Probabilistic Voting and the Importance of Centrist Ideologies in Democratic Elections

James M. Enelow
SUNY–Stony Brook

Melvin J. Hinich
University of Texas at Austin

This study indicates that a probabilistic voting model allowing voters to abstain from alienation or indifference leads to a transitive ordering over candidate ideologies. Further, a weighted form of the mean most preferred ideology of the voters is the optimal ideology for both candidates in a two-candidate election. Centrist ideologies are attractive to the major parties in a two-party system, and probabilistic voting provides an important part of the explanation.

Introduction

What may be described as the classical spatial model of electoral competition originates with the work of Davis and Hinich (1966, 1967). Building on the work of Downs (1957), this model has several prominent features. Electoral competition is viewed as taking place in a multidimensional Euclidean space that represents the issues of the campaign. Voters are represented in this space by ideal points denoting their most preferred positions on the issues. Each candidate is commonly perceived by the voters as a point in this same space, corresponding to the positions taken by the candidate on these issues. The basis for electoral

* This work was supported by the National Science Foundation under Grant SES-8111411.

competition consists of each candidate's efforts to adopt a set of positions closest to some winning number of voters.

Empirical studies demonstrate that voter-survey data can be used to recover both the electoral space postulated by the classical model and the positions of voters and candidates in this space. Beginning with Weisberg and Rusk (1970) and continuing with Rabinowitz (1973, 1978), Cahoon, Hinich, and Ordeshook (1978), and others, various metric and nonmetric scaling techniques have been used to obtain good fitting maps of voters and candidates in recent American presidential elections.

This empirical work raises certain questions about the adequacy of the classical spatial model. Typically, two-dimensional maps are sufficient to represent the data given by the voters, suggesting either that there have been only two salient issues in each recent American presidential election or that a difference exists between the issues and the dimensions of the campaign. The second conclusion appears more reasonable and empirically justifiable, but it raises several questions for the classical model. If the dimensions and issues of the campaign define two separate spaces, then how are these two spaces related? Are voter perceptions of each candidate based on the candidate's position on the issues or the dimensions of the campaign? This latter point is important because the classical model contains the implicit assumption that voters have direct knowledge of the issue positions of all the candidates. Such an assumption posits a level of voter information far above what is revealed by voter surveys.

Beginning with the papers of Hinich (1978) and Hinich and Pollard (1981), a new approach to modeling elections has been developed that addresses these questions. This approach is based on two fundamental principles. First, the distinction between the set of campaign issues and the much smaller set of underlying dimensions is explicitly recognized. Second, these underlying dimensions are viewed as predictive devices used by voters to estimate candidate positions on the set of salient campaign issues. The labels used to describe candidate positions on the predictive dimensions are commonly shared among the voters, but the issue positions estimated on the basis of each label are a matter of individual perception. In contrast to the classical model, the voter may assign issue positions to each candidate on only a subset of the total number of issues.

This new model of voters and candidates combines several features. First, it rationalizes the underlying order in voter attitudes toward the candidates revealed by multidimensional scaling with the perceptual variation in a given candidate's issue positions evident in the raw survey data. Second, it explains how voters are able to judge candidates on the basis of their issue stands without investing enormous resources acquiring information. Finally, the model can be viewed as a synthesis of the

Downsian and the Davis-Hinich models. The Downsian model postulates an ideological continuum along which voters and candidates are arranged. The Davis-Hinich spatial model postulates an issue space in which voters and candidates are located. The perceptual model sees these two spaces as distinct but closely connected. The space postulated by Downs forms a basis by which voters can estimate candidate locations in their own versions of the space postulated by Davis and Hinich.

This perceptual model has already been employed to study political campaign dynamics. Enelow and Hinich (1982a) have used this model to illuminate the role of ideology in election campaigns. In Enelow and Hinich (1981), the connection between predictive labeling and voter uncertainty is explored. Finally, in their later work (1982b, 1984), the effects of nonpolicy issues on optimal candidate location in the predictive space are examined.

One of the most interesting themes characterizing the results that grow out of this model is that a number of conditions give rise to centrist election outcomes. Widely divergent views of nonspatial candidate characteristics, familiarity with the future policies expected of ideologically centrist candidates, and the expectation of large policy changes under noncentrist candidates all contribute to centrist election outcomes.

This paper adds to our understanding of the importance of the center of voter opinion in democratic elections. Hinich, Ledyard, and Ordeshook (1973) develop an individual-level model of probabilistic voting in which a voter may abstain due to alienation or indifference. We build the assumption that voters can abstain from alienation or indifference into the perceptual model to measure the effects of probabilistic voting. An equilibrium result is obtained, stronger than that in Hinich, Ledyard, and Ordeshook (1973) and with a much simpler proof.

The focus of this paper is on the location of the optimal candidate position in the underlying predictive space. We discuss our results in one predictive dimension representing left-right ideology to contrast our findings with the classic median-voter result and earlier work on nonpolicy issues. But the results we obtain hold for multiple dimensions. Our major result is that the optimal location for the candidates is a weighted mean of the most preferred points of the voters on the predictive dimension. If the electorate is homogeneous in all other respects, the optimal location for the candidates is the mean most preferred point of the voters.

This result says that if the electorate is composed of probabilistic voters, the candidate who is closest to the center of the predictive space stands the best chance of winning. The closer the candidate is to the center, the better. We argue that given candidate uncertainties about how they are perceived, probabilistic voting is a reasonable assumption for the candidates to make. Thus, we conclude that probabilistic voting provides an explanation for the attractiveness to the candidates of the center of voter opinion.

It has often been observed that the center of voter opinion exerts a powerful force over election outcomes. Surprisingly, spatial theory has offered little in the way of an explanation for this regularity. Even though empirical results support the conclusion that the center of the electorate is attractive to candidates, standard spatial theory is preoccupied with disequilibrium results. We maintain that election theory must explain what we observe. The center of voter opinion is important in democratic elections, and this paper provides part of the explanation.

The Perceptual Model

We begin by outlining the perceptual model, and then we introduce probabilistic voting. Because a fuller exposition of the perceptual model can be found in Hinich and Pollard (1981) and Enelow and Hinich (1984) our discussion of the model will be somewhat brief. We will exposit the perceptual model for the case of a single predictive dimension. Thus the discussion will be kept simpler and will be consistent with the way we develop our results later in the paper. It should be remembered throughout, though, that all of our results generalize to multiple dimensions.

Assume two candidates, Theta and Psi. Both Theta and Psi are known to the voters by the labels they have, denoting their positions on the underlying predictive dimension Π. To motivate the discussion, assume this predictive dimension is left-right ideology. The candidate's predictive label is, then, the position obtained on this scale. This position is generally established over a long period of time through interest group endorsements, party affiliation, and, to some extent, through the candidate's own efforts. However, it is important to keep in mind that politicians cannot change labels at will. For over ten years Nelson Rockefeller attempted to modify his label as a liberal Republican, with little or no success. Changing predictive labels is not easy.

Let π_θ and π_ψ be the ideologies of Theta and Psi. The voters then use the ideologies as a shorthand device to estimate each candidate's positions on $n>1$ real issues of the campaign (such as defense spending, fiscal policy, government jobs programs, abortion, and busing) that interest them. This is precisely the way ideology is viewed by Downs (1957). Each voter arrives at a personal estimate of these positions. If voter i uses a linear translation function that maps ideological positions into positions on the n issues of the campaign, we can represent i's perception of Theta and Psi's positions on campaign issue j by

$$\theta_{ij} = b_{ij} + \pi_\theta v_{ij} \text{ and } \psi_{ij} = b_{ij} + \pi_\psi v_{ij}.$$

The coefficient v_{ij} represents the difference in position on issue j associated with a unit difference in ideology. This is a matter of individual perception. A voter may believe that a unit difference in ideology translates into a large, small, positive, or negative difference on issue j. The answer depends on the voter and the issue. The intercept b_{ij} represents i's perception of the position on issue j associated with $\pi = 0$. It is easiest to interpret $\pi = 0$ as the ideological position of the incumbent, in which case b_{ij} is voter i's perception of the status quo policy on issue j.

The perceptual linkage between issues and ideology need not work in a single direction. The voter may directly perceive θ_{ij} and estimate π_θ by $(\theta_{ij} - b_{ij})/v_{ij} = \hat{\pi}_\theta$. This estimate of Theta's ideology may then be used to arrive at estimates of Theta's positions on other issues. In short, the voter may directly perceive Theta's position on one important issue and, on this basis, infer the same candidate's positions on a list of others.

As in classical spatial theory, each voter has an ideal point $\underline{x}_i = (x_{i1}, \ldots, x_{in})'$ denoting the most preferred positions on the n issues of the campaign that are of personal concern. The voters evaluate Theta and Psi according to the weighted Euclidean distance between themselves and the candidates on these issues. Letting A_i be the $n \times n$ positive definite matrix of salience weights describing the relative importance of the n issues to voter i and the way in which trade-offs are made between these issues, the distance between i and Theta is $\|\underline{\theta}_i - \underline{x}_i\|A_i$, where $\underline{\theta}_i = (\theta_{i1}, \ldots, \theta_{in})'$. The distance between i and Psi is $\|\underline{\psi}_i - \underline{x}_i\|A_i$, where $\underline{\psi}_i = (\psi_{i1}, \ldots, \psi_{in})'$. Voter i then prefers Theta to Psi if and only if $\|\underline{\theta}_i - \underline{x}_i\|A_i < \|\underline{\psi}_i - \underline{x}_i\|A_i$. As Hinich and Pollard (1981) show, this inequality holds if and only if

$$|z_i - \pi_\theta| < |z_i - \pi_\psi|, \tag{1}$$

where $z_i = (\underline{v}_i' A_i \underline{y}_i)/(\underline{v}_i' A_i \underline{v}_i)$ is i's most preferred point on the ideological dimension and $\underline{y}_i = \underline{x}_i - \underline{b}_i$. Expression (1) is a reformulation of Black's (1958) classic median-voter result. A linear mapping from a single predictive dimension into the larger space of campaign issues induces indirect preferences for each voter on the predictive dimension that are single peaked. Thus the median z_i is the optimal location for Theta and Psi on the predictive dimension. Viewing a candidate's ideology as immutable, at least for the duration of a single election period, this result means that the winning candidate will be the one whose ideology is closest to the median z_i.

Campaign tactics consist of candidate efforts to alter the b_{ij}'s and v_{ij}'s of the voters. If successful, these efforts result in movement of voter most preferred points (and the median z_i) on the underlying dimension. Thus, unlike classical spatial theory, in the perceptual model the candidates remain fixed during the election and the voters move in the underlying space.

Probabilistic Voting

Hinich, Ledyard, and Ordeshook (1973) develop a model in which each individual votes for one of two candidates or abstains, using a probabilistic decision rule. We will explain this model and show how it can be recast in terms of the perceptual model.

Alienation and indifference are two policy-related reasons for abstention (Hinich and Ordeshook, 1969; Ordeshook, 1970). Voters abstain from alienation if their utility for their favorite candidate fails to exceed what they consider a minimum level. In other words, even a favorite candidate leaves the voter cold. A voter abstains from indifference if the utility difference between the two candidates fails to exceed some minimal amount. In other words, the voter feels there just isn't a big enough difference between the candidates to make voting worthwhile.

Alienation and indifference are separate phenomena, but they may occur simultaneously. If the voter abstains because both candidates are equally obnoxious, this abstention is due to both alienation and indifference. If the voter abstains because both candidates are found to be equally acceptable, the abstention is due to indifference but not alienation, and if abstention occurs because the voter mildly dislikes one candidate but finds the other despicable, alienation but not indifference causes the behavior.

Alienation and indifference are rational, policy-related reasons for not voting. Both are based on voter utilities for the two candidates. Distance can be viewed as a (negative) utility function. Thus, we can express the hypotheses of abstention from alienation and indifference in terms of the weighted Euclidean distance preference rule developed in the previous section. Denote voter i's utility for Theta by $u(\underline{\theta}_i)$ and i's utility for Psi by $u(\underline{\psi}_i)$. Assume that

$$\begin{aligned} u(\underline{\theta}_i) &= c_i - \phi(\|\underline{\theta}_i - \underline{x}_i\|A_i) \\ u(\underline{\psi}_i) &= c_i - \phi(\|\underline{\psi}_i - \underline{x}_i\|A_i), \end{aligned} \tag{2}$$

where ϕ is a monotonically increasing function of its argument and $\phi\,(0) = 0$. The term c_i represents i's maximum utility level. This function may be asymmetric and idiosyncratic to the voter. None of the more restrictive assumptions about utility is required for our existence results. We need specify the functional form of ϕ only for purposes of characterizing this result. Thus, (2) should be viewed as a general distance representation of voter utility.

Let $p_{i\theta}[u(\underline{\theta}_i), u(\underline{\psi}_i)]$ be a general representation of the probability that i will vote for Theta, and $p_{i\psi}[u(\underline{\theta}_i), u(\underline{\psi}_i)]$ a general representation of the probability that i will vote for Psi. The more restrictive formulation used by Hinich, Ledyard, and Ordeshook (1973), hereafter referred to as HLO, to allow for abstention from alienation is

$$\begin{aligned} p_{i\theta} &= a_i[u(\underline{\theta}_i)] \\ p_{i\psi} &= a_i[u(\underline{\psi}_i)], \end{aligned} \tag{3}$$

where $\partial a_i \,/\, \partial u \geq 0$, $\partial^2 a_i \,/\, \partial u^2 \leq 0$ so that a_i is either a strictly concave function or a positive constant. For convenience, we will assume the latter.

Expression (3) postulates that $p_{i\theta}$ increases as $u(\underline{\theta}_i)$ increases, with $p_{i\psi}$ exhibiting the same type of behavior. Thus, the probability of voting for one's favorite candidate (which is necessarily greater than the probability of voting for the other candidate) declines as the utility associated with that candidate declines, in keeping with the intuitive concept underlying abstention from alienation. Further, this probability declines either at a constant or increasing rate as the utility for one's favorite candidate drops off. Again, an intuitive justification underlies this assumption, since it is reasonable to view alienation as a progressive condition.

Expression (3) should be seen as a probabilistic decision rule that weights three alternatives (vote for Theta, vote for Psi, abstain) in such a way that abstention becomes increasingly likely as alienation sets in. However, (3) presents a few difficulties. The probability of voting for Theta is not a function of $u(\underline{\psi}_i)$ nor is the probability of voting for Psi a function of $u(\underline{\theta}_i)$. If (3) were a complete specification of our probabilistic voting model, this assumption would require some justification.

We assume that voters can also abstain from indifference. As formulated by HLO, abstention from indifference can be permitted by setting

$$\begin{aligned} p_{i\theta} &= h_i[u(\underline{\theta}_i) - u(\underline{\psi}_i)], \text{ if } u(\underline{\theta}_i) > u(\underline{\psi}_i) \\ &= 0 \text{ otherwise;} \\ p_{i\psi} &= h_i[u(\underline{\psi}_i) - u(\underline{\theta}_i)], \text{ if } u(\underline{\psi}_i) > u(\underline{\theta}_i) \\ &= 0 \text{ otherwise,} \end{aligned} \tag{4}$$

with h_i a positive constant.

Expression (4) postulates that the probability of voting for one's favorite candidate increases as the utility difference between the two candidates increases. The probability of voting for a less preferred candidate is zero in expression (4). Observe that this is not the case in expression (3). We will justify this discrepancy shortly.

Like expression (3), expression (4) is viewed as a probabilistic decision rule that in this case specifies a positive probability of voting for one's more preferred candidate, a zero probability of voting for one's less preferred candidate, and a residual probability of not voting that increases as the utility difference between the two candidates decreases. Thus, (4) is a probabilistic voting rule that allows for abstention on the basis of indifference.

We can now formulate a complete model of probabilistic voting. Let $\epsilon_i (0 \leq \epsilon_i \leq 1)$ be the importance of alienation to voter i as a policy-related reason for nonvoting and $(1-\epsilon_i)$ the importance of indifference to voter i as a policy-related reason for nonvoting. We will henceforth drop the voter subscript for ϵ, a, and h to avoid cluttering up the algebra. The reader should remember that we continue to regard these parameters as idiosyncratic.

Again, borrowing from HLO, if alienation and indifference are causes of nonvoting, then

$$\begin{aligned} p_{i\theta} &= \epsilon\, a[u(\underline{\theta}_i)] + (1-\epsilon)\, h[u(\underline{\theta}_i) - u(\underline{\psi}_i)] \text{ if } u(\underline{\theta}_i) > u(\underline{\psi}_i) \\ &= \epsilon\, a[u(\underline{\theta}_i)] \text{ otherwise} \\ p_{i\psi} &= \epsilon\, a[u(\underline{\psi}_i)] + (1-\epsilon)\, h[u(\underline{\psi}_i) - u(\underline{\theta}_i)] \text{ if } u(\underline{\psi}_i) > u(\underline{\theta}_i) \\ &= \epsilon\, a[u(\underline{\psi}_i)] \text{ otherwise.} \end{aligned} \quad (5)$$

Thus, i's overall probability of voting [when $u(\underline{\theta}_i) > u(\underline{\psi}_i)$] is

$$p_{i\theta} + p_{i\psi} = \epsilon\, a[u(\underline{\theta}_i) + u(\underline{\psi}_i)] + (1-\epsilon)\, h[u(\underline{\theta}_i) - u(\underline{\psi}_i)].$$

If $\underline{\theta}_i = \underline{\psi}_i : p_{i\theta} + p_{i\psi} = 2\epsilon a[u(\underline{\theta}_i)]$. Since $p_{i\theta} + p_{i\psi} \leq 1$, it must be true that $\epsilon a[u(\underline{\theta}_i)] \leq 1/2$. Since $u(\underline{\theta}_i) = c_i$ is the maximum value of u, it is sufficient to assume that $\epsilon a c_i \leq 1/2$. Of course, it is also assumed that $p_{i\theta}$ and $p_{i\psi}$ are constrained to lie in the [0,1] interval.[1]

Examining (5), we see the effects of both alienation and indifference. $p_{i\theta}$ increases as $u(\underline{\theta}_i)$ increases and, for $u(\underline{\theta}_i) > u(\underline{\psi}_i)$, decreases as $u(\underline{\theta}_i) - u(\underline{\psi}_i)$ decreases. As is true for expression (3), expression (5) allows the voter to vote for the least preferred candidate. Our justification for allowing this lies in the existence of exogenous factors, specifically

[1] To interpret the constraint $\epsilon a c_i \leq 1/2$, let us 0–1 normalize i's utility function and set $c_i = 1$. If $a = 1$, then the constraint reduces to $\epsilon \leq 1/2$, which means that alienation cannot be more important to i as a reason for nonvoting than indifference. Examining (5), suppose $u(\underline{\theta}_i) = 1$ and $u(\underline{\psi}_i) = 0$. Then, $p_{i\theta} = \epsilon + (1-\epsilon)h$. If $h = 1$, then $p_{i\theta} = 1$ and $p_{i\psi} = 0$, which makes sense.

nonspatial candidate characteristics. Note in (2) that utility is formulated solely in spatial terms. In Enelow and Hinich (1982b), we subscripted the c term by the candidate as well as the voter, allowing this term to represent the nonpolicy value of the candidate to voter i. The second term in the utility function then represents the policy value of the candidate to voter i. In our present model we can achieve the same result by assuming a non-zero probability that i votes for the candidate whose policy value is less than that of the other candidate. Thus, (5) is a very general formulation, permitting abstention for policy-related reasons and allowing candidate choice to be influenced by nonspatial considerations, such as candidate personality and ascriptive traits.

Define $p_{i\theta} - p_{i\psi}$ as i's bias for Theta over Psi. The expected plurality for Theta over Psi as a fraction of the total number of voters is simply $E(p_{i\theta} - p_{i\psi})$, the average bias among the voters. Voter i's bias function is termed separable if there exists a function f_i such that for any $\underline{\theta}_i$, $\underline{\psi}_i$; $p_{i\theta} - p_{i\psi} = f_i(\underline{\theta}_i) - f_i(\underline{\psi}_i)$. To understand the meaning of a separable bias function, suppose $u(\underline{\theta}_i)$ increases and $u(\underline{\psi}_i)$ remains fixed. To be separable, the increase in i's bias for Theta over Psi must be independent of $u(\underline{\psi}_i)$. In other words, the increase in $p_{i\theta} - p_{i\psi}$ must not depend on the fixed value of $u(\underline{\psi}_i)$.

We now derive our first result. From (5), it follows, regardless of whether $u(\underline{\theta}_i) > u(\underline{\psi}_i)$ or $u(\underline{\psi}_i) > u(\underline{\theta}_i)$, that

$$p_{i\theta} - p_{i\psi} = \epsilon\, a[u(\underline{\theta}_i) - u(\underline{\psi}_i)] + (1 - \epsilon)\, h[u(\underline{\theta}_i) - u(\underline{\psi}_i)]. \tag{6}$$

Defining $f_i = [\epsilon a + (1 - \epsilon)h]u_i$, expression (6) implies that $p_{i\theta} - p_{i\psi} = f_i(\underline{\theta}_i) - f_i(\underline{\psi}_i)$. In addition, $p_{i\psi} - p_{i\theta} = f_i(\underline{\psi}_i) - f_i(\underline{\theta}_i)$. Thus, i's bias function is separable. Substituting for $u(\underline{\theta}_i)$ and $u(\underline{\psi}_i)$, we can rewrite (6) as

$$\begin{aligned} p_{i\theta} - p_{i\psi} &= g_i[u(\underline{\theta}_i) - u(\underline{\psi}_i)] \\ &= g_i[\phi(\|\pi_\psi \underline{v}_i - \underline{y}_i\|A_i) - \\ &\quad \phi(\|\pi_\theta \underline{v}_i - \underline{y}_i\|A_i)] \end{aligned} \tag{7}$$

where $\underline{y}_i = \underline{x}_i - \underline{b}_i$ and $g_i = [\epsilon a + (1 - \epsilon)h]$. We subscript g to make it clear that it is idiosyncratic.

We can use (7) to establish a ranking of positions on the predictive dimension based on the expected plurality of each position over that of any other. Assuming $\underline{v}_i$, A_i and $\underline{y}_i$ to be fixed parameters, we can define for any $\pi_1, \pi_2 \epsilon \Pi$,

$$\gamma(\pi_1, \pi_2) = E(p_{i1} - p_{i2}) = E\big[g_i[\phi(\|\pi_2\underline{v}_i - \underline{y}_i\|A_i) - \phi(\|\pi_1\underline{v}_i - \underline{y}_i\|A_i)]\big].$$

Thus, $\gamma(\pi_1, \pi_2)$ is the expected plurality for a candidate whose ideology is π_1 over a candidate whose ideology is π_2.

We can then say that π_1 is ranked higher than π_2 if and only if $\gamma(\pi_1, \pi_2) > 0$. The following result is due to Hinich and Ordeshook (1973): if $p_{i1} - p_{i2}$ is separable for all i, a transitive ranking of positions on Π is established.

We have shown that the probabilistic voting rule given by (5) is sufficient for candidate positions on the predictive dimension to be ranked from best to worst in terms of the expected plurality associated with any two positions. In fact, Hinich and Ordeshook's (1973) result is independent of the dimensionality of the candidate space. Suppose, then, that there are $p < n$ predictive dimensions underlying the n issues of the campaign. Assuming (5) as a model of voter choice in no way precludes multiple predictive dimensions. Thus, each voter's bias for any candidate over any other is separable and, consequently, a transitive ranking over positions in the multidimensional predictive space exists.

It is important to contrast our result with the equilibrium theorem derived by HLO. Two differences are notable. First, the HLO theorem guarantees a best position for each candidate but not a transitive ordering over all positions. Second, HLO require concave utility. It is clear that the voter bias function expressed by (6) is separable, regardless of the shape of the voter's utility function.

At this point it may be wondered how we have sidestepped the disequilibrium results that loom so large in the spatial modeling literature. The work of Plott (1967), McKelvey (1976, 1979), and others establishes that in multidimensional choice spaces, two-candidate competition is rarely characterized by a best position for the candidates. Only under the unlikeliest configuration of voter preferences does a best position exist; otherwise, candidate positions can be arrayed in a single-cycle set that subsumes the entire choice space.

These results depend on deterministic voting. If each voter casts a vote with certainty for the candidate for whom the highest utility is perceived, the disequilibrium results cited above are unavoidable. Probabilistic voting models avoid these results by eliminating the discontinuity in the candidate's expected plurality function that exists under deterministic voting. The works of Coughlin and Nitzan (1981a, 1981b), Hinich, Ledyard, and Ordeshook (1972, 1973), and Denzau and Kats (1977) all demonstrate that probabilistic voting models of the type described in this paper are characterized by a much higher degree of candidate stability than is found in deterministic voting models.

Abstention from alienation, particularly in a polarized society, is associated with more than a single optimal candidate location if one exists. However, our model leads to a single optimal location. How can this discrepancy be explained? The answer lies in the joint postulation of alienation and indifference as causes of abstention. Alienation as the sole cause of abstention can lead to multiple optima for the candidates (for example, the modes of voter opinion). Alienation and indifference as joint causes of abstention are invariably associated with a unique candidate equilibrium when one exists (Riker and Ordeshook, 1973).

Optimal Candidate Location

We will now solve for the optimal candidate location on the ideological dimension. It should be kept in mind that we have more than an equilibrium result. Separability of voter bias functions is sufficient to rank ideologies from best to worst in terms of expected pluralities.

Let there be $k \geq 2$ homogeneous groups in the electorate, called voting blocs or groups. For the members of each voting bloc v_i, A_i, y_i, and g_i will be assumed to be identical. All members of the bloc have the same preferences, the same translation weights, and attach the same importance to alienation and indifference. In other words, all members of a given voting bloc view politics the same way.

This is a heuristic device used for purposes of motivating the discussion. We might average bias functions over individual voters, but this detracts from our understanding. Politicians simplify their environment by viewing the electorate as composed of aggregates: farmers, blacks, union members, Jews, etc. Our assumption of homogeneous voting blocs captures the essence of this practice. There is nothing to be gained by politicians in attempting to decipher the attitudinal complexities of various voting bloc members. Instead, as a shorthand device, it is much simpler for politicians to lump voters together into politically significant groups of like-minded members. Campaigns are devised on this basis.

Let N_i be the fraction of the total number of voters in voting bloc i so that $\sum_{i=1}^{k} N_i = 1$. For reasons of specificity, assume that u is a quadratic function, so that $\phi(\|\underline{\theta}_i - \underline{x}_i\| A_i) = \|\underline{\theta}_i - \underline{x}_i\|^2 A_i$. Returning to (7), we see that

$$p_{i\theta} - p_{i\psi} = g_i[(\pi_\psi^2 - \pi_\theta^2)(\underline{v}_i' A_i \underline{v}_i) + 2(\pi_\theta - \pi_\psi)(\underline{v}_i' A_i \underline{y}_i)], \quad (8)$$

so that

$$\frac{\partial E(p_{i\theta} - p_{i\psi})}{\partial \pi_\theta} = \sum_{i=1}^{k} N_i g_i[-2\pi_\theta(\underline{v}_i' A_i \underline{v}_i) + 2(\underline{v}_i' A_i \underline{y}_i)]. \quad (9)$$

Letting π^* be the optimal ideology, we can solve for π^* by setting (9) equal to zero, which yields

$$\pi^* = \frac{\sum_{i=1}^{k} N_i g_i (\underline{v}_i' A_i \underline{y}_i)}{\sum_{i=1}^{k} N_i g_i (\underline{v}_i' A_i \underline{v}_i)} . \tag{10}$$

It is simple to see that the second-order condition for a maximum is satisfied since

$$\frac{\partial^2 E(p_{i\theta} - p_{i\psi})}{\partial \pi_\theta^2} = -2 \sum_{i=1}^{k} N_i g_i (\underline{v}_i' A_i \underline{v}_i),$$

where A_i is positive definite and $g_i > 0$, and thus $E(p_{i\theta} - p_{i\psi})$ is a symmetric, concave function of π_θ (the second partial derivative is negative). The symmetry of $E(p_{i\theta} - p_{i\psi})$ establishes the transitive ranking of positions on Π. A position π_θ has a positive expected plurality over a position π_ψ if and only if π_θ is closer to π^* than π_ψ. Thus, the closer the candidate's position is to π^*, the better.

Recall that $z_i = (v_i' A_i y_i)/(v_i' A_i v_i)$ is i's most preferred candidate ideology. We can use this result to reexpress (10) as

$$\pi^* = \frac{\sum_{i=1}^{k} N_i g_i z_i (\underline{v}_i' A_i \underline{v}_i)}{\sum_{i=1}^{k} N_i g_i (\underline{v}_i' A_i \underline{v}_i)} . \tag{11}$$

Since $p_{i\psi} - p_{i\theta} = g_i[u(\underline{\psi}_i) - u(\underline{\theta}_i)]$, π^* is also optimal for Psi.

Discussion

Let us now interpret this result. The easiest way to do so is with a concrete example. The example we will use is from Enelow and Hinich (1984).

Assume there are two campaign issues: public land use and pollution control. The electorate consists of two voting blocs — a pro-growth group and a no-growth group. The first group wishes to see a large proportion of publicly owned land opened to private development and believes that air and water pollution standards should be considerably eased. The second group has diametrically opposite views, but both voting blocs stop short of total radicalism.

If issue 1 is public land use, let this issue be measured on a scale that represents the proportion of all federally owned lands that conceivably can be opened to private development. If issue 2 is pollution control, let this issue be measured on a scale that represents the proportion of some ideal level of air and water quality (defined in terms of the amounts of various types of pollutants) that is attainable by government regulation.

For sake of argument, assume the ideal point of group 1 to be $\underline{x}_1 = (.35, .65)'$. The pro-growth group most prefers that 35 percent of available federal lands be opened to private development (they do want a few national parks left). The same group also most prefers that 65 percent of attainable air and water quality be mandated by government (they do, after all, have to breathe the air and drink the water).

Let $\underline{x}_2 = (.2, .8)'$ be the ideal point of the no-growth group (they do recognize that the economy has recently been in a recession). In addition, assume $A_i = I$ for all voters in the electorate (i.e., both issues are equally important and preferences are separable by issue). We assume this simply for convenience.

From the candidates' viewpoint, the data we have described can be obtained through polls. It is also possible to obtain survey data on voter perceptions of the incumbent's policy on a campaign issue (i.e., b_{ij}) and the perceived policies of other candidates on the same issue (i.e., θ_{ij} or ψ_{ij}). Since $(\theta_{ij} - b_{ij})/\pi_\theta = v_{ij}$, this data permits estimation of the v_{ij}'s of the voters.

Assume then that $\underline{b}_1 = (.2, .8)'$ and $\underline{v}_1 = (.3, -.3)'$ for the pro-growth group, while $\underline{b}_2 = (.3, .7)'$ and $\underline{v}_2 = (.2, -.2)'$ for the no-growth group. The no-growth group believes that present policies are closer to the ideal policies of the pro-growth group than the pro-growth group believes they are. The no-growth group also believes that a unit move to the right ideologically will result in less change in these policies than the pro-growth believes will occur. Both groups believe that the election of a challenger one unit to the right of the incumbent will result in the same actual policies (i.e., [.5, .5]).

We need only specify N_1, N_2, g_1, and g_2 to compute the optimal candidate ideology. N_1 and N_2 are estimable from polls. Set $N_2 = 2N_1$ for the sake of argument. There are twice as many no-growth as pro-growth voters in the electorate. Estimating g_1 and g_2 appears more difficult. As a result, we will set $g_1 = g_2$. If the candidates can't obtain this information, it is reasonable to assume they will discount any differences in this parameter across groups.

We can now compute π^*. First, we see that

$$z_1 = (\underline{v}_1' A_1 \underline{y}_1)/(\underline{v}_1 A_1 \underline{v}_1) = 1/2 \text{ and } z_2 = -1/2.$$

Thus, the most preferred ideology of the pro-growth group is one unit to the right of the most preferred ideology of the no-growth group. If all voters voted with certainty for the candidate closest to them on the set of campaign issues, the median z_i would be the optimal ideology for both candidates. Since $N_2 = 2N_1$, the median $z_i = -1/2$ would therefore be optimal. The ideological favorite of the no-growth group would also be the ideal candidate in a two-candidate contest.

However, $-1/2$ is not an optimal ideology under our model of probabilistic voting. It is easily seen that

$$\pi^* = \frac{N_1 z_1(v_{11}^2 + v_{12}^2) + N_2 z_2(v_{21}^2 + v_{22}^2)}{N_1(v_{11}^2 + v_{12}^2) + N_2(v_{21}^2 + v_{22}^2)} = .03 \quad (12)$$

is the optimal ideology. Probabilistic voting makes centrist ideologies attractive labels for the candidates (where the center is defined relative to the most preferred ideologies of the voting groups in the electorate). Assume $a = h = 1$. Then a candidate Theta whose ideology is .03 running against a candidate Psi whose ideology is $-1/2$ will receive a positive *expected plurality* of

$$E(p_{i\theta} - p_{i\psi}) = \sum_{i=1}^{2} N_i[.25\,(\underline{v}_i' \underline{v}_i) + 1.06\,(\underline{v}_i' \underline{y}_i)] = .032$$

or 3.2 percent of the total vote. Clearly, the centrist candidate is expected to do better at the polls.

It is apparent from (11) and (12), however, that π^* is also affected by the $\underline{v}$'s of the various voting blocs. π^* is not strictly a function of relative group size. If we were to set $N_1 = N_2$, $\pi^* = .19$, which is closer to the most preferred ideology of the pro-growth group. The magnitude of the vector $\underline{v}_i$ ($\|\underline{v}_i\|$) denotes the policy distance as perceived by members of group i between two candidates one unit apart on the ideological dimension. In a direct sense, $\|\underline{v}_i\|$ measures the policy significance of ideology to members of group i since it represents the size of the policy difference associated with a unit difference in ideology. What is clear is that $\|\underline{v}_i\|$ bears a direct relation to π^*. As $\|\underline{v}_i\|$ increases, π^* moves toward z_i. We can see this most easily in (12). However, it is unnecessary for A_i to equal I to get this result. If λ_i is the largest eigenvalue of A_i, $\lambda_i \, \underline{v}_i' \underline{v}_i \geq \underline{v}_i' A_i \underline{v}_i$, an inequality which allows us to use the same type of proof we will employ.

For simplicity, let $\underline{v}_2 = c\underline{v}_1$, where c is a scalar. Thus $\underline{v}_2$ is some multiple of $\underline{v}_1$. We can then reexpress (12) as

$$\pi^* = \frac{N_1 z_1 \|\underline{v}_1\|^2 + N_2 z_2 c^2 \|\underline{v}_1\|^2}{N_1 \|\underline{v}_1\|^2 + N_2 c^2 \|\underline{v}_1\|^2} = \frac{N_1 z_1 + N_2 z_2 c^2}{N_1 + N_2 c^2} . \tag{13}$$

Now suppose that $\|\underline{v}_1\|$ becomes very large relative to $\|\underline{v}_2\|$. Then c^2 gets very small, approaching zero. It is plain from (13) that as this happens π^* approaches z_1. If $\|\underline{v}_1\|$ increases but $\underline{y}_1 = \underline{x}_1 - \underline{b}_1$ remains the same, z_1 will get smaller and π^* will increase. For example, if $\underline{v}_1$ is multiplied by 2, then $\|\underline{v}_1\|$ increases by a factor of 4, $z_1 = .25$, and π^* increases from .03 to .11.

Recalling our assumption that candidate ideology is fixed during an election period, this result suggests a way by which disadvantaged candidates can improve their prospects. Suppose that one candidate is closer to π^* than the other. Instead of simply accepting defeat, the disadvantaged candidate can attempt to increase $\|\underline{v}_i\|$ for whichever group i is nearest on the predictive dimension. In this manner, π^* is moved closer to the candidate.

To increase $\|\underline{v}_i\|$, the candidate must convince the members of group i that the ideological difference with the competing candidate is of great significance in terms of the policies that will be enacted. For example, suppose that $z_2 \leq \pi_\psi < \pi_\theta \leq \pi^*$. Psi must then convince the no-growth group (group 2) that the election of Theta will result in policies disastrous to public lands and environmental quality. Of course, Theta should attempt to counter these charges.

The policy significance of ideology is therefore an important factor in determining the optimal candidate ideology. A voting bloc that sees great policy significance in ideology will exert a stronger influence over the election outcome than a bloc that sees ideology as bearing little significance for policies.

One conclusion we can draw from this result is that centrist outcomes are promoted by widely shared views of the policy significance of ideology. If $\|\underline{v}_i\|$ is the same for each voting bloc i (as well as A_i and g_i), then π^* is the mean z_i. This result also provides insights into conditions under which noncentrist outcomes are promoted. Differences among voting blocs in the policy significance attributed to ideology have been mentioned as one reason why π^* may deviate from the mean z_i. Another reason for divergence may be differential turnout rates among various voting groups.

Differential turnout can be represented in our model through differences in g_i across voting blocs. Suppose, other things being equal, that a member of group 2 is twice as likely to vote as a member of group 1. We can achieve this effect by setting $a_2 = 2a_1$ and $h_2 = 2h_1$ in expression (5), so that $g_2 = 2g_1$. Recomputing π^* in expression (12), we find that π^* decreases from .03 to $-.14$. Thus, if members of group 2 are more likely to vote than members of group 1, π^* will be biased in favor of group 2's most preferred ideology.

Rabinowitz (1978) and Poole and Rosenthal (1982) find that candidates in recent American presidential elections do not converge to the position of the mean voter and, in fact, are rather far apart in the recovered space. While our model does not necessarily specify the mean voter's location as optimal for the candidates, it does specify a single optimal location. Why, then, are presidential candidates not closer together?

Recall our basic premise that it is the voters and not the candidates whose positions change in the underlying space in the course of an election campaign. Poole and Rosenthal (1982) provide empirical support for this assertion. In scaling the 1980 Major Panel File, they find "little or no deterioration in fit in assuming that *candidates have constant spatial positions.*" In other words, it is quite reasonable to argue that candidate positions are relatively fixed by the time the campaign begins.

As discussed in Enelow and Hinich (1984), candidate positions in the predictive space evolve over a relatively long period of time from a number of sources. Party affiliation is perhaps the most important source, with voters consistently viewing Democratic politicians as left of center in terms of economic ideology and Republican politicians as right of center (Enelow and Hinich, 1984). Given a candidate's need to appeal to primary voters and a party's support groups, this finding is not surprising.

Association with organized interest groups is a related factor contributing to the voters' perception of a candidate's predictive label. Endorsements by the AFL-CIO or the National Conservative Political Action Committee (NCPAC) are also likely to convince voters that the candidate is not located at the center of the predictive space.

For these reasons, American presidential candidates are not likely to be found at the center of the electorate. However, we have argued elsewhere (Enelow and Hinich, 1984) that relative to the distribution of voters, American presidential candidates are still relatively close to the mean voter. In any event, the center of the candidate cluster is never far from the center of the voters.

The importance of π^* can then be understood in two ways. First, to the extent that π^* is seen by the candidates as fixed, it acts as a centripetal force counteracting the pull of party support groups and other biased constituencies. Second, to the extent that π^* is seen as movable, through one of the tactics described above, each candidate will compete in an effort to move π^* closer. In either case, the importance of π^* is clear. A single location in the predictive space is optimal for both candidates engaged in two-candidate competition.

It is interesting to compare the optimal candidate ideology under probabilistic voting with the optimal candidate ideology under deterministic voting, with a nonspatial term in the voter's utility function representing the nonpolicy value of each candidate. This latter model is employed by Enelow and Hinich (1982b, 1984). As mentioned earlier, if the c term in expression (2) is subscripted by voter and candidate, $c_{i\theta}$ can be seen as voter i's assessment of the nonpolicy value of Theta. The difference $c_{i\psi} - c_{i\theta}$ is then a random variable among the members of voting bloc i, since perceptions of the nonpolicy difference between two candidates can be expected to vary.

Let s_i be the standard deviation for voting bloc i of $c_{i\psi} - c_{i\theta}$. Then, assuming s_i to be sufficiently large for i = 1, 2 to satisfy a lower bound derived in Enelow and Hinich (1984),

$$\pi^* = \frac{N_1 s_1^{-1} z_1(v_{11}^2 + v_{12}^2) + N_2 s_2^{-1} z_2(v_{21}^2 + v_{22}^2)}{N_1 s_1^{-1}(v_{11}^2 + v_{12}^2) + N_2 s_2^{-1}(v_{21}^2 + v_{22}^2)} . \tag{14}$$

Interestingly, if $s_1 = s_2$, then (14) is identical to (12).

To gain some insight as to why (14) and (12) are so similar, we can write the distribution function for the nonpolicy model as

$$\begin{aligned} &F(\|\underline{\psi}_i - \underline{x}_i\|^2 - \|\underline{\theta}_i - x_i\|^2) = \\ &P_r(c_{i\psi} - c_{i\theta} < \|\underline{\psi}_i - \underline{x}_i\|^2 - \|\underline{\theta}_i - \underline{x}_i\|^2), \end{aligned} \tag{15}$$

which is the probability that a voter from bloc i will prefer Theta to Psi. Refer, now, to expression (4), which formalizes probabilistic voting with abstention from indifference. Substituting for $u(\underline{\theta}_i)$ and $u(\underline{\psi}_i)$ with the squared distance representation of utility, we can rewrite $p_{i\theta}$ for the case where $u(\underline{\theta}_i) > u(\underline{\psi}_i)$ as

$$h_i(\|\underline{\psi}_i - \underline{x}_i\|^2 - \|\underline{\theta}_i - \underline{x}_i\|^2). \tag{16}$$

If the standard deviation of $c_{i\psi} - c_{i\theta}$ is very large, the distribution function given by (15) will approximate the linear function given by (16). In short, (16) and (15) are very much alike.

We see that probabilistic voting can have the same effect on election outcomes as deterministic voting when voter decisions are influenced by nonspatial candidate characteristics. This is an unexpected result since the two models are quite different. Under the deterministic model, everyone votes. Further, voters must have sufficiently diverse views of the nonpolicy difference between the candidates for an optimal ideology to exist. The probabilistic model permits abstention and yields an optimal ideology without any statement about the distribution of views concerning the nonpolicy values of the candidates. The result for the deterministic model is also explicitly developed under the assumptions of concave utility and a single predictive dimension. Unlike the result we have derived for the probabilistic model, a generalization for the deterministic model would not be easy.

Conclusion

We have shown that if voters use a probabilistic voting model that permits abstention from alienation and indifference, an optimal candidate ideology exists. In fact, candidate ideologies can be ranked in a transitive ordering with respect to the expected plurality of each ideology over that of any other. This result holds for multiple predictive dimensions and utility functions of any shape.

The optimal ideology for both candidates is a weighted mean of the most preferred ideologies of the voters. Other things being equal, the optimal candidate ideology is weighted toward the voting bloc for which ideology has the greatest policy significance. If all blocs see ideology as having the same policy significance, weight issues identically, and attach the same importance to alienation and indifference, the mean most-preferred ideology among the voters is the optimal candidate ideology.

Finally, probabilistic voting, while based on another set of behavioral assumptions, can have the same effect on election outcomes as deterministic voting with widely differing views of the difference in nonpolicy values attached to the candidates. Both types of voting can lead to centrist election outcomes.

As stated earlier, an important concern of the candidates is how they are perceived by the voters. Offhand comments can be distorted by the press or printed out of context. A seemingly innocent remark about ethnic neighborhoods can be blown up to appear as evidence of racism. "Shooting oneself in the foot" is a concern that leads many campaign managers to muzzle their candidates. In short, candidates and those who work for them are constantly worried about how their statements and behavior will appear to the electorate. The probabilistic voting model is consequently a reasonable one for the candidates to employ. It assumes that citizens cast votes both on the basis of how much they like each candidate and on the basis of how much more they like one candidate than another. In addition, the model permits the voter to abstain for policy-related reasons and to vote for a less-preferred candidate for nonpolicy reasons. In short, almost every worry the candidate may have about the possible behavior of the voter is captured by the probabilistic voting model. It therefore is a good tool to employ.

An important theme of this paper is that the center does exert a strong force over election outcomes. The center we have described is the ideological center of the electorate. As we have seen, the optimal candidate ideology is not simply a zero point on the ideological scale. Instead, it is a function of several parameters besides relative group size: the policy significance of ideology to the voters, the way voters weight issues, the importance voters attach to alienation and indifference, differential rates of turnout, etc. But the central point should not be lost. The optimal candidate ideology is likely to be close to the center of the ideological spectrum. The conditions required to drag π^* far away from the mean z_i are unlikely to be met in democratic societies.

Thus, we conclude as we began this paper, on an optimistic note. The repeated finding of multidimensional scaling analysis that the major party candidates in democratic elections tend to cluster around the center of voter opinion is supported by theory. It is possible to construct a general theoretical explanation for the center as an attractive force in democratic electoral politics. To do so is to perform what should be the obvious goal of election theory: to explain what actually happens.

References

Black, D. (1958). *Theory of Committees and Elections.* Cambridge: Cambridge University Press.

Cahoon, L. S., M. J. Hinich, and P. C. Ordeshook (1978). "A Statistical Multidimensional Scaling Method Based on the Spatial Theory of Voting." In P. C. Wang (ed.), *Graphical Representation of Multivariate Data.* New York: Academic Press.

Coughlin, P., and S. Nitzan (1981a). "Directional and Local Electoral Equilibria with Probabilistic Voting." *Journal of Economic Theory* 24: 226–39.

——— (1981b). "Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Maxima." *Journal of Public Economics* 15: 113–21.

Davis, O. A., and M. J. Hinich (1966). "A Mathematical Model of Policy Formation in a

Democratic Society." In Joseph Bernd (ed.), *Mathematical Applications in Political Science, II.* Dallas: Southern Methodist University Press.

_________ (1967). "Some Results Related to a Mathematical Model of Policy Formation in a Democratic Society." In J. L. Bernd (ed.), *Mathematical Applications in Political Science, III.* Charlottesville: University of Virginia Press.

Denzau, A., and A. Kats (1977). "Expected Plurality Voting Equilibrium and Social Choice Functions." *Review of Economic Studies* 44: 227–33.

Downs, A. (1957). *An Economic Theory of Democracy.* New York: Harper and Row.

Enelow, J. M., and M. J. Hinich (1981). "A New Approach to Voter Uncertainty in the Downsian Spatial Model." *American Journal of Political Science* 25: 483–93.

_________ (1982a). "Ideology, Issues, and the Spatial Theory of Elections." *American Political Science Review* 76: 493–501.

_________ (1982b). "Nonspatial Candidate Characteristics and Electoral Competition." *Journal of Politics* 44: 115–30.

_________ (1984). *The Spatial Theory of Voting: An Introduction.* Cambridge: Cambridge University Press.

Hinich, M. J. (1978). "Spatial Voting Theory When Voter Perceptions of Candidates Differ." Virginia Tech, CE 78–11–4.

Hinich, M. J., J. Ledyard, and P. C. Ordeshook (1972). "Nonvoting and the Existence of Equilibrium Under Majority Rule." *Journal of Economic Theory* 4: 144–53.

_________ (1973). "A Theory of Electoral Equilibrium: A Spatial Analysis Based on the Theory of Games." *Journal of Politics* 35: 154–93.

Hinich, M. J., and P. C. Ordeshook (1969). "Abstentions and Equilibrium in the Electoral Process." *Public Choice* 7: 81–106.

_________ (1973). "A Sufficient Condition for an Election Equilibrium and the Implications of Alternative Empirical Models." Business and Economic Statistics Section, *Proceedings of the American Statistical Association.*

Hinich, M. J., and W. Pollard (1981). "A New Approach to the Spatial Theory of Electoral Competition." *American Journal of Political Science* 25: 323–41.

McKelvey, R. D. (1976). "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control." *Journal of Economic Theory* 12: 472–82.

_________ (1979). "General Conditions for Global Intransitivities in Formal Voting Models." *Econometrica* 47: 1085–1111.

Ordeshook, P. C. (1970). "Extensions to a Mathematical Model of the Electoral Process and Implications for the Theory of Responsible Parties." *Midwest Journal of Political Science* 14: 43–70.

Plott, C. R. (1967). "A Notion of Equilibrium and Its Possibility Under Majority Rule." *American Economic Review* 57: 787–806.

Poole, K., and R. Rosenthal (1982). "The Estimation of Spatial Voting Models." Carnegie-Mellon University, unpublished paper.

Rabinowitz, G. (1973). *Spatial Models of Electoral Choice: An Empirical Analysis.* Chapel Hill, NC: Institute for Research in Social Science.

_________ (1978). "On the Nature of Political Issues: Insights from a Spatial Analysis." *American Journal of Political Science* 22: 793–817.

Riker, W., and P. Ordeshook (1973). *An Introduction to Positive Political Theory.* Englewood Cliffs, NJ: Prentice-Hall.

Weisberg, H., and J. Rusk (1970). "Dimensions of Candidate Evaluation." *American Political Science Review* 64: 1167–85.