Campaign Effects with Ambiguity-Averse Voters*

Scott Ashworth†

August 1, 2005

Abstract

A voter is *ambiguity averse* when she dislikes acting on beliefs that are not backed up by hard information. This ambiguity aversion can lead voters to abstain when they have low quality information. Electoral candidates face a tradeoff between mobilization of ambiguity-averse independents and persuasion of already committed partisans of the other party. In a simple theoretical model, mobilization is a more attractive strategy than is persuasion. Negative campaigns, aimed at demobilization, can be useful, and the model makes predictions about when this is a desirable strategy. Furthermore, races for less visible offices will be more oriented towards "base politics". These theoretical result are related to the empirical literature on campaign effects and the "minimal effects" hypothesis.

*I received helpful comments Tom Palfrey and from seminar participants at Washington University and Princeton's CSDP lunch.

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

Perhaps the most surprising finding of the empirical literature on voter behavior is the *minimal effects hypothesis*—political campaigns do not lead many citizens to change their intended vote. Why do campaigns appear ineffective in this way? An obvious answer looks to voter psychology, and concludes that citizens are simply difficult to persuade. But a complete account of campaign effects must account for candidate strategy as well. After all, candidates have beliefs about how the various groups of voters will respond to campaign messages, and they will target those groups whom they expect to respond most favorably. It seems intuitive that these strategic effects will work against the psychological barriers to persuasion—shouldn't candidates target the most responsive voters? In a theoretical model of the campaign targeting problem, I show that this intuition is wrong. Instead, the candidates have incentives to target the *least* responsive voters. Thus we could see minimal effects of campaigns even if a sizeable fraction of the electorate was actually easy to persuade.

This conclusion does not mean that researchers are wrong to implicate the voters' information processing. Indeed, it is precisely the *interaction* of campaign strategy and the way voters deal with information that leads to the results. I show how a simple, psychologically motivated modification of standard models leads the parties to decide to target voters who are more partisan and who are less responsive to information. In both cases, this targeting strategy implies that information will have a hard time changing many vote decisions, even if there are many responsive voters in the population. In addition, the model makes novel predictions about when candidates "play to their base" or when they chase "swing voters", about the choice to mount a positive or a negative campaign, and about the relationship between voter polarization and the scale of campaigns.

The main innovation of the paper is that the voters are allowed to be averse to *ambiguity*—"uncertainty about probability, created by missing information that is relevant and could be known" (Frisch and Baron, 1988). The role of ambiguity in voter behavior 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]

The model builds on two ideas expressed in this passage. First, Posner could not rank the two candidates. Second, he abstained in response to this ambiguity. Section 1 reviews the extensive empirical evidence that these qualities are not limited to just this one voter.

The standard assumption about information processing in formal theory is that voters are *Bayes rational*—their beliefs can be represented by a probability distribution, and learning is equivalent to updating this distribution with Bayes's rule. However, this model is inconsistent with the considerations Posner mentions. 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. That is, a Bayesian voter treats her subjective probabilities as known. A Bayesian voter cannot decide to abstain simply because she has "low quality" information—she is indifferent to ambiguity.

What about risk-aversion? Can't that account for ambiguity of the kind described by Posner? No, it cannot. Ghirardato and Katz (2002) show that a Bayesian voter with no cost of voting almost surely votes. This result holds independently of any assumptions about the voter's risk tolerance. To see the intuition for this result, notice that abstaining is also a risky act for an outcome-motivated voter. 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, as in Alvarez (1997) and Bartels (1986). 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.

Also note that a positive cost of voting will not help the Bayesian model capture the effects I ascribe to ambiguity. A cost of voting implies abstention when the difference in the expected utility of the two candidates is low relative to the cost. The voter's overall level of information has no systematic relationship to the magnitude of this difference. In the quadratic utility case, for example, adding the same constant to the variance of the voter's beliefs about each candidate has no impact on the magnitude of this difference.

To capture ambiguity aversion, I use a model of Bewley (2002) that is based on incomplete preferences. This means that the voter *cannot* always decide which of two candidates is better, in contrast to the Bayesian voter discussed above. This is a natural modification of the standard model for applications to voter behavior, since it allows the voter to answer "I don't know" to questions about how she plans to vote.

Incompleteness arises naturally from ambiguity. Letting the voter have a *set* of beliefs and assuming only that she chooses an undominated alternative implies incompleteness. For example, the voter might entertain the belief that the Democratic candidate is better than the Republican candidate with probability 2/3 *and* that the Republican candidate is better than the Democratic candidate with probability 2/3. In this case, the voter cannot rank the two candidates.

When the voter cannot rank the two candidates, her preferences alone do not determine what she does. In such cases, the model is closed with an "inertia" assumption—there is a status quo action and the decision maker sticks with the status quo unless some other action is better according to all of the beliefs. I assume that the status quo is abstention, so a voter who cannot rank the two candidates does not vote at all. This is a natural assumption given the evidence that poorly informed voter do not turnout, as discussed above.

Another natural choice would be that a voter who cannot rank the two candidates votes for the candidate of a specific party, as in the "standing decision" model of party identification. In section 3.3, I show that this status quo leads to a voter who is observationally equivalent to a Bayesian with "extra" leanings toward the favored party.

This is not the first paper to look at ambiguity aversion in voter behavior. Ghirardato and Katz (2002) study a formal model of a voter who can costlessly vote in a two-alternative election. They show that a Bayesian generically has a strict preference to turnout, but that a voter who is ambiguity averse may have a strict preference to abstain for an open set of parameter values. They also work with a multi-prior representation of preferences, but they use Gilboa and Schmeidler's (1989) maximin expected utility model. In this model, the voter has complete preferences, so 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. In this paper, by contrast, the possibility of abstention is trivial, and the main purpose of the paper is to integrate ambiguity-averse voters into a model of campaign strategy.

I show that, when voters are ambiguity averse and may abstain, a candidate faces a tradeoff between mobilizing voters who lean her way, and persuading voters who favor the other candidate to switch. It turns out that the greatest contribution to a candidate's vote margin comes from attempting to mobilize voters who are relatively unresponsive to new information. This result provides a strategic foundation for minimal effects.

In addition to giving these foundations for minimal effects, the model makes several testable predictions. For example, the model predicts that, the lower is the initial level of voter information, the more partisan will be the targeted voters. In other words, candidates "play to their base" in low information environments, while they "play to the center" in high information environments. The model also makes predictions about the choice between positive and negative campaigns, and about the relationship between voter polarization and the scale of campaigns.

Of course, these predictions are based on a specific model. The conclusion discusses some assumptions that might be problematic, and discusses opportunities for future research.

1 Empirical Evidence on Ambiguity and Incompleteness

Of course, the model will be more valuable the greater is the evidence for the psychological traits it embodies. What is the evidence for ambiguity aversion in political behavior?

1.1 Evidence for Ambiguity Aversion

The seminal work showing that some individuals dislike acting on ambiguous information was Daniell Ellsberg's (1961) urn experiments. He showed 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 thought 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, since the two expectations would have to sum to 100. (Risk aversion cannot account for these results since there are only 2 prizes.)

Ellsberg's result has been replicated many times, often with real payoffs. (Camerer and Weber (1992) review 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 are averse to ambiguous bets about the outcomes of political and sporting events. They also find that the degree of ambiguity aversion varies with the level of the subject's background knowledge—the aversion to ambiguous bets is greater the less informed the subject is about the issue. 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". Applied to elections, both the competence hypothesis and the comparative ignorance hypothesis imply that voters will be more ambiguity averse in elections for less visible offices, since these elections are characterized by less background information.

Several important studies of voter behavior suggest that ambiguity may be an important attitude for some voters. A classic illustration of the importance of ambiguity in voter decision making is given by Bartels's (1988) study of presidential primaries. He finds that survey respondents are less likely to place candidates on issue scales the less informed they are, and that respondents who refuse to place candidates on issue scales are less likely to participate in politics. Furthermore, he finds that one of the main things that happens over the course of the campaign is that citizens learn more about the candidates, and so 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). These findings have been reinforced by the observational study of Lassen (2005) and the experimental study of Horiuchi, Imai and Taniguchi (2005), both of whom find that citizens who are randomly selected to receive information about the candidates in an election are more likely to vote than are citizens who do not get the information.

1.2 Evidence for Incomplete Preferences

We've seen that ambiguity aversion is a robust feature of human decision making, and that is plays a role in the specific context of political decision making. But our model will go beyond just ambiguity aversion—the specific expression of his ambiguity aversion was incomplete preferences over the candidates. Next I review the extant empirical evidence that links these two features of decision-making.

A few papers in political science have tried to directly test the hypothesis that voters have complete, transitive preferences. The seminal work in this literature is Brady and Ansolabehere (1989) (also see the follow-up work by Hansen (1998) and Alvarez and Kiewiet (2005)). They use data from a poll of 1034 CA adults in March 1976 and a self-administered questionnaire completed in March 1983 by 170 University of California at Berkeley students, faculty, and staff. In each case, the respondents were asked about their preferences over each pair of candidates in the then current Democratic presidential primary contest. They find that between 60% and 70% of the respondents have fully rational preferences and less than 10% have intransitive strict preferences.[1] So just under 1/3 of the respondents have strict preferences that are transitive, but not negatively transitive. They write:

> We do not think that this means that people are irrational about their interests; rather it means that people sometimes do not know enough about their interests to make a commitment one way or another. [p. 158]

To test this idea about voter information, they constructed measures of uncertainty based on self-reported knowledge and on willingness to place the candidates on issue scales. They find that more uncertainty leads to less likelihood of negative transitivity. To give a sense of the substantive magnitude of this information effect, they simulate the preferences over the Mondale–Hart pair under two counterfactual scenarios: each voter's knowledge of Hart raised to that of Mondale (HIGH), and each voter's knowledge of Mondale reduced to that of Hart (LOW). The results are:

[1] More specifically, they show that 60% to 70% have negatively transitive strict preferences. Since the strict preferences are asymmetric by the construction of the surveys, this means that the weak preferences are complete and transitive. See Proposition 2.2 of Kreps (1990).

	LOW	Actual	HIGH
Prefer Hart	40.1%	21.6	38.3
Prefer Mondale	24.1	62.3	58.6
No Preference	35.8	16.1	3.1

Thus the effect of knowledge on full rationality is substantively large, with a swing from 35.8% to 3.1% complete. Also, this simulation shows the different effects of ambiguity aversion and risk aversion. Moving from LOW to actual to HIGH monotonically increases overall information, and the probability of no preference decreases monotonically, as implied by ambiguity aversion. On the other hand, the same change has a nonmonotonic effect on the *relative* information advantage of Mondale. And just as risk aversion predicts, Mondale's relative popularity is nonmonotonic.

Brady and Ansolabehere force the subjects' preferences to be complete by declaring them indifferent when they say "I'm really not sure" (these are the exact words from their 1983 survey).[2] However, they recognize that what they call indifference may be a form of incompleteness:

> When indicating indifference a person may be making an equal commitment to each alternative or indicating insufficient knowledge to make a commitment. Indifference may indicate equal preference or it may indicate no preference. Intransitive preferences, then, suggest a deep-seated confusion about one's interests whereas intransitive indifference may only indicate temporary ignorance about one's interest. (p. 147)

On semantic grounds, it seems to me that incompleteness better fits their question than does indifference. However, this is not an issue to be settled by arguing about words—it can be resolved by more experiments, experiments that allow subjects to express both indifference or incompleteness (as in the Curley, Yates and Abrams (1986) study mentioned above). I adopt the incompleteness interpretation because it leads to a tractable model that explicitly links the degree of incompleteness to the informational level.

Brady and Ansolabehere are following the dominant tradition in revealed preference theory when they impose completeness. However, foundations for incomplete preferences have recently 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

[2]See Mandler (2005) on the relationship between incompleteness and intransitivity with this convention.

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.)

2 The Model

2.1 Candidates and Voters

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 $1 \leq n \leq N$. These groups are internally homogeneous.

A voter's payoff to voting for L is normalized to 0, while her payoff to voting for R is $\theta \in \mathbb{R}$. The voter does not know θ. Since the treatment of this uncertainty is the primary novelty of the paper, we defer its discussion to the next subsection.

The payoff from abstaining is between the two payoffs of voting for one or the other candidate; for concreteness, we take it to be $\delta\theta$, where $0 < \delta < 1$. These payoffs are independent across voters. Voting is costless, so if the voter knew θ, then she would have a dominant strategy to vote for R if $\theta > 0$, would have a dominant strategy to vote for L if $\theta < 0$, and would be indifferent over all three options if $\theta = 0$.

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.

Notice that our voter is risk neutral—as emphasized in the introduction, risk aversion on its own could not produce our results.

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. Adding a cost of voting, constant across elections, would not affect our results.

Each candidate can send messages targeted to specific groups. If a group is targeted, each member gets a signal on θ. These signals are "independent" across the voters within a group.[3] 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.

2.2 Voter Beliefs

A 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 might have incomplete preferences. Following Bewley (2002), 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 set of priors is a singleton, then the voter has complete preferences.

2.3 A Voter's Decision Rule

As is typical in models with a continuum of voters, equilibrium does not pin down voter behavior, since no voter is ever pivotal. We follow standard practice and simply assume that voters use a strategy which would be undominated in a large, finite electorate. We elaborate on this a bit more than is usual, since the decision theory is nonstandard.

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] We follow the usual practice of extending the idea of independence to a continuum of random variables by making the law of large numbers into a definition.

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.

The appendix outlines the decision-theoretic background of this type of multi-prior model. Readers who are not interested in axiomatic decision theory can skip that discussion without any loss of continuity.

2.4 Information and Updating

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, §3.3) treatment of information, our voters also receive a signal with known likelihood, and beliefs are updated prior-by-prior, using Bayes's rule.[4]

Formally, 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 $\mathcal{P} \mid s$. 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 $\mathcal{P} \mid s$.

As Ghirardato and Katz (2002) point out, Bayesian updating prior-by-prior is inappropriate with the maximin EU form of multiprior utility—that updating rule can lead to time-inconsistency in that framework. Since Bewley's model keeps the independence axiom, this time consistency problem does not arise for us, and prior-by-prior updating is appropriate.

[4]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.

Most applications of Bayesian learning by voters use the particularly simple normal learning model (Zechman, 1979; Achen, 1992; Alvarez, 1997). The next section adapts the normal learning model to our multi-prior context, and derives some preliminary results that will help characterize the equilibrium of the game between the two candidates.

3 Preliminary Results

3.1 Normal Learning

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

$$\theta \mid s \sim \mathcal{N}\left(m, \lambda\sigma^2\right),$$

where

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

and

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

We will consider a simple extension of this model to the case of multiple priors. Each member of group n has beliefs $\mathcal{P}_n$ about the utility differential that arise as follows: Let $\mathcal{G}_n$ be a finite set of normal distributions with common variance σ_n^2, and let $\mathcal{P}_n$ be the convex hull of $\mathcal{G}_n$. Define the *upper expectation* by $\overline{\mu}_n = \sup_{p\in\mathcal{P}_n} \mathbb{E}_p(\theta) = \max_{p\in\mathcal{G}_n} \mathbb{E}_p(\theta)$ and the *lower expectation* by $\underline{\mu}_n = \inf_{p\in\mathcal{P}_n} \mathbb{E}_p(\theta) = \min_{p\in\mathcal{G}_n} \mathbb{E}_p(\theta)$. These two numbers provide a summary of the voter's priors that is sufficient to study her decisions. Although the prior distributions are common within groups, the actual valances are "independent".

The two parties have common beliefs over the signals. These are normal, with mean $\alpha_n = \frac{1}{2}\overline{\mu}_n + \frac{1}{2}\underline{\mu}_n$. Since all of the normal distributions in the set have the same variance σ_n^2, this means that the signals are distributed $\mathcal{N}(\alpha_n, \sigma_n^2 + \sigma_\epsilon^2)$. 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, $\Delta_n = \overline{\mu}_n - \alpha_n = \alpha_n - \underline{\mu}_n$. The midpoint measures partisanship, while the radius measures ambiguity aversion.

We will suppress the n subscripts when no confusion will result.

Lemma 1 *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 is because when the prior is a convex combination of simpler priors, then the posterior is a convex combination of the simpler posteriors. (The weights change, in general.)

3.2 Voter Decisions Revisited

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

$$\begin{array}{rl} \text{vote } R \text{ if} & \underline{m}_n > 0 \\ \text{vote } L \text{ if} & \overline{m}_n < 0 \\ \text{abstain} & \text{otherwise} \end{array}$$

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

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

or

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

Similarly, the voter votes R if

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

We collect these results in a proposition.

Proposition 1 *The voter chooses*

$$\begin{array}{ll} L & \textit{if } s < -\frac{\sigma_\epsilon^2}{\sigma_n^2}\overline{\mu}_n \\ \textit{abstain} & \textit{if } -\frac{\sigma_\epsilon^2}{\sigma_n^2}\overline{\mu}_n \leq s \leq -\frac{\sigma_\epsilon^2}{\sigma_n^2}\underline{\mu}_n \\ R & \textit{if } s > -\frac{\sigma_\epsilon^2}{\sigma_n^2}\underline{\mu}_n. \end{array}$$

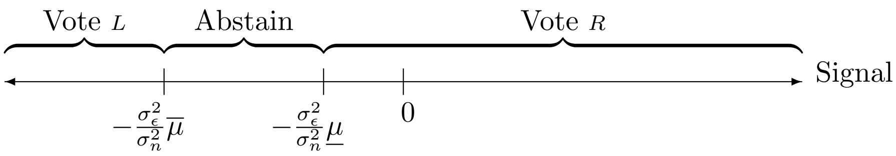

Figure 1: The voter has upper expectation $\overline{\mu}$ and lower expectation $\underline{\mu}$. These are both positive, indicating a voter who leans towards the R candidate. He votes L for low values of the signal, abstains for moderate values, and votes R for high values of the signal.

Given this decision rule and a probability distribution over s, we can calculate the probabilities of all three choices. Figure 1 gives a graphical depiction of the voting rule for a specific example. To calculate the probabilities, we would calculate the probabilities of the indicated intervals, using the assumed normal distribution of the signals.

An important implication of the decision rule 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 $\sigma^2 = \sigma_\epsilon^2$. 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.

3.3 The "Standing Decision" Model

As mentioned in the introduction, another plausible choice for a voter's status-quo action is to vote for one of the two parties. This reflects the classical idea that party identification represents a standing decision to vote for the party unless race-specific factors intervene. It turns out that such a voter is observationally equivalent to a voter who is not ambiguity averse at all, but leans more strongly toward his favored party.

To see this consider a voter with upper expectation $\overline{\mu}$ and lower expectation $\underline{\mu}$, and who votes for R when he cannot rank the two candidates. His signal-contingent decision rule is to vote

$$\begin{array}{ll} L & \text{if } s < -\frac{\sigma_\epsilon^2}{\sigma_n^2}\overline{\mu}_n \\ R & \text{if } s > -\frac{\sigma_\epsilon^2}{\sigma_n^2}\overline{\mu}_n. \end{array}$$

But this is identical to the decision rule of a Bayesian with prior mean $\overline{\mu}$. Thus allowing for status-quo actions consistent with the standing decision model does not add anything, since we already allow voters to be ambiguity neutral.

4 Equilibrium

4.1 Definition of 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.[5] 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, given the voter decision rule above.

4.2 Characterization of Pure-Strategy Equilibrium

The next result will allow us to derive a sharp characterization of pure-strategy equilibrium. Say that a profile is *specialized* if each group is targeted by at most one candidate. Formally,

Definition 1 *An equilibrium is* specialized *if $n \in \mathcal{L}$ implies $n \notin \mathcal{R}$ and $n' \in \mathcal{R}$ implies $n' \notin \mathcal{L}$.*

This may seem restrictive, but the next result says it is not, at least if we restrict attention to pure strategies.

Proposition 2 *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.*

The idea of the proof is simple. Except for knife-edged cases, one candidate does better when group n gets only one signal instead of two. Thus a profile in which both candidates target n is unstable—the candidate who does better with only one signal will deviate.

[5] For any set X, the cardinality of X is denoted by $|X|$.

To use this result, we need to derive the probabilities. We saw above that a voter votes L if and only if

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

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

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

Using the identity $\overline{\mu}_n = \alpha_n + \Delta_n$, this probability becomes

$$\Pr\left(s < -\frac{\sigma_\epsilon^2}{\sigma_n^2}\overline{\mu}_n\right) = \Phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_n^2} + 1)\alpha_n - \frac{\sigma_\epsilon^2}{\sigma_n^2}\Delta_n}{\sqrt{\sigma_n^2 + \sigma_\epsilon^2}}\right).$$

Similarly, the probability she votes for R is

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

The comparative static results will follow from analyzing these expression.

Given the specialization result, we can give a sharp characterization of equilibrium. To do so, we need a few definitions. Let ρ_n be R's net plurality in group n when that group gets one signal:

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

When deciding whom to target, the candidate compares this plurality with that the group gives her without any information, denoted γ_n. This is

$$\gamma_n = \begin{cases} 1 & \text{if } \alpha - \Delta > 0 \\ 0 & \text{if } \alpha + \Delta \geq 0 \geq \alpha - \Delta \\ -1 & \text{if } 0 > \alpha + \Delta \end{cases}.$$

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 3 *Assume there are no ties. Then there is at most one specialized equilibrium. In such an 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. $n \in \mathcal{R}$ *implies* $\delta_n > 0$*, and*

3. $n \in \mathcal{R}$ *and* $n' \notin \mathcal{R}$ *imply that* $\beta_n \delta_n > \beta_{n'} \delta_{n'}$*.*

The equilibrium strategy of L is the same, but with the inequalities reversed in (2) and (3).

This proposition is illustrated in the following example. Figure 2 plots the net expected contribution to R's plurality for giving one signal to each of four groups, and labels the sets of targeted groups, $\mathcal{R}$ and $\mathcal{L}$. Assume that $M = 2$. The R candidate targets groups 1 and 2. She would like to also target group 3, but her budget does not allow it, reflecting condition 1 of the proposition. The L candidate targets group 4. She can afford to target an additional group, but chooses not to. This is because all of the other groups move towards R in expectation, and her decision reflects condition 2 of the proposition. Thus conditions 1 and 2 give the two different ways that targeting is limited in equilibrium. Finally, condition 3 shows up in R's choices: she could switch from 1 or 2 and target group 3, but this would reduce her expected plurality. Thus each candidate is targeting those groups that add the most to their plurality.

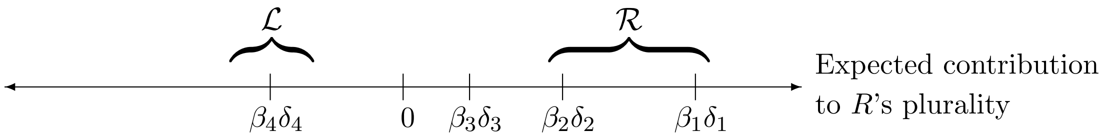

Figure 2: Group n's expected contribution to R's plurality when they get one signal is plotted as $\beta_n \delta_n$. The L candidate targets group 4, and the R candidate targets groups 1 and 2. Proposition 3 says that this is the only possible specialized equilibrium.

4.3 An Algorithm

Notice that Figure 2 illustrates an algorithm for finding a specialized equilibrium. 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 $\delta_n > 0$. This will produce a profile satisfying the conditions listed in Proposition 3.

When does this algorithm find an actual equilibrium? Well, the only thing that can go wrong is that one candidate wants to deviate and target a group already targeted by the other party. This possibility depends on how the expected plurality varies with the signal variance. The next Proposition gives a sufficient condition for this deviation to not be profitable. As a preliminary step, we state a lemma that is useful both in that proposition and in the next section.

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.*

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.

Using the description of the probabilities above, we have the following:

Proposition 4 *If the algorithm above produces a profile in which neither candidate targets a group n with α_n in the intervals $(-(\lambda+1)\Delta, -\Delta)$ or $(\Delta, (\lambda+1)\Delta)$, then that profile is a pure-strategy equilibrium.*

This condition is sufficient, but not necessary. Even if a group in one of these intervals is targeted, the other party might choose not to respond. After all, we've only established that the incremental payoff to responding is positive, not that it is greater than the alternatives.

5 Who is Targeted?

Some predictions about which groups are targeted are immediate from the characterization in the previous section. For example, all else equal, a candidate prefers to target the larger of two groups. To draw more refined predictions, we must use look at the plurality functions in more detail. The basic lesson is that the best groups to target are large and are on the verge of supporting the party, without too great a probability of being driven away or driven to the other party. The following results flesh out this claim by listing some specific comparative static results, followed by their interpretations as empirical predictions.

5.1 Comparative Statics

To confront the empirical work on campaign effects, a crucial question is the relative efficacy of attempts to mobilize voters who will not otherwise vote versus persuading voters who currently plan to vote for the opponent to change their plan. This section explores the tradeoffs that underlie this choice. In principle, proposition 3 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 1

1. *Consider 2 groups that differ only in their partisan leanings, (α), 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.*
2. *Consider 2 groups that differ only in their partisan leanings, (α), both of which are in $(-\infty, -\Delta]$. If the R candidate can target only one of the two groups, it will choose the less partisan one.*

Thus when a candidate targets groups that lean her way, the best groups to target are most partisan among those who are not yet planning to vote, while when targeting groups that initially plan to vote for the other candidate, the best groups to target are the least partisan.

We've just seen that, when targeting abstainers who lean in her direction, a candidate does best by targeting a group just on the threshold of turning out. Given a choice between two such groups, which will be chosen?

Fact 2 *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.*

With limited resources, parties first "secure their base"—targeting the most ideologically sympathetic voters who are not yet convinced to turnout. Intuitively, this is because members of that group are less likely to be inadvertently mobilized for the other candidate.

Fact 3 *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.

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

Fact 4 *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.

Notice that Facts 4 and 5 refer to pure-strategy equilibria in which existence is problematic, according to Proposition 4. However, that proposition does not say that such equilibria are impossible.

Furthermore, the candidates' incentives to target initial abstainers are stronger than their incentives to target partisans of the other candidate, which mitigates against the nonexistence problem.

Fact 5 *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.

5.2 Empirical Predictions

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. Recall that the radius, $\Delta = \overline{\mu} - \alpha$, measures how averse the voter is to the ambiguity he perceives. When the radius is small, there is a good chance that giving information 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.

5.2.1 Minimal Effects

Minimal effects come from a combination of Facts 2 and 3. The candidates want to target voters who are unresponsive in the sense of being a great distance away from being mobilized for the other party, and conditional on finding such voters, they want the ones who are least moved by new information.

5.2.2 Securing the Base vs. Swing Voters

We can also interpret Fact 2 in terms of a comparison of different types of elections. Based on Heath and Tversky's (1991) competence hypothesis, we expect voters to be more ambiguity averse the worse is their background knowledge about the race. Thus, 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).

5.2.3 Positive vs. Negative Campaigns

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 or when the electorate is more polarized ideologically. This follows from Fact 5, which implies that candidates will turn towards a negative campaign only if there are no otherwise similar latent partisans.

5.2.4 Equilibria when funds have alternative uses

We can derive some more empirical predictions by slightly augmenting the model. Let the candidates maximize $P + \eta y$, where y is the amount of cash not spent on the campaign and η is a parameter that reflects the relative importance of the plurality and cash on hand. This cash could be valued for several reasons. First, the candidate could anticipate running additional races in the future, in which case η is the shadow value of a campaign warchest. Second, the candidate could donate funds to party leaders. The leaders would use this cash to help more marginal candidates, and then reward the donor with good committee assignments. REFS

It's not too hard to show that Proposition 3 carries over to this case, with the only change being that condition 2 becomes $\beta_n \delta_n \geq \eta c$. Thus the characterization carries over except that groups with small net contributions to the plurality are not targeted, even when the candidate can afford to do so.

Which groups have small net contributions to plurality? There are two examples. Groups that are close to being evenly balanced between the parties ($\alpha \approx 0$) will have a small impact on the plurality because voters are close to equally likely to move toward either party. Groups that are strongly in favor of one of the parties ($|\alpha| \gg 0$) will have a small impact on the plurality because voters are likely to vote for their favored party no matter what happens during the campaign.

Now consider the effects of increasing voter polarization. In a minimally polarized electorate, all groups have α close to zero, and the scale of campaigns will be small. If we increase polarization by increasing the absolute values of some of the α, then more groups will have $\beta_n \delta_n \geq \eta c$, and the scale of campaigns will increase. However, if we continue increasing polarization to the point where most groups have αs that are very large in absolute value, then the scale of campaigns will decrease again. Thus we can have a nonmonotonic relationship between voter polarization and the scale of campaigns.

5.3 Comparison to "Swing Voter" Models

The above results differ greatly from the kinds of results we see in more traditional models. The campaign stage of the model is related to work by Lindbeck and Weibull (1987) and Dixit and Londregan (1996). In these models, parties target their resources to "swing voters"—voters who are 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.[6] They also find that 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

We have seen that adding ambiguity-aversion to an otherwise standard model of targeting in an electoral campaign has a big impact on the predictions. In this framework, voters will not turnout if they perceive that it's possible for either candidate to be better in expectation. This opens up new possibilities for candidate strategies. In particular, they can mobilize voters by giving them unbiased information. This is particularly effective when the voter leans strongly toward the party, and is not too responsive to new information. Such a voter is overwhelmingly more likely to move in the right direction if he moves at all. A more responsive voter, on the other hand, has a real chance of moving the "wrong" way, and enough so that he actually turns out for the other party. This gives the parties an incentive to target voters who are relatively unresponsive to new information, and provides a strategic explanation for the minimal effects of campaigns.

The key to getting these results is that the voters do not conform to the

[6] 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.

standard model of Bayesian rationality. There are many other deviations from the standard model of voters to consider—limited memory and biased information processing, for example. There are also other important models of elite behavior with which these models could be combined. This suggests a potentially very fruitful research agenda linking traditional formal theory, voter behavior, and behavioral economics.

More concretely, the model considered in this paper is based on several specific assumptions in addition to just ambiguity aversion. In particular, the normal distribution puts a great deal of structure on the learning process. Alternative models, such as Bernoulli uncertainty or normal distributions with unknown variance, could introduce new effects, such as strategies aimed at *increasing* ambiguity.

Finally, we might want a notion of "persuasion" that is more directional—the L candidate undertakes some action attempting to move the voter in the left direction. While the model captures some of this, since the voter's actions do drift left, we are still missing something. In particular, there is no directionality in the drift in beliefs. This is because we are sticking close enough to the Bayesian model to activate the martingale property of beliefs. Moving further away from the Bayesian case is left for future research. I don't think that this will affect the results, since more directionality would be equally useful for a candidate engaged in mobilization/demobilization.

A 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 the standard Anscombe and Aumann (1963) axioms, except that completeness is weakened to hold only for constant acts.[7] 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. We have:

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 the case with no uncertainty. Then preferences can be represented by a utility function if and only if they are complete and transitive.[8] 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). This is because any reflexive and transitive order is equal to the intersection of all complete and transitive orders that contain it. (Zorn's lemma implies that the latter set is nonempty regardless of the cardinality of the set of alternatives.)

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.

[7] 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.

[8] Modulo topology.

A comparison to social choice theory may help 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, and treating that "society" as a description of an individual's preferences.

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 (2002) builds on 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.

B Proofs

Proof of Lemma 1

This follows from a standard result in statistics (Williams, 2001, p. 213):

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

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

□

Proof of Proposition 2 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. $\psi_n < \rho_n$
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. □

Proof of Proposition 3 First we show that an equilibrium strategy for R must satisfy the three properties. Property 1 is just feasibility. If 3 is not satisfied, then R can increase its plurality by ceasing to target group n. Finally, 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.

Now consider two sets $\mathcal{R} \subset \mathcal{R}'$ that both satisfy all three properties. (This is possible because the budget constraint is a weak inequality rather than an equality.) By property 2, the groups in $\mathcal{R}' \setminus \mathcal{R}$ all make a positive contribution to R's plurality. Thus no such 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 3 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. $\square$

Proof of Lemma 2 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 $\alpha < -\Delta$. $\square$

Proof of Proposition 4 To simplify the notation, we define

$$\phi^+ = \phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_n^2} + 1)\alpha_n + \frac{\sigma_\epsilon^2}{\sigma_n^2}\Delta}{\sqrt{\sigma_\epsilon^2 + \sigma_n^2}}\right)$$

and

$$\phi^- = \phi\left(\frac{-(\frac{\sigma_\epsilon^2}{\sigma_n^2} + 1)\alpha_n - \frac{\sigma_\epsilon^2}{\sigma_n^2}\Delta}{\sqrt{\sigma_\epsilon^2 + \sigma_n^2}}\right).$$

We have

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

If $\alpha > 0$, this expression is positive if

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

or if

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

Thus an extra signal decreases the R plurality if $\alpha > 0$ and $\alpha > (\lambda+1)\Delta$, and an extra signal increases the R plurality if $\alpha \in (0, (\lambda+1)\Delta)$. Thus we have the following three cases:

1. If R targets an $\alpha \in (0, \Delta)$, then L does not respond.
2. If L targets an $\alpha > (\lambda+1)\Delta$, then R does not respond.
3. If L targets an $\alpha \in (\Delta, (\lambda+1)\Delta)$, then R might want to respond.

Only in the third case can the algorithm fail. □

Proof of Fact 2 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 4 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 5 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 3 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.$\square$

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.

Alvarez, R. Michael and D. Roderick Kiewiet. 2005. "Rationality and the Recall." Unpublished paper, Caltech.

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.

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

Bartels, Larry M. 1986. "Issue Voting Under Uncertainty: An Empiricial Test." *American Journal of Political Science* 30(4):709–728.

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.

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. 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.

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

Hansen, John Mark. 1998. "Individuals, Institutions, and Public Preferences over Public Finance." *American Political Science Review* 92(3):513–531.

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

Horiuchi, Yusaku, Kosuke Imai and Naoko Taniguchi. 2005. "Designing and Analysing Randomized Experiments." Unpublished paper, Princeton University.

Kreps, David M. 1990. *A Course in Microeconomic Theory.* Princeton: Princeton University Press.

Lassen, David Dreyer. 2005. "The Effect of Information on Voter Turnout: Evidence from a Natural Experiment." *American Journal of Political Science* 49(1):103–118.

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

Mandler, Michael. 2005. "Incomplete Preferences and Rational Intransitivity of Choice." *Games and Economic Behavior* 50(2):255–277.

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

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.