Campaign Effects with Ambiguity-Averse Voters*

Scott Ashworth†

March 14, 2005
PRELIMINARY AND INCOMPLETE

Abstract

I study a model of campaign strategy in which the voters may be ambiguity averse. The model of ambiguity aversion is based on the extensive empirical literatures in cognitive psychology and voter behavior. Since voters ambiguity can lead voters to abstain, the parties face a tradeoff between mobilization of ambiguity averse independents and persuasion of already committed partisans of the other party. I find that mobilization is a more attractive strategy than is persuasion, and that races for less visible offices will be more oriented towards "base politics". I also consider the possibility of negative campaigns aimed at demobilization, and derive some comparative statics about the importance of negative campaigns. I relate these theoretical result to the empirical literature on campaign effects and the "minimal effects" hypothesis.

*I received helpful comments from participants at Princeton's CSDP lunch, especially Larry Bartels.

†Department of Politics, Princeton University. Email: `sashwort@princeton.edu`

1 Introduction

A lively debate in the study of voter behavior asks what effects electoral campaigns have on voter decisions about participation and whom to vote for. Most of this literature concludes that campaigns have "minimal effects". In particular, a robust finding of the survey-based empirical literature is that campaigns are ineffective at persuasion. But a complete account of campaign effects must consider that candidates make predictions about how different groups of voters would respond to a campaign, and then target those groups they expect will respond in the most favorable way. For example, the relative importance of mobilization and persuasion in campaigns depends not only on the psychological attributes of the voters, but also on the candidates' decisions about which strategy to emphasize. This paper attempts to develop a theoretical understanding of this process. The model suggests that we should observe limited persuasion not because voters are necessarily difficult to persuade, but because a focus on persuasion is a bad strategy for candidates.

One of the themes emphasized in the campaign effects literature is the role of information in shaping voter choices. The standard approach to handling informational issues in formal political theory is based on *Bayesian rationality*: Voters understand their environment, maximize a well defined expected utility, and update their beliefs with Bayes's rule.[1] The implications of Bayesian updating for voter behavior were worked out by Zechman (1979) and Calvert (1986). Perhaps surprisingly, this model has had some success explaining data on voter behavior and public opinion (Achen, 1992; Alvarez, 1997; Bartels, 1988; Gerber and Green, 1998). However, the literature on campaign effects emphasizes some important regularities the Bayesian model cannot capture.

For example, the classic study of the 1940 Presidential campaign by Lazarsfeld, Berelson and Gaudet (1948) found that:

> [w]hat the political campaign did, so to speak, was not to form new opinions but to raise old opinions over the thresholds of awareness and decision. Political campaigns are important primarily because they *activate* latent predispositions.

It is quite difficult to get such an effect in a model with Bayesian voters—such voters do not have "thresholds of awareness and decision".

[1] See Gilboa, Postlewaite and Schmeidler (2004) for more discussion of the strengths and weaknesses of the Bayesian model.

Similarly, Bartels (1988) finds that low information respondents are less likely to place candidates on issue or quality scales, and that respondents who refuse to answer such questions are less likely to participate in politics. Furthermore, Bartels finds that one of the main things that happens during the campaign is that voters learn more about the candidates, and consequently become more likely to turnout.[2] Bayesian voters, however, do not recognize differences in the "quality" of beliefs about events. For a Bayesian, any assessment that one candidate is better than the other with probability p is equal to any other, even if one is based on detailed information and the other is pulled from the air.

This paper presents a model of campaign strategy and voter behavior when voters are *ambiguity averse*—they distinguish between different qualities of information. As emphasized by Bartels, voters with low quality information are less likely to participate than are voters with better information. Candidates can target informative campaign messages to various groups of voters, where the groups differ in their receptivity to information, their partisanship, and their degree of ambiguity aversion. The introduction of some ambiguity-averse voters can have large impact on the parties' equilibrium campaign strategies. In particular, when parties can target their campaigns to specific groups of voters, the introduction of ambiguity-aversion leads the parties to target more extreme voters than in the standard Dixit-Londregan-Lindbeck-Weibull "swing voter" model. Furthermore, the parties prefer to focus on mobilization at the expense of persuasion.

Inspired by its close fit to traditional ideas in the political science literature on voter behavior, I model ambiguity aversion by endowing voters with incomplete preferences, à la Bewley (2002). The voter's beliefs are represented by a *set* of probabilities, and she prefers action a to action b if and only if her expected utility of a is greater than of b for *all* of her priors. If two priors disagree, then the voter cannot rank the two actions. The model is closed with an inertia assumption—there is a status quo action that the voter takes unless some alternative dominates it. Based on the evidence from Bartels (1988), we assume that the status quo is abstention.

The inability of voters to rank some candidates is the model's departure from standard models. After all, one of the foundational assumptions of the standard model is that voters have complete preferences—they can always make a choice between any two alternatives. (Although indifference is allowed.) But many poll or survey respondents "don't know" whom they

[2]Ansolabehere and Iyengar (1995) and Jamieson (2001) also find that more campaign information and exposure to advertising increases the chances someone votes.

would vote for in an immediate election. In an experiment, Brady and Ansolabehere (1989) found that $\approx 20\%$ of subjects had incomplete preferences over the candidates in the 1976 and 1984 Democratic primaries. (This experiment is discussed in more detail in the next section.) These empirical observations motivate the particular form of ambiguity aversion we study in this paper.

The most closely related paper to the voter behavior section of this paper is Ghirardato and Katz (2002). They study a citizen who can costlessly vote in a two alternative election. They show that a Bayesian voter generically has a strict preference to vote rather than abstain, but that an ambiguity averse voter may strictly prefer abstention for a generic set of beliefs. Like this paper, they work with a multi-prior model of ambiguity aversion, but they use Gilboa and Schmeidler's (1989) maximin expected utility, which is based on complete preferences. In their framework, just showing that abstention may be optimal is already an impressive feat, and they do not go on to discuss the dynamics of a campaign or the problem of integrating their voters with strategic candidates.

The rest of the paper is organized as follows. Section 2 describes some of the empirical background from cognitive psychology and voter behavior in more detail, and section 3 outlines a simple model of voters that embodies these empirical considerations. Section 3 also discusses the model's background in axiomatic decision theory. Section 4 derives some comparative statics of voter behavior with respect to exogenous variation in campaign information. Section 5 considers a full equilibrium model of the campaign, and section 6 concludes.

2 Empirical Evidence on Ambiguity and Incompleteness

The "psychology" of the voters I study is captured well by Richard Posner (2003), discussing his own decision to not vote in the 2000 presidential election:

> People have a pretty good idea of their own interests, or at least a better idea than officials do. But often they have a poor idea of how those interests will be affected by the forthcoming election. That was my own situation in regard to the 2000 election, and I am better informed about political matters than the average American. I did not have a clear sense of which candidate was on

> balance likely to deliver more of the things that I seek from the federal government, and so I didn't bother to vote. [pp. 168–9]

2.1 Evidence for Ambiguity Aversion

Posner refrained from voting because he knew that his information about the candidates was too poor to let him decide which he preferred. This kind of uncertainty is different than that usually included in formal models—it is *ambiguity*—"uncertainty about probability, created by missing information that is relevant and could be known" (Frisch and Baron, 1988). In contrast to *risk*, ambiguity cannot be represented by a single probability measure.

The seminal work showing that some individuals dislike betting on ambiguous information was Daniell Ellsberg's (1961) urn experiments. He showed (in hypothetical choices) that people preferred to bet on a draw from a urn with known proportions than from one with unknown proportions. Here is the simplest version of Ellsberg's original experiment. The subject is offered the chance to draw a ball from one of two urns. She is told that the first urn contains exactly 50 red and 50 black balls. The second urn also contains 100 balls, but the subject is not told what fraction are red and what fraction are black. (She does know that every ball is either red or black.) In one stage of the experiment, the subject wins $10 if she draws a red ball. In the second stage, she wins $10 if she draws a black ball. Most subjects strictly prefer to draw from the first urn in both treatments. This is inconsistent with probabilistic beliefs (assuming more money is preferred to less). A strict preference for urn 1 in the first treatment would imply that the subject's subjective expectation of the number of red balls in the second urn is less than 50, while a strict preference for urn 1 in the second treatment would imply the subjective expectation is greater than 50. But both of these cannot be true if the subject's beliefs about the composition of the second urn are probabilistic.

Ellsberg's result has been replicated many times, often with real payoffs. (See Camerer and Weber (1992) for a review of this literature.) Slovic and Tversky (1974) show that explaining the normative arguments against ambiguity aversion does not reduce Ellsberg outcomes, and Curley, Yates and Abrams (1986) show that people are strictly ambiguity averse—the results cannot be explained by assuming that subjects are indifferent.

Most of the early studies on ambiguity were based on experiments with physical randomization devices, like drawing balls from an urn. But aversion to ambiguity has also been found in studies based on real-world uncertainty. Heath and Tversky (1991) find that subjects aversion to ambiguous bets is greater the less informed the subject is about the subject matter. They call this the "competence hypothesis". Fox and Tversky (1995) find that subjects are more averse to ambiguous bets when they know that more informed people are making similar choices. They call this the "comparative ignorance hypothesis". Both the competence hypothesis and the comparative ignorance hypothesis imply that voters will be more ambiguity averse in elections for less visible offices.

Several important studies of voter behavior suggest that ambiguity may be an important attitude for some voters. For example, the aversion to acting on imprecise information is similar to the finding of Bartels (1988) that low information respondents are less likely to place candidates on issue or quality scales. This suggests that ambiguity may be electorally important, since Bartels also finds that respondents who refuse to answer such questions are less likely to participate in politics. Furthermore, Bartels finds that one of the main things that happens during the campaign is that voters learn more about the candidates, and consequently become more likely to turnout. He says "Voters do not cast their ballots for candidates they do not feel that they know, at least superficially" (p. 57).

2.2 Evidence for Incomplete Preferences

In the model I develop below, voters who face ambiguity are unable to rank some of the candidates. This is based on evidence like that in Bartels, as well as other sources. In an experiment, Brady and Ansolabehere (1989) find that $\approx 20\%$ of subjects had incomplete preferences over the candidates in the 1976 and 1984 Democratic Presidential primaries. This cannot be explained simply as some people not thinking about politics much, because roughly 20% had no preference over a and b or over b and c, but could rank a and c. In the experiment, less informed voters were more likely to have incomplete preferences.

3 The Voter

Consider an election with two candidates, L and R. The voter's payoff to voting for L is normalized to 0, while her payoff to voting for R is $\theta \in \mathbb{R}$. Her payoff from abstaining is between these two payoffs; for concreteness, we take it to be $\delta\theta$, where $0 < \delta < 1$. These payoffs are independent across voters. If the voter knew θ, then she would vote for R if $\theta > 0$, would vote for L if $\theta < 0$, and would be indifferent over all three options if $\theta = 0$.

The voter's motivation to act this way can be interpreted quite broadly. It may reflect the voter's personal evaluation of the candidates, or it may reflect satisfaction of a civic duty to vote, or they may be based on a Kantian group expected utility calculation, as in Feddersen and Sandroni (2002).[3] What is important is that any of these motivations, Kantian as well as Downsian, implies responsiveness to information—all of these voters want to vote *correctly*.

Notice that having a payoff for abstaining lie between the payoff from voting for the better candidate and the worse candidate arises in most models of turnout. For example, in the classical model of outcome-oriented voters, the voter ranks actions by their probability of producing the favored candidate, conditional on the voter being pivotal. This probability is greatest when voting for the favored candidate, and least when voting for the worse candidate.

We assume there is no cost to voting. This assumption, which is standard in models that focus on informational issues (e.g. Feddersen and Pesendorfer (1996) and Ghirardato and Katz (2002)), lets us isolate the effects of ambiguity and ambiguity aversion on voting behavior. We will see that there are interesting comparative statics with respect to reasonably sized changes in informational variables, while comparative statics based on changes in costs or benefits require implausible magnitudes in standard turnout models (Palfrey and Rosenthal, 1985). Adding a cost of voting, constant across elections, would not affect our results.

The voter's problem is complicated by the fact that she does not know the true payoff to voting for the left candidate, θ. The standard approach to uncertainty about θ is to assume that the voter believes θ is distributed according to some probability density f, and votes for R if and only if $\int \theta f(\theta)\, d\theta > 0$. This representation implies that the voter has a strict preference with probability 1 (Ghirardato and Katz, 2002). In contrast, we assume that the voter has incomplete preferences. We model this by representing her beliefs by a convex *set* of probability measures, $\mathcal{P}$. Specifically, the voter prefers R to L if $\mathbb{E}_p(\theta) > 0$ for all $p \in \mathcal{P}$, she prefers L to R if $\mathbb{E}_p(\theta) < 0$ for all $p \in \mathcal{P}$, and she is unable to rank the candidates if neither condition holds.

If the priors in $\mathcal{P}$ are not unanimous and the voter cannot rank the two candidates, then the preferences alone do not determine her actions. In this case, Bewley's model assume that the voter takes a *status-quo* action. We

[3]See Feddersen (2004) for a survey of different models of turnout.

take the status-quo to be abstaining, so the voter's decision rule is:

$$\begin{array}{rl} \text{vote } R \text{ if} & \mathbb{E}_p(\theta) > 0 \;\; \forall p \in \mathcal{P} \\ \text{vote } L \text{ if} & \mathbb{E}_p(\theta) < 0 \;\; \forall p \in \mathcal{P} \\ \text{abstain} & \text{otherwise.} \end{array}$$

Note that we are not assuming that the payoff to abstaining is zero—it is still $\delta\theta$. What's going on is that, if there exist $\overline{p}$ and $\underline{p}$ such that $\mathbb{E}_{\overline{p}}(\theta) > 0$ and $\mathbb{E}_{\underline{p}}(\theta) < 0$, then convexity of $\mathcal{P}$ implies that there is a p such that $\mathbb{E}_p(\theta) = 0$. In this case, abstaining is optimal for one of the beliefs, and the inertia assumption implies that it is selected.

Notice that our voter is risk neutral—risk aversion on its own could not produce our results. Formally, this follows from Ghiardato and Katz's result that a Bayesian voter with no cost of voting turns out with probability one. The intuition is based on the fact that abstaining is itself a risky option—abstaining might cause the favored candidate to lose. Of course, with risk aversion, the voter's evaluation of one or both candidates will be worse when information is limited, and this very well might affect the voter's preference between the two candidates. Still, a Bayesian voter can always decide which of the two candidates is better, and once she has done so, she will perceive abstaining as bearing the risk of causing the less favored candidate to win. To model voters who abstain in response to their uncertainty about candidates, we must move beyond the standard Bayesian model.

The following subsection outlines the decision-theoretic background of this type of multi-prior model. Readers who are not interested in axiomatic decision theory can skip to section 4 without any loss of continuity.

3.1 Representing Incomplete Preferences

A ballot is what decision theorists call an *act*—a map from states of the world (θ) to consequences (utility of the selected candidate). We follow Bewley and assume standard (Anscombe-Aumann) axioms, except that completeness is weakened to hold only for constant acts.[4] Intuitively, the voter would have complete preferences were she certain of the candidates' attributes, but she may be unable to rank candidates because she is uncertain about the map between candidates and attributes (as Judge Posner was in 2000). We have:

[4] Ryan (2003) provides a useful survey of the Anscombe-Aumann model and various modifications, including the Bewley model that we use and the maximin expected utility model that Ghirardato and Katz (2002) use.

Bewley's Theorem *If preferences satisfy all of the Anscombe and Aumann (1963) axioms except that completeness need hold only for constant acts, then the preferences are represented by a (cardinal) utility u and a* set *of probabilistic beliefs $\mathcal{P}$ in the sense that*

$$a \succ b \quad \textit{iff} \quad \mathbb{E}_p(u(a)) > \mathbb{E}_p(u(b)) \textit{ for all } p \in \mathcal{P}.$$

Why is this true? The best way to build some intuition for the result is to think first of a case with no uncertainty. The voter's preferences are represented by a utility function if and only if those preferences are complete and transitive.[5] What happens to the classical theory if we drop completeness? Under some regularity conditions, transitive and reflexive preferences are represented by a *set* of utilities $\mathcal{U}$, where $a \succsim b$ if and only if $u(a) \geq u(b)$ for all $u \in \mathcal{U}$ (Ok, 2002). As an example, consider the alternatives $\{a, b, c\}$, and $\mathcal{U} = \{u, v\}$, where

$$u(a) = 0 \quad u(b) = 1 \quad u(c) = 2$$

and

$$v(a) = 1 \quad v(b) = 0 \quad v(c) = 2.$$

Then $c \succ a$ and $c \succ b$ since u and v agree on these comparisons, while a and b are not ranked.

A comparison to social choice theory may be helpful to explain this result. For any society of people with complete and transitive preferences, we can define the unanimity order, which consists of the preferences that all citizens hold in common. In general, this order will be incomplete—there will be two alternatives and two citizens such that those citizens disagree about how to rank the alternatives. The representation essentially comes from recognizing that *any* reflexive, transitive order is the unanimity order for some society[6], and treating that "society" as a description of an individual's preferences.[7]

Moving from ordinal utility to expected utility, Aumann (1962) shows that the mixture space axioms, minus completeness, imply that the preference can be represented by a set of linear utility functions. Bewley builds on

[5] Modulo topology.

[6] Any reflexive and transitive order is equal to the intersection of all complete and transitive orders that contain it. (To see that the latter set is nonempty for an infinite set of alternatives, use Zorn's lemma.)

[7] Link to other multi-selves approaches . . . Camerer mentions a paper (by Nehring I think) that explicitly takes a multiselves approach.

Aumman's theorem to show that adding back a limited form of completeness (completeness over constant acts) is sufficient to mimic the Anscombe and Aumann (1963) proof and pin down the utility uniquely.

In this case, the voter's preferences can be represented by a Bernoulli utility (unique up to an affine transformation) over outcomes and a closed, convex set of probability measures over states. An act a is better than b if and only if it leads to a greater expected utility according to all of the probability measures.

When preferences are complete, we assume that the DM chooses an alternative that is *optimal* in the sense that it is weakly preferred to all other alternatives. (If there are multiple optima, then the DM is indifferent between them, and is happy to choose any of them.) With incomplete preferences, we cannot use this assumption, since optimal choices might not exist. (This will be the case in our model whenever two beliefs lead to different best choices—no act dominates a *and* no act dominates b.) Instead, we assume that the DM chooses an alternative that is *maximal* in the sense that there is no other alternative that is preferred to the choice. We cannot be as flippant about multiple maximal elements as we could about multiple optima—the DM is not indifferent between the different maximal elements. Thus, to complete the model, we need to know what the voter does when not all alternatives can be ranked. Bewley assumes that there is some status quo act, and that the voter chooses this status quo unless some act dominates it. He calls this assumption "inertia". In our voting application, a natural choice for the status quo is to abstain unless some candidate dominates.

Recently, foundations for incomplete preferences and status quo bias have been presented in a revealed preference framework. Eliaz and Ok (2005) present a condition called the *weak axion of revealed non-inferiority* (WARNI), which is strictly between Sen's condition α and WARP in strength. They show that a choice function satisfies WARNI if and only if it is the maximal set of a binary relation that is reflexive, transitive, and satisfies a regularity condition. (It's easy to check that the data in Brady and Ansolabehere (1989) satisfy WARNI.) Similarly, Masatlioglu and Ok (2005) give a revealed preference foundation for the inertia assumption in Bewley's model. In both papers, the authors go on to consider uncertainty, and show that the addition of a natural adaptation of the independence axiom leads to a representation by a set of linear utilities.

4 Learning

Observing the campaign allows voters to refine their beliefs about the candidates, and so can change their preferences over candidates. This section will derive the comparative statics of this learning, as a prelude to the next section's discussion of equilibrium targeting of messages by candidates.

In the standard Bayesian model, new information is modeled as the realization of a signal whose likelihood is known up to the unknown parameter, and beliefs are updated with Bayes's rule. Following Bewley's (2002) and Halpern's (2003) treatment of information, our voters also receive a signal with known likelihood, and beliefs are updated prior-by-prior, using Bayes's rule.[8]

So assume that the voter gets a signal s with conditional density $f(s \mid \theta)$. She uses Bayes's rule to update her beliefs from $\mathcal{P}$ to $\mathcal{P}'$. (We use primes to denote updated objects.) In particular, if the signal is s and $p(\cdot)$ is in $\mathcal{P}$, then

$$p(\cdot \mid s) = \frac{p(\cdot) f(s \mid \cdot)}{\int p(\theta) f(s \mid \theta)\, d\theta}$$

is in $\mathcal{P} \mid s = \mathcal{P}'$.

Most applications of Bayesian learning by voters use the particularly simple normal learning model. The rest of this section adapts the normal learning model to our multi-prior context, and derives some comparative statics results.

4.1 Normal Learning

Results for the single-prior case are well-known (DeGroot, 1970). The prior is that $\theta \sim \mathcal{N}(\mu, \sigma_\theta^2)$, and the signal is conditionally normal: $s = \theta + \epsilon$ where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. In this case, Bayes's rule gives the posterior

$$\theta \mid s \sim \mathcal{N}\left(\lambda s + (1 - \lambda)\mu, \lambda \sigma_\theta^2\right),$$

where $\lambda = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\epsilon^2}$.

[8] In the most general formulation of Bewley's model, we would start with beliefs written as a set of joint measures over states and signals. These joint distributions could all be expressed as a prior over states and a conditional distribution of the signals given the state. We are assuming that all of these conditional distributions are the same. With this assumption, the Blackwell and Dubins (1962) theorem implies that a voter with a infinite amount of information would have complete preferences. This is reasonable if lack of information is the only reason for incompleteness.

One nice feature of the normal-normal example is that it has linear conditional expectations. This means we can write the posterior mean m as a linear combination of the signal and the prior mean μ:

$$m = \lambda s + (1-\lambda)\mu.$$

We will consider a simple extension of this model to the case of multiple priors. Let $\mathcal{G}$ be a finite set of normal distributions with common variance σ_θ^2, and let $\mathcal{P}$ be the convex hull of $\mathcal{G}$. Define the *upper expectation* by $\overline{\mu} = \sup_{p\in\mathcal{P}} \mathbb{E}_p(\theta) = \max_{p\in\mathcal{G}} \mathbb{E}_p(\theta)$ and the *lower expectation* by $\underline{\mu} = \inf_{p\in\mathcal{P}} \mathbb{E}_p(\theta) = \min_{p\in\mathcal{G}} \mathbb{E}_p(\theta)$. These two numbers provide a summary of the voter's priors that is sufficient to study her decisions.

In particular, the posterior beliefs are represented by a convex set with extreme expectations $\overline{m}$ and $\underline{m}$, where

$$\begin{aligned} \overline{m} &= \lambda s + (1-\lambda)\overline{\mu} \\ \underline{m} &= \lambda s + (1-\lambda)\underline{\mu}, \end{aligned}$$

and $\lambda = \frac{\sigma_\theta^2}{\operatorname{var} s}$. This follows from the following lemma (Williams, 2001, p. 213):

Lemma 1 *If each π_k is a distribution on Θ, if p is a probability measure on the finite set $\{1,\ldots,K\}$, and if we use the prior $\sum_k p(k)\pi_k$, then the posterior given y is*

$$\pi(\cdot \mid y) = \sum_k p(k \mid y)\pi_k(\cdot \mid y).$$

Thus given a prior that is a convex combination of simpler priors, the posterior is a convex combination of the simpler posteriors. (The weights change, in general.)

Given this updating rule, we can write the decision rule conditional on the new information as:

vote R if	$\underline{m} > 0$
vote L if	$\overline{m} < 0$
abstain	otherwise

It will be convenient to reformulate this decision rule in the signal space. The condition to vote for L can be written as

$$\overline{m} = \lambda s + (1-\lambda)\overline{\mu} < 0,$$

or

$$s < -\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu}.$$

Similarly, the voter votes R if

$$s > -\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\underline{\mu}.$$

Given this decision rule and a probability distribution over s, we can calculate the probabilities of all three choices.

NEED PICTURE HERE.

Choosing the distribution of signals is a nontrivial matter. In a standard model of Bayesian voters, we would assume that the agent's beliefs are correct, and evaluate the distribution of signals with respect to the subjective beliefs of the agent. Here, however, this does not lead to a unique distribution, since each of the agent's priors will lead to a different distribution over s. Since we want to use the model of voters later as a component of a model of party strategy, we will use the parties' distribution over s, which is assumed unique. To stick with the spirit of the correct beliefs assumption, we will assume that the parties' distribution over s corresponds to combining the conditional distribution of signals with one of the distributions over θ that the voter entertains. For simplicity, we assume that this reference distribution is normal, with mean $\alpha = \frac{1}{2}\overline{\mu} + \frac{1}{2}\underline{\mu}$. Since all of the normal distributions in the set have the same variance σ_θ^2, this means that the signals are distributed $\mathcal{N}(\alpha, \sigma_\theta^2 + \sigma_\epsilon^2)$. When we use this midpoint of the voter's prior expectations, it will be convenient to complete the description with the *radius* of the interval of prior expectations, $\Delta = \overline{\mu} - \alpha = \alpha - \underline{\mu}$. The midpoint measures partisanship, while the radius measures ambiguity aversion.

4.2 Neutral News

Say that a signal is *neutral* if it leaves the decision maker's prior mean beliefs unchanged. An important difference between a fully Bayesian voter and an ambiguity averse voter is that there can be no neutral signals for a voter who is strictly ambiguity averse. Even a signal that leaves the midpoint of the interval of expectations fixed will reduce the difference $\overline{\mu} - \underline{\mu}$. This makes the decision maker more willing to act.

An important implication of this is that information that does not move the midpoint of the beliefs can still mobilize a voter. Consider, as an example, a voter with prior expectations $[-1, 5]$ who has $\sigma_\theta^2 = \sigma_\epsilon^2$. If this voter observes a signal $s = 2$, then her interval of posterior means is $[\frac{1}{2}, \frac{7}{2}]$. Even though the midpoint is fixed at 2, the signal has caused the voter to transition from abstention to turning out for R.

4.3 Comparative Statics

Now we turn to the comparative statics of the voter's decisions. We allow four dimensions of heterogeneity: Different voters can vary in partisanship, ambiguity aversion, prior variances, and signal variances.

We saw above that the voter votes L if and only if

$$s < -\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu}.$$

Since the signal is distributed $\mathcal{N}(\alpha, \sigma_\epsilon^2 + \sigma_\theta^2)$, the probability that she votes L is

$$\Pr\left(s < -\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu}\right) = \Phi\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu} - \alpha}{\sqrt{\sigma_\theta^2 + \sigma_\epsilon^2}}\right).$$

Using the identity $\overline{\mu} = \alpha + \Delta$, this probability becomes

$$\Pr\left(s < -\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu}\right) = \Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2} + 1)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\theta^2 + \sigma_\epsilon^2}}\right).$$

The comparative static results will follow from analyzing this expression.

Say that voter i is *more right-leaning than* voter j if the midpoint of i's interval of prior means is greater than that of voter j: $\alpha_i > \alpha_j$. The first result is not very profound, but the model would be very suspect were it not true.

Fact 1 *More left-leaning voters are more likely to vote L and more right-leaning voters are more likely to vote R. The probability of abstaining is decreasing in the absolute value of partisanship.*

The proof of this, and of all subsequent facts, is in the appendix. To see the intuition, consider a moderate voter and one who is more right-leaning. Compared to the moderate, the right-leaning voter will vote R with a weaker signal in favor of R. At the same time, the distribution of signals is shifted to the right for the right-leaning voter. These effects combine to make her more likely to vote R.

Fact 2 *More ambiguity-averse voters are less likely to vote for either candidate. Very ambiguity-averse voters turnout only if they are very partisan.*

An increase in ambiguity aversion means that the interval of abstention in the figure expands. For a fixed distribution of signals, this increases the probability of abstention. Since downballot races have less information, this fact combined with Heath and Tversky's (1991) competence hypothesis suggests that those races will have lower turnout, and that the electorates in those races will be more partisan.

The comparative statics with respect to the variances are more subtle, and depend on the voter's initial leanings.

Fact 3 *Voters are more likely to switch from abstain or L to R the more precise is their signal. Voters are more likely to switch from abstain or R to L the more precise is their signal. Intended voters who are not too partisan are more likely to stay with their initial choice the more precise is their signal, while more partisan intended voters are less likely to stay the more precise is their signal.*

A more precise signal has more impact on the voter—her beliefs move more, on average, and the radius of her interval of expectations shrinks more. For a voter who does not initially intend to vote L, these changes work together and unambiguously make her more likely to vote L. For an initial L voter, on the other hand, the two effects work in opposite directions. For a voter who is not too partisan ($\alpha \in (-k\Delta, \Delta)$ for some $k > 1$), the dominant effect is that the radius shrinks, making it less likely that a move takes her out of the range where she vote for L. For a more partisan voter, the dominant effect is that her beliefs are more likely to move in an R-friendly direction.

Fact 4 *Initial partisans are more likely to stay the more precise are their prior beliefs. Initial abstainers who lean strongly towards one party are more likely to turnout for that party the more precise are their prior belief, while initial abstainers who do not lean strongly towards one party are more likely to abstain the more precise are their prior beliefs.*

More precise prior beliefs make the voter less receptive to new information—her beliefs move less, on average, and the radius of her interval of expectations shrinks less. For an initial partisan, the dominant effect is the reduced movement of beliefs, and the voter is more likely to stick with her initial plan. Similarly, an initial abstainer who does not lean strongly towards one of the parties is less likely to move enough to decide to vote the more precise are her prior beliefs. For an abstainer who leans sufficiently strongly towards one of the two parties, however, more precise priors make her more likely to ultimately turn out, since she is less likely to move away from the region where she is mobilized by neutral news.

The key to these comparative static results in that there is a range of signals that lead to abstention. The role of ambiguity aversion is to justify this interval and to provide an interpretation of the comparative statics with respect to that interval. At this level, other interpretations, such as a cost of voting, might work as well. However, the ambiguity aversion interpretation has certain advantages. First, it is quite natural and provides links to the empirical literature on voter behavior. For example, ambiguity aversion was crucial to our earlier finding that neutral news could mobilize a voter. Second, the interpretation in terms of uncertainty aversion provides a natural way to compare different kinds of races in terms of information, for example high-information Presidential races vs. low-information local elections, and explains why party strategy will differ systematically across these different types of elections.

4.4 Discussion

We will defer comparison of the model's predictions and the empirical literature until the next section on party strategy, since observed voter behavior is in response to the equilibrium campaign strategies of the candidates. Still, the results on exogenous information flows allow us to make one comment on empirical matters.

If this model is right, there is an ambiguity in the interpretation of "uncertainty" in surveys. A high value for the prior variance increases the likelihood of a transition to voting. This is because a voter with a high prior variance is very responsive to new information—the high prior variance is a reflection of a lack of faith in her existing beliefs. A large prior radius, on the other hand, reduces the likelihood of a transition to voting. This is because a large radius reflects a high degree of aversion to ambiguity—this voter is reluctant to act without a great deal more information.

This result means that there are no clear comparative statics concerning a one-dimensional measure of "information"—Δ and σ^2 measure different ways that a voter may perceive uncertainty, and the comparative statics of these two uncertainties are different. To see how this works, consider two voters who we might think of as uncertain—one with a high variance of a single prior, and one with a large number of different, very precise priors. The first voter will have large responses to campaign messages, and is likely to have volatile vote intentions. The second voter, on the other hand, will have beliefs that respond very little to campaign events, and she is unlikely to change her plans. She is also unlikely to vote. A voter with small Δ and small σ_θ^2 will also have little change over the campaign, but she will be quite likely to vote.

This distinction may provide a compelling test of the model, once empirical referents for the two different kinds of uncertainty are identified. However, these comparative statics also raise the question of how to measure the various components. The best referent for the prior mean radius might be Bartels's measure based on item nonresponse. Direct questions about certainty might also be relevant. For the prior variance, one might use the question "how likely are you to change your mind?", or a measure of prior knowledge.

5 Strategically Targeted Campaign Messages

The previous section derived differences in the way that different voters respond to campaign information. On their own, these results are not enough to guide our thinking about observed differences in behavior across voters during a campaign. Strategic candidates will foresee that voters will differ in their responses to campaigns, and will therefore be selective in how they target voters, focusing on those voters who will respond in the most favorable way.

This section addresses this lacuna. We will use the comparative statics results to determine the optimal targeting strategies of two office-seeking parties, and then use these strategies to look at some comparative static results about *equilibrium* voter behavior.

5.1 The Model

Two candidates compete for an elected office. One represents the R party and the other the L party.

There is a continuum of voters, with measure 1. These voters are divided into N targetable groups, and group n has measure β_n for $1 \leq n \leq N$. Each member of group n has beliefs about the utility differential between the R and L candidates represented by a set of beliefs with upper expectation $\overline{\mu}_n$, lower expectation $\underline{\mu}_n$, and variance σ_n^2. Although the prior distributions are common within groups, the actual valances are "independent".[9]

[9]We follow the usual practice of extending the idea of independence to a continuum of

Each candidate can send messages targeted to specific groups. If a group is targeted, each member gets a signal on θ with variance σ_ϵ^2. These signals are "independent" across the voters within a group. Without a signal, a voter learns nothing during the campaign.

Sending a message to a group costs c, and each candidate has total resources cM, where $M < N$. Thus the candidates must choose which groups to target and which to ignore.

We assume that each candidate maximizes her plurality. This is a concession to absence of any other learning—if the voter got a continuously distributed signal (however imprecise) independent of the actions, then a candidate who maximized the probability of winning would take actions just like those we derive below.

5.2 Equilibrium

Let $\mathcal{L}$ denote the set of groups that candidate L targets and let $\mathcal{R}$ denote the set of groups that R targets. A strategy $\mathcal{L}$ is *feasible for* L if $|\mathcal{L}| \leq M$, and similarly for R.[10] An *equilibrium* is a pair $(\mathcal{L}^*, \mathcal{R}^*)$ such that:

1. Both $\mathcal{L}^*$ and $\mathcal{R}^*$ are feasible, and
2. for neither candidate is there a feasible strategy that leads to a greater plurality.

The next result will allow us to derive a sharp characterization of equilibrium in the generic case. Say that a profile is *specialized* if each group is targeted by at most one candidate. Formally, we have $n \in \mathcal{L}$ implies $n \notin \mathcal{R}$ and $n' \in \mathcal{R}$ implies $n' \notin \mathcal{L}$.

Proposition 1 *If there is a pure strategy equilibrium, then there is a specialized pure strategy equilibrium. For generic values of the parameters, every pure strategy equilibrium is specialized.*

Proof Let ψ_n be R's plurality from group n when it gets two signals, and let ρ_n be R's plurality from group n when it gets one signal. There are three cases to consider:

1. $\psi_n > \rho_n$
2. $\psi_n < \rho_n$

random variables by making the law of large numbers into a definition.

[10]For any set X, we denote the cardinality of X by $|X|$.

3. $\psi_n = \rho_n$

Consider some profile in which both groups target group n. In case 1, L can increase its plurality by deviating and not targeting n. In case 2, R can increase its plurality by deviating and not targeting n. In either case, the profile is not an equilibrium.

Now consider case 3, and assume that there is an equilibrium in which both candidates target group n. Since $\psi_n = \rho_n$, if L switches to not targeting n, its plurality is unaffected. Furthermore, L has no strict incentive to use the freed-up funds to target some new group—at an equilibrium, neither candidate can strictly increase his plurality by switching from group n to some other, untargeted, group. Thus the new profile is also an equilibrium. Repeating this procedure for each group targeted by both parties in the initial equilibrium produces a specialized equilibrium.

ADD PROOF THAT EQUALITY IS NON-GENERIC. □

Given the specialization result, we can give a sharp characterization of equilibrium. With one signal, R's net plurality in group n is

$$\rho_n = 1 - \Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha + \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) - \Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

When deciding whom to target, the candidate compares this plurality with that the group gives her without any information, denoted γ_n. This is $\gamma_n = 1$ if $\alpha - \Delta > 0$, $\gamma_n = 0$ if $\alpha + \Delta \geq 0 \leq \alpha - \Delta$, and $\gamma_n = -1$ if $0 > \alpha + \Delta$. Define the difference in these pluralities as $\delta_n = \rho_n - \gamma_n$. Say that there are *no ties* if $n \neq n'$ implies $\beta_n\delta_n \neq \beta_{n'}\delta_{n'}$ and $\delta_n \neq 0$ for all n. When there are no ties, the groups are strictly ordered by their net contributions to R's plurality.

When there are no ties, we can give a sharp characterization of equilibrium.

Proposition 2 *Assume there are no ties. Then there exists a unique specialized equilibrium. In this equilibrium, the R candidate targets the set of groups $\mathcal{R}$ which is maximal with respect to the properties that*

1. *$|\mathcal{R}| \leq M$ (budget balance),*
2. *$n \in \mathcal{R}$ and $n' \notin \mathcal{R}$ imply that $\beta_n\delta_n > \beta_{n'}\delta_{n'}$, and*
3. *$n \in \mathcal{R}$ implies $\delta_n > 0$.*

The equilibrium strategy of L is the same, but with the inequalities reversed in (ii) and (iii).

Proof First we show that an equilibrium strategy for R must satisfy the three properties. Property 1 is just feasibility. If 2 is not satisfied, then there are two groups, n and n' such that n is targeted and n' is not, but $\beta_{n'}\delta_{n'} > \beta_n\delta_n$. (Equality is ruled out by the no ties assumption.) Then R can strictly increase its plurality by switching from n to n'. Thus the initial strategy could not be part of an equilibrium. Finally, if 3 is not satisfied, then R can increase its plurality by ceasing to target group n.

Now consider two sets $\mathcal{R} \subset \mathcal{R}'$ that both satisfy all three properties. By property 3, the groups in $\mathcal{R}' \setminus \mathcal{R}$ all make a positive contribution to R's plurality. Thus no profile $(\mathcal{R}, \mathcal{L})$ can be an equilibrium—R does strictly better by deviating to $\mathcal{R}'$.

So far, we've seen that any equilibrium strategy must be maximal with respect to the three properties. But property 2 and the no ties assumption imply that the sets which satisfy the properties are linearly ordered by set inclusion. Thus there is a unique maximal set.

All that's left is to show that the maximal set actually is an equilibrium strategy. But this is obvious. □

Notice that this proposition gives us an algorithm for finding specialized equilibria. Number the groups so that $\beta_1\delta_1 > \cdots > \beta_N\delta_N$. Construct $\mathcal{R}$ by adding groups in order, starting with $n = 1$. Continue until either $|\mathcal{R}| = M$ or $\delta_n < 0$. Construct $\mathcal{L}$ in the same way, but start with $n = N$ and stop if $\delta_n > 0$.

5.3 Who is Targeted?

Some predictions about which groups are targeted are immediate from this characterization. For example, all else equal, a candidate prefers to target the larger of two groups. To draw more refined predictions, we need to undertake a closer study of the δ as functions of the underlying parameters.

Lemma 2 *The net responsiveness δ is increasing in α except for jumps down at $\alpha = -\Delta$ and $\alpha = \Delta$. It is positive for $\alpha \in (-\infty, -\Delta) \cup (0, \Delta]$, is zero for $\alpha = 0$, and is negative otherwise.*

Proof Differentiate ρ with respect to α to get

$$\frac{\frac{\sigma_\epsilon^2}{\sigma_\theta^2} + 1}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}(\phi^+ + \phi^-) > 0.$$

Thus the contribution to R's net plurality is greater the more right-wing is the group. The function crosses 0 at $\alpha = 0$, and has limits 1 as $\alpha \to \infty$ and -1 as $\alpha \to -\infty$.

Comparing the two plurality functions, we see that giving information improves the R plurality if and only if $\alpha \in (0, \Delta]$ or $\alpha < -\Delta$. □

This result is intuitive. Voters who are initially supporters of the candidate can only be pushed away by campaign activity, so targeting them has a negative return. Similarly, voters initially supporting the other candidate can only be pushed in the right direction, so targeting them has a positive return. Initial abstainers might be pushed toward either party, so the value of targeting them depends on a comparison of the likelihoods. For leaners toward a party, they are more likely to move toward the party. Leaners toward the other party are more likely to move away. Finally, the maximal impact is from groups who are on the threshold of switching. Thus the best groups to target are large and have many voters on the verge of supporting the party, without too many offsetting voters who might be driven away or driven to the other party.

5.3.1 Mobilization vs. Persuasion

When the R candidate targets a group with $\alpha \in (0, \Delta]$, the targeting is aimed at *mobilization*. When he targets a group with $\alpha < -\Delta$, the interpretation of targeting is sensitive to the degree of ambiguity aversion. When the radius is small, there is a good chance that giving info to a partisan of the other party will cause them to switch their vote. This is like *persuasion*. If the radius is large, on the other hand, the more likely outcome is that the voter becomes *demobilized*. This is like negative campaigning.

To confront the empirical work on campaign effects, a crucial question is the relative efficacy of attempts to mobilize voters who will not otherwise turnout vs. persuading voters who currently plan to vote for the opponent to change their plan. This section explores the tradeoffs that underly this choice. In principle, proposition 2 gives a complete answer to this question, but at that level of generality it's hard to see how the individual factors work. Here we look at the comparative statics as we vary one parameter at a time across the groups, to try and learn some general lessons.

The first result follows directly from Lemma 2.

Fact 5 *Consider 2 groups that differ only in their alpha's, both of which are in* $(0, \Delta]$*. If the R candidate can target only one of the two groups, it will choose the more partisan one.*

Thus the best voters to try to mobilize are those who are most partisan among those who are not yet planning to vote.

We've just seen that a party does best by attempting to mobilize a group just to on the threshold of turning out. Given a choice between two such groups, which differ only in their degree of ambiguity aversion, which will be chosen?

Fact 6 *Consider two groups that are identical except that $\alpha_i = \Delta_i > \alpha_j = \Delta_j$. If the R candidate can afford to target only one of the groups, it will choose group i, the more partisan one.*

Intuitively, this is because members of that group are less likely to be inadvertently mobilized for the other candidate.

With limited resources, parties first "secure their base"—targeting the most ideologically sympathetic voters who are not yet convinced to turnout. Since voters perceive more ambiguity the less they know about about the race (Heath and Tversky's (1991) competence hypothesis), this suggests some comparative statics across campaigns. As the background informational condition varies, we expect the ideological leanings of the marginal voters to change as well. In Presidential elections, there are high-information public signals, so we expect that more ideological voters will not be marginal. On the other hand, in races for lower offices, information will be lower and the marginal voters will be more partisan. This will produce swings in the degree that parties chase swing voters (α close to 0) versus playing to their base (α far from 0).

Next, consider two potential targets for persuation/demobilization with $\alpha = -\Delta$.

Fact 7 *Consider two groups that are identical except that $\alpha_i = -\Delta_i > \alpha_j = -\Delta_j$. If the R candidate can afford to target only one of the groups, it will choose group i, the less partisan one.*

Intuitively, members of this group are more likely to be persuaded rather than just demobilized.

Finally, consider a candidate who must decide to use mobilization or persuasion.

Fact 8 *Consider two groups with the same values of σ_θ^2 and Δ, but one leans towards R ($\alpha = \Delta$) and the other is just voting for L ($\alpha = -\Delta$). If the R candidate can target only one of these groups, it will choose to mobilize the group that leans toward it.*

This means that candidates prefer mobilization to persuasion.

We can think of the choice between mobilization and demobilization as a choice between positive and negative campaigns. Doing so allows us to predict that a party is more likely to mount a negative campaign when it has relatively few latent supporters, when the electorate is more partisan, and when the electorate is more polarized ideologically.

5.3.2 Responsive Voters?

Another way that candidate choices can affect the magnitude of campaign effects is by choosing more or less receptive voters to target.

Fact 9 *If a candidate can choose to target one of two groups that are identical except for their prior variances, the candidate prefers to target the one with more precise prior beliefs if they are abstainers who lean strongly in the candidate's direction, and prefers to target the one with less precise prior beliefs otherwise.*

Thus when a candidate attempts to mobilize voters who lean strongly in his direction, he prefers voters who are less responsive to new information.

5.4 Equilibria when funds have alternative uses

To be written.

5.5 Comparison to "Swing Voter" Models

The campaign stage of the model builds on work by Lindbeck and Weibull (1987) and Dixit and Londregan (1996). In these models, parties target their resources to "swing voters"—voters who evenly balanced between the two parties. These voters are the most likely to change their vote decision. Strömberg (2002) finds supportive evidence for U.S. Presidential elections. Ansolabehere and Snyder (2003) test this prediction with data on the distribution of expenditures by state governments. Contrary to the prediction, they find that the party in power skews the resource allocation in favor of districts which traditionally support the party, rather than to districts which have supported both parties at different times.[11] They also find that

[11] Dixit and Londregan (1996) also consider a "core support" version of their model. It is based on a stylized depiction of big-city machine politics. In this model, one of the parties controls the bureaucracy, and consequently has an advantage in targeting pork. Furthermore, the machine has better connections to its base of support than to swing voters. They suggest this as a model of pre-civil service municipal politics, so it doesn't apply to the kind of evidence that Ansolabehere and Snyder (2003) find.

regions that receive extra funds have greater turnout in the subsequent election. This suggests that the basic logic of the swing voter model is correct—parties target the groups where the marginal increase in net votes is greatest, but it suggests that turnout is a more relevant margin in state-level elections. Cox and McCubbins (1986) present a model in which a party's "home base" is more responsive to spending, but they do not derive this assumption from a model of voters. This paper provides microfoundations for such an assumption. This allows me to link the voters in the model of the campaign with evidence about voter behavior, and to derive comparative statics based on *endogenous* changes in the responsiveness of turnout to spending.

6 Conclusion

To be written.

A Proofs

We start with a preliminary result.

Lemma 3 *If $y > 0$, then the expression $\phi(x+y) - \phi(x-y)$ is positive if and only if $x < 0$.*

Proof We have

$$\phi(x+y) - \phi(x-y) > 0$$

if and only if

$$\frac{\phi(x+y)}{\phi(x-y)} > 1.$$

By the symmetry of the normal density, we have

$$\frac{\phi(y)}{\phi(-y)} = 1.$$

This and the MLR property of the normal density imply the result. $\square$

Proof of Fact 1 Consider the probability of voting for L. This is

$$\Pr\left(s < -\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu}\right) = \Phi\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu} - \mu}{\sqrt{\sigma_\theta^2 + \sigma_\epsilon^2}}\right).$$

It's easy to see that this expression is decreasing in $\overline{\mu}$. If the voter becomes more right-wing, so the both $\underline{\mu}$ and $\overline{\mu}$ decrease, then the voter becomes more likely to vote L. Results are symmetric for the probability of voting R.

The probability of abstention is

$$\Phi\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\underline{\mu} - \mu}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right) - \Phi\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu} - \mu}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right).$$

Differentiate with respect to α to get

$$\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right)(\phi^+ - \phi^-).$$

Lemma 3 implies that this derivative is positive if and only if $\alpha < 0$. This shows that the probability of abstention is decreasing in the absolute value of partisanship. $\square$

Proof of Fact 2 The probability that a voter turns out is

$$1 - \Phi\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\underline{\mu} - \mu}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right) + \Phi\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\overline{\mu} - \mu}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right).$$

Differentiate with respect to Δ to get

$$\left(\frac{-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right)(\phi^+ + \phi^-) < 0.$$

Thus the voter becomes less likely to vote. □

Proof of Fact 3 The probability a voter votes L is

$$\Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2} + 1)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right).$$

Differentiate with respect to σ_ϵ^2 to get

$$\phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2} + 1)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2 + \sigma_\theta^2}}\right)\left(\frac{-\frac{\alpha+\Delta}{\sigma_\theta^2}\sqrt{\sigma_\theta^2 + \sigma_\epsilon^2} + \left(-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2} + 1)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta\right)\frac{1}{2\sqrt{\sigma_\theta^2+\sigma_\epsilon^2}}}{\sigma_\theta^2 + \sigma_\epsilon^2}\right),$$

which has the same sign as

$$-(\alpha + \Delta)(\sigma_\epsilon^2 + \sigma_\theta^2) + \frac{1}{2}(\alpha + \Delta)\sigma_\epsilon^2 + \frac{1}{2}\alpha\sigma_\theta^2.$$

Thus the derivative is positive if and only if

$$\alpha < -\frac{\sigma_\epsilon^2 + 2\sigma_\theta^2}{\sigma_\epsilon^2 + \sigma_\theta^2}\Delta < -\Delta.$$

If $\alpha > -\Delta$, then the voter initially plans to either abstain or vote R. Thus all such voters become less likely to vote L as σ_ϵ^2 increases. Similarly, initial L voters with $\alpha > -k\Delta$ become less likely to vote L, where $k = \frac{\sigma_\epsilon^2+2\sigma_\theta^2}{\sigma_\epsilon^2+\sigma_\theta^2}$. The other L voters become more likely to vote L. Since a more informative campaign corresponds to a lower value of σ_ϵ^2, the result follows. □

Proof of Fact 4 Differentiate the probability of voting L to get

$$\phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)\left(\frac{\frac{\sigma_\epsilon^2}{\sigma_\theta^4}(\alpha+\Delta)\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}+\left(\left(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1\right)\alpha+\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta\right)\frac{1}{2\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}}{\sigma_\epsilon^2+\sigma_\theta^2}\right).$$

This has the same sign as

$$\frac{\frac{\sigma_\epsilon^2}{\sigma_\theta^4}(\alpha+\Delta)\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}+\left(\left(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1\right)\alpha+\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta\right)\frac{1}{2\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}}{\sigma_\epsilon^2+\sigma_\theta^2}.$$

Thus the derivative is positive if

$$\sigma_\epsilon^2(\sigma_\epsilon^2+\sigma_\theta^2)(\alpha+\Delta)+\frac{1}{2}(\sigma_\epsilon^2\sigma_\theta^2+\sigma_\theta^4)\alpha+\frac{1}{2}\sigma_\epsilon^2\sigma_\theta^2\Delta>0,$$

or

$$\alpha>-\frac{\frac{3}{2}\sigma_\epsilon^2\sigma_\theta^2+\sigma_\epsilon^4}{\frac{1}{2}\sigma_\theta^4+\frac{3}{2}\sigma_\epsilon^2\sigma_\theta^2+\sigma_\epsilon^4}\Delta\equiv-k\Delta.$$

For an initial L voter, we have $\alpha<-\Delta<-k\Delta$. Such a voter is more likely to vote L the smaller is σ_θ^2, that is, the more precise is her prior distribution. □

Proof of Fact 6 Recall that R's net plurality from a group is

$$\rho_n=1-\Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha+\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)-\Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha-\frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

Since the group is just on the cusp of turning out for R, we have $\alpha=\Delta$, so the plurality simplifies to

$$\rho_n=1-\Phi\left(\frac{-\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)-\Phi\left(\frac{-(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

Differentiate with respect to α to get

$$\phi\left(\frac{-\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)\frac{1}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}+\phi\left(\frac{-(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)\frac{(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}>0.$$

Thus the candidate prefers to target the more ambiguity averse (and more partisan) group. □

Proof of Fact 7 Candidate R's plurality simplifies to

$$\rho_n = 1 - \Phi\left(\frac{(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) - \Phi\left(\frac{\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

Differentiate with respect to Δ to get

$$-\phi\left(\frac{(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)\frac{(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}} - \phi\left(\frac{\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)\frac{1}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}} < 0.$$

Thus the candidate prefers to target the less ambiguity averse (and less partisan) group. □

Proof of Fact 8 If the R candidate targets its latent partisans, it gets

$$\rho_n = 1 - \Phi\left(\frac{-\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) - \Phi\left(\frac{-(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

If he targets the L voters, he gets

$$\rho_n = 1 - \Phi\left(\frac{(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) - \Phi\left(\frac{\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

Since $\Delta > 0$ and Φ is increasing, we have

$$\Phi\left(\frac{\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) > \Phi\left(\frac{-\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)$$

and

$$\Phi\left(\frac{2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) > \Phi\left(\frac{-2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

Thus

$$1-\Phi\left(\frac{-\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)-\Phi\left(\frac{-(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) > 1-\Phi\left(\frac{(2\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)-\Phi\left(\frac{\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

□

Proof of Fact 9 Recall that group n's contribution to R's plurality is

$$\rho_n = 1 - \Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha + \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right) - \Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right).$$

Differentiate with respect to σ_θ^2 to get

$$\begin{aligned}\frac{\partial\rho}{\partial\sigma_\theta^2} &= -\phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha + \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)\left(\frac{\frac{\sigma_\epsilon^2}{\sigma_\theta^4}(\alpha-\Delta)\sqrt{\sigma_\epsilon^2+\sigma_\theta^2} + \left(\left(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1\right)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta\right)\frac{1}{2\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}}{\sigma_\epsilon^2+\sigma_\theta^2}\right) \\ &\quad -\phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta}{\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}\right)\left(\frac{\frac{\sigma_\epsilon^2}{\sigma_\theta^4}(\alpha+\Delta)\sqrt{\sigma_\epsilon^2+\sigma_\theta^2} + \left(\left(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1\right)\alpha + \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta\right)\frac{1}{2\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}}{\sigma_\epsilon^2+\sigma_\theta^2}\right).\end{aligned}$$

Consider first the case of a group which leans right, so $\alpha > 0$. In this case, the final factor is positive, so the second term is negative. Thus the overall derivative will be negative if the final factor on the first line is positive. This is true if

$$\frac{\frac{\sigma_\epsilon^2}{\sigma_\theta^4}(\alpha-\Delta)\sqrt{\sigma_\epsilon^2+\sigma_\theta^2} + \left(\left(\frac{\sigma_\epsilon^2}{\sigma_\theta^2}+1\right)\alpha - \frac{\sigma_\epsilon^2}{\sigma_\theta^2}\Delta\right)\frac{1}{2\sqrt{\sigma_\epsilon^2+\sigma_\theta^2}}}{\sigma_\epsilon^2+\sigma_\theta^2} > 0,$$

or

$$\sigma_\epsilon^2(\sigma_\epsilon^2+\sigma_\theta^2)(\alpha-\Delta) + \frac{1}{2}(\sigma_\epsilon^2\sigma_\theta^2+\sigma_\theta^4)\alpha - \frac{1}{2}\sigma_\epsilon^2\sigma_\theta^2\Delta > 0,$$

or

$$\alpha > \frac{\frac{3}{2}\sigma_\epsilon^2\sigma_\theta^2+\sigma_\epsilon^4}{\frac{1}{2}\sigma_\theta^4+\frac{3}{2}\sigma_\epsilon^2\sigma_\theta^2+\sigma_\epsilon^4}\Delta \equiv k\Delta.$$

Thus for $\alpha \in (k\Delta, \Delta)$, R prefers to target groups with smaller prior variances, while for $\alpha \in (0, k\Delta)$, R prefers to target groups with larger prior variances.

By symmetry, when the R candidate targets L supporters in an attempt to demobilize them, he prefers to target groups with large prior variances.□

References

Achen, Christopher H. 1992. "Social Psychology, Demographic Variables, and Linear Regression: Breaking the Iron Triangle in Voting Research." *Political Behavior* 14(3):195–211.

Alvarez, R. Michael. 1997. *Information and Elections.* Ann Arbor: University of Michigan Press.

Anscombe, F. J. and R. J. Aumann. 1963. "A Definition of Subjective Probability." *Annals of Mathematical Statistics* 34:199–205.

Ansolabehere, Stephen and James M. Snyder, Jr. 2003. "Party Control of State Government and the Distribution of Public Expenditures." MIT working paper.

Ansolabehere, Stephen and Shanto Iyengar. 1995. *Going Negative: How Political Advertisements Shrink and Polarize the Electorate.* New York: Free Press.

Aumann, Robert J. 1962. "Utility Theory without the Completeness Axiom." *Econometrica* 30:445–462.

Bartels, Larry M. 1988. *Presidential Primaries and the Dynamics of Public Choice.* Princeton, NJ: Princeton University Press.

Bewley, Truman F. 2002. "Kightian Decision Theory: Part I." *Decisions in Economics and Finance* 25:79–110.

Blackwell, David and Lester Dubins. 1962. "Merging of Opinions with Increasing Information." *The Annals of Mathematical Statistics* 33(3):882–886.

Brady, Henry E. and Stephen Ansolabehere. 1989. "The Nature of Utility Functions in Mass Publics." *American Political Science Review* 83(1):143–163.

Calvert, Randall L. 1986. *Models of Imperfect Information in Politics.* New York: Harwood Academic Publishers.

Camerer, Colin and Martin Weber. 1992. "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity." *Journal of Risk and Uncertainty* 5:325–370.

Cox, Gary W. and Mathew D. McCubbins. 1986. "Electoral Politics as a Redistributive Game." *Journal of Politics* 48(2):370–389.

Curley, Shawn P., Frank J. Yates and R. A. Abrams. 1986. "Psychological Sources of Ambiguity Avoidance." *Organizational Behavior and Human Decision Processes* 38:230–256.

DeGroot, Morris H. 1970. *Optimal Statistical Decisions.* New York: McGraw-Hill.

Dixit, Avinash and John Londregan. 1996. "The Determinants of Sucess of Special Interests in Redistributive Politics." *Journal of Politics* 58(4):1132–1155.

Eliaz, Kfir and Efe A. Ok. 2005. "Indifference or Indecisiveness? Choice-Theoretic Foundations of Incomplete Preferences." Unpublished paper, NYU.

Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics* 75:643–669.

Feddersen, Timothy J. 2004. "Rational Choice Theory and the Paradox of Not Voting." *Journal of Economic Perspectives* 18(1):99–112.

Feddersen, Timothy J. and Alvaro Sandroni. 2002. "A Theory of Participation in Elections with Ethical Voters." Unpublished paper.

Feddersen, Timothy J. and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review* 86:408–424.

Fox, Craig R. and Amos Tversky. 1995. "Ambiguity Aversion and Comparative Ignorance." *Quarterly Journal of Economics* 110:585–603.

Frisch, Deborah and Jonathan Baron. 1988. "Ambiguity and Rationality." *Journal of Behavioral Decision Making* 1:149–157.

Gerber, Alan and Donald P. Green. 1998. "Rational Learning and Partisan Attitudes." *American Journal of Political Science* 42(3):794–818.

Ghirardato, Paolo and Jonathan N. Katz. 2002. "Indecision Theory: Quality of Information and Voting Behavior." Unpublished paper.

Gilboa, Itzhak, Andrew Postlewaite and David Schmeidler. 2004. "Rationality of Belief, Or: Why Bayesianism is Neither Necessary nor Sufficient for Rationality." PIER working paper.

Gilboa, Itzhak and David Schmeidler. 1989. "Maximin Expected Utility with a Non-Unique Prior." *Journal of Mathematical Economics* 18:141–153.

Halpern, Joseph Y. 2003. *Reasonong About Uncertainty.* Cambridge: MIT Press.

Heath, Chip and Amos Tversky. 1991. "Preference and Belief: Ambiguity and Competence in Choice under Uncertainty." *Journal of Risk and Uncertainty* 4:2–28.

Jamieson, Kathleen Hall. 2001. *Everything You Think You Know About Politics.* New York: Basic Books.

Lazarsfeld, Paul, Bernard Berelson and Hazel Gaudet. 1948. *The People's Choice: How the Voter Makes up his Mind in a Presidential Campaign.* New York: Columbia University Press.

Lindbeck, Assar and Jörgen W. Weibull. 1987. "Balanced-Budget Redistribution as the Outcome of Political Competition." *Public Choice* 52:273–297.

Masatlioglu, Yusufcan and Efe A. Ok. 2005. "Rational Choice with Status Quo Bias." *Journal of Economic Theory* 121:1–29.

Ok, Efe A. 2002. "Utility Representation of an Incomplete Preference Relation." *Journal of Economic Theory* 104:429–449.

Palfrey, Thomas R. and Howard Rosenthal. 1985. "Voter participation and Strategic Uncertainty." *American Political Science Review* 79:62–78.

Posner, Richard A. 2003. *Law, Pragmatism, and Democracy.* Cambridge: Harvard University Press.

Ryan, Matthew J. 2003. "Generalizations of SEU: A Geometric Tour of Some Non-Standard Models." Unpublished paper.

Slovic, Paul and Amos Tversky. 1974. "Who Accepts Savage's Axiom?" *Behavioral Science* 19:368–373.

Strömberg, David. 2002. "Optimal Campaigning in Presidential Elections: The Probability of Being Florida." IIES working paper.

Williams, David. 2001. *Weighing the Odds: A Course in Probability and Statistics.* New York: Cambridge University Press.

Zechman, Martin J. 1979. "Dynamic Models of the Voter's Decision Calculus." *Public Choice* 34:297–315.