A Unified Model of Spatial Voting*

Nathan A. Collins
Santa Fe Institute
1399 Hyde Park Road
Santa Fe, NM 87501
nac@santafe.edu

September 7, 2010

Abstract

Experimental research shows that while most voters have some form of spatial preferences, individuals differ in the type of spatial preferences they have: many voters prefer candidates closer to themselves in a policy space (proximity voting), others prefer candidates that are simply on the same side of an issue as themselves (directional voting), and still others prefer those who will move policy closest to them (discounted proximity voting). No existing theory explains this variation. I propose a theory based on the idea that that people categorize candidates and have preferences defined over categories. As a voter gains political experience, she makes finer distinctions between candidates, and the set of categories grows. In this way, voters move from either-or conceptions of politics that approximate directional preferences toward more detailed conceptions consistent with proximity preferences, with some cases approximating discounted proximity voting as well. I show that the categorization model accurately predicts the observed frequencies of different voting types as well as observed comparative statics. I also show that the comparative statics results explain observed differences in the distribution of voting types across different policy areas.

*I thank John Bullock, Jonathan Bendor, Matt Levendusky, and Ken Shotts for helpful comments.

Introduction

Spatial models of voting have been an important component of research on voter decision making and candidate behavior in the half-century since Downs (1957) proposed his version. In that time, two different classes of spatial model have emerged: proximity models (e.g., Downs 1957; Grofman 1985) suppose that voters prefer candidates whose policy positions are closer to their own, while directional models (e.g., Matthews 1979; Rabinowitz and MacDonald 1989) suppose voters care primarily about whether candidates are on the same side of an issue as they are. A long-running debate exists regarding whether voters are proximity or directional voters (Lewis and King 1999; Merrill and Grofman 1997, 1999; Macdonald, Rabinowitz and Listhaug 1998, 2001; Rabinowitz and MacDonald 1989; Westholm 1997, 2001), but recent experimental work (Tomz and Van Houweling 2008; Claasen 2009) shows that many people are proximity voters, some are directional voters, and still others are so-called discounted proximity voters. Furthermore, there is variation in the distribution of voting types across different policy areas (Claasen 2009). For instance, behavior consistent with the directional model is more frequent on the abortion issue than on others. The experimental findings pose a stark question: what accounts for the qualitative variation in individuals' preferences over political candidates? No existing theory answers this question.

The central contribution of this paper is to explain this variation in terms of categorization. Different types of spatial voting arise when voters break the policy space into different numbers of more or less narrowly defined categories and have preferences defined over the categories rather than over individual points in a policy space. I show that the categorization-based model accurately predicts the observed prevalences of different spatial voting models as well as comparative statics consistent with experimental observations (see Tomz and Van Houweling 2008). Furthermore, I show that the comparative statics results actually explain observed differences across issue areas (see Claasen 2009).

There are several ideas at work here. The first, that people categorize objects in the world, is

well established, whether the objects are pictures in a laboratory experiment or candidates in an upcoming election. Categorization is a basic, necessary feature of human psychology that makes inference possible and allows people to simplify and organize the world (Estes 1994; Rosch 1978; Smith 1990). Economists have begun to use categorization to explain phenomena such as group decision making, stock market pricing, and stereotyping (Hong and Page 2001; Mullainathan 2002; Mullainathan, Schwartzstein and Shleifer 2006; Fryer and Jackson 2008).

The second idea is that people have preferences over categories rather than individual policy positions in a policy space. In order to choose between candidates, a voter must make some sort of inference or prediction, e.g., a prediction of how happy she will be if the candidate wins. As with any other prediction task, categorization simplifies this process by summarizing and organizing one's experience with similar politicians. Individuals will, however, differ in how they categorize. Voters who use many categories to organize these predictions make fine-grained distinctions that closely approximate proximity preferences. Near the other extreme, voters who use just two categories mainly differentiate between their side and the other, closely approximating directional preferences. As I will demonstrate, intermediate cases sometimes approximate discounted-proximity preferences, and in general there is a range of preference types, each corresponding to a different number and arrangement of categories.

The remaining issue is why voters would differ in the number and arrangement of their categories. There are a variety of possible explanations, including (perhaps) the reduced costs of making distinctions that might accompany greater educational attainment or political sophistication. However, an important feature of categorization is that it depends on experience, i.e., the number of times a person has performed a given task. For example, people make fewer classification errors as they gain experience (e.g. Anderson 1991; Nosofsky, Gluck, Palmieri, McKinley and Glauthier 1994), and there is good reason to believe that people begin with simple categorization schemes and increase complexity as they gain experience (see Love, Medin and Gureckis 2004). These observations mean that voters will make finer distinctions, use more categories, and have preferences that more closely approximate proximity preferences as they gain experience, i.e., as they observe

and categorize more political figures. Furthermore, the experience-based categorization model I develop does a better job of explaining the Tomz and Van Houweling (2008) and Claasen (2009) experimental observations, particularly with regard to the observed comparative statics.

There are a few potential criticisms that should be addressed up front. First, it is well known that voters think in terms of political parties (e.g., Jackman and Sniderman 2002; Lodge and Hamill 1986), so one may be tempted to think that voters think in terms of two categories, one for each major party. If that were the case, however, we could not explain any of experimental observations. In particular, it is not clear how one could derive comparative statics predictions consistent with the experimental findings without allowing a variable number of categories. Furthermore, if voters thought entirely in terms of parties, spatial voting itself would not make any sense — a conclusion at odds with both experimental and survey research on spatial preferences.

Second, I derive many of the results of this paper using computer simulations of a mathematical model. I do so because the model is adaptive in nature and categorizations schemes in the model depend on a moderate but not large number of stochastically-determined experiences (voters' observations of politicians). As a result, one cannot make use of limiting cases or large-number approximations, so it is a challenge at best to derive analytical predictions. If that is not a satisfying defense of the approach, then one should note two other attributes of the model: the model does a very good job of predicting what Tomz and Van Houweling (2008) and Claasen (2009) observed, and *it is the only model that explains these observations at all.*

Third, many view the debate between proximity and directional voting as resolved, and many of those appear to view the proximity model as having been vindicated. The experimental results contradict this position. Some people are proximity voters and others are directional voters, and indeed there is variation based on the issue under consideration. The data, therefore, are posing questions that existing models cannot answer. This means — in no uncertain terms — that we do not yet fully understand spatial voting and that we should not view the spatial voting debate as settled.

Finally, some have voiced concerns regarding the external validity of the Tomz and Van Houwel-

ing (2008) and Claasen (2009) experiments because they represent candidates' positions numerically rather than verbally. However, it is a fundamental assumption of all spatial voting models and most categorization models that people think — at least subconsciously — in terms of spatial and therefore numerical representations. In other words, a conversion to a numerical representation must happen at some stage, and it is probably not that important whether it happens at the presentation or cognition stages. (I will have a bit more to say on this point below.)

In the next section, I review the standard spatial voting models and observational and experimental research on the topic. I then briefly state the model, which is a simplified version of the SUSTAIN model of categorization (Love, Medin and Gureckis 2004) combined with preferences over categories. The model predicts a range of voting types with proximity, directional voting, and discounted proximity voting as special cases. I then show that the model predicts behavior consistent with the Tomz and Van Houweling (2008) experiments under a fairly wide range of assumptions about the distribution of experience levels in the population. Because I use a computational model to develop these predictions, I then demonstrate that the model's predictions are robust to changes in parameters and other modifications. I also show that while alternative, more economically-motivated categorization models produce somewhat similar results, they do no better at predicting the overall prevalences of different voting types and are especially bad at reproducing observed comparative statics. Finally, I discuss Claasen's 2009 results, which are interesting because they indicate differences in the distribution of proximity and directional voters that depend on the issue under consideration. I conclude with a summary and a discussion of additional testable predictions.

Approaches to Spatial Voting

Spatial voting is the idea that candidates and voters have positions in a policy space and that these positions determine the voter's preferences. Formally, there is a policy space $\mathbb{P}$, typically a continuous subset of $\mathbb{R}^n$. For simplicity, I specialize to a one-dimensional policy space, e.g., a liberal-conservative dimension. Each voter has an ideal point $v \in \mathbb{P}$ and candidate i has an ideal

point $c_i \in \mathbb{P}$, and the voter's and candidate's ideal points determine the voter's preferences over candidates. In keeping with most voting models, individuals vote for the candidate they prefer over other candidates. Spatial models vary in how they define the policy space, how they compare the various policy positions, and whether and how they take into account other policy-relevant information.

A voter is a *proximity voter* if she prefers candidates with positions more similar to her own (Downs 1957). Formally, a voter strictly prefers candidate 1 to candidate 2 if and only if $|v - c_1| < |v - c_2|$. *Discounted proximity voting* is similar to proximity voting, except that voters compare likely policy outcomes instead of candidates' policy positions. (In economics "discounting" usually refers to temporal discounting; here it refers to spatial discounting.) Voters may do so because they know politicians do not always get what they say they want. In one version of discounted proximity voting (Grofman 1985), a voter's expected outcome for candidate i combines c_i and the status quo Q linearly: $p_i = \alpha Q + (1 - \alpha)c_i$, $\alpha \in [0, 1]$. A voter strictly prefers candidate 1 to candidate 2 if and only if $|v - p_1| < |v - p_2|$.

A third kind of spatial voting is *directional voting*, in which voters care primarily about what side of an issue a candidate is on relative to either a neutral point N or a status quo Q. *Matthews directional voters* (Matthews 1979) prefer one side of the status quo to the other, i.e., a voter prefers candidate 1 to candidate 2 if and only if $(v - Q)(c_1 - Q) \geq 0$ and $(v - Q)(c_2 - Q) \leq 0$. (The preference is strict if we replace one of the weak inequalities with a strict inequality.) Matthews's idea was that voters would focus on the direction in which policy moved because, among other things, they would not be able to make precise judgements about where policy would end up. In a second version, *Rabinowitz-MacDonald (RM) directional voting*, a voter strictly prefers candidate 1 to candidate 2 if and only if $(v - N)(c_1 - N) > (v - N)(c_2 - N)$, where N is the *policy neutral point*. Here, the policy space represents two sides of an issue and the intensity with which candidates take sides. Voters, in turn, prefer one side to the other but also prefer candidates who take their side more intensely, i.e., more extreme candidates.

There is an ongoing and contentious debate regarding which of these theories (if any) is correct.

For the most part, this debate relies on survey data, and much of it focuses on determining the form of voters' utility functions, operationalized as voters' ratings of various candidates as functions of the candidate and voter locations on seven-point ideology scales. Survey-based research, however, forces several methodological choices, and results seem to depend largely on which choices one makes (Lewis and King 1999). For example, one must decide how to measure candidate locations. Concerned about measurement error and projection bias, Rabinowitz and MacDonald (1989) use the mean perceived candidate location for all voters, which favors the directional model. Westholm (1997) argues that only voter perceptions of these locations matter, and his approach favors the proximity model.

A second issue is that the survey approach relies on some kind of interpersonal utility comparison, for which there is no economic or psychological justification. Macdonald, Rabinowitz and Listhaug (1998) argue that problems of interpersonal comparison vanish if there are sufficiently many voters and find support for their directional model. Westholm (1997) argues strenuously against that approach, but he implicitly assumes that individuals' utility functions differ only by additive constants, which have no effect on choice in any utility model, and ideal points.

A much better approach is to measure choices in a controlled experiment rather than try to measure utilities and compare them interpersonally. Tomz and Van Houweling (2008) showed that one can use observed choices to distinguish directional, proximity, and discounted proximity voters using certain configurations of v, c_1, c_2, N, and Q to isolate one of the voting models as predicting a different choice than the other two. Using this "critical test" approach, Tomz and Van Houweling (2008) showed that 57.7 percent of their subjects were proximity voters, 27.6 percent were discounted proximity voters, and 14.7 percent were directional voters.

The Tomz and Van Houweling (2008) result is the main observation I will explain, but two others are worth mentioning. The first is research demonstrating variation across issues in the distribution of proximity versus directional voters. Claasen (2009), using an experimental approach similar to Tomz and Van Houweling, examined military spending, abortion, and general ideological dispositions (Tomz and Van Houweling focussed on health care) and found that behavior consistent

with directional voting was more prevalent on the abortion issue. I describe these results — and why they are consistent with this paper's model — in greater depth below.

Second is experimental research conducted by Lacy and Paolino (2005) which appears to favor the proximity model. A shortcoming of this research is that it reaches its conclusions by regressing candidate ratings on candidate and subject ideal points and looking for a given coefficient to be statistically significant or not. In particular, the authors find that in most cases there is a statistically significant quadratic component to voters' utility functions (as measured by the candidate ratings), which they interpret as evidence in favor of the proximity model. However, the test really only indicates that *most* voters are proximity voters, and it is not clear that one can use this approach to rule out the presence of any directional voters. Because there exists evidence that there are both proximity voters and directional voters in the population, the Lacy and Paolino results must be viewed somewhat skeptically, and I focus instead on the Tomz and Van Houweling (2008) and Claasen (2009) results.

Overview of the Model and Key Features

The motivation for the model is an observation about the difference between Matthews directional and proximity voting in one dimension. Matthews voting can be thought of as dividing the policy space into two categories, i.e., a group of candidates like the voter and one unlike the voter. Proximity voting can be thought of as dividing the policy space into a large number of categories, one for each possible candidate. If voters' preferences are defined over these categories and if voters vary in how finely they divide the policy space, then voters will vary qualitatively in the kinds of spatial preferences they have.

To be a bit more precise, we can think of a voter's set of categories as a *perceived policy space*. This is essentially a finite set of points that correspond to, for example, estimates of how happy a person will be with candidates she places in the categories. (Categorizing candidates reduces the size of the policy space from a continuous interval to a finite set, thus reducing the cognitive complexity of choosing which candidate to vote for.) Now, we assume that voters have an ideal

category in the perceived policy space and that voters have proximity preferences over categories, i.e., a person votes for the candidate she places in a category closest to her ideal category. If different voters have different numbers of categories, then they will vary in what sorts of voters they appear to be. For instance, a voter may use two categories and prefer one to the other. Assuming that these two categories roughly correspond to the two sides of the status quo, then this voter will have preferences that closely approximate Matthews directional preferences. Voters with many categories will have approximately proximity preferences, since they are more likely than not to place candidates in categories near the candidates' policy positions. As I discuss further below, the model also generates preferences consistent with discounted-proximity and Rabinowitz-MacDonald directional voting.

Now, why voters would have different numbers of categories? The essential idea is that as voters gain experience with political candidates, they become better at differentiating them and hence use more categories. At a conceptual level, there are a variety of ways to justify this idea. One could, for example, suppose that constructing categories is (mentally) costly, so that greater education or simply more time would lead voters to use more categories. I discuss the viability of such models later on. It happens, however, that a central feature of categorization is that it depends on experience: people make fewer classification errors as they gain experience (e.g. Anderson 1991; Nosofsky et al. 1994), and people appear to start with very simple categorization schemes and increase complexity as they gain experience (Love, Medin and Gureckis 2004). Since there are experimentally well-tested models of categorization that implement this idea and in order to ground the model in well-understood psychology, I focus on developing a model of voter preferences in the context of an established categorization model, SUSTAIN (Love, Medin and Gureckis 2004).

The final issues concern how people process and categorize candidates. First — and this is *absolutely vital* to understanding the model — people may have *perfect, complete* information about a candidate and yet only retain a memory of the category the candidate belongs to. In fact, this is a necessary feature of human cognition: we can not maintain a veridical representation of the world, so we simplify it by placing objects (such as politicians) in categories. The basic idea

is not revolutionary; although Lodge, McGraw and Stroh (1989) did not mention categorization, they similarly proposed that people maintain candidate evaluations rather than the thoughts and considerations that led to those evaluations. In the context of the Tomz and Van Houweling (2008) experiment, subjects observe precise policy positions but, in the model, think in terms of which category a politician belongs in.

Second, for present purposes it is not particularly important what *kind* of information people process. A concern about the Tomz and Van Houweling experiment and others like it (Claasen 2007, 2009) is that candidate positions are represented as numbers but that real policy positions are stated verbally. Although I focus on numerical representations in the present model, this is not a key feature. The key issue is whether people think spatially, and in fact this is not really in dispute, at least within the relevant literature. Indeed, it is a fundamental assumption of all spatial voting models and most categorization models that people think in terms of spatial — and therefore numerical — representations. It may be presenting positions as numbers has some effect, but the conversion to a numerical representation must happen at some point if the spatial model is to be believed at all. One should not therefore view this issue as a serious challenge to the external validity of the experiments I address here.

The Model

The model comprises three parts. The first is a policy space and a distribution of political candidates, which should be thought of as an input to the model. I normalize the policy space to the interval $[0, 1]$ for convenience. The distribution of politicians' positions on the policy space is not by itself very important. In the simulations I report below, it is the sum of two normal distributions with means $x_R = 0.3$ and $x_D = 0.7$ and variances $\sigma^2 = 0.01$. Each normal distribution represents the politicians from one of two political parties. Formally, the distribution is $\Psi(c) \propto \Phi(x_D, \sigma^2) + \Phi(x_R, \sigma^2)$, where $\Phi(x, \sigma^2)$ is a normal distribution with mean x and variance σ^2. The distribution is truncated so that the density of politicians is zero outside the interval $[0, 1]$.

The second component is the model of categorization. There are many possible models, includ-

ing more traditionally economic models in which voters pay some sort of mental cost to construct more fine-grained categories. Although developing such a model may lead to similar results, doing so would also reduce the value of this paper in several ways. First, it would unnecessarily add another categorization model to the already large set of established models. Of greater concern, it would add additional unmeasured parameters to the model. That implies greater parametric flexibility and therefore weaker conclusions when I compare the model's predictions with experimental results. I will, however, explore such models briefly near the end of this paper.

Therefore, instead of constructing a new model of categorization, I use a simplified version of the SUSTAIN model (Love, Medin and Gureckis 2004). SUSTAIN is a well-supported and generally-applicable model of categorization based on sound principles and strong psychological regularities. Furthermore, its parameters have been estimated through fits to a variety of experimental data, so that there are experimental constraints on the values of these parameters. From a purely theoretical point of view, SUSTAIN is appropriate because it has an attention mechanism that controls how finely people distinguish policy positions (or other objects), and because it has an explicit mechanism for constructing new categories when it encounters distinctly new policy positions, so that there is a natural means by which different voters would use different numbers of categories.

Voters in the model observe a sequence of politicians sampled from the politician distribution Ψ and attempt to place each in a category, which is a point k in a finite set of categories $\mathbb{K} \subset [0,1]$. To do so, voters use the similarity $H_k(c)$ between c and k as a guide:

$$H_k(c) = e^{-\lambda|c-k|}, \tag{1}$$

where λ quantifies attention, i.e., one's sensitivity to policy differences or the degree to which one makes fine distinctions between different policy positions. Note that exponentially decaying similarity is among the strongest regularities in psychology (Shepard 1987). Given the similarity function H, the following rule determines the process of categorizing politicians:

- If a voter observes a politician c and $\mathbb{K}$ is empty, she creates a first category $k_1 = c$. Now $\mathbb{K} = \{k_1\}$.

- If $\max_{k \in \mathbb{K}} H_k(c) > \tau$, where τ is an exogenous threshold similarity, the voter places politician c in the category most similar to the politician, i.e., $\arg\max_{k \in \mathbb{K}} H_k(c)$.

- Otherwise, the voter creates a new category $k' = c$, and $\mathbb{K} \rightarrow \mathbb{K} \cup \{k'\}$.

It is important to emphasize that a set of categories $\mathbb{K}$ is a finite set of points, not a partition of the policy space. Nor does the set of politicians satisfying $\max_k H_k(c) > \tau$ necessarily cover the policy space. If it did, the model would never generate new categories and would fail to explain variation in voters' preference types.

The third component is learning. A categorization decision provides two pieces of new information. First, it provides new information about the proper location of the category. If the voter placed politician c in category k, then

$$k \rightarrow (1 - \eta)\, k + \eta c, \tag{2}$$

where η is a learning rate. This rule means that k is approximately the mean position of these politicians. Second, the voter has new information about how sensitive she should be to differences in policy positions and so updates λ:

$$\lambda \rightarrow \lambda + \eta e^{-\lambda |c-k|}\, (1 - \lambda |c - k|), \tag{3}$$

where η is the same learning rate parameter. One can understand the origins of this rule as follows. We imagine a "receptive field" around a category prototype's location, with a response function $\alpha(|c - k|)$ that decays exponentially as we move away from the prototype (Love, Medin and Gureckis 2004; Shepard 1987). The receptive field has some total amount of response that must be distributed across the entire field, i.e., $\int_0^\infty \alpha(x) dx$ is fixed. Setting the fixed value to one and noting $\alpha(x) \propto \exp(-\lambda x)$, we find

$$\alpha(x) = \lambda e^{-\lambda x}.$$

Incrementally maximizing this expression at the most recent prototype-to-politician distance yields Equation (3). As a practical matter, Equation (3) tunes λ according to the typical variation of

candidates that belong in a given category (see Love, Medin and Gureckis 2004, 314-316 for further discussion). In the case that an experimenter decides what belongs in a category and gives subjects feedback regarding their choices, λ may eventually settle down. In the present model, there is no feedback, and on average the rule increases attention as a voter makes more and more observations. (This, incidentally, is why the set $\{c : H_k(c) > \tau, k \in \mathbb{K}\}$ does not in general cover the policy space.)

The final component of the model is that voters have preferences over the set of categories $\mathbb{K}$ rather than the entire policy space. The motivation for this idea is that people do not make distinctions between candidates in the same category — if two candidates are in the same category, a voter predicts the same policy outcome, happiness, etc., for both. Thus, she should be indifferent between them. Elsewhere I discuss how preferences over categories might evolve. For present purposes, I assume that voters have an ideal category $\bar{k}$ and that they have assigned each candidate c_i to a category k_i. Then, a voter strictly prefers candidate 1 to candidate 2 if and only if $|\bar{k} - k_1| < |\bar{k} - k_2|$. That is, voters have *proximity preferences over categories*. I assume that voters use the positions of k_1 and k_2 after categorizing both candidates. Since categories typically move after categorization, this choice prevents voters from having a strict preference over two candidates in the same category. As with other models, voters vote for the candidate they most prefer and randomize uniformly if they prefer two (or more) candidates equally.

Because λ increases with the number of politician observations, the set of points satisfying $H_k(c) > \tau$ for some $k \in \mathbb{K}$ decreases. Hence the number of categories also increases. Figure 1 presents utility representations of voters' preferences at two experience levels. Low-experience voters have few categories and have preferences roughly consistent with Matthews directional voting, since they prefer one side of the policy space to the others. Higher-experience voters use more categories and have preferences that begin to approximate proximity preferences.

[Figure 1 about here.]

The model also generates behavior consistent with each of the other standard preference models. Behavior consistent with Rabinowitz-MacDonald directional voting may result when the number of

categories is greater than or equal to two but still small, since it is then possible to find arrangements of categories such that a voter prefers a more extreme candidate even though there is a closer candidate on the same side of the neutral point. Similarly, there are candidates and sets of categories that generate behavior consistent with discounted proximity voting. Figure 2 presents an example; as the figure shows, the category positions k_i function similarly to perceived policy outcomes p_i. Finally, a voter may behave consistently with the proximity model when he has a small number of categories if he creates a new category for one or both candidates. For instance, a young voter may place one candidate in his ideal category and create a new category for the other candidate. In most of these cases, the candidate that ends up in the ideal category took a position closer to it than the other candidate. Thus, although this young voter is in a sense making a distinction between a candidate like himself and one not like himself, he will appear to be a proximity voter.

[Figure 2 about here.]

Predicted Prevalence of the Voting Types and Comparative Statics: Tomz and Van Houweling (2008)

Tomz and Van Houweling (2008) found that about 57.7 percent of their subjects were proximity voters, 27.6 percent were discounted proximity voters, and 14.7 percent were directional voters. As I explain above, the model does not predict any of these behaviors exactly but can generate behavior consistent with each. To make predictions about the frequency of the various preference types, I simulated the model with voters of varying experience levels, selected two candidates for each voter to choose between, and then used the simulated choices and Tomz and Van Houweling's critical-tests approach to compute the frequencies. This approach identifies scenarios — configurations of v, c_1, c_2, the neutral point N, status quo Q, and status quo weight α — that discriminate between different voting models. The scenarios Tomz and Van Houweling used for their estimations are listed in Table 1. (Note that these scenarios may be reflected, in which case the choices are also reversed). Let π_S be the fraction of (simulated) voters that choose c_2 under scenario S (or c_1 under

the reflected scenario). Tomz and Van Houweling show that

$$\begin{aligned}\pi_{\text{dir}} &= (\pi_I - \pi_{II})/(1 - 2\pi_{II}) \\ \pi_{\text{disc}} &= (\pi_{VI} - \pi_{II})/(1 - 2\pi_{II}) \\ \pi_{\text{prox}} &= 1 - \pi_{\text{dir}} - \pi_{\text{prox}}.\end{aligned} \tag{4}$$

It is important to emphasize again that the model does not predict any of these preference types; I use this estimation procedure to make predictions about the *apparent* distribution of standard voting types and to compare the model's predictions with experimental estimates.

[Table 1 about here.]

In the simulations, the neutral point N is always the policy midpoint, i.e., $N = 0.5$ by definition. Since the status quo Q does not enter the categorization-based voting model, its choice is fairly arbitrary; I set $Q = 0.4$. Similarly, α is unobservable, but only its value relative to $\alpha^* = (v - \bar{c})/(Q - \bar{c})$, where $\bar{c} = (c_1 + c_2)/2$, matters (see Tomz and Van Houweling 2008, Proposition 1). I generated α randomly for each voter — again reasonable since it does not enter the categorization model — and use cases in which $\alpha > \alpha^*$. For comparison, Tomz and Van Houweling, unable to measure or choose α, focus on cases in which $\alpha^* < 0.1$, so that most likely $\alpha > \alpha^*$. I chose each voter's ideal category by one of the voter's categories at random with uniform probability. This assumption does affect the predictions, as I discuss below in relation to the ideology comparative statics.

To generate population-level predictions, I must also make assumptions about the distribution of experience levels, i.e., the number of politician observations. I assumed that each simulated voter had n opportunities to observe politicians and that at each opportunity the probability of actually observing a politician was p. Thus, the distribution of the number of politician observations across a population was binomial. To check the robustness of the model's predictions, I sampled 100 (n, p) pairs from a uniform distribution with $n \in \{1, 2, 3, \ldots, 100\}$ and $p \in (0, 1]$. I describe results using alternative distributions below.

Regarding SUSTAIN parameters, I follow Love, Medin and Gureckis (2004) in setting $\tau = 0.5$,

$\eta = 0.2$, and initializing at $\lambda = 1$. The choice of τ is somewhat arbitrary; $\eta = 0.2$ is the estimate Love, Medin and Gureckis arrived at using a wide range of categorization tasks. (Note, however, that these categorization tasks were all laboratory experiments, so that the real-world learning rate may be lower due to less frequent repitition of a particular task.) Initializing $\lambda = 1$ means that voters initially make few distinctions among different points in the policy space. For a given number of total politician observations, initializing λ at higher levels will typically increase the frequency of proximity and discounted proximity voting. I address the predictions' sensitivity to these parameters below.

I ran 100 simulations with sampled values of (n, p) and 100,000 voters for each simulation. For each sample of the distributional parameters, I computed the estimates in Eq. (4). These estimates have non-zero variance, but with 100,000 simulated voters it is small compared to the variance of the experimental estimates, so I ignore this aspect.

Results

Before stating the results of the simulations, it is important to clarify expectations. The model should of course explain the observed mix of voting types — 57.7 percent proximity, 27.6 percent discounted-proximity, and 14.7 percent directional voters. In addition, the model should explain the observed comparative statics (see Tomz and Van Houweling 2008, 313-4). First, increasing education (from those without to those with a college degree) decreases the frequency of directional voting from 18.6 percent to 8.5 percent, while the frequency of discounted proximity voting stays roughly constant. Second, increasing partisanship (from independents and moderates to strong partisans) significantly reduces the frequency of discounted-proximity voting from about 39 percent to 22 percent while keeping the frequency of proximity voting roughly constant. Third, increasing ideological strength increases the frequency of proximity voting, but largely at the expense of discounted-proximity voting.

We should therefore expect the predictions to vary in some systematic way, i.e., as some parameter varies, the predictions should form a path in the space of possible predictions and, if the model

is correct, this path should come very near the experimental prevalence estimates. The key variable is np, the mean number of politician observations, which serves as a proxy for both education and, to a lesser extent, partisanship. Increasing education most likely increases the frequency with which one observes and thinks about politicians and therefore increases np and with it the frequency of proximity voting. Increased partisanship is likely to reduce the number of politician statements a voter observes and thinks about because voters are more likely to reject statements inconsistent with their views, thus reducing the number of messages they receive and the frequency of proximity voting relative to similarly-involved but less partisan voters (cf. Zaller's 1992 opinion formation model). The predicted partisanship effect should not be as close to the experimental results as the education results, however, since partisanship likely influences the distribution of ideal points in a way that I have not modeled. For this reason, I also study the effects of making ideal points more extreme, which should be correlated with increased ideological strength and which helps explain both the partisanship and ideology comparative statics.

[Figure 3 about here.]

Predicted Prevalence of Voting Types. Figure 3 presents a comparison of the predicted and experimentally-determined prevalences of the three major voting types using, as Tomz and Van Houweling did, a ternary plot (essentially a simplex plot). The model makes a range of predictions that fall neatly along a path in the space of preference distributions (see Figure 3). This path passes right through the overall experimental estimates, and for values of np roughly between five and 15, the predictions are within about 10 percent of the experimental estimates, which corresponds roughly to the 95 percent confidence region Tomz and Van Houweling (2008) placed on their estimates.[1] The best predictions occur for $np \approx 7.5$, where they fall within about 1 to 2 percent of the experimental estimates.[2] The model also predicts a substantial fraction of discounted proximity

[1]The authors used bootstrap sampling and convex hull peeling to determine the confidence region, but do not provide a detailed description of the region that results, so it is difficult to be precise about whether a given prediction falls inside the confidence region or not. On the other hand, given the nature of the predictions, whether a given prediction falls inside or outside the confidence region around the experimental estimate is not of primary importance.

[2]If these seem like small values of np, note that I used a value of the learning rate η derived from laboratory experiments — contexts in which subjects are likely to learn much faster than in the real world.

voters, a particularly striking result given that one of the dominant arguments for discounting is voter sophistication, something that is not explicitly present in the model. For $np \gtrsim 5$, the model predicts the main qualitative finding, that proximity voting is more frequent than discounted-proximity voting and both are g more frequent than directional voting

Education and Partisanship Comparative Statics. As I discussed above, education should be correlated with increased numbers of politician observations, i.e., higher np, while one aspect of partisanship should be to decrease the number of politician observations (or, at least, the number of observations that one thinks about and categorizes.) As np increases, the frequency of proximity voting increases, while the frequency of directional voting decreases and the frequency of discounted-proximity voting stays roughly constant. These predictions are in line with the experimentally-observed comparative statics for education and partisanship, indicated by the solid and dashed arrows in Figure 3, although the match is better for the education results (see below for further discussion of the partisanship results).

Ideology Comparative Statics. Because ideological strength is likely to be correlated with ideal point extremity, I ran additional simulations in which I modified the ideal category distribution. I chose each voter's ideal category by generating a normal random variable X with mean $\bar{x}$ and variance $\sigma^2 = 0.01$ and identifying the category nearest X as the ideal category. For simplicity, all voters observed seven politicians (conditional on the model parameters, roughly the number that makes predictions closest to the overall experimental estimates). This process produced ideal categories that on average were within a distance of about 0.1 from the policy neutral point. I then examined the prevalence of different voting types as a function of distance between the mean ideal category and the policy neutral point. Consistent with the experimental results, I found that as one moved the mean ideal point from the neutral point to roughly $\bar{v} = 0.4$, the frequency of proximity voting increased from about 60 to about 65 percent, discounted-proximity voting decreased from about 24 to about 14 percent, and directional voting increased from about 15 to about 22 percent. For comparison, Tomz and Van Houweling found that as ideological strength increased, proximity voting increased from 51.2 to 63.9 percent, discounted-proximity voting decreased from 37.9 to 18.6

percent, and directional voting increased from 10.9 to 17.5 percent. Although the magnitude of the predicted effect is not across the board as strong as the experimental result, it is clearly in the right direction.

These predictions may also help account for the partisanship comparative statics, since increasing partisanship is correlated with increasing ideological strength.[3] By combining increased ideological strength and a reduced number of mean politician observations, one finds a trajectory more or less in line with the observed partisanship comparative statics.

Sensitivity Analysis and Alternative Distributions

The model has a fair number of parameters, some of which have been estimated in other contexts but all of which may have some bearing on the predictions. I therefore examined how varying these parameters affects the simulation results. Overall, these checks indicate that the predictions are largely insensitive to changes in the model parameters. I first considered varying the SUSTAIN parameters. Varying the learning rate η from 0 to 1 and the threshold activation τ from 0 to 1 makes essentially no difference provided τ is not too small. If $\tau \lesssim 0.3$, very few simulated voters ever create new categories, so that most voters are either indifferent or essentially directional voters. For $\tau \gtrsim 0.3$ and η arbitrary, the predictions follow the same path as they did in the simulations I reported above, although because the learning rate varies, the mean number of candidate observations that brings the predictions closest to the Tomz and Van Houweling (2008) results does vary.

Varying auxiliary parameters likewise makes little difference. Varying the status quo point Q from 0 to 1 does not affect the path along which predictions lie, but with the SUSTAIN parameters it does change the number of candidate observations at which the predictions most closely match the experimental observations. The same is true of varying the policy position variance σ^2. Varying the separation between the two parties again had little influence on the predictions, although there was greater variance relative to the predictions of Figure 3, probably because varying the separation between the parties while fixing within-party variance affects the typical range of category locations

[3] In the 2004 American National Election Studies data, at least, the correlation between party identification and ideology was 0.63.

when the number of candidate observations is fairly small.

I also ran simulations with several other distributions of the number of candidate observations in order to analyze how sensitive the data were to the functional form of the distribution. Using the same parameters and distribution of policy positions as the simulations reported in the last section, I first examined cases in which all simulated voters observed the same number of politicians n. These simulations produced results identical to those for the simulations with binomial distributions.

I next considered normal distributions with mode $n \in [0, 40]$ and variance $\sigma^2 \in [0, 500]$ (truncated so that the number of politicians any simulated voter observed was positive) and uniform distributions over intervals $[1, n]$ with $n \in \{1, 2, \ldots 80\}$. These distributions produced predictions that were generally similar to the cases already considered, although with generally higher levels of both directional and proximity voting. This also results in more variable levels of discounted-proximity voting, so that while the comparative statics predictions regarding education and partisanship are roughly in line with experimental results, they do not come as close as those for the binomial distribution. The vast majority of cases for the normal and uniform distributions predict discounted-proximity voting is more frequent than directional voting and proximity voting is more frequent than both.

Variation in Spatial Voting Across Issues: Claasen (2009)

Using a similar technique to Tomz and Van Houweling (2008), Claasen (2009) found that there is variation in the distribution of proximity and directional voting that depends on the issue under consideration. Like Tomz and Van Houweling, Claasen experimentally manipulated candidate locations, but rather than posing a choice between two candidates, he asked subjects to evaluate single candidates on a five-point scale and regressed these evaluations on candidate and subject policy locations. His regression had the form

$$E = \beta_0 + \beta_1|v - c| + \beta_2|v|, \tag{5}$$

where E is the subject's evaluation and the third term on the right is included to control for the extremity of the subject's ideological position. The focus of the analysis was on the second term. As Claasen observes, smaller β_1 indicates a higher proportion of proximity voting, since under proximity voting, increased distance leads to less favorable evaluations, while under directional voting increased distance may lead to more or less favorable evaluations.

Claasen (2009) considered military spending, abortion, and general ideological positions in his experiments and found behavior more consistent with directional voting (i.e., positive β_1) on the abortion issue and more consistent with proximity voting (negative β_1) on general ideology and military spending (although β_1 was not statistically significant at conventional levels for military spending).

In this section, I present a qualitative analysis of Claasen's results. One might wonder about simulating data and using this data to replicate Claasen's 2009 regression results in a manner similar to the replication of the Tomz and Van Houweling (2008) results above. Although such a replication would provide further evidence in favor of the model, it would also require building an additional model of candidate ratings — a step not needed in the previous section — based on the underlying preference model which would constitute at least a short paper by itself. Furthermore, to do so would introduce more parameters into the model with less guidance on how to set them. For these reasons, I will not carry out this analysis here and instead focus on Claasen's qualitative findings.

We can understand these results in terms of the model's comparative statics related to differences in ideal point distributions across policy areas. Although Tomz and Van Houweling (2008) studied the comparative statics of partisanship, ideology, and education, I explained each of these in terms of variation in the number of candidate observations and in terms of the extremity of the ideal point distribution. Recall that increased ideological extremity led to an increase in proximity voting, a decrease in discounted-proximity voting (for an overall decrease in both kinds of proximity voting taken together), and an increase in directional voting. Therefore, if people take somewhat more extreme positions — and recall from the simulations above that the increase in ideological extremity

need not be large to make a difference — on abortion than on a general ideology scale and military spending, the model predicts more directional voting on the abortion issue, consistent with the Claasen (2009) results.

Fortunately, Claasen (2009) provides the distribution of subject self-placements on 11-point scales (-5 to 5) for each of the issues he examines (see his Appendix A). Based on this data, I computed the mean deviation from the centrist position for each issue. (Let the centrist position be x_0 and let one of Claasen's subject's positions be v. I computed the mean value of $|v - x_0|$ across all subjects.) On the general ideology dimension, the mean deviation is 2.22, and on the military spending issue, the mean deviation is 2.49. In constrast, on the abortion issue the mean deviation is 3.25.

Because the mean deviation is higher for the abortion issue, the earlier comparative-statics discussion suggests that we should observe higher frequencies of behavior consistent with directional voting on the abortion issue than on general ideological concerns, just as Claasen (2009) found. Likewise, the mean deviations for a comparison between general ideological views and defense spending are similar, so we should expect similar levels of proximity and directional voting-consistent behavior on these issues, again just as Claasen (2009) found, though again with the caveat that the β_1 coefficient for defense spending is insignificant.

To summarize, there is every reason to expect variation in the prevalence of proximity versus directional voting across issues since we have observed variation in individuals' ideal points and in the distribution of ideal points across issues. Furthermore, the earlier discussion of comparative statics indicates that increasing ideal point extremity implies higher frequencies of directional voting and lower frequecies of (generalized) proximity voting. This observation leads to a testable prediction that, if we observe greater ideal-point extremity on an issue, we should also observe greater frequencies of directional-voting consistent behavior on that issue. The available data support this hypothesis. For example, the slightly increased level of ideological extremity on the abortion issue goes along with an increased frequency of directional-voting consistent behavior on this issue.

Alternative Models

Many readers are likely to wonder how much the particular categorization model matters. Earlier I argued that it was best to use an existing, successful categorization model from cognitive psychology for two reasons. First, it grounds the subsequent model of preferences in well-understood and well-tested psychology. Second, when it comes time to make predictions about the prevalence of different voting types, using an existing model is more disciplined because the model has been tested in other domains and its parameters have been estimated. This latter point constrains subsequent predictions, making it harder to make predictions near the experimental estimates simply by varying the model's parameters (although, as I have just discussed, the SUSTAIN-based model's predictions are not very sensitive to such changes).

It is nonetheless useful to try to construct an alternative categorization model with the aim of understanding whether predictions depend on *how* people categorize or simply *whether* they categorize. I focus on a simple, economically-motivated model that I refer to as the resource-based model. Suppose voters' sets of categories partition the policy space and that they must pay some sort of cost to construct new categories. Let each voter have some amount r of a mental resource, and let the construction of a new category cost c units of this resource. Assume that voters initially have one category (encompassing every point in the policy space) and construct as many categories as they can, i.e., the number of categories is the largest integer smaller than $1 + r/c$. Finally, suppose that the category locations are spread out roughly evenly over the policy space, i.e., they are uniformly spaced but then a small amount of noise is added to their positions. (It happens that nonsensical predictions are much too frequent if the spacing is too regular.)

Already, one should start to see the theoretical problems with this kind of model. Neither the number nor the locations of the categories has any obvious mechanistic justification. Nor is it clear what the "resource" is. In constrast, the model I analyzed above has clear psychological foundations and empirical support independent of the present problem. The ultimate judge, however, is to make predictions and compare these with the experimental estimates. I therefore ran new simulations

with r distributed according to either the binomial, normal, fixed, or uniform distributions I used to model the number of politician observations in the earlier simulations. In all other respects the simulations were the same as those reported earlier.

The simulations indicate that the model does not do as a good job as the SUSTAIN-based model at predicting behavior consistent with experimental results, although the model's qualitative predictions are in line with experimental results: the frequencies of the different voting types are in most cases ordered consistently with the experimental findings, and the frequency of proximity voting increases with r/c. The best predictions in these models occur for small mean values of r/c — about 3 — but are not any better than those for the SUSTAIN-based model.

Of greatest concern are the comparative statics. While increasing r/c increases the frequency of proximity voting, it also decreases the frequencies of both discounted-proximity and directional voting, and these changes are in near-perfect proportion to each other. Furthermore, most predictions fall such that directional voting is just slightly less prevalent than discounted-proximity voting. Therefore, regardless of how one interprets the experimental findings or the predictions, there is no obvious connection to the observed comparative statics, since these comparative statics typically involve one voting type's prevalence staying roughly fixed while the other two vary.

Alternative specifications of the distribution of the category locations actually worsen the results. One plausible modification is that categories' locations should be biased somewhat toward the central positions of the parties, which tend to be somewhat moderate (though not as moderate as the average member of the mass public). Doing so, however, seems to dramatically increase the frequency of directional voting in about half of the cases I studied, so that this model fails to predict even the central qualitative observations.

To summarize, while the results are in some respects qualitatively similar to those for the SUSTAIN-based model, the resource-based model is less likely to make predictions in line with both the quantitative and qualitative experimental findings. In particular, these models have little hope of matching the experimentally-observed comparative statics and do no better at predicting the frequencies of different voting types than the model I developed above. When combined with

the other theoretical and experimental shortcomings relative to the SUSTAIN-based model — that the models do not have particularly strong economic or psychological foundations and have not been tested in other contexts — there is not much reason to pursue these models further.

Finally, a word is in order about the possibility of a Bayesian model of voter preferences, which seems to be the political economist's model of choice when it comes to learning. We might suppose, for instance, that (1) voters have proximity preferences characterized by a linear or quadratic utility function, (2) their beliefs about candidates' positions are normally distributed, and (3) they become more certain about a candidates' locations as they learn more about politics, become more educated, etc. In this way, voters might be more like directional voters when they are less educated even though their underlying preferences are proximity preferences. (This position is conceptually similar to Matthews's 1979 justification of directional voting.) However, it is straightforward to show that under such a model if a voter is equally uncertain about all candidates, uncertainty has no effect on the voter's preferences. While me may be able to find a utility function or belief distribution that leads us to predictions in line with the experimental results, the utility function and belief distribution are likely to be contrived and without any clear psychological justification. In contrast, the SUSTAIN-based model has clear psychological foundations and does a good job explaining the experimental observations, so it is does not seem worthwhile to pursue a Bayesian model further in this paper.

Discussion

This paper laid out a categorization-based model of spatial preferences that explains (1) why voters would have (or would appear to have) qualitatively different preferences and (2) makes precise predictions about both the apparent overall levels of different voting types in the population as well as comparative statics that agree with experimental observations. Specifically, the model's predictions trace a path in the space of possible distributions that comes within a few percent of the experimental observations, and the path itself yields comparative statics that explain two additional experimental observations: increasing education and decreasing partisanship — both

of which should be correlated with an increased number of observed politicians — increases the prevalence of proximity voting mainly at the expense of directional voting. In addition, making the ideal point distribution more or less extreme leads to comparative statics in line with the observed effects of ideological strength and also contributes to an explanation of the partisanship effects. Additional simulations indicate that the predictions are largely insensitive to changes in its underlying parameters or in the functional form of the distribution of voter experience with politics. Finally, the comparative statics results combined with difference in the observed distribution of ideal points explain at least qualitatively the observed variation across policy areas.

The model makes an additional prediction I have not yet discussed: the frequency of indifference decreases with education, since the likelihood voters will place two candidates in the same category will decrease with the number of categories. This result is consistent with the observation that the incidence of indifference between two candidates declines over the course of a campaign (Brady and Ansolabehere 1989). One could more explicitly test this prediction using reaction-time experiments common in cognitive and social psychology, i.e., when choosing whom to vote for, less experienced, less educated people should take longer on average to make their decisions.

Finally, since the last section in particular focused on choices between two candidates, it is worth pointing out that the model applies to voters in any democracy regardless of the number of candidates or parties. Arguably, one could apply it to study political preferences in non-democracies. The basic insight applies regardless of the form of government: people categorize political figures, they differ in how finely they categorize, and these differences imply qualitative differences in their preferences and choices. Applications outside the United States may require modifications to the model but are likely straightforward, and I plan to pursue them in the future.

References

Anderson, John R. 1991. "The Adaptive Nature of Human Categorization." *Psychological Review* 98:409–429.

Brady, Henry E. and Steven Ansolabehere. 1989. "The Nature of Utility Functions in Mass Publics."

American Political Science Review 83:143–163.

Claasen, Ryan. 2007. "Ideology and Evaluation in an Experimental Setting: Comparing the Proximity and Directional Models." *Political Research Quarterly* 60(2):263–273.

Claasen, Ryan L. 2009. "Direction Versus Proximity: Amassing Experimental Evidence." *American Politics Research* 37(2):227–253.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper.

Estes, William K. 1994. *Classification and Cognition.* New York: Oxford University Press.

Fryer, Roland G. and Matthew O. Jackson. 2008. "A Categorical Model of Cognition and Biased Decision-Making." *The Berkeley Electronic Press Journal of Theoretical Economics (Contributions)* 8.

Grofman, Bernard. 1985. "The Neglected Role of the Status Quo in Models of Issue Voting." *The Journal of Politics* 47:230–237.

Hong, Lu and Scott E. Page. 2001. "Problem Solving by Heterogeneous Agents." *Journal of Economic Theory* 97:123–163.

Jackman, Simon D. and Paul M. Sniderman. 2002. The Institutional Organization of Choice Spaces: A Political Conception of Political Psychology. In *Political Psychology*, ed. Kristen Monroe. Mahway, New Jersey: Lawrence Erlbaum pp. 209–224.

Lacy, Dean and Philip Paolino. 2005. "Testing Proximity Versus Directional Voting Using Experiments." Manuscript.

Lewis, Jeffrey B. and Gary King. 1999. "No Evidence on Directional vs. Proximity Voting." *Political Analysis* 8:21–33.

Lodge, Milton, Kathleen M. McGraw and Patrick Stroh. 1989. "An Impression-Driven Model of Candidate Evaluation." *American Political Science Review* 83:399–419.

Lodge, Milton and Ruth Hamill. 1986. "A Partisan Schema for Political Information Processing." *The American Political Science Review* 80:505–520.

Love, Bradley C., Douglas L. Medin and Todd M. Gureckis. 2004. "SUSTAIN: A Network Model of Category Learning." *Psychological Review* 111:309–332.

Macdonald, Stuart Elaine, George Rabinowitz and Ola Listhaug. 1998. "On Attempting to Rehabilitate the Proximity Model: Sometimes the Patient Just Can't Be Helped." *The Journal of Politics* 60:653–690.

Macdonald, Stuart Elaine, George Rabinowitz and Ola Listhaug. 2001. "Sophistry versus science: On further efforts to rehabilitate the proximity model." *The Journal of Politics* 63:482–500.

Matthews, Steven A. 1979. "A Simple Direction Model of Electoral Competition." *Public Choice* 34:141–156.

Merrill, Samuel and Bernard Grofman. 1997. "Directional and Proximity Models of Voter Utility and Choice: A New Synthesis and an Illustrative Test of Competing Models." *Journal of Theoretical Politics* 9:25–48.

Merrill, Samuel and Bernard Grofman. 1999. *A Unified Theory of Voting*. Cambridge University Press.

Mullainathan, Sendhil. 2002. "Thinking Through Categories." mimeo.

Mullainathan, Sendhil, Joshua Schwartzstein and Andrei Shleifer. 2006. "Coarse Thinking and Persuasion." NBER Working Paper No. W12720.

Nosofsky, Robert M., Mark A. Gluck, Thomas J. Palmieri, Stephen C. McKinley and Paul Glauthier. 1994. "Comparing models of rule-based classification: A replication of Shepard, Hovland, and Jenkins (1961)." *Memory and Cognition* 22:352–369.

Rabinowitz, George and Stuart Elaine MacDonald. 1989. "A Directional Theory of Issue Voting." *American Political Science Review* 83:93–121.

Rosch, Elanor. 1978. Principles of Categorization. In *Cognition and Categorization*, ed. Elanor Rosch and Barbara B. Lloyd. Hillsdale, New Jersey: Erlbaum.

Shepard, Roger N. 1987. "Toward a Universal Law of Generalization for Psychological Science." *Science* 237:1317–1323.

Smith, Edward E. 1990. Categorization. In *An Invitation to Cognitive Science*, ed. Daniel N. Osherson and Edward E. Smith. MIT Press.

Tomz, Michael and Robert P. Van Houweling. 2008. "Candidate Positioning and Voter Choice." *American Political Science Review* 102(3):303–318.

Westholm, Anders. 1997. "Distance versus Direction: The Illusory Defeat of the Proximity Theory of Electoral Choice." *American Political Science Review* 91:865–883.

Westholm, Anders. 2001. "On the return of epicycles: Some crossroads in spatial modeling revisited." *The Journal of Politics* 63:436–481.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion.* Cambridge University Press.

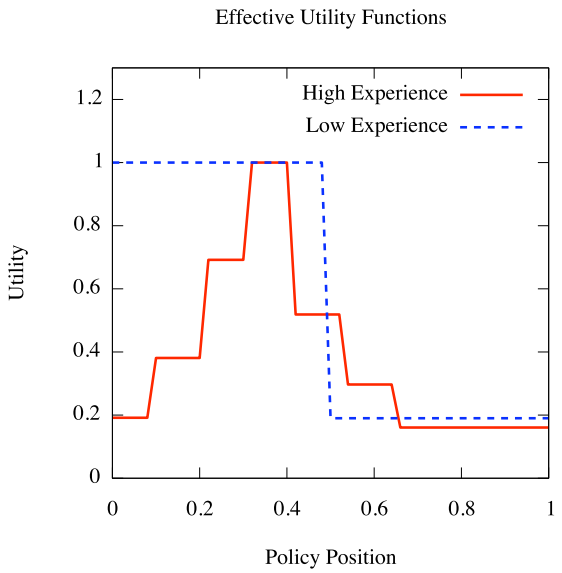

Figure 1: Effective utility functions for simulated voters with low and high levels of experience with politics. The solid line is the utility function of a voter who has observed 50 political figures; the dashed line represents a voter who has observed 500. Both voters observed political figures sampled from the same distribution of the form described earlier in this section: $x_D = 0.4$, $x_R = 0.6$, and $\sigma^2 = 0.1$. I used $H_\kappa(k)$ as the utility function, though as noted in the text this choice is not particularly important.

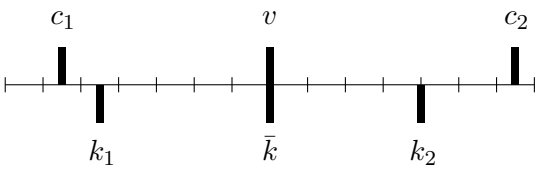

Figure 2: An example that generates behavior consistent with discounted proximity but not proximity voting. With this arrangement of candidates, a proximity voter prefers candidate 1. The category locations are effectively perceived policy outcomes, and since $|v - k_2| < |v - k_1|$, the voter chooses candidate 2, consistent with discounted proximity voting and inconsistent with proximity voting.

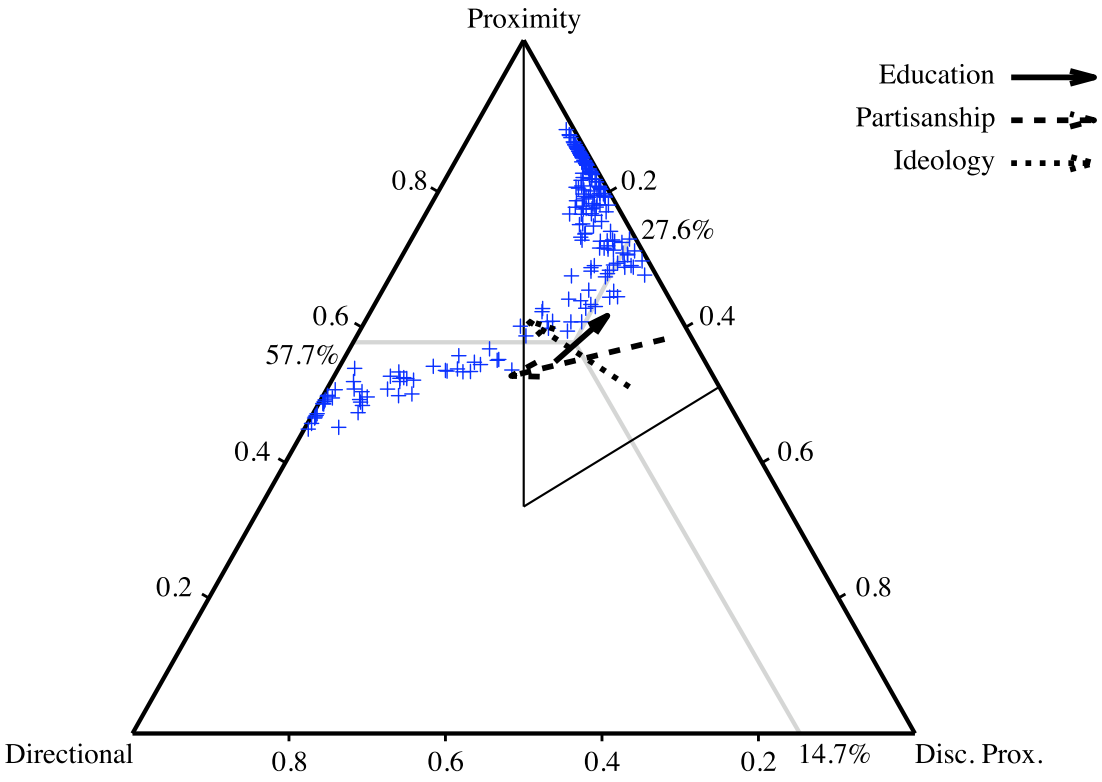

Figure 3: Distribution of predicted prevalence of voting types, with the Tomz and Van Houweling (2008) estimate indicated with light lines. The interior solid lines form the boundary of the region for which proximity voting is more frequent than discounted proximity and both are more more frequent than directional voting. The arrows indicate experimentally-observed comparative statics; see text for details.

Scenario	Conditions	Dir.	Disc.	Prox.
I	$N < v < \bar{c},\, v \leq Q$	c_1	c_1	c_2
II	$v < \bar{c}, N,\, v \leq Q$	c_1	c_1	c_1
VI	$N, Q < v < \bar{c},\, \alpha > \alpha^*$	c_1	c_2	c_2

Table 1: Discriminating scenarios used in the estimations. Note $\bar{c} = (c_1 + c_2)/2$ and $c_1 < c_2$ in these statements.