundaries for each case can be difficult. When the sufficient conditions for neither case hold, the equilibrium may be similar to one of the cases or may differ from both, having a mix of equilibrium types.

Given sufficiently diffuse observation errors, a unique equilibrium exists in this model.

**Theorem 1:** Under probabilistic voting (restrictions (8) and (9)), a unique equilibrium exists which is in pure strategies.

Although the formal model has only two media markets, the result in Theorem 1 extends straightforwardly to situations with any number of media markets. The model is not limited to the special case of two markets.

When is the equilibrium on a boundary and when is it interior? To get some insight into this question, consider the special case of $h$ given in (4). Recall that $E_L \leq E_R$ is assumed as a convention. Add shift parameters $T_n$ and $T_m$ in each community with the objective function $\hat{V}$ in (7) now equal to $\hat{V} = n_1 V(-\theta_n + g(x_n) - g(y_n) + T_n) + n_2 V(g(x_n) - g(y_n) + T_n) + n_3 V(\gamma_n + g(x_n) - g(y_n) + T_n) + m_1 V(-\theta_m + g(x_m) - g(y_m) + T_m) + m_2 V(g(x_m) - g(y_m) + T_m) + m_3 V(\gamma_m + g(x_m) - g(y_m) + T_m)$. Theorem 2 presents the possible types of boundary equilibria and gives some necessary conditions for their existence.
**Theorem 2:** For the specification of \( h \) given in (4), the only possible boundary equilibria (those with at least one candidate at a corner) have \( y_i = 0 \) and \( 0 \leq x_i < E_R \). Necessary conditions for such equilibria are:

(A) \( g'(0) < \infty \)

(B) \( \partial \hat{V} / \partial T_i < \partial \hat{V} / \partial T_j \) evaluated at \( y_i = 0 \) and \( x_i \)

(C) if \( x_i > 0 \), then \( x_i < E_R - E_L \)

To gain some understanding of the conditions in Theorem 2, start with a situation with \( E_R = E_L \) so that from (C), the only possible boundary equilibrium is \( x_i = y_i = 0 \). In this case, \( \partial \hat{V} / \partial T_n = n_1 \phi(\theta_n) + n_2 \phi(0) + n_3 \phi(-\gamma_n) \) and \( \partial \hat{V} / \partial T_m = m_1 \phi(\theta_m) + m_2 \phi(0) + m_3 \phi(-\gamma_m) \). Then \( \partial \hat{V} / \partial T_m \) will increase in the \( m_i \)'s and, if \( \phi \) is unimodal, will decrease in \( \theta_m \) and \( \gamma_m \). Thus, candidates are more likely to expend no resources in a market if it is smaller and has more intense partisans than the other market. Of course, if \( g'(0) = \infty \) so that (A) is violated, then the equilibrium must be at an interior regardless of the relative sizes or preference intensities of the two markets.\(^{21}\)

Next, consider increasing \( E_R \) above \( E_L \), with \( x_n = 0 \) held fixed. Since

\[
\partial \hat{V} / \partial T_m = m_1 \phi(\theta_m - g(E_L) - g(E_R)) + m_2 \phi( g(E_L) - g(E_R)) + m_3 \phi(-\gamma_m + g(E_L) - g(E_R)),
\]

then \( \partial^2 \hat{V} / \partial T_m \partial E_R = - [m_1 \phi'(\theta_m) + m_2 \phi'(0) + m_3 \phi'(-\gamma_m)] \). If \( \phi \) is unimodal, then \( \phi'(\theta_m) < 0 \), \( \phi'(-\gamma_m) > 0 \), and \( \phi'(0) = 0 \). If so, \( \partial^2 \hat{V} / \partial T_m \partial E_R \) is ambiguous in size. The increase in \( E_R \) could reinforce the \( x_n = y_n = 0 \) equilibrium if \( m_1 \phi'(\theta_n) > m_3 \phi'(-\gamma_n) \) or could weaken it,

\(^{21}\) Lindbeck and Weibull (1987) make an analogous assumption on marginal utilities of income to rule out corner solutions.
leading to the wealthier candidate advertising in both markets if the inequality were reversed.

(IV) Comparative Statics

We next derive comparative statics results for some of the parameters. This gives insight into the nature of probabilistic voting in this context, and also develops some testable implications of the model. For interior equilibria, let $\mu$ be any parameter in the model. Then totally differentiating the system of interior first order conditions for $x_n$ and $y_n$ ($\hat{V}_1 - \hat{V}_2 = 0, -\hat{V}_3 + \hat{V}_4 = 0$) yields the comparative statics equations:

$$\begin{bmatrix}
\frac{\partial}{\partial \mu} x_n / \frac{\partial \mu}{\partial \mu} \\
\frac{\partial}{\partial \mu} y_n / \frac{\partial \mu}{\partial \mu}
\end{bmatrix} =
\begin{bmatrix}
\frac{\partial V_1}{\partial \mu} + \frac{\partial V_2}{\partial \mu} \\
-\frac{\partial V_3}{\partial \mu} - \frac{\partial V_4}{\partial \mu}
\end{bmatrix}$$

From the concavity and convexity of $\hat{V}$ in $x_n$ and $y_n$, respectively, the determinant of the matrix in (10) is positive. Hence, solving (10) using Cramer’s Rule yields:

$$\text{sign} \left( \frac{\partial x_n}{\partial \mu} \right) =$$

$$\text{sign} \left[ (-\hat{V}_{33} - \hat{V}_{44})(\partial \hat{V}_2 / \partial \mu - \partial \hat{V}_1 / \partial \mu) + (\hat{V}_{13} + \hat{V}_{24})(\partial \hat{V}_4 / \partial \mu - \partial \hat{V}_3 / \partial \mu) \right] \quad (11)$$

$$\text{sign} \left( \frac{\partial y_n}{\partial \mu} \right) =$$

$$\text{sign} \left[ (\hat{V}_{11} + \hat{V}_{22})(\partial \hat{V}_3 / \partial \mu - \partial \hat{V}_4 / \partial \mu) + (\hat{V}_{13} + \hat{V}_{24})(\partial \hat{V}_2 / \partial \mu - \partial \hat{V}_1 / \partial \mu) \right] \quad (12)$$

To interpret (11) and (12), consider a starting point with the two markets being identical: $n_i = m_i, i = 1, 2, 3, \theta_n = \theta_m,$ and $\gamma_n = \gamma_m$. Then the equilibrium is $x_n = x_m = \frac{1}{2}E_R$ and $y_n = y_m = \frac{1}{2}E_L$. In this circumstance, $\hat{V}_{11} = \hat{V}_{22}, \hat{V}_{33} = \hat{V}_{44},$ and $\hat{V}_{12} = \hat{V}_{24}$. By

---

22 If $x_n \neq x_m$ or $y_n \neq y_m$ multiple equilibria would exist contradicting Theorem 1.
considering μ to be some change that makes the markets differ from each other, we can consider how candidates would tend to allocate resources when facing asymmetric markets. Assume μ enters only market n. Hence, \( \partial \hat{V}_2 / \partial \mu = \partial \hat{V}_4 / \partial \mu = 0 \). Conditions (11) and (12) then reduce to:

\[
\text{sign}\left( \frac{\partial x_n}{\partial \mu} \right) = \text{sign}\left[ \frac{\partial^2 h(h_{1\text{ER}}, h_{1\text{EL}})}{\partial x_n \partial y_n} \right] \quad (13)
\]

\[
\text{sign}\left( \frac{\partial y_n}{\partial \mu} \right) = \text{sign}\left[ \frac{\partial^2 h(h_{1\text{ER}}, h_{1\text{EL}})}{\partial x_n \partial y_n} \right] \quad (14)
\]

In order to sign some of the comparative statics, it is helpful to place bounds on the magnitude of \( \frac{\partial^2 h(h_{1\text{ER}}, h_{1\text{EL}})}{\partial x_n \partial y_n} \). These bounds essentially specify the magnitude of clutter effects when candidates divide their resources equally. Consider two possible conditions bounding \( h_{xy} \):

\[
(h_x / h_y) h_{yy} < h_{xy} < (h_x / h_y) h_{xx} \quad (15)
\]

\[
h_x h_y \varphi_3' < h_{xy} < h_x h_y \varphi_i' / \varphi_i \quad (16)
\]

Given (9), the bounds in (16) imply those in (15). The condition in (16) is only valid if \( \varphi \) is strictly single peaked, strengthening weak single peakedness specified in (1b).

Given strict single peakedness, \( \varphi_3' > 0 \) since it is evaluated at \( \varepsilon < 0 \) while given (6), \( \varphi_1' < 0 \) since it is evaluated at \( \varepsilon > 0 \). Condition (15) is satisfied for a wide variety of special cases. The case specified in (4) with \( h_{xy} = 0 \) satisfies the condition since the lower bound is negative and the upper bound positive. It is straightforward to check that the example with clutter \( h((x_n - y_n)/(x_n + y_n)) \) also satisfies the condition.

---

23 In what follows, the derivatives are generally taken in market n so x and y are presumed to be \( x_n \) and \( y_n \).

For simplicity, we will define the values of \( \varphi \) in the three submarkets of market n by \( \varphi_1 = \varphi[\theta_n - h (\frac{1}{2} E_R, \frac{1}{2} E_L)], \varphi_2 = \varphi[-h (\frac{1}{2} E_R, \frac{1}{2} E_L)], \) and \( \varphi_3 = \varphi[-\gamma_n - h (\frac{1}{2} E_R, \frac{1}{2} E_L)]. \)
Theorem 3: Assume that the two markets are initially identical and that \( \mu \) is a parameter which enters only market \( n \) and not through \( h(x_n, y_n) \). If \( (15) \) holds, then
\[
\text{sign} \left( \frac{\partial x_n}{\partial \mu} \right) = \text{sign} \left( \frac{\partial y_n}{\partial \mu} \right) = \text{sign} \left( \frac{\partial z}{\partial \mu} \right).
\]

There are two crucial results contained in Theorem 3. First, for parameters which do not enter the function \( h \), the two candidates respond to differences between markets in the same way.\(^{24}\) If one candidate spends more on ads in market \( n \) than in market \( m \), then so will the other candidate. Second, determining the market in which both candidates spend more reduces to a relatively simple condition. If \( \frac{\partial y_n}{\partial \mu} \) is positive then both spend more in market \( n \) while if it is negative, they spend more in \( m \). Based on this result, the effects on allocation of a variety of parameters not directly entering \( h \) are presented in Table 1.
(Insert Table 1)

In the following subsections we discuss the comparative statics for polarization, partisan leanings, nonmonetary resources, and market size.

a. Increased polarization

One market might be more polarized than the other in two different ways. Either the two markets have partisans of the same intensity but one of them has fewer swing voters and more partisans than the other, or the two markets have submarkets of the same size but the partisan individuals are more intense in one of them. Result (III) considers the first possibility. Given that the resource difference between the candidates is bounded

\(^{24}\) These comparative statics give the effects of small changes starting from a situation of complete symmetry between the markets. The results will hold for a neighborhood around symmetry, but may not be valid if the markets are very different. Deriving the effects for other situations is difficult, involving ambiguous factors.
by condition (6), if the error distribution is not uniform, then the candidates spend less in the market with more partisan voters. Result (IV) considers the second possibility. The effect of such increased polarization is ambiguous. If the error distribution is uniform, then there is no difference in allocation. If the error distribution is strictly concave, then candidates spend less in the more polarized community. If the error distribution is strictly single peaked (but as in the normal distribution need not be concave) and if the difference in the resources of the candidates is not too large, then again they spend less in the more polarized community. Note that the sufficient condition bounding the resource difference, $h(\frac{1}{2} E_R, \frac{1}{2} E_L) < \frac{1}{2} \theta_n$ is stronger than (6) but is clearly not necessary. Nevertheless, if the resource difference between the candidates is sufficiently large and $\phi$ has regions of nonconcavity, then it is possible for the candidates to spend more in the community more polarized prior to any campaign expenditures. See Figure 4 for an example of this. When $h(\frac{1}{2} E_R, \frac{1}{2} E_L)$ is sufficiently large, $\theta_n - h$ is near zero, which given the bell shape of $\phi$ makes $\phi_1$ sufficiently larger than $\phi_2$ and $\phi_3$. After the expenditure of resources by the candidates, the partisans of $L$ become close to being swing voters. The community with more such partisans is a better target for candidate activity.

b. Differential partisanship

One way for one market to lean more toward a candidate than does the other market is if it has relatively more partisans of that candidate. Result (V) considers this by assuming that in market $n$ the number of partisans of $L$ is larger and the number of partisans of $R$ is smaller than in market $m$, keeping the total populations in the two markets the same. The effect depends on whether $\phi_1$ is larger or smaller than $\phi_3$. For the
uniform error distribution, these are equal so despite the different population mixes, candidates spend the same in both markets. Consider the case of a strictly single peaked error distribution. The term $2h + \gamma_n - \theta_n$ will be positive unless the intensity of individuals favoring L exceeds that of those favoring R with the needed difference in intensity larger the larger is the resource difference. There is thus some tendency for the candidates to spend more in the market that leans relatively toward the poorer candidate unless the poorer candidate’s partisans are sufficiently more intense.

c. Nonmonetary resources

In addition to their paid advertising, candidates often have a variety of types of nonmonetary resources that sway voters, such as unpaid volunteers and media endorsements. These nonmonetary resources might interact with the paid advertisements in different ways. Result (VI) considers some cases where the effect of these expenditures is independent of that of advertisements, and thus the effects enter outside the function $h$. Different types of nonmonetary resources are covered by different values of the parameters $k_i$ in result (VI) of Table 1. These allow the effects of changing these nonmonetary resources to differ across submarkets. First, consider a candidate receiving a newspaper endorsement read by the entire market that has an effect independent of additional advertising by the candidate. This can be modeled by assuming that $k_1 = k_2 = k_3$, so that all submarkets are affected equally. Under the uniform distribution of errors, $\varphi' = 0$ at all $\varepsilon$ so $\partial x_n / \partial \mu$ and $\partial y_n / \partial \mu$ will be 0 and candidates do not alter their own activities in response. Consider $\varphi$ strictly single peaked. For $\mu$ positive all submarkets in market $n$ are shifted in favor of R. The effect of this is ambiguous. If $E_R = E_L$, $\theta_n = \gamma_n$, and $n_1 = n_3$, then there is no initial change in the allocation by either candidate as $\mu$
increases from 0. Large changes in $\mu$ would reduce $x_n$ and $y_n$ if $\phi$ were strictly concave. If $\phi$ is bell shaped, the effect would be ambiguous depending upon the values of $\theta_n$ and $\gamma_n$ relative to the concave and convex regions of $\phi$. Deviations from these symmetry conditions could cause the initial changes in $x_n$ and $y_n$ to be of either sign. Thus, whether candidates respond to such endorsements by more or less advertising is ambiguous.

Second, consider candidate R having an additional resource that is not a substitute for advertising and that particularly affects R’s partisans. Union membership for a Democratic candidate and evangelical Christian membership for a Republican candidate could be examples of this. Such members could be directly contacted by mail, phone calls or personal visits without spillovers to others. Such contacts might have different effects than an additional ad. This can be modeled by setting $k_1 = k_2 = 0$, so that only submarket 3 is affected. In this case, both candidates respond to this advantage for one of the candidates by spending less in market $n$. This analysis was based upon the candidates having additional resources in only one market. However, it could be extended to their having additional resources in both, with $\mu$ indicating the difference in resources between the two markets. This would require only a slight reinterpretation of the bound in (6) since the additional resources are formally equivalent to a change in the intensities of support.

Consider an alternative specification of such additional resources in which they are perfect substitutes for ordinary advertising. Assume in market $n$, the candidates can directly approach their partisans at no cost through unions or churches. Then the expected vote for R in market $n$ is $n_1 V(-\theta_n + h(x_n, y_n + \mu_L)) + n_2 V(h(x_n, y_n)) + n_3 V(\gamma_n + h(x_n + \mu_R, y_n))$. As before, market $m$ is assumed identical to market $n$ when $\mu_L = \mu_R = 0$. 
Theorem 4: Starting from initially identical markets including \( \mu_R = \mu_L = 0 \), the following comparative statics results hold:

(A)  \( \partial x_n / \partial \mu_R < 0 \) and \( \partial y_n / \partial \mu_L < 0 \) either if (16) holds or if \( \varphi \) is uniform.

(B)  \( \partial x_n / \partial \mu_R = 0 \) if \( \varphi \) is uniform and \( \partial x_n / \partial \mu_L = 0 \) if (15) holds and \( \varphi \) is strictly single peaked

(C)  \( \partial y_n / \partial \mu_R \) can have any sign.
   (i)  \( \partial y_n / \partial \mu_R = 0 \) if \( \varphi \) is uniform
   (ii) \( \partial y_n / \partial \mu_R < 0 \) if (15) holds and either if \( E_L = E_R \) and \( \varphi \) is strictly single peaked or if \( E_R > E_L \) and \( \varphi \) is quadratic or normal, or if \( n_2 \) is sufficiently small relative to \( n_1 \).
   (iii) \( \partial y_n / \partial \mu_R > 0 \) if (15) holds, \( \varphi' \varphi_3 / \varphi_3 < \varphi' \varphi_2 / \varphi_2 \), and \( n_1 \) is small relative to \( n_2 \).

Condition (A) gives two sufficient conditions for \( \partial x_n / \partial \mu_R \) and \( \partial y_n / \partial \mu_L \) to be negative. These seem to go in different directions. One (\( \varphi \) uniform) involves small values of \( \varphi' \) with no restrictions on \( h_{xy} \) while the other through condition (16) involves limits on the magnitude of the cross effect \( h_{xy} \) which are weaker when \( \varphi' \) is larger. The explanation for this is that the expressions giving the signs of these comparative static effects are quadratic in \( h_{xy} \). For the expressions to be positive reversing the comparative statics requires \( h_{xy} \) to violate (16) and for the peak of the quadratic to be above 0. When \( \varphi' \) is small the peak is reduced, while when \( \varphi' \) is big the range of \( h_{xy} \) for which the
expression might be positive is small. Hence, it will be only for a limited set of values of $\varphi'$ and $h_{xy}$, if any, for which the signs will be reversed.

Overall, in a wide range of circumstances, if either candidate has nonmonetary resources targeted toward the candidate’s base in one market which are perfect substitutes for paid expenditures, then that candidate will spend less in that market and more in the other. If it is the candidate with less monetary resources who has the extra nonmonetary resources, then the other candidate also spends less in that market relative to the other. For the most part, these results are similar to those in result (VI) of Table 1, when the nonmonetary resources are not perfect substitutes with monetary expenditures since typically the candidates spend less in the market with the nonmonetary resources. There are some subtle differences, however. If the error distribution is uniform, then in the non perfect substitute case, the presence of nonmonetary resources does not induce unequal expenditures across markets by either candidate. With perfect substitutes, a candidate with nonmonetary resources does deviate from equal expenditures across markets but the other candidate does not. With non-uniform error distributions, under perfect substitutes, a low resource disadvantaged candidate may actually spend more in the market where his opponent has an advantage. The conditions for this seem unlikely to hold. From (Ciii), for candidate L to spend more in the market where R has extra nonmonetary resources, R must also have more monetary resources, the error distribution must be relatively flat in the tail, and $n_1$ cannot be large relative to $n_2$.

d. Market size and ad prices

The effect of market size on candidate allocations depends upon whether everything else is held fixed when population differs or, as would seem more reasonable,
other factors --- especially advertising prices --- also vary with population. Results (I) and (II) show that both candidates spend more in the larger market either if only some of the submarkets are larger or if all submarkets are proportionally larger when ad prices are independent of community size. In fact, ad prices tend to increase with market size. The ad price may be proportional to market size so that the cost per person is constant, or it might rise less than proportionally so that cost per person declines. To analyze this, we first consider the direct effect of differential ad prices by considering communities that are identical except that the ad price is not the same in them. Since \( x_i \) and \( y_i \) are the dollar expenditures on ads, \( x_i/\mu_i \) and \( y_i/\mu_i \) would be the number of ads in market \( i \). The function \( h \) should depend upon the number of ads and not the dollars spent with 

\[ h(x_i/\mu_i, y_i/\mu_i). \]

Assume \( \mu_m = 1 \) as a normalization and consider increases in \( \mu_n \) above this level.

**Theorem 5:** Starting from identical markets and normalizing the price of an ad to 1 in market \( m \), if (15) is satisfied then:

(i) \[
-\frac{1}{2}x_n = \frac{\partial (x_n/\mu_n)}{\partial \mu_n} + \frac{1}{2}x_n < 0
\]

(ii) \[
-\frac{1}{2}y_n = \frac{\partial (y_n/\mu_n)}{\partial \mu_n} + \frac{1}{2}y_n < 0
\]

In the community with the higher ad price, the candidates will run fewer ads but may spend more or less in total. To explore this further, assume the special case of \( h \) in (4). If \( E_L = E_R \) or if \( \varphi \) is uniform, then 

\[
-\frac{\partial}{\partial \mu_1} \frac{\partial}{\partial \mu_2} + (\frac{\partial}{\partial \mu_1})^2 = A^2 g''(x_n) g''(y_n).
\]

Substituting yields 

\[
\frac{\partial x_n}{\partial \mu_n} = \frac{g'(x_n)}{(2g''(x_n))} [1 + x_n \frac{g''(x_n)}{g'(x_n)}] \text{ and } \frac{\partial y_n}{\partial \mu_n} =
\]
\[ (g'(y_n)/(2g''(y_n))) \left[ 1 + \frac{y_n g''(y_n)}{g'(y_n)} \right]. \]

The terms \( zg''(z)/g'(z) \) are curvature measures of \( g \) analogous to the coefficient of relative risk aversion and could be greater or less than \(-1\), showing that in this case the ambiguity of \( \partial x_n/\partial \mu_n \) and \( \partial y_n/\partial \mu_n \) depend in part on the degree of curvature of \( g \).

Finally, assume that ad prices are proportional to market size. Again starting from otherwise identical markets, \( n_i = \mu_i \) and, in market \( n \), the ad effectiveness function is \( h(x_n/\mu, y_n/\mu) \). Changes in \( \mu \) now incorporate ad price changes as well as the direct effect of population size.

**Theorem 6**: Starting with identical markets, let \( \mu \) denote an increase in population in market \( n \) that affects all submarkets proportionally and raises ad prices proportionally. Then

\[
\begin{align*}
\partial x_n/\partial \mu &= -\partial(x_n/\mu)/\partial \mu = \frac{1}{2} x_n \\
\partial y_n/\partial \mu &= -\partial(y_n/\mu)/\partial \mu = \frac{1}{2} y_n
\end{align*}
\]

In contrast to the result when ad prices are independent of market size given in Table 1, candidates run fewer instead of more ads in larger markets when ad prices are proportional to market size. They spend more in the larger market but the spending rises less than does population and hence the ad price. If ad prices vary with population but not necessarily proportionally, then depending upon the relationship between price and population, candidates could run more, less, or about the same number of ads in larger markets.

(V) **Testable implications**
In the formal analysis above, candidates only had two markets between which to allocate resources. If this assumption were significant, our ability to empirically test the model would be limited since there are few campaigns dealing with exactly two media markets. However, because the equilibrium is in pure strategies, in a situation with more than two markets, similar results will hold looking at any pair of markets. Fix expenditure in all but two of the multiple markets. There are no spillovers across markets in the effects of spending, and the objective functions of maximizing expected vote are additive across markets. This means the decisions over any pair of markets are independent of how resources are allocated in other markets, given the total spending allocated to the pair of markets. Thus we can test the results in contests with more than one market.

Significant additional difficulties exist for empirical testing of the comparative statics results. Not all the parameters in the theoretical model specified above are easily observable and, because of the complexity of the model, some factors which are easily observable do not relate to unambiguous comparative static predictions. The only direct model parameters that are observable are the total sizes of the different media markets N and M. Parameters such as voter intensities $\theta_i$ and $\gamma_i$ and the numbers of the different voter types $n_i$ and $m_i$ are not directly observable. One way to deal with this is to find observable proxies for the unobservable parameters. For example, the fraction of a market’s population that is Black and the fraction that is evangelical Christian are observables that are significantly correlated with voter intensities and submarket sizes and thus can be used as proxies. Observables for which the model does not make unambiguous predictions, such as partisan leaning, cannot be used to test the model.
However, the empirical analysis can still have great interest in indicating which of the counteracting tendencies dominates.

One clear hypothesis for the model is that the effect of any observable, including those whose direction is ambiguous, should be the same for both candidates. As shown in Theorem 3, this is always true for factors that affect campaign allocations but do not enter the effectiveness function $h$. Even when a factor enters $h$, such as nonmonetary resources that are substitutes for monetary resources as considered in Theorem 4, the candidates differ in how their allocations compare across markets only in exceptional cases. That is, if one market differs from another in some dimension, both parties will tend to spend more in the same market relative to the other market. This is a nonobvious prediction of the model since it might be thought that a Democrat would spend more in the market more favorable to Democrats and the Republican would do the reverse.

Two additional hypotheses follow from the proxies fraction Black and fraction evangelical Christian. African Americans vote overwhelmingly for Democrats, with little variability across elections. Controlling for the overall fraction of the market’s population supporting the Democratic candidate, a higher fraction Black would tend to indicate the media market has a relatively large number of individuals with a high intensity of support for Democratic candidates. The comparative statics would thus indicate that both candidates would spend less in such media markets. Given a reasonable shape of the error distribution, there will be a smaller chance of swaying voters in such markets. Thus, a second testable implication is that both candidates should devote fewer resources to the market with a higher fraction of African Americans.
In recent years, evangelical Christians have been viewed as a block reliably supporting Republicans similar to the African American support of Democrats. In addition, there have been arguments that Republicans have utilized evangelical support in a way similar to that in which Democrats use union resources to mobilize voters. Either of these cases would lead to the third testable implication that both candidates should spend less in a market with a higher fraction of evangelical voters.25

To summarize, our empirical model will be used to test the following hypotheses:

1. Any difference in a market characteristic will affect both the Republican and Democratic candidate in the same direction. That is, the effect of any market characteristic should have the same sign for both candidates.

2. Both candidates should advertise less in markets with a larger fraction of African Americans, *ceteris paribus*.

3. Both candidates should advertise less in markets with a larger fraction of evangelical Christians, *ceteris paribus*.

**Empirical insights from the model**

These three implications allow us to test the validity of the model. We can also use the data to clarify some of the model’s predictions, which are ambiguous because of countervailing factors. The empirical results give insight as to which of these factors dominate.

As noted above, the effect of market size on advertising strategies depends upon how ad prices vary with population. The dataset we use to measure candidate advertising

---

contains an ad price variable, but it is not considered to be very reliable (see Goldstein and Freedman 2002). We consider it here in an attempt to get some idea, even if a crude one, of the relation between advertising price and population across markets. We calculate the average price per advertising second for each candidate in each market, and regress the natural log of this price on the log of the market’s population. State dummies are included to control for differences in pricing across states. For the entire sample, the elasticity of market price with respect to population is 0.63 (with a robust t-statistic of 24), suggesting that advertising price is not independent of population size, but varies less than proportionally with population. This variation in ad price is very much between the case in Table 1 result (II) where candidates would advertise more in the larger market, and that in Theorem 6 where the candidates would advertise less in the larger market. However, the theory does not determine at what elasticity the two effects would cancel each other out. Thus, it does not provide an unambiguous prediction for how market size should affect candidates’ advertising allocation decisions.

Some other variables that are observable, including the partisan leaning of a market or the closeness of the contest, have been used in a variety of empirical analyses. These variables clearly relate to the model here but not in any simple way,

26 The adjusted R squared statistic for this regression is 0.82. The result is robust to several changes in specification. In addition to this regression, we use the average over candidates and races so the sample size is smaller, and the average price for only prime time ads to reduce simultaneity by candidates choosing more ads in cheaper time slots. In all cases the estimated elasticities fall between 0.61 and 0.64, with t-statistics of at least 13. In addition, because we will consider the South and the rest of the country separately below, we also estimate this relation separately for each region. There does not appear to be a large difference between the Southern and non-Southern elasticities of ad price with respect to population. The estimated elasticities range from 0.55 in the non-South to 0.65 in the South, and are both statistically different from one at the 99.9% level of confidence. Thus, in both cases, ad price appears to vary less than proportionally with population.

27 To avoid simultaneity problems when estimating the effectiveness of campaign expenditures on vote shares, these variables are often measured by proxies such as the vote outcome in the Presidential election closest in time to the election being analyzed, as in Erickson and Palfrey (1998, 2000). Other authors (for
and their comparative statics are more ambiguous than they might appear on the surface. Consider closeness. It would seem straightforward that the candidates would spend more in the market where the contest is closer. Under probabilistic voting, however, this is not necessarily true since votes can be gained for a candidate from individuals who will vote for the opponent with probability greater than one half.\textsuperscript{28} There is no consistent prediction from the model for the effect of closeness on the spending patterns of the candidates. If closeness is correlated with spending decisions, it is because of correlations between closeness and unobserved intensities or subgroup populations. For example, if the number of initially neutral individuals happens to be larger in more equally divided markets, the candidates would spend more in markets where the contest is closer. Although we have no predictions for its effect, we include closeness in some of the empirical specifications below as it has been an important component of other analyses of campaign spending. If candidates spend more in more evenly divided markets, it would indicate that closer markets also happen to be less partisan.

A final relevant demographic in this context would be union membership. Two potential effects exist, which may work in opposite directions. One is that union resources tend to go to Democratic candidates. This would effectively be like having

\textsuperscript{28} To see this, consider two identical markets in which R gets a higher expected vote. This could be because initially $\gamma_i > \theta_i$; that is, because partisans for R are more intense than those for L. It could also be because $n_3 = m_3 > m_1 = n_1$, so that there are more partisans for R than for L in both markets. In the first case, the contest in market n could be made closer by raising $\theta_n$ or by lowering $\gamma_n$ relative to the values in market n. From result (3) of Table 1, both candidates will spend more in market n if $\gamma_n$ is decreased but less if $\theta_n$ is increased. Similarly, they will spend more if the contest is made closer through an increase in $n_1$ but less if it is through a decrease in $n_3$. In all cases, the result is not due to an effect through changing the closeness of the election, but to the direct effect of the parameters. An increase in $n_1$ makes N bigger than M and thus market n gets more advertising by the candidates and the reverse for a decrease in $n_3$. Similarly, increasing $\theta_n$ reduces spending because market n is more polarized while lowering $\gamma_n$ reduces polarization.
extra resources available to Democratic candidates, such as those considered in result (VI) of Table 1 or in Theorem 4. This effect would tend to lead both candidates to spend less in markets with stronger unions. A second effect relates to the nature of union voters. Historically, they were viewed as strong Democratic voters, but more recently some view them as swing voters whose views on various issues align with different parties. They might favor Democrats on economic issues and Republicans on social and security issues, and thus might be persuaded to support either party depending on which issues they consider more salient. If they actually are swing voters, both candidates would tend to spend more in markets with higher union membership.29 Because of these potentially countervailing effects, the model does not have an unambiguous prediction about the effect of union membership on campaign allocation among markets. Whether the effect is positive or negative sheds light on the candidates’ beliefs about the current behavior of union members. This sign indicates whether candidates view union members as swing voters or, if they are, whether the candidates believe the resource effect or the “swingability” effect dominates.

As discussed below, we will also use the model to see if the South differs significantly from the rest of the country in its political behavior.

**(VI) Empirical evidence**

a. **Data**

---

29 There is not a consensus among politicians and commentators as to the nature of union voters. In post 2002 election commentary, Dunham (2002) in Business Week argues that Karl Rove had tried with some success to make them swing voters. On the other hand, the Pew Research Center (2005) analyzes polling data and concludes that while Republicans have made gains generally among middle income voters, “the labor movement has done a very good job of getting those members it has to the polls, and keeping them loyal to the Democratic Party.”
We will test these hypotheses by examining candidate advertising behavior in the 2002 races for governor and U.S. Senate. Our advertising data are for the 2002 gubernatorial and U.S. Senate races and come from the Campaign Media Analysis Group (CMAG), which collects information on broadcast political advertisements at the individual ad level. These data were made available to us by the Wisconsin Advertising Project, and include information on the content and length of each ad, as well as the station on which it was aired and at which time.30,31 We use only ads from the 2002 elections, because this is the only non-Presidential election year for which the data are currently available.32 In years where there are Presidential elections, coattail effects may mean that the extent of clutter exceeds the bounds in conditions 15 and 16 that affect many of our comparative statics results.

Examination of candidate allocations suggests that candidates are almost always at an interior solution.33 To further test our model of probabilistic voting, we consider whether various market characteristics affect resource allocation in ways that are consistent with our comparative statics predictions. Our general empirical strategy is to

---

31 The advertising data were obtained from a project of the Wisconsin Advertising Project, under Professor Kenneth Goldstein and Joel Rivlin of the University of Wisconsin-Madison, and include media tracking data from the Campaign Media Analysis Group in Washington, D.C. The Wisconsin Advertising project was sponsored by a grant from The Pew Charitable Trusts. The opinions expressed in this article are those of the authors and do not necessarily reflect the views of the Wisconsin Advertising Project, Professor Goldstein, Joel Rivlin or The Pew Charitable Trusts.
32 CMAG data are not available from the Wisconsin Advertising Project for 1998 or before 1996, and the most recent year for which data are available is 2002.
33 In our sample of two-party races in states with at least two markets, there is only one case in which a candidate allocated no advertisements to a market. In the Alabama race for U.S. Senate, Parker --- who ultimately lost the race --- played 302 ads in Birmingham, 201 in Huntsville, and none in Mobile. His opponent Sessions, on the other hand, played 1809 ads in Birmingham, 1705 in Huntsville, and 1314 in Mobile. In fact, the allocations are surprisingly symmetric across markets. If we consider the pairs of markets faced by a candidate, in fewer than 8% of the cases do candidates play twice as many ads in one market in the pair as in the other. This relative symmetry is inconsistent with a model of deterministic voting but is consistent with the probabilistic voting model considered here (see Fletcher and Slutsky (2008b)).
examine the relation between the quantity of political advertising for a given candidate in the markets he faces and the relevant characteristics of those markets.

Our model was developed only for races with two candidates. For this reason, we consider only those advertisements that ran after the primary. Elections in Louisiana in 2002 were held by “nonpartisan blanket primary,” where at the first stage, all candidates of all parties ran in the same contest, then, if no candidate received more than 50% of the vote, the top two vote-getters ran against each other in the general election, even if they were in the same party. Because this means there are many candidates in each race, we drop Louisiana races from our analysis. We also drop the governor’s races in Arizona, New York, and Oklahoma, as there were three serious candidates in each of these races. Several other races had a third candidate who ran only a few ads or received 1% or less of the vote; these candidates are dropped but those races are otherwise included.

We drop the U.S. Senate races in Michigan and Illinois because CMAG lists advertisements by only one candidate. We also drop several races because the state has only one media market. Finally, we drop a few markets because they cross state lines; if only a small proportion of the market crosses state boundaries, we include the market for the state that has most of the market and drop the market for the state with only a small fraction of the market. This leaves us with 32 markets in 11 states for Senate candidates, and 53 markets in 14 states for gubernatorial candidates. As there are two candidates in each race and we consider the allocations of each candidate, this gives us 170 observations.

Based upon the testable implications in Section V, we will consider how the number of advertisements for a candidate in the markets he faces are affected by the
markets’ fractions Black, fractions evangelical Christian, voting age population, fractions belonging to a union, fractions that voted for Gore in 2000 and, in some specifications, the “competitive distances” of the races. Voting age population is the population aged 19 and older in the market. Data for this variable as well as fraction Black are from the 2000 Census of Population and Housing at the MSA level. Union membership data at the MSA level are from Hirsch and Macpherson (2003). We divide the number of union members by the population aged 19 and older for the fraction union.34

Religious affiliation data at the county level are available from Jones et al (2002), who survey churches in each county for the number of their adherents. As the data are survey-based and the definition of membership and adherence varies across denominations and congregations,35 substantial noise in this variable is likely. The Association of Religion Archives (TheARDA) (2006) categorizes denominations by type including Protestant evangelical and presents the Jones (2002) data at the MSA level. We derive fraction evangelical by dividing the number of Protestant evangelicals by the total population in the MSA as given by TheARDA.36

The fraction that voted for Gore in 2000 is included as a control for the level of Democratic party support. With this control, fraction Black and fraction evangelical are included to indicate the intensity of support (analogous to $\theta$ and $\gamma$, respectively, in our

34 Another measure of union presence would be to divide union members by the employed population. Our results are robust to this alternative measure. All signs in Table 3 are unchanged and significance levels change only slightly.
35 In fact, membership and adherence of traditionally Black churches were available for 1990 but not for 2000. The authors state that they stopped reporting Black church membership because those figures are only rough estimates and not very reliable.
36 We also construct fraction evangelical by aggregating the Jones et al (2002) county data to the MSA level and dividing by the total population in the MSA as provided by the U.S. Census Bureau from the 2002 Census of Population and Housing. With the exception of the single market of Greenville, whose fraction evangelical we calculated to be greater than one, the difference between the two measures is at most three decimal places out, and the correlation between them is 1.000. As a result, the signs and significance of our results are the same with either measure if Greenville is excluded as an outlier.
theoretical model), holding the number of partisans of each candidate (the submarket sizes in our theoretical model) constant. Finally, in some specifications we include a variable for the competitive distance in the 2000 Presidential race, which is equal to the absolute value of 0.5 minus the fraction of the market’s vote for Gore. Data for these two variables are from Leip (2006) at the county level, and we aggregate from the county to the MSA.

Historically, the South voted in very different ways from the rest of the country and as a consequence of this, many studies in the voting literature include a Southern dummy. The main reason given for the special nature of Southern voting is race, which seems to have trumped economic and other factors. Recently, however, there has been debate as to whether the South continues to be different or whether voting patterns there can now be explained by the same factors as elsewhere, with race no longer playing the same defining role. By including a Southern indicator variable interacted with the variables for market characteristics, we may be able to gain some insight into whether the South continues to be fundamentally different from the rest of the Union.

(Insert Table 2)

37 Note that the value of this variable will be small in close elections and will become larger as the race becomes less competitive.
38 We use the U.S. Census Bureau’s county and MSA codes, which can be found at http://www.census.gov/population/estimates/metro-city/99mfips.txt.
39 See, for example, Erikson and Palfrey (1998, 2000).
40 The traditional view of the predominance of race in Southern voting was initiated by the classic book by Key (1949). Knuckey (2005) argues that racial factors reasserted themselves in southern politics in the 1990’s after becoming not a significant predictor of voting in the period prior to that. The main recent challenge to this traditional view is Johnston and Shafer (2006) who argue that class and economic factors rather than race explain southern voting patterns post the civil rights era of the 60’s.
41 Recall that there is only one year (2002) of data available from a non-Presidential year. Due to the resulting small number of observations, we have not included other demographic variables such as income and educational attainment that might serve as controls. In any case, we would have no priors about the direction of the effect of such variables.
Table 2 shows summary statistics for levels of each of the variables used in our analysis for each market.\textsuperscript{42} We present summary statistics for the full sample, as well as separating the sample into Southern and non-Southern markets to gain some insight to possible differences between these areas. Some differences are immediately apparent in the Table. The mean size of the voting age population in the non-Southern states is 1.4 times that in the Southern states, but much of this is due to the large population in Los Angeles. The median values are 798,243 and 658,297, respectively. Both the mean and median fraction Black as well as the mean and median fraction evangelical Christian in the South are approximately twice that in the non-South. The mean and median fraction union in the non-South, on the other hand, are about 2.5 times that in the South.

b. Estimation

The theoretical model predicts how a given candidate will allocate resources across markets, given the different characteristics of those markets. We estimate the following ordinary least squares equation:

\[ (ADS)_i = \beta(X)_i + \alpha(SENATE)_i + \delta(SOUTH)_i + \nu(SOUTH*X)_i + \eta(STATE)_i + u_i \]  

where \((ADS)_i\) is the number of advertising occurrences shown in market i, and \(X_i\) is a vector of market characteristics, as discussed above and shown in Table 2. \(SENATE\) is a dummy variable equal to 1 if the race is for the U.S. Senate to account for possible spending differences between Senate and gubernatorial races. In order to test whether the Southern states are significantly different from the rest of the nation, we include a dummy variable \(SOUTH\), as well as a vector of interactions between the market

\textsuperscript{42} We assume that ads placed on behalf of a candidate by another group are perfect substitutes for advertising by the candidate, and thus include all advertising for a candidate, regardless of its source. However, our results are unchanged if only ads sponsored by the candidate’s campaign are included.
characteristics and the Southern dummy, SOUTH*X. Finally, STATE is a full set of state dummies, which is intended to control for differences in spending across states. Recall that the advertising variable is for the television broadcast market and the market characteristics are based on the MSA; these two areas do not perfectly overlap. As a result, there will be some measurement error for each market, which will persist for all candidates whose race includes that market. Standard errors are clustered by market to capture these correlations in the error term.

(Insert Table 3)

The results of our ordinary least squares regressions are shown in Table 3. The first column shows results for all candidates; the second is for Democrats only and the third is for Republicans only. Columns 4 – 6 are the same as 1 – 3 except that the variable for competitive distance is added.

The three hypotheses given in Section V are generally supported by our empirical results, although some of the estimates are not statistically significant. Given our small sample size, the results conform closely to the predictions of the model, yielding strong support for the probabilistic voting model.

Our first prediction is that any difference in a market characteristic will cause both candidates to devote more resources to the same market, so that the sign of the effect of each variable should be the same in the Democrat-only specifications (columns 2 and 4) as in the Republican-only specifications (columns 3 and 6). This is true for all variables except the fraction that voted for Gore in 2000, for which the model does not provide a sign prediction. This variable has a mixed sign across specifications and is not generally statistically significant.
The variables for voting age population, fraction belonging to a union, and competitive distance have the same sign across specifications, but are statistically significant in only a few of the specifications. Thus, the results for these variables offer weak support for hypothesis 1.

The effects of the fractions Black and evangelical are negative and statistically significant across all specifications, offering stronger support for hypothesis 1. These results also support hypotheses 2 and 3, that both candidates should devote fewer resources to markets with larger fractions of Black voters and evangelicals. The coefficients without the interactions reflect allocations in the non-Southern states. When the interaction with the southern dummy is included, the net effect of fraction Black in the South remains negative, but the overall effect of fraction evangelical switches sign, becoming positive. Thus, hypothesis 2 is supported in both the South and non-South, but hypothesis 3 is only supported in the non-South.

In addition to testing the model, our empirical analysis offers some insight to the effect of market size, given its possible relation to the price of advertising, to the way candidates view union members, and to differences between the Southern states and the rest of the country. Market size has a statistically insignificant effect in four of the six specifications without the interaction, and the effect in the non-South is small. A one standard deviation increase in voting age population (which more than doubles the population of the average-sized market) is associated with an increase of at most 129 advertising occurrences, a 6% increase from the mean number of ads. The small magnitude of this effect is consistent with the evidence that advertising prices do vary with population but at a significantly less than proportional rate. In the South, the effect
is larger: a one standard deviation increase in voting age population is associated with an increase of about 369 ads, or about 15% of the mean number of ads.

In the non-South, the fraction of the population belonging to a union has a negative effect in all specifications, but the effect is statistically significant in only half of the specifications. In the South, the difference is only significant in the specifications including competitive distance, but the effect of union membership becomes positive in those specifications. This mixed evidence may suggest that in the non-South, candidates do not view union members as swing voters, or that the resource effect dominates if they do. In the South, however, union members appear to be viewed more as swing voters than as a resource to Democrats.

There are several statistically significant differences between the South and the non-South. The most striking difference is in the effect of fraction evangelical. The difference is statistically significant and so large that the effect switches sign. Another large difference can be seen in the Senate variable. While Senate candidates allocate significantly more resources than gubernatorial candidates in the non-South, this is reversed in the south. There is also a statistically significant difference in the effect of voting age population, strengthening the positive effect of market size on the number of ads. There are significant differences in some specifications in the effects of fraction Black, fraction union and fraction that voted for Gore. These factors clearly indicate the existence of fundamental differences between the South and the rest of the country. There is some evidence that the treatment of race continues to be an important difference in how campaign resources are allocated. The most striking difference, though, is in how
evangelical Christians increase spending in the South. How to explain this in the context of the probabilistic voting model is an open question.

(VII) Conclusions

We have modified the probabilistic voting model usually specified in a spatial context to apply instead in a non-Downsian context where the candidates’ strategies are their campaign activities instead of their issue positions. As candidates rarely switch platform positions during a campaign, this modification is more realistic than the standard model, and has important empirical implications. Empirically, focusing on observable campaign activities such as the level of advertising in different media markets instead of on issue positions that are harder to observe or quantify opens up the model to tests of its validity. Looking at statewide races in the U.S. during the 2002 election cycle, we find that candidate behavior, at least in the non-South, seems to correspond in large part to the comparative statics predictions of the model. Probabilistic voting is not only an intuitively plausible approach; it is also consistent with campaign decisions.

In the model we have developed, a crucial assumption is that all individuals vote but that the candidates are uncertain about how an individual will vote even after observing some traits of the individual. An important and realistic consideration left out is whether individuals will vote or will abstain, a factor underlying many probabilistic voting models in the spatial context. Having advertising influence voter turnout, especially when candidates are uncertain as to who the voters will be, would be an important extension although one which would considerably complicate the analysis. Turnout considerations might help explain the significant differences we find between candidate behavior in the South and in the rest of the country. Historically, the South has been the region with the
lowest turnout of voters; while this is still true, the differential is decreasing with time (McDonald and Popkin, 2001). Whether the reason candidates spend more, instead of less, in media markets with intense partisans is to affect turnout is a question worth further exploration.
References


Dunham, R.S., 2002. Are Democrats losing their grip on the union hall?, Business Week, http://www.businessweek.com/magazine/content/02_48/c3810079.htm


Friedman, J., 1990. Game Theory with Applications to Economics. Cambridge, Massachusetts: Oxford University Press


Reported for 149 Religious Bodies, Nashville, TN: Glenmary Research Center: © 2002 Association of Statisticians of American Religious Bodies. (All rights reserved)


Patty, J., 2005. Generic difference of expected vote share and probability of victory maximization in simple plurality elections with probabilistic voters. Harvard University working paper


Figure 1.

\[
\begin{array}{c}
n_1 & n_2 & n_3 \\
-\theta_n + h(x_n) - h(y_n) & h(x_n) - h(y_n) & \gamma_n + h(x_n) - h(y_n) \\
m_1 & m_2 & m_3 \\
-\theta_m + h(x_m) - h(y_m) & h(x_m) - h(y_m) & \gamma_m + h(x_m) - h(y_m)
\end{array}
\]

Figure 2.

\[
V(I)
\]
h = h(½E_R, ½E_L) and θ_n = γ_n. If polarization is greater in community n because n_1 and n_2 are bigger than m_1 and m_3 while n_2 is smaller than m_3, then the candidates spend more in market n.
Table 1: Comparative statics predictions for parameters not entering h

<table>
<thead>
<tr>
<th>I</th>
<th>( \mu )</th>
<th>( \frac{\partial \phi}{\partial \mu} )</th>
<th>Sign of ( \frac{\partial x_n}{\partial \mu} ) and ( \frac{\partial y_n}{\partial \mu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( n_i = n_i^0 + \mu ), some i</td>
<td>( \phi h_x )</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>( n_i = \mu n_i^0 ), all i</td>
<td>( \phi_i )</td>
<td>+</td>
</tr>
<tr>
<td>III</td>
<td>( \mu = \theta_n ) or ( \mu = \gamma_n )</td>
<td>( \frac{\partial \phi}{\partial \theta_n} = n_1 h_x \phi'_1 ) ( \frac{\partial \phi}{\partial \gamma_n} = -n_3 h_x \phi' )</td>
<td>0 if ( \phi ) is uniform (-) if ( \phi ) is strictly single peaked</td>
</tr>
<tr>
<td>IV</td>
<td>( n_1 = n_1^0 + \mu/2 ), ( n_2 = n_2^0 - \mu ), ( n_3 = n_3^0 + \mu/2 )</td>
<td>( \frac{1}{2}(\phi_1 + \phi_3) - \phi_2 ) ( h_x )</td>
<td>0 if ( \phi ) is uniform (-) if ( \phi ) is strictly concave (a) (-) if (-\frac{1}{2}\theta_n + h &lt; 0) and ( \phi ) is strictly single peaked (b)</td>
</tr>
<tr>
<td>V</td>
<td>( n_1 = n_1^0 + \mu ), ( n_3 = n_3^0 - \mu )</td>
<td>( [\phi_1 - \phi_3] h_x )</td>
<td>0 if ( \phi ) is uniform () Sign (2h + \gamma_n - \theta_n) if ( \phi ) is strictly single peaked (c)</td>
</tr>
<tr>
<td>VI</td>
<td>( n_1 V(-\theta_n + h + k_1 \mu) + n_2 V(h + k_2 \mu) + n_3 V(\gamma_n + h + k_3 \mu) )</td>
<td>(-[k_1 n_1 \phi'_1 + k_2 n_2 \phi'_2 + k_3 n_3 \phi'_3] h_x )</td>
<td>0 if ( \phi ) is uniform (-) if ( k_1 = k_2 = 0 ) and ( \phi ) is strictly single peaked</td>
</tr>
</tbody>
</table>

Note that \( h \) and \( h_x \) are evaluated at \((\frac{1}{2}E_R, \frac{1}{2}E_L)\) and \( \phi_1 \equiv \phi(\theta_n - h), \phi_2 \equiv \phi(-h), \) and \( \phi_3 \equiv \phi(-\gamma_n - h) \).

(a) Let \( \alpha \equiv \min[\theta_n, \gamma_n] \). Strict concavity and single peakedness imply strict single peakedness. Then \( \phi_1 + \phi_3 = \phi(-\theta_n + h) + \phi(\gamma_n + h) < \phi(-\alpha + h) + \phi(\alpha + h) < 2\phi(h) \) where the last inequality follows from strict concavity. \( \frac{\partial V_i}{\partial \mu} < 0 \) then follows.

(b) If \(-\frac{1}{2}\theta_n + h < 0\) then \( h < \theta_n - h \) which implies that \( \phi(h) > \phi(\theta_n - h) \). From symmetry, \( \phi_2 \equiv \phi(-h) = \phi(h) > \phi(\theta_n - h) \equiv \phi_1 \). Since \( h < \gamma_n + h \), \( \phi_2 > \phi_3 \). These imply \( \phi_1 + \phi_3 < 2\phi_2 \) with \( \frac{\partial V_i}{\partial \mu} < 0 \).

(c) From symmetry and strict single peakedness, \( \phi_1 > \phi_3 \) iff \( \theta_n - h \) is closer to 0 than \( \gamma_n + h \), which in turn is equivalent to \( 2h + \gamma_n - \theta_n > 0 \).
Table 2: Summary statistics: Means and standard deviations for advertising and market characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>All states</th>
<th>Non-Southern states only</th>
<th>Southern states only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total number of ads shown in the market</td>
<td>2,336</td>
<td>2,215</td>
<td>2,444</td>
</tr>
<tr>
<td></td>
<td>(1,508)</td>
<td>(1,576)</td>
<td>(1,445)</td>
</tr>
<tr>
<td>Fraction Black</td>
<td>0.142</td>
<td>0.089</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td>(0.103)</td>
<td>(0.056)</td>
<td>(0.111)</td>
</tr>
<tr>
<td>Fraction evangelical Christians</td>
<td>0.198</td>
<td>0.124</td>
<td>0.264</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.082)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>Population aged 19 and older (thousands)</td>
<td>1,272</td>
<td>1,555</td>
<td>1,020</td>
</tr>
<tr>
<td></td>
<td>(1,703)</td>
<td>(2,206)</td>
<td>(1,028)</td>
</tr>
<tr>
<td>Fraction belonging to a union</td>
<td>0.064</td>
<td>0.096</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.041)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Absolute value of 0.5 - fraction Gore in 2000</td>
<td>0.094</td>
<td>0.07</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.052)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Fraction that voted for Gore in 2000</td>
<td>0.427</td>
<td>0.456</td>
<td>0.401</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.075)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>Senate dummy</td>
<td>0.376</td>
<td>0.275</td>
<td>0.467</td>
</tr>
<tr>
<td></td>
<td>(0.486)</td>
<td>(0.449)</td>
<td>(0.502)</td>
</tr>
<tr>
<td>Observations</td>
<td>170</td>
<td>80</td>
<td>90</td>
</tr>
</tbody>
</table>
Table 3: Determinants of campaign advertising levels

<table>
<thead>
<tr>
<th></th>
<th>(1) All candidates</th>
<th>(2) Democrats</th>
<th>(3) Republicans</th>
<th>(4) All candidates</th>
<th>(5) Democrats</th>
<th>(6) Republicans</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction Black</td>
<td>-3,676**</td>
<td>-3,795**</td>
<td>-3,557**</td>
<td>-3,204**</td>
<td>-3,454*</td>
<td>-2,953**</td>
</tr>
<tr>
<td></td>
<td>(1,408)</td>
<td>(1,897)</td>
<td>(1,414)</td>
<td>(1,322)</td>
<td>(1,849)</td>
<td>(1,318)</td>
</tr>
<tr>
<td>Fraction evangelical Christian</td>
<td>-6,954***</td>
<td>-7,314**</td>
<td>-6,593***</td>
<td>-5,733**</td>
<td>-6,433*</td>
<td>-5,032**</td>
</tr>
<tr>
<td></td>
<td>(2,398)</td>
<td>(3,527)</td>
<td>(2,144)</td>
<td>(2,522)</td>
<td>(3,635)</td>
<td>(2,276)</td>
</tr>
<tr>
<td>Voting age population</td>
<td>0.050**</td>
<td>0.076**</td>
<td>0.025</td>
<td>0.035</td>
<td>0.065</td>
<td>0.006</td>
</tr>
<tr>
<td>(in thousands)</td>
<td>(0.020)</td>
<td>(0.035)</td>
<td>(0.020)</td>
<td>(0.024)</td>
<td>(0.044)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Fraction belonging to a union</td>
<td>-2,710*</td>
<td>-4,172</td>
<td>-1,247</td>
<td>-4,074**</td>
<td>-5,156</td>
<td>-2,991**</td>
</tr>
<tr>
<td></td>
<td>(1,562)</td>
<td>(2,669)</td>
<td>(1,333)</td>
<td>(1,829)</td>
<td>(3,282)</td>
<td>(1,465)</td>
</tr>
<tr>
<td>Competitive distance</td>
<td></td>
<td></td>
<td></td>
<td>-2,427</td>
<td>-1,751</td>
<td>-3,102**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1,609)</td>
<td>(2,526)</td>
<td>(1,441)</td>
</tr>
<tr>
<td>Fraction that voted for Gore in 2000</td>
<td>444</td>
<td>1,988</td>
<td>-1,100</td>
<td>307</td>
<td>1,889</td>
<td>-1,276</td>
</tr>
<tr>
<td></td>
<td>(1,000)</td>
<td>(1,503)</td>
<td>(907)</td>
<td>(1,056)</td>
<td>(1,615)</td>
<td>(883)</td>
</tr>
<tr>
<td>Senate</td>
<td>2,103**</td>
<td>2,542**</td>
<td>1,664*</td>
<td>2,103**</td>
<td>2,542**</td>
<td>1,664*</td>
</tr>
<tr>
<td></td>
<td>(913)</td>
<td>(1,180)</td>
<td>(880)</td>
<td>(920)</td>
<td>(1,202)</td>
<td>(897)</td>
</tr>
<tr>
<td>South</td>
<td>2,156</td>
<td>404</td>
<td>-560*</td>
<td>3,533**</td>
<td>1,221</td>
<td>1,092</td>
</tr>
<tr>
<td></td>
<td>(1,606)</td>
<td>(1,583)</td>
<td>(1,249)</td>
<td>(1,630)</td>
<td>(1,686)</td>
<td>(1,391)</td>
</tr>
<tr>
<td>South*fraction Black</td>
<td>2,864*</td>
<td>2,946</td>
<td>2,783***</td>
<td>1,895</td>
<td>2,273</td>
<td>1,517</td>
</tr>
<tr>
<td></td>
<td>(1,576)</td>
<td>(2,122)</td>
<td>(1,608)</td>
<td>(1,489)</td>
<td>(2,088)</td>
<td>(1,497)</td>
</tr>
<tr>
<td>South*fraction evangelical</td>
<td>8,593***</td>
<td>9,469**</td>
<td>7,717***</td>
<td>7,550***</td>
<td>8,708**</td>
<td>6,393**</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------</td>
<td>--------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td>(2,644)</td>
<td>(3,847)</td>
<td>(2,445)</td>
<td>(2,714)</td>
<td>(3,920)</td>
<td>(2,519)</td>
</tr>
<tr>
<td>South*voting age pop (thousands)</td>
<td>0.246***</td>
<td>0.221**</td>
<td>0.272***</td>
<td>0.263***</td>
<td>0.233**</td>
<td>0.293***</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.087)</td>
<td>(0.073)</td>
<td>(0.058)</td>
<td>(0.094)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>South*fraction union</td>
<td>10,138</td>
<td>12,249</td>
<td>8,027</td>
<td>15,039**</td>
<td>15,597*</td>
<td>14,481**</td>
</tr>
<tr>
<td></td>
<td>(6,410)</td>
<td>(7,800)</td>
<td>(7,695)</td>
<td>(5,873)</td>
<td>(7,906)</td>
<td>(6,554)</td>
</tr>
<tr>
<td>South*fraction Gore</td>
<td>1,168</td>
<td>649</td>
<td>1,687***</td>
<td>-1,006</td>
<td>-797</td>
<td>-1,215</td>
</tr>
<tr>
<td></td>
<td>(1,451)</td>
<td>(2,021)</td>
<td>(1,640)</td>
<td>(1,502)</td>
<td>(2,131)</td>
<td>(1,810)</td>
</tr>
<tr>
<td>South*Senate</td>
<td>-3,877***</td>
<td>-5,182***</td>
<td>-2,571</td>
<td>-3,877***</td>
<td>-5,182***</td>
<td>-2,571**</td>
</tr>
<tr>
<td></td>
<td>(967)</td>
<td>(1,250)</td>
<td>(959)</td>
<td>(974)</td>
<td>(1,274)</td>
<td>(976)</td>
</tr>
<tr>
<td>South*competitive distance</td>
<td>-3,156</td>
<td>-1,981</td>
<td>-4,331</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2,222)</td>
<td>(3,017)</td>
<td>(2,657)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1,053</td>
<td>3,496***</td>
<td>3,079***</td>
<td>1,058</td>
<td>3,603***</td>
<td>3,268***</td>
</tr>
<tr>
<td></td>
<td>(1,325)</td>
<td>(1,039)</td>
<td>(670)</td>
<td>(1,341)</td>
<td>(1,174)</td>
<td>(644)</td>
</tr>
<tr>
<td>R squared</td>
<td>0.5488</td>
<td>0.8787</td>
<td>0.6472</td>
<td>0.5552</td>
<td>0.8806</td>
<td>0.6741</td>
</tr>
<tr>
<td>Observations</td>
<td>170</td>
<td>85</td>
<td>85</td>
<td>170</td>
<td>85</td>
<td>85</td>
</tr>
</tbody>
</table>

Robust standard errors are in parentheses, and are clustered by market. A full set of state dummies is also included. The unit of observation is the market. Columns 1 and 4 include observations for both Democratic and Republican candidates, while columns 2 and 5 include only Democrats and columns 3 and 6 include only Republicans. The Southern states are defined as those states that were in the Confederacy.

Voting age population and fraction Black are from the 2000 Census of Population and Housing. Union membership data are from Hirsch and Macpherson (2003). Religious affiliation at the county level is from Jones et al (2002). Competitive distance is equal to the absolute value of (0.5 – the fraction of the market’s vote for Gore). Data for this variable and the fraction that voted for Gore in 2000 were provided by David Leip at the county level. We aggregate county-level variables to the MSA level.

*** Statistically significant at the 99% level
** Statistically significant at the 95% level
* Statistically significant at the 90% level
Appendix: Proofs

Proof of Lemma 1: From (8), $0 < V < 1$ holds in every submarket over the entire range of $x_n$ and $y_n$. Thus, if $\hat{V}(x_n, x_m, y_n, y_m)$ is locally concave in the $x_i$ and locally convex in the $y_i$ at every feasible value, then it globally satisfies the asserted properties. As given in (7), $\hat{V}$ is the sum of six terms. Consider the term $V(-\theta_n + h(x_n, y_n))$ with the other terms following similarly. Given (1), (2) and (3), for all $x_n$ or $y_n$, $V$ is differentiable with:

$$\frac{\partial V}{\partial x_n} = \varphi(\theta_n - h(x_n, y_n))(h_x)$$

$$\frac{\partial^2 V}{\partial x_n^2} = -\varphi'(\theta_n - h(x_n, y_n))(h_x)^2 + \varphi(\theta_n - h(x_n, y_n))(h_{xx})$$

$$\frac{\partial V}{\partial y_n} = \varphi(\theta_n - h(x_n, y_n))(h_y)$$

$$\frac{\partial^2 V}{\partial y_n^2} = -\varphi'(\theta_n - h(x_n, y_n))(h_y)^2 + \varphi(\theta_n - h(x_n, y_n))(h_{yy})$$

By (3), $h_{xx} < 0$ and $h_{yy} > 0$. Given the weak single peakedness of $\varphi$ in (1), $\varphi'$ is nonpositive if $-I = \theta_n - h(x_n, y_n) \geq 0$ and is nonnegative if $-I \leq 0$. Then, from (1) and (3), $\frac{\partial^2 V}{\partial x_n^2} < 0$ when $-I < 0$ and $\frac{\partial^2 V}{\partial y_n^2} > 0$ when $-I > 0$. From (9), when $-I > 0$, then $\frac{\partial^2 V}{\partial x_n^2} < 0$ and when $-I < 0$, then $\frac{\partial^2 V}{\partial y_n^2} > 0$ even though the $\varphi'$ term has the opposite sign in each second derivative. Hence, $V(-\theta_n + h(x_n, y_n))$ is strictly concave in $x_n$ and strictly convex in $y_n$. The overall objective function, as the weighted sum of strictly concave functions, is strictly concave in $x_n$ and $x_m$ and, as the weighted sum of strictly convex functions, is strictly convex in $y_n$ and $y_m$.

Q.E.D.

Proof of Theorem 1: The strategy spaces in (5) are compact and convex. The objective function in (7) is defined for all feasible $x_n$, $x_m$, $y_n$ and $y_m$ and is concave in $x_n$ and $x_m$ and convex in $y_n$ and $y_m$ from Lemma 1. Existence follows from the standard Nash existence theorem.
To show uniqueness, use (5) to eliminate \( x_m \) and \( y_m \) from the objective function (7) which becomes \( \hat{V} (x_n, E_R - x_n, y_n, E_L - y_n) \). The first order conditions which determine the best replies are \( \hat{V}_1 - \hat{V}_2 = 0 \) and \( -\hat{V}_3 + \hat{V}_4 = 0 \). The Jacobian of this system is:

\[
J(x_n, y_n) = \begin{bmatrix}
V_{11} + V_{12}, & V_{13} + V_{14} \\
-V_{21} - V_{22}, & -V_{33} - V_{44}
\end{bmatrix}
\]

This follows from the additive structure of \( \hat{V} (x_n, x_m, y_n, y_m) \) for which

\[
\hat{V}_{12} = \hat{V}_{14} = \hat{V}_{21} = \hat{V}_{32} = \hat{V}_{34} = \hat{V}_{41} = \hat{V}_{43} = 0 .
\]

Then:

\[
J(x_n, y_n) + J(x_n, y_n)^T = \begin{bmatrix}
2(V_{11} + V_{22}) & 0 \\
0 & -2(V_{33} + V_{44})
\end{bmatrix}
\]

Since \( \hat{V} \) is strictly concave in \( x_n \) and strictly convex in \( y_n \) as shown above, the diagonals are negative and this matrix is negative semidefinite. Therefore, \( J(x_n, y_n) \) is negative quasidefinite. All the conditions of the Rosen uniqueness theorem as given in Friedman (1990, page 86) are satisfied and the equilibrium is unique.

Q.E.D.

Proof of Theorem 2: Consider the conditions for \( x_n \) or \( y_n \) to equal 0 with similar results for \( x_m \) or \( y_m \). The Kuhn-Tucker corner conditions are \( x_n = 0 \) if \( \hat{V}_1 - \hat{V}_2 \leq 0 \) and \( y_n = 0 \) if \( -\hat{V}_3 + \hat{V}_4 \leq 0 \). These reduce to

\[
x_n = 0 \text{ if } g'(0)(\partial \hat{V} / \partial T_n) - g'(E_R) (\partial \hat{V} / \partial T_m) \leq 0
\]

\[
y_n = 0 \text{ if } g'(0)(\partial \hat{V} / \partial T_n) - g'(E_L) (\partial \hat{V} / \partial T_m) \leq 0
\]

where the derivatives with respect to \( T_i \) are taken around \( T_i = 0 \). Since \( \partial \hat{V} / \partial T_n = \)
n_1V'(-\theta_n + g(x_n) - g(y_n)) + n_2V'(g(x_n) - g(y_n)) + n_3V'(\gamma_n + g(x_n) - g(y_n)), then \partial \hat{V} / \partial T_n > 0
and similarly \partial \hat{V} / \partial T_m > 0. Assume that x_n = 0 and y_n > 0 at an equilibrium. Then g'(0) \partial \hat{V} / \partial T_n \leq g'(E_R) \partial \hat{V} / \partial T_m and g'(y_n) \partial \hat{V} / \partial T_n = g'(E_L - y_n) \partial \hat{V} / \partial T_m. For these both to hold, (g'(0) - g'(y_n)) \partial \hat{V} / \partial T_n \leq (g'(E_R) - g'(E_L - y_n)) \partial \hat{V} / \partial T_m. Since g'(0) > g'(y_n), but g'(E_R) < g'(E_L - y_n), this is inconsistent. Hence, at an equilibrium, if x_n = 0, then y_n = 0 must also hold. Thus, y_n = 0 and x_m > 0 must hold as required. (A) follows immediately since the corner conditions cannot hold if g'(0) = \infty while (B) follows from the corner condition for y_n given g'(0) < g'(E_L). For (C), (g'(0) - g'(x)) (\partial \hat{V} / \partial T_i) \leq (g'(E_L) - g'(E_R - x_i)) (\partial \hat{V} / \partial T_i) must hold where y_n = 0 and x_n > 0. The right hand side must be positive so E_L < E_R - x_n must hold.

Q.E.D.

**Proof of Theorem 3:** Given the specification of h,

\[ \hat{V}_1 = Ah_x, \hat{V}_3 = Ah_y, \hat{V}_{11} = A'(h_x)^2 + Ah_{xx}, \]

\[ \hat{V}_{13} = A'h_x h_y + Ah_{xy}, \text{ and } \hat{V}_{33} = A'(h_y)^2 + Ah_{yy} \] where

\[ A = n_1\phi(\theta_n - h(x_n, y_n)) + n_2\phi(- h(x_n, y_n)) + n_3\phi(-\gamma_n - h(x_n, y_n)) \text{ and} \]

\[ A' = -n_1\phi'(\theta_n - h(x_n, y_n)) + n_2\phi'(- h(x_n, y_n)) + n_3\phi'(-\gamma_n - h(x_n, y_n)). \] Then

\[ \partial \hat{V}_1 / \partial \mu = (\partial A / \partial \mu) h_x \text{ and } \partial \hat{V}_3 / \partial \mu = (\partial A / \partial \mu) h_y . \] Substituting these into (13) and (14) yields:

\[ \text{sign}(\partial x_n / \partial \mu) = \text{sign}[A(h_{yy} - (h_y / h_x) h_{xy}) \partial \hat{V}_1 / \partial \mu] \]

\[ \text{sign}(\partial y_n / \partial \mu) = \text{sign}[A((h_y / h_x) h_{xx} - h_{xy}) \partial \hat{V}_1 / \partial \mu] \]
Since $A > 0$, condition (15) ensures that the expressions multiplying $\partial \hat{V}_1 / \partial \mu$ in both conditions are positive. Hence, the signs equal the sign of $\partial \hat{V}_1 / \partial \mu$ as asserted.

Q.E.D.

Proof of Theorem 4: (A) Under the assumptions given in the Theorem, the sign of $\partial x_n/\partial \mu_R$ is given in (13) where $(\partial \hat{V}_1 / \partial \mu_R) = n_3 [-\varphi'_3(h_x)^2 + \varphi_3 h_{xx}]$, $(\partial \hat{V}_3 / \partial \mu_R) = n_3 [-\varphi'_3 h_x h_y + \varphi_3 h_{xy}]$, $\hat{V}_{33} = A'(h_y)^2 + A h_{yy}$, and $\hat{V}_{13} = A' h_y h_x + A h_{yx}$, for $A$ and $A'$ defined in the proof of Theorem 3. Substituting into (13) and manipulating yields sign

$(\partial x_n/\partial \mu_R) = \text{sign} \left[ A(h_{xy} - (h_x/h_y)h_{yy})/(h_x h_y) - (h_{xy} - (h_y/h_x)h_{xx})/(h_x h_y) \right].$

If (16) is assumed, then $(h_{xy}/(h_x h_y) - \varphi'_3/\varphi_3) < 0$ and since (15) is then implied,

$(h_{xy} - (h_x/h_y)h_{yy}) > 0$ and $(h_{xy} - (h_y/h_x)h_{xx}) < 0$. From (9), $A' + (h_{yy}/(h_y)^2)A > 0$.

Hence, sign $(\partial x_n/\partial \mu_R) < 0$ in this case as asserted. If $\varphi$ is uniform, then $\varphi'_3 = A' = 0$ and

sign $(\partial x_n/\partial \mu_R) = \text{sign} \left[ A((h_{xy})^2 - h_{xx} h_{yy})/(h_x h_y) \right]$. Since $h_{xx} < 0, h_{yy} > 0, h_x > 0$ and $h_y < 0$, then sign $(\partial x_n/\partial \mu_R) < 0$ in this case.

Using (14) and $\partial \hat{V}_3 / \partial \mu_L = n_1 (-\varphi'_3 h_y^2 + \varphi_3 h_{yy})$, $\partial \hat{V}_1 / \partial \mu_L = n_1 (-\varphi'_3 h_x h_y + \varphi_3 h_{xy})$, $\hat{V}_{11} = A'(h_x)^2 + A h_{xx}$, and $\hat{V}_{13} = A' h_x h_y + A h_{xy}$ yields sign $(\partial y_n/\partial \mu_R) = \text{sign} \left[ A(h_{xy} - (h_y/h_x)h_{xx})/(h_x h_y) - (h_{xy} - (h_x/h_y)h_{yy})/(h_x h_y) \right].$
From (16), \( h_{xy}/(h_x h_y) - (\varphi''_1/\varphi_1) > 0 \), from (15), \( h_{xy} - (h_y/h_x)h_{xx} < 0 \) and \\
h_{xy} - (h_x/h_y)h_{yy} > 0 \), and from (9), \( A' + (h_{xx}/(h_x)^2)A < 0 \) making \( \partial y_n/\partial \mu_R < 0 \). If instead, \( \varphi'_1 = A' = 0 \) under the uniform distribution of \( \varphi \), then again \( \partial y_n/\partial \mu_R < 0 \).

\( \text{(B)} \) Substituting \( \partial \hat{V}_1 / \partial \mu_L \) and \( \partial \hat{V}_3 / \partial \mu_L \) into (13) yields \( \partial x_n/\partial \mu_L = \)
\[
\text{sign} \left[ (h_{xy} - (h_x/h_y)h_{yy})(A'\varphi_1 + A'\varphi'_1) \right].
\]
From (15), \( h_{xy} - (h_x/h_y)h_{yy} > 0 \) so sign \\
(\( \partial x_n/\partial \mu_L \)) = sign(A'\varphi_1 + A'\varphi'_1). Substituting for \( A' \) and \( A \) gives sign (\( \partial x_n/\partial \mu_L \)) = sign \\
\[
\left[ (n_2\varphi_2)((\varphi'_1/\varphi_1) - (\varphi'_2/\varphi_2)) + (n_3\varphi_3)((\varphi'_1/\varphi_1) - (\varphi'_3/\varphi_3)) \right].
\]
Under the uniform distribution, \( \varphi'_i = 0 \), all \( i \), so \( \partial x_n/\partial \mu_L = 0 \). Given (6), strict single peakedness implies that \\
\( \varphi'_1 = \varphi'(\theta_n - h(\frac{1}{2}E_R, \frac{1}{2}E_L)) < 0 \), that \( \varphi'_2 \geq 0 \) with strict inequality if \( E_R > E_L \), and that \( \varphi'_3 \)
\[= \varphi'(-\gamma_n - h(\frac{1}{2}E_R, \frac{1}{2}E_L)) > 0. \]
Then \( \partial x_n/\partial \mu_L < 0 \) follows.

\( \text{(C)} \) Analogous substitutions into (14) yield the sign (\( \partial y_n/\partial \mu_R \)) = 
\[
\text{sign} \left[ n_1\varphi_1((\varphi'_1/\varphi_1) - (\varphi'_3/\varphi_3)) + n_2\varphi_2((\varphi'_2/\varphi_2) - (\varphi'_3/\varphi_3)) \right].
\]
Again, for \( \varphi \) uniform, the right hand side is 0. If \( \varphi \) is strictly single peaked, then \\
\( \varphi'_1 < 0 \) and \( \varphi'_2 > 0 \). Hence, all terms are negative except for \( \varphi'_2/\varphi_2 \). If \( E_L = E_R \), \( \varphi'_2 \) is taken at the peak of the error distribution and must equal 0, making the right hand side negative. If \( E_L < E_R \), then \( \varphi'_2 > 0 \). The right hand side is still negative if \\
\( \varphi'_1/\varphi_3 > \varphi'_2/\varphi_2 \). For the quadratic distribution, \( \varphi'(\varepsilon)/\varphi(\varepsilon) = -2\varepsilon/(c^2 - \varepsilon^2) \) which for \\
\( -c < \varepsilon < 0 \) is decreasing in \( \varepsilon \) so this holds. For the normal distribution \\
\( \varphi'(\varepsilon)/\varphi(\varepsilon) = -\varepsilon/\sigma^2 \) which again increases in \( \varepsilon \) as required. However, there can exist \\
single peaked distributions for which \( \varphi'_2/\varphi_2 > \varphi'_3/\varphi_3 \) provided \( \gamma_n \) and \( E_R - E_L \) are
sufficiently large and that \( \varphi' \) becomes small for \( \varepsilon \) far from 0. The two terms on the right hand side are then of opposite sign with the positive term dominating for \( n_2 \) sufficiently bigger than \( n_1 \) and being dominated for \( n_2 \) sufficiently smaller than \( n_1 \).

**Q.E.D.**

**Proof of Theorem 5:** Solving (10) yields:

\[
\frac{\partial x_n}{\partial \mu_n} = \frac{P_{33}(\partial P_1/\partial \mu_n - P_{13}(\partial P_2/\partial \mu_n))}{2[-P_{11}P_{22} + (P_{12})^2]} \quad \text{where} \quad \frac{\partial P_1}{\partial \mu_n} = -A(h_x + h_{xx}x_n + h_xy y_n) - A'(h_x h_{xy}y_n + h_{xx} x_n) - A'(h_x y_n) + (h_y)^2 y_n,
\]

and \( \hat{\nu}_{33}, \hat{\nu}_{11}, \hat{\nu}_{13}, A, \) and \( A' \) are defined in the proof of Theorem 3. Substituting these and simplifying yields:

\[
\frac{\partial x_n}{\partial \mu_n} = \frac{1}{2} x_n + \frac{1}{2}(h_x h_{xy} - h_x h_{yy}) / (-\hat{\nu}_{11} \hat{\nu}_{33} + (\hat{\nu}_{13})^2) \quad \text{since} \quad \hat{\nu}_{11} < 0 \quad \text{and} \quad \hat{\nu}_{33} > 0, \quad \text{the denominator is positive. From (15),} \quad h_x h_{xy} - h_x h_{yy} < 0 \quad \text{showing} \quad \frac{\partial x_n}{\partial \mu_n} - \frac{1}{2} x_n < 0. \quad \text{Since} \quad \frac{\partial (x_n/\mu_n)}{\partial \mu_n} = \frac{\partial x_n/\partial \mu_n}{x_n} \quad \text{at} \quad \mu_n = 1, \text{then} \quad \frac{\partial (x_n/\mu_n)}{\partial \mu_n} + \frac{1}{2} x_n \quad \text{as required for (i).} \quad \text{A similar analysis yields} \quad \frac{\partial y_n}{\partial \mu_n} = \frac{1}{2} y_n + \frac{1}{2}(h_x h_{xy} - h_y h_{xx}) / (-\hat{\nu}_{11} \hat{\nu}_{33} + (\hat{\nu}_{13})^2) = \frac{\partial (y_n/\mu_n)}{\partial \mu_n} + y_n. \quad \text{From (15),} \quad h_x h_{xy} - h_y h_{xx} < 0 \quad \text{yielding (ii).}
\]

**Q.E.D.**

**Proof of Theorem 6:** The proof follows identically to that in Theorem 5 except that \( \nu_2 = Ah_x(x_n/\mu, y_n/\mu) \) and \( \nu_3 = Ah_y(x_n/\mu, y_n/\mu) \).

**Q.E.D.**