A Unified Model of Proximity and Directional Preferences

Nathan A. Collins
Department of Political Science
Stanford University
nac@stanford.edu

January 2, 2008

Abstract
Recent survey and experimental research suggests 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 preferences), but others prefer candidates that are simply on the same side of an issue as themselves (directional preferences). Current theory does not explain why some people have proximity preferences and others have directional preferences. In this paper, I propose a theory that explains this observation based on the idea that people categorize candidates and that preferences are 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.

1 Introduction

Spatial models of voting have been an important component of research on both voter decision making and party behavior in the half-century since Downs (1957) proposed his version. In that time, two different types of spatial models have emerged: proximity models (e.g., Downs 1957; Grofman 1985) suppose that voters care about how close candidates’ policy positions are 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. Recent survey research suggests voters as a group use some mix of proximity and directional preferences (Merrill and Grofman 1999), and current experimental research suggests that many voters have

*I thank Jonathan Bendor and Kenneth Shotts for helpful comments on earlier versions, John Bullock for helpful discussions, and Michael Alvarez for encouraging me to pursue this project.
proximity preferences but that some have directional preferences (Tomz and Van Houweling 2007).
The experimental finding in particular poses a stark question for political scientists: if different
people have qualitatively different preferences over political candidates, what accounts for these
differences? No existing theory of political preferences answers this question.

The essential claim of this paper is that different types of spatial preferences arise when voters
break the policy space into more or less narrowly defined categories. Voters with less narrowly
defined categories also have fewer categories. Some voters will divide the policy space into just two
categories: policy positions similar to those of the voter and policy positions different from those
of the voter. Others divide the space into a large number of categories. If we suppose that voters
still prefer policy positions similar to their own, but that they vary in how finely they divide the
policy space, qualitatively different preferences arise. Those who divide the policy space into two
groups have essentially directional preferences, since they distinguish only between positions like
and unlike their own, while those who divide the policy space into many groups have essentially
proximity preferences, since they make much finer distinctions.

What accounts for coarser or finer grouping of candidates? Such differences arise naturally
as a byproduct of learning to categorize candidates. When citizens first start to learn about
politics, they have little experience and are not able to make fine distinctions between different
policy positions, and as a result they likely collect politicians into just one or two groups. As they
gain experience, citizens make finer distinctions and divide politicians into more groups. In order
to make distinctions regarding preferences, citizens first must make distinctions regarding policy;
those with more experience make finer distinctions and are more likely to make choices consistent
with proximity preferences.

The main purpose of this paper is to propose a new model of spatial preferences and spa-
tial voting based on categorization that has as special cases both proximity and (approximately)
directional preferences. A second purpose is to show how the psychology of categorization has
far-reaching consequences for political behavior and public choice. Elsewhere I show how cate-
gorization psychology unifies many different findings in political behavior, including memory for
political candidates, proximity preferences, and interactions between preferences and perceptions of political candidates (Collins 2007). The specific aim of this paper is to show how categorization explains variation in the structure of preferences and to generate sharp, testable predictions about who should have what kind of preferences.

The rest of this paper is organized as follows. In the next section, I briefly review and formally define the different types of spatial preferences that have been proposed. I then introduce a new model of preference based on categorization, and I discuss the kinds of preferences that emerge from this model. Following that I derive some observable patterns of behavior the model predicts. I employ techniques that Tomz and Van Houweling (2007) developed to experimentally investigate spatial voting. These techniques allow us to determine whether particular voters — real or modelled — exhibit behavior consistent with proximity or directional models. I investigate the relationship between political experience and observed choices as well as the relationship between experience and indifference between two candidates. Not surprisingly, the prevalence of indifference and behavior consistent with directional preferences declines with political experience. I conclude the paper by discussing how one could test the model and suggesting some new theoretical directions.

2 Approaches to Spatial Voting

Spatial voting is based on the idea that candidates take positions on one or more issues; a comparison of these positions and a voter’s own position determines the voter’s preferences. Abstractly, voters and candidates are located somehow in a policy space and voters’ preferences are determined by these locations. Spatial models vary in how they define the policy space, how they compare various policy positions, and whether and how they take into account other policy-relevant information, e.g., the status quo position and the ability of candidates to move the status quo toward their own positions. In this section, I discuss proximity preferences, discounted proximity preferences, and two forms of directional preferences.

All spatial models suppose the existence of a policy space — typically but not always a continuous subset of of $\mathbb{R}^n$ — and represent voters and candidates with their most preferred positions.
The spatial preferences most familiar to political scientists are proximity preferences, somewhat better known as Downsian preferences (Downs 1957). Conceptually, a voter has proximity preferences when she prefers candidates with positions more similar to her own over other candidates. Formally, a voter strictly prefers candidate 1 to candidate 2 if and only if $|v - c_1| > |v - c_2|$. Note that this is a symmetric relationship; the sign of $v - c_i$ does not play a role in determining preferences. We can of course represent proximity preferences with a utility function, in which case the utility function is symmetric about $v$ and declines monotonically as $|v - c|$ increases.

Discounted proximity\(^1\) models are essentially the same as proximity models, except that voters compare likely policy outcomes instead of candidates’ policy positions. While politicians may state particular policy positions, in democratic societies they do not necessarily get exactly what they say they want. Furthermore, voters may know this and as a result discount candidates’ stated positions. Instead of paying attention to the stated positions, voters instead pay attention to the likely policy outcomes that would result if a particular candidate is elected. As Grofman (1985) points out, Downs had this in mind when he conceived of his proximity model.

In Grofman’s version of the discounted proximity model, voters believe that policy outcomes will be convex combinations of the status quo $q$ and a candidate’s policy position $c_i$. Let $p_i$ be the policy outcome associated with candidate $i$. Then $p_i = \alpha Q + (1 - \alpha)c_i$, and a voter prefers candidate 1 to candidate 2 if and only if $|v - p_1| > |v - p_2|$. Voters who use this rule vote differently from voters who use the simple proximity voting rule when $\alpha$ is sufficiently large and $v$ is between the status quo and the candidate midpoint, $\bar{c} = (c_1 + c_2)/2$. Equivalently, voters using the two rules vote differently when $\bar{p} < v < \bar{c}$ or $\bar{c} < v < \bar{p}$, where $\bar{p} = (p_1 + p_2)/2$. Figure 1(a) presents an example in which a person with ideal point $v$ votes differently under simple and discounted

---

\(^1\)The term “discounted” may cause some confusion, since it usually refers to discounting the future. Here, it refers to the idea that the policy outcome of a particular candidate’s election depends on both the candidate’s position and the status quo.
Figure 1: (a) A situation in which a person with ideal point $v$ votes differently depending on whether she uses a simple or discounted proximity voting rule. Under proximity voting, this person votes for $c_1$, but under discounted proximity voting, the person votes for $c_2$. (b) Under proximity voting, a voter with ideal point $v$ votes for $c_1$, but under Matthews directional voting, the voter votes for $c_2$.

It is worth pointing out that the distinction between proximity voting and discounted proximity voting concerns not the nature of preferences themselves but instead the objects of preferences. As Downsian proximity voting is usually conceived, preferences are defined over candidates’ stated policy positions. The only difference between these and discounted proximity preferences is that the latter are defined over expected policy outcomes. Abstractly, discounted proximity models map candidates’ positions into new positions, and citizens have proximity preferences over these new positions.

A third kind of spatial preferences are directional preferences, which can be subdivided into two classes. The basic idea behind directional preferences is that voters care primarily about what side of an issue a candidate is on relative to either a neutral point or a status quo point. The justifications for directional preferences differ considerably, however. First, Matthews (1979) argued that (1) voters may believe a winning candidate can only marginally shift the status quo regardless of campaign promises (cf. discounted proximity), (2) voters may have imperfect information about candidates’ policy beliefs, so that voters can at most extract the direction in which a particular candidate would move policy, and (3) voters may not have well-defined preferences outside a small neighborhood of the status quo. For these reasons (and a few others that concern what messages
candidates have incentives to send), voters have preferences over the direction a candidate would shift policy rather than over points in a Euclidean policy space.

While Matthews assumed voters have preferences over directions, it is possible to represent Matthews preferences without explicit reference to candidates moving policy in a particular direction. In particular, there is a natural mapping between points in the policy space and directions: in one dimension, candidates whose policy positions are to the left of the status quo would move policy to the left if elected, while candidates whose policy positions are to the right of the status quo would move policy to the right. Figure 2 illustrates this map.

Therefore, it is possible to formulate Matthews voting in one dimension as follows: a voter will prefer candidate 1 to candidate 2 if and only if \( v \) and \( c_1 \) are on the same side of \( Q \) and \( c_2 \) is on the opposite side. (See Merrill and Grofman (1999) for a similar formulation.) Matthews directional voting differs from proximity voting when \( v \) is closer to one candidate but is on the opposite side of the status quo from this candidate. Figure 1(b) presents an example in which voters using a proximity rule vote differently from those using a Matthews directional rule.

In contrast to the Matthews approach, Rabinowitz and MacDonald (1989) begin from a symbolic-politics conception of political attitudes and from the observation that citizens are unlikely to understand policy proposals in sufficient detail to be able to order them. As a result, citizens perceive of policy proposals as having a direction — do I feel positively or negatively toward the proposal? — and an intensity — how strongly do I feel? — rather than conceive of proposals as belonging to an ordered set. Rabinowitz and MacDonald propose that the quantity \( (v-N)(c_i-N) \), where \( N \) is a policy neutral point, be used to encode preferences. The neutral point represents zero intensity, i.e.,
the midpoint between strongly positive and strongly negative evaluations of a particular political idea. A voter prefers candidate 1 to candidate 2 if and only if \((v - \mathcal{N})(c_1 - \mathcal{N}) > (v - \mathcal{N})(c_2 - \mathcal{N})\).

Rabinowitz and MacDonald’s conception of the policy space differs from the usual conception in that points in the space represent positive or negative intensities of belief rather than specific, ordered policy proposals. It is therefore worth pointing out that we can define both proximity and Matthews directional preferences on such spaces. For example, a voter may care how intensely a candidate takes a stand and care in such a way that she has an ideal intensity level on some or all issues. That these different conceptions of the policy space are not mutually exclusive or even inherently different is implicit in recent experimental and survey-based studies of spatial voting (Tomz and Van Houweling 2007; Merrill and Grofman 1999).

The Rabinowitz and MacDonald model has one particularly curious feature: if a person takes (say) a moderately intense position, she will nonetheless prefer candidates who take the same position more strongly than she does over those with the same position and intensity as her. This result leads to the counterintuitive prediction that moderate candidates would vote for more extreme candidates, i.e., a particular candidate may prefer a candidate other than herself. To correct for this, the authors define a “region of acceptability” centered on the neutral point; voters prefer candidates inside the region to those outside (Rabinowitz and MacDonald 1989, 108).

Some effort has been made to study mixed models. Merrill and Grofman (1999) observed that survey thermometer ratings of various candidates plotted as a function of respondents’ policy position seemed to have features of both proximity and directional preferences. The authors therefore constructed an empirical model that mixed proximity and directional preferences and found support for this model. They did not, however, explain directional and proximity preferences in terms of some underlying model of preference; that is what I will do in the next two sections.

Lewis and King (1999) studied a mixed model in an effort to examine the methodological assumptions various authors used to conclude that voters had either directional or proximity preferences. Their research is particularly important because it shows that the conclusions of survey research on the topic depend largely on one’s assumptions and that survey research generally cannot
distinguish between different spatial voting models.

3 Overview of the Model

The motivation for the model that follows is a subtle observation about the differences between Matthews and proximity voting in one dimension. Recall that one can formulate Matthews voting in terms of candidate positions $c_i$ and ideal points $v$ instead of policy directions. The reason this is possible is that there is a natural way to map policy positions into directions: all candidates who position themselves to the left of the status quo would move policy to the left if elected, while all candidates to the right of the status quo would move policy to the right if elected. (See Figure 2.)

By combining this map with preferences over directions, we can formulate Matthews preferences as follows: collect all of the points in the policy space into two groups, one for points to the left of the status quo $Q$ and one for points to the right, and then determine which group a voter prefers — naturally, a voter prefers the left group if $v < Q$ and the right group if $v > Q$. A voter then prefers candidates in her group — i.e., on her side — over those in the other group, and she is indifferent between any two candidates in the same group. Put another way, the map from positions to directions is just a particular, coarse way of grouping points in the policy space, and Matthews voting results from defining preferences over the groups instead of over policy points. We can formulate proximity preferences similarly: voters have preferences over groups, but each group corresponds to a single point in the policy space.

With that in mind, the basic idea of this paper is that behavior that approximates proximity and directional spatial voting results from finer or coarser grouping of points in a policy space. The set of groups is a perceived policy space, with coarser groupings resulting in smaller perceived policy spaces. For example, instead of understanding the policy space as the continuous interval $[0, 1]$, a voter might group together points in the intervals $[0, 0.4)$ and $[0.4, 1]$, making no distinctions between any two politicians who fall in the same interval. Such a voter perceives a policy space that consists of just two points. If we embedded this smaller space in the interval $[0, 1]$, we could represent it as the set of points $\{0.2, 0.7\}$, i.e., the midpoints of the intervals $[0, 0.4)$ and $[0.4, 1]$.
Now, suppose that all voters have proximity preferences, but these are defined only on the voter’s perceived policy space. A voter who perceives the policy space as the interval $[0, 1]$ will have what we usually call proximity preferences. A voter who perceives the space as the set $\{0.2, 0.7\}$ will have preferences that approximate Matthews directional preferences, since this voter only makes distinctions between positions that are on one or the other side of the space. If her ideal point is 0.2 instead of 0.7, she perceives every politician in $[0, 0.4)$ as like her, and she prefers these politicians to those in $[0.4, 1]$, who she perceives as unlike her.

A main theoretical claim of this paper is that categorization psychology explains why some voters would group candidates coarsely while others group them finely. When voters are inexperienced with politics, they do not make very fine distinctions between different policy positions, so they use just one or two categories to group politicians. As they gain experience, they group politicians more finely. Throughout, they prefer candidates who take positions closer to them. When they are inexperienced, however, closeness is not a graded concept — there is just “like me” and “unlike me.” As they gain experience, they make finer distinctions, use more categories, and have preferences that more closely approximate what we think of as proximity preferences.

4 The Model

Formally, the model consists of three parts. First, there is the political environment from which candidates come. The environment is not strictly part of the model — one should think of it as the input to the model — but it is necessary to describe it in order to perform the analysis in the next section. Second, there is the model of categorization. This is the part of the model that allows voters to collapse the policy space into a relatively small number of groups. Third is the definition of preferences over categories and a method of translating preferences over the categories into preferences over candidates and eventually into choices. I begin with the environment.
The Environment

People in the model categorize candidates, and the first order of business is to define what candidates they categorize. As with other spatial models, a candidate is a point in a policy space, which I take to be the interval $[0, 1]$. I assume that each voter sequentially observes $N$ candidates sampled from a distribution on this policy space; $N$ represents the amount of political experience a person has (or will eventually have) and varies between voters. The distribution is the sum of two normal distributions with means $x_R$ and $x_D$ and variances $\sigma^2$. Each normal distribution represents the candidates from one of two political parties. The parties each have some average policy position, and there is variance in the positions of a party’s individual candidates. Formally, the distribution of the position of each candidate a voter observes is proportional to

$$\Psi(c) = \frac{1}{2} \Phi(x_D, \sigma^2) + \frac{1}{2} \Phi(x_R, \sigma^2),$$

where $\Phi(x, \sigma^2)$ is a normal distribution with mean $x$ and variance $\sigma^2$. The distribution is truncated so that the density of candidates is zero outside the interval $[0, 1]$. I present an example of this distribution with $x_D = 0.3$, $x_R = 0.7$, and $\sigma^2 = 0.02$ in Figure 3.

Categorization

To model the process of categorizing politicians, I use the SUSTAIN model of categorization (Love, Medin, and Gureckis 2004). SUSTAIN is convenient because it has an explicit attention mechanism, which controls how finely people distinguish policy positions (or other objects of categorization), and because it has an explicit mechanism for constructing new categories when it encounters distinctly new policy positions (or other objects). It is also based firmly in principles of cognitive psychology, so while other clustering algorithms may achieve similar goals, employing SUSTAIN keeps the study of political behavior close to the study of human behavior more generally.

The process of categorizing a candidate proceeds in three steps. First, SUSTAIN — that is, a voter in the model — observes a candidate sampled from the political environment. Then, the voter attempts to place the candidate in a category. To do so she determines if there is a category
sufficiently similar to the candidate. If there is, she puts the candidate in the category most similar to the candidate. If there is no sufficiently similar category, she creates a new category to represent this apparently new kind of political candidate.

As with candidates and voters, categories are points \( k \) in the policy space, and the similarity of a candidate \( c \) to a category \( k \) is defined in terms of the activation \( H_k(c) \):

\[
H_k(c) = e^{-\lambda|c-k|}.
\] (2)

(See Shepard (1987) for a discussion of the exponential form.) Here, \( \lambda \) is the attention one pays to politics. In this context, attention is a measure of how finely a person makes distinctions between different policy positions — if a person pays more attention, she can make finer distinctions. Let \( K \) be the set of categories. The following rule determines the process of categorizing candidates:

- If a voter observes a candidate \( c \) and \( K \) is empty, she creates a first category \( k_1 = c \). Now \( K = \{k_1\} \).
If \( \max_{k \in K} H_k(c) > \tau \), where \( \tau \) is an exogenous threshold activation, the voter places candidate \( c \) in category \( k \).

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

Once the voter categorizes the candidate, two things happen. First, placing a candidate in a category provides new information about the location of the category. Thus if the voter placed the candidate in an existing category, she updates the location of the category to take into account this new information. If the voter placed candidate \( c \) in category \( k \), then

\[
k \rightarrow (1 - \eta) k + \eta c, \tag{3}
\]

where \( \eta \) is a learning rate parameter. Second, the voter updates attention \( \lambda \):

\[
\lambda \rightarrow \lambda + \eta e^{-\lambda \sigma} (1 - \lambda \sigma), \tag{4}
\]

where \( \eta \) is the same learning rate parameter. If the policy space had multiple dimensions, this rule would eventually draw most attention to those dimensions that were most useful for making decisions about which category something belonged in. In one dimension, the rule increases attention with experience, and attention increases more rapidly the less within-category variance there is.

Preferences and Choice

The essential idea of this model of preferences is that instead of having preferences or a utility function defined on a continuous policy space, people have preferences defined over categories. The process of categorization means that for the most part people do not make distinctions between candidates they place in the same category; as a result they will not make preference distinctions either.

How then do candidate positions get translated into preferences? Suppose for the moment that each category \( k \in K \) has an associated utility \( u_k \) which defines a map \( u : K \rightarrow \mathbb{R} \). The categorization process defines a map \( h : [0,1] \rightarrow K \) from points in the policy space to categories. (Note that categorizing a candidate may expand the set \( K \), so it may not be quite right to call
Given $u_k$ we can construct a map $h \otimes u : [0, 1] \to \mathbb{R}$, which defines an effective utility for each point in the policy space.

Elsewhere, I explore how voters could learn the utilities $u_k$ (Collins 2007). For the purposes of this paper, it is enough to suppose that voters have an ideal category $\kappa$, and that voters prefer categories closer to $\kappa$ over categories farther away. For a given category $k$, we can use the activation $H_\kappa(k) = \exp(-\lambda|k - \kappa|)$ as a utility function. Then, the model could be stated succinctly this way: voters have proximity preferences, but their proximity preferences are defined over a finite policy space — the space of categories, which is embedded in the interval $[0, 1]$ — that begins small and grows as they gain experience with politics. It is important to realize that the choice of utility function is not terribly important. What matters is that a voter has an ideal category and that a voter does not make distinctions of preference between two candidates she puts in the same category. In particular, it makes little difference whether distance from the ideal category is measured by the number of categories or the distance in the policy space.

Now, as a person gains experience and attention increases, the activation of any particular category by a candidate decreases: $dH_k(c)/d\lambda < 0$. Therefore, the set of candidates $c$ satisfying $H_k(c) > \tau$ shrinks, and citizens create more categories as they gain experience. As described above, using this observation in combination with preferences expressed as category utilities $u_j$ defines a map from points in the policy space to utilities, i.e., an effective utility function. Figure 4 illustrates effective utility functions at different experience levels. The main point is that those with less political experience have blockier utility functions that approximate directional preferences, while those with more experience have utility functions that more closely approximate proximity preferences.

It is important to note that the present model does not predict directional or proximity preferences exactly; it predicts preferences that are approximately Matthews directional or proximity preferences, depending on the level of experience. Matthews preferences, for example, divide the policy space into points to the left and to the right of the status quo, while for voters with a small amount of experience the present model divides the space into categories, one to the left and one to
Figure 4: Effective utility functions for 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 $\sigma^2 = 0.1$. 
the right with a dividing line somewhere in the middle. However, the dividing lines will typically be at moderate positions in the policy space \([0, 1]\), i.e., both the status quo and the dividing line between the categories of those with little experience will be near each other, so that the present model approximates Matthews preferences. Similarly, voters with a great deal of experience will have many categories, so that their effective utility functions more closely approximate proximity preferences, but they will never have a completely smooth utility function.

Let the effective utility of candidate \(c\) be \(u(c)\). Given the effective utilities of two or more candidates, I model choice in the usual fashion: a citizen votes for a candidate in the set \(\arg \max_c u_c\). If this set has more than one element, I assume voters choose one of the candidates in the set at random, each with equal probability.

5 Candidate Choices

In the last section, I defined a new spatial model of preferences and choice based on the idea that citizens categorize political figures and that people vary in how finely they divide political figures into categories. I illustrated the kinds of effective utility functions that result from the model; the key observation is that as voters gain political experience they shift from preferences that resemble directional preferences to those that more closely resemble proximity preferences. Of course, utility functions are only a representation of preferences. Furthermore, we cannot directly observe preferences, but we can observe choices. Therefore, we want to understand what kinds of choices voters make.

In this section, I analyze what choices voters make and how their choices depend on political experience. My approach is to identify what kinds of preferences voters appear to have given their choices and other information, such as the positions of candidates and the status quo. I use techniques Tomz and Van Houweling (2007) originally developed to study experimentally what kinds of preferences real voters have. These techniques use voters’ choices to discriminate between proximity, discounting, directional, or some alternative preferences, and they can be used to estimate the frequencies of particular preferences types.
I performed three analyses of the behavior of simulated voters. It may be helpful to think of these analyses as simulated experiments with simulated voters as subjects, keeping in mind that there are no real voters and no data on real choices. In each analysis, simulated voters developed categories and then voted for one of two candidates. First, because voters in the model move from having approximately Matthews directional preferences to having proximity preferences as they gain experience, I examined the frequency of vote choices consistent with Matthews and proximity preferences as a function of experience. Second, because indifference between two candidates is common in the model, I examined the frequency of indifference as a function of experience. These first two analyses show that the frequency of behavior consistent with proximity voting increases with experience, while the frequency of indifference declines. Finally, because simulated voters may behave as if they had Matthews directional, Rabinowitz-MacDonald directional, proximity, or even discounted proximity preferences, I used the Tomz and Van Houweling techniques to estimate the frequency of these preferences in the simulated voter population. Of course, none of the simulated voters actually have these kinds of preferences — they have the preferences defined in the previous section. What this analysis will show is that voters may nonetheless appear to have such preferences.

Because the analyses I report indicate that the apparent distribution of preference types depends on political experience, it is worth pointing out that MacDonald, Rabinowitz, and Listhaug (1995) specifically argue that the frequency of directional voting does not depend on political sophistication, which one would expect to be correlated with experience. In fact, MacDonald et al. argue that by and large all voters use directional voting rules. This result is in conflict with experimental research (Tomz and Van Houweling 2007) and with other survey research that reaches the opposite conclusion (e.g., Westholm 1997). As noted above, Lewis and King (1999) called into question whether survey research like that of MacDonald et al. or Westholm could properly distinguish between the two voting models. Lewis and King note that different conclusions about spatial voting depend largely on different methodological assumptions. Another problem is that much survey research on the subject tries to elicit and then characterize voters’ utility functions, which are both hard to elicit and difficult to compare between voters. The upshot of all this is that
(1) it remains interesting to develop predictions about the frequency of different voting rules as a
function of experience and (2) it is important to develop predictions about voters’ decisions, since
we can observe decisions but we cannot directly observe utilities.

**Framework**

The basic framework for the analyses that follow was developed by Tomz and Van Houweling
(2007). One can get a feel for the approach by examining Figure 1. As that figure suggests, there
are particular configurations of ideal point, candidate positions, status quo, and status quo weight
that lead voters using different decision rules to different choices. For example, suppose \( \bar{c} < v < \bar{p} \)
and \( v < Q \), where \( \bar{c} \) is the candidate midpoint and \( \bar{p} = \alpha Q + (1 - \alpha)\bar{c} \) is the policy midpoint.
Under these circumstances, a voter with proximity preferences will vote differently than voters
with discounted proximity or Rabinowitz-MacDonald directional preferences.

Given a particular configuration of a voter’s ideal point \( v \), candidate positions, and so forth,
the Tomz and Van Houweling approach allows us to estimate upper bounds on the frequency of
particular preference types. Consider the set of configurations such that a voter with proximity
preferences will vote differently than one with discounting or directional preferences. The fraction of
voters in this set who make a choice consistent with proximity preferences is an estimate of the upper
bound of the fraction of voters who correctly use proximity preferences. Let a voter’s choice be \( c \), and
let the choice consistent with preference type \( t \) be \( c_t \), where \( t \in \{ \text{proximity}, \text{discounting}, \text{directional} \} \)
and let \( P(t) \) be the probability that a randomly selected voter has preferences of type \( t \) and uses
them correctly to make a choice. One can show that

\[
P(t) \leq P(c = c_t | c_t \neq c_s \forall s \neq t). \tag{5}
\]

We can estimate the right-hand side using the subset of (simulated) data satisfying \( c_t \neq c_s \forall s \neq t \).
The estimate is just the fraction of voters in this subset that make the same choices a person with
preferences of type \( t \) would, i.e., the fraction in the subset satisfying \( c = c_t \). We can also establish
a lower bound on the fraction who use some alternative preferences (including those who have
proximity, discounting, or directional preferences but who make errors when they vote):

\[ P(\text{alternative}) \geq P(c \neq c_t | c_t = c_s \forall s, t). \]  \hspace{1cm} (6)

**Simulations**

To study whether voters in the model make choices consistent with proximity, directional, or discounting preferences, I ran a series of simulations. In each simulation, a voter observed a series of \( N \) candidates, with \( N \in \{20, 40, 60, \ldots, 400\} \), sampled from the distribution of Equation (1) with \( \sigma^2 = 0.02, x_D = 0.3, x_R = 0.7 \). As the voter observes candidates, she develops a set of categories according to the model described above. I set the threshold category activation to \( \tau = 0.3 \). Consistent with past work employing the SUSTAIN algorithm (Love et al. 2004), I set the learning rate to \( \eta = 0.2 \). I ran the model 5000 times for each \( N \).

At the end of a run, each of which represented one voter, I chose a category to serve as the ideal category. To choose the ideal category, I sampled a random point uniformly from the interval \([0, 1]\) and used the voter’s categories and attention \( \lambda \) to categorize the point. The resulting category served as the ideal category, and the location of the ideal category served as the voter’s ideal point \( v \) for the purposes of determining choices under proximity, discounted proximity, and directional models. I chose the voter’s status quo weight \( \alpha \) from a uniform distribution on \([0, 1]\). I set the status quo at \( Q = 0.4 \) and the policy neutral point at \( N = 0.5 \), i.e., the midpoint of the underlying policy space \([0, 1]\). These points were only used to determine how voters with particular ideal points would vote if they actually had directional or discounted proximity preferences. Therefore voters did not categorize these points. Instead, they were used as-is to determine how individuals would vote if they had pure or discounted proximity preferences or directional preferences. This choice does not affect any of the results qualitatively and makes only minor quantitative difference. On the other hand, it makes it possible to study how inexperienced voters would vote under different rules. For example, it would be impossible to distinguish proximity, discounted proximity, and directional voting if a voter had only two perceived policy points — corresponding to two categories — and five positions \( v, c_1, c_2, Q, \) and \( N \) that may matter.
As the choice of $Q$, $N$, $\alpha$, and so forth may seem somewhat ad hoc, it is important to clarify the role these parameters play in what follows. Note that this paper’s model does not incorporate the status quo, the neutral point, or discounting in any way. Furthermore, voters in the model do not use proximity voting or discounted proximity voting or anything other than the rule defined in Section 4. Assigning these parameters just defines regions in which, for example, voters with proximity preferences would make different choices than those with discounting or directional preferences. Thus, the purpose is not to identify whether voters have one of these preference types. The simulated voters do not have any of these preference types. Instead, the purpose of assigning $Q$, $N$, $\alpha$, and so on is to examine whether the simulated voters make choices that could be interpreted as consistent with one or another of the types described above.

With these definitions in hand, I then chose a pair of candidates and had the simulated voter choose which candidate to vote for. Voters made their decisions by categorizing candidates and translating their categorization decisions into preferences according to the model laid out above. I also recorded which candidate the voter would choose under proximity, discounted proximity, and both kinds of directional voting. Following Tomz and Van Houweling (2007), I chose candidate positions to maximize the ability to discriminate between different kinds of voting. The details of the selection process are not particularly important, since what matters are the subsets of choices such that, for example, $c_{\text{prox}} \neq c_{\text{disc}}$ and $c_{\text{prox}} \neq c_{\text{dir}}$. Choosing candidates carefully simply improves the number of cases that allow us to distinguish between different kinds of preferences.

**Matthews vs. Proximity Voting**

The main conceptual prediction the model makes is that voters move from choices approximately consistent with Matthews directional voting to proximity voting as they gain experience. The first task is to study whether simulated voters behave consistently with this assertion. For each experience level, I computed estimates of $P(c = c_{\text{prox}}|c_{\text{prox}} \neq c_{\text{Matthews}})$ and $P(c \neq c_{\text{prox}}|c_{\text{prox}} = c_{\text{Matthews}})$, i.e., the upper bound on the frequency of proximity voting and the lower bound on the frequency of voting inconsistent with both proximity and Matthews voting. Note that the upper
Figure 5: Estimated upper bounds on Matthews directional and proximity voting and the estimated lower bound on alternative voting rules as a function of experience.

Bound on Matthews voting is just one minus the upper bound on proximity voting, so there is no need to separately estimate this bound. I used only choices made by voters who were not indifferent between the two candidates under either Matthews, proximity, or this paper’s preference models, and I only present bounds when there were at least 50 choices that could be used to estimate the bounds.²

Figure 5 presents the results. As the figure shows, the frequency of choices consistent with proximity voting increases dramatically as voters become more experienced. Choices inconsistent with both proximity and Matthews voting are generally rare. Among the most experienced voters in the simulations, the rate of voting consistent with proximity voting is about 80 percent, while among the least experienced voters it is about 10 percent.

Indifference

Indifference does not easily arise in either proximity or Rabinowitz-MacDonald directional preferences, but it does arise in the present model and in Matthews directional preferences. In the Matthews (1979) model, voters are indifferent between any two candidates on the same side of an

²With few exceptions, there were at least 100 and more typically several hundred or more choices in any particular subset of interest.
issue. In the present model, voters are indifferent between any two candidates they place in the same category. With Figure 4 in mind, voters less experienced with politics should be considerably more likely to be indifferent between two candidates. This is conceptually similar to the finding that indifference between presidential primary candidates is more prevalent early in the campaign — i.e., before voters have much experience with these particular candidates (Brady and Ansolabehere 1989). While the time-scale of the present model is meant to cover voters’ lifetimes, one could get an effect of this sort if we allow for substantial forgetting between election seasons, in which case we might expect voters’ ability to discriminate between candidates to degrade significantly.

Figure 6 shows the frequency of indifference among the simulated voters as a function of political experience. Indifference declines from initially very high levels — those with little experience do not differentiate between any two candidates — to roughly 30 percent among more experienced voters. In other words, almost all of the most inexperienced voters think that all politicians are the same. Though this claim does not appear to have been studied directly, it is consistent with the observations that inexperienced voters have difficulty correctly recalling candidate positions and that these voters are less stable in their choices (e.g. Campbell, Converse, Miller, and Stokes 1960; Lodge and Hamill 1986), both of which suggest that inexperienced voters have difficulty distinguishing between candidates. As voters gain experience, they gain the ability to differentiate between different candidates, though it remains possible that they will be indifferent in some cases.

Other Preferences

None of the conclusions described above addresses the apparent frequency of discounted proximity or Rabinowitz-MacDonald directional voting. The model does not obviously predict preferences consistent with either of these decision rules. However, voters may behave as if they had such preferences even if they did not. To study this possibility, I performed a study similar to that of Tomz and Van Houweling (2007), who studied the frequency of proximity, discounted proximity, and Rabinowitz-MacDonald directional preferences. Using the simulated voters, I estimated upper bounds on the frequency of these three kinds of voting and a lower bound on the frequency of alter-
Figure 6: The frequency of indifference in the model as a function of experience.

I used only choices made by voters who were not indifferent under Rabinowitz-MacDonald directional, proximity, discounted proximity or this paper’s preference models, and again I only present bounds when there were at least 50 choices that could be used to estimate the bounds.

Figure 7 presents the results. As voters gain experience, they make choices that are increasingly consistent with proximity voting and increasingly inconsistent with other preference types. However, there are voters who make choices consistent with each of the other preference types. That some voters appear to have discounted proximity preferences is particularly important because none of them actually takes the status quo into account when making their decisions. Thus, voters may appear to use sophisticated decision rules despite having decidedly less sophisticated preferences.

Because real voters may be indifferent but nonetheless choose to vote for one or another candidate — particularly in experimental settings — and because of the relatively high rate of indifference in the model, it is important to examine what kinds of preferences all voters, including indifferent voters, appear to have. Figure 8 includes choices made when simulated voters were indifferent between the two candidates. As the figure shows, including indifferent voters in the analysis does not change the basic prediction: the frequency of voting consistent with proximity preferences increases as a function of experience, while the frequency of voting consistent with other types of preferences
Figure 7: The upper bounds on the frequencies of different preference types, excluding cases in which the simulated voters were indifferent between the two candidates.

decreases.

6 Discussion

This paper makes two important contributions to the study of spatial voting. First, it explains in terms of categorization why different voters would have or would appear to have qualitatively different preferences: voters differ in the way they categorize candidates, and these differences show up in their preferences and observed choices. In particular, the model explains why some voters appear to have directional preferences while others appear to have proximity preferences, and it predicts that some voters will appear to have discounted proximity preferences even when they do not actually take the status quo into account when making voting decisions. Second, the model makes a prediction about who will appear to have what kind of preferences: the frequency of proximity voting should increase as voters’ political experience increases while the frequency of other types declines.

This last statement suggests experimental work that could shed light on the model’s validity. A key prediction is that the frequency of proximity voting increases as experience increases, while the apparent frequency of other voting rules and the frequency of indifference decline. The decline
Figure 8: Upper bounds on the frequency of proximity, discounted proximity and Rabinowitz-MacDonald directional voting, and the lower bound on the frequency of alternative voting methods (including mistakes) as functions of experience. Experience is measured as the number of political figures a voter has observed.

is most dramatic for Matthews directional voting. If we focus attention on just Matthews and proximity voting and exclude indifferent voters, the frequency of choices consistent with Matthews voting declines from close to 100 percent to less than 20 percent over the range of political experience levels I considered. Likewise, indifference declines dramatically from over 95 percent to less than 30 percent. Experiments designed to study the frequency of proximity and Matthews voting as a function of experience are clearly possible, and experimental studies of reaction times may shed light on the relationship between political experience and indifference. Because experience — i.e., the number of candidates a person has observed — may difficult to measure, it may be more practical to study preferences as a function of age, education, or measures of civic involvement. While experience is not perfectly correlated with any particular demographic feature, these variables are probably easier to measure reliably than the number of candidates a person has heard about. Likewise, since preferences over specific candidates change over the course of a campaign (Brady and Ansolabehere 1989), it may be useful to study spatial voting as a function of political advertising exposure, keeping in mind the multiple steps and subtleties of political communication (see Price and Zaller 1993; Freedman, Franz, and Goldstein 2004; Ashworth and Clinton 2007, for examples)
Regardless of what these experiments show about the present model, they will provide valuable insight into how spatial voting works.

Finally, when any new model of preferences arises, it is important to study what the model can tell us about candidate behavior, since candidate behavior is one of the things we are ultimately interested in. One of the main ingredients in the median voter theorem is the continuity of the policy space, or at least that voters differentiate between any two positions a candidate can take. However, if a substantial number of voters do not perceive more than a few distinguishable policy positions, candidates may not need to locate themselves precisely at the median in order to secure votes. Thus, the present model may help explain why candidates often do not run to the center.

References


Jeffrey B. Lewis and Gary King. No evidence on directional vs. proximity voting. Political Analysis, 8:21–33, 1999.


