Dynamic models of the voter's decision calculus: Incorporating retrospective considerations into rational-choice models of individual voting behavior*

MARTIN J. ZECHMAN

Abstract

Since Kramer's article (1971), a growing body of literature indicates that U.S. national elections can be viewed as referenda on the performance of incumbent administrations. Retrospective considerations, however, have not been explicitly incorporated into a spatial model of party competition.

The model of voting behavior presented in this paper provides a mechanism for the inclusion of these retrospective considerations into spatial models. Borrowing liberally from the concepts of Bayesian decision theory, this model allows the voter to use all of his political information in estimating a party's future program. Retrospective considerations are represented by the voter's estimate of a party's future policies prior to the campaign. The voter's initial expectations are revised during the campaign as he acquires additional information. His decision is based upon these revised estimates.

During the past twenty years, the formal structure of spatial models of party competition has evolved considerably. Downs' (1956) recognition of the interdependence of the policies candidates advocate and the decisions voters make remains the cornerstone of these models. The structure of current models, however, bears little resemblance to his one-dimensional spatial model in which everyone is assumed to vote and to vote rationally. For example, the model's single issue dimension has been replaced by a multidimensional issue space. The assumption that everyone must vote has also been substantially modified. Citizens have been allowed to abstain because they were either indifferent between the candidates' policies or alienated by them. The consequences of varying the candidates' objectives have been considered. Some complications associated with the necessity of appealing for support from two distinct groups have been evaluated (see McKelvey, 1972). Even the possibility that the candidate might benefit from deliberately adopting an ambiguous campaign platform has been investigated (see Shepsle, 1972).

Throughout this period, the assumption of *perfect spatial mobility* has remained inviolate; that is, no constraints have been placed upon a candi-

Public Choice 34 (1979) 297-315.

date's freedom to manipulate the electorate's perceptions of his campaign platform. Voters are assumed not to be predisposed in favor of one of the parties. Substantively, this means that any candidate is free to adopt any set of issue positions as his campaign platform. A Democrat can adopt a strong anti-union stand as easily as he can advocate pro-union measures. Moreover, if two candidates exchange platforms, they must also exchange their share of the vote. Formally, these conditions are sufficient to induce symmetry into the competitive relations between the candidates (Aranson, Hinich and Ordeshook, 1974), p. 138). Neither candidate is favored initially. Both candidates have an equal opportunity to influence their electoral fortunes.

Symmetry in the competitive relations between candidates plays a crucial role in the formal development of traditional spatial models – those models that include the assumption of perfect spatial mobility. Spatial models represent a two-candidate election as a two-person zero-sum game. The assumption of perfect spatial mobility is required in order to demonstrate that this game is also *symmetric* (McKelvey, 1972, pp. 27-39). Letting S_i represent Candidate i's set of possible strategies ($i = 1, 2$), then a two-person zero-sum game is defined to be symmetric only if S_1 and S_2 are equivalent and whenever the candidates exchange strategies, they exchange payoffs.[1] Since it can be proven that in a symmetric zero-sum game the payoff to each candidate associated with an equilibrium pair of strategies is necessarily zero, this property greatly simplifies the search for optimal campaign strategies (McKelvey, 1972, p. 22). The analyst can confine his search to those strategies that at worst guarantee a candidate a tie in the election. This, in fact, is what has generally been done. Consequently, eliminating the assumption of perfect spatial mobility from the formal structure of spatial models would require the alteration of existing interpretations of the conclusions of most of this literature's major theorems.[2] If citizens are predisposed to favor one of the candidates, it is not necessarily an optimal strategy for either candidate.

The principle focus of empirical objections to the assumption of perfect spatial mobility is its requirement that citizens cannot be predisposed to favor one of the candidates prior to the start of the campaign. This requirement implicitly entails constraints on the information that a voter can use in his comparative evaluation of the candidates. He is limited to that information that can be completely controlled by a candidate. Since a candidate can only manipulate the information reaching the electorate during the course of the campaign, this assumption effectively wipes his slate clean. Thus, voters must ignore the institution of political parties (Riker and Ordeshook, 1973, p. 334). A citizen's knowledge of each party's record prior to the campaign cannot be used in deciding which candidate he prefers. Similarly, a voter must ignore any information pertaining to the candidate's performance prior to the campaign, even if he acquires this information during the campaign. In short, the assumption of perfect spatial mobility necessitates the isolation of a campaign in time.

By effectively limiting a voter's choice to his evaluation of each party's campaign platform, the assumption of perfect spatial mobility contradicts so much of what we have learned about individual voting behavior over the past thirty-five years. Generally, people are predisposed in favor of one of the candidates. In their discussion of the meaning of this assumption, Riker and Ordeshook (1973) acknowledge that

> . . . this assumption does not typically characterize the real world. . . . [C] itizens in a secure Democratic constituency favoring liberal labor legislation may remain unconvinced that a Republican is in fact more pro-labor than the Democrat. In such a case the Republican is not free to adopt any position, simply because the voters' biases will not permit him to do so (pp. 333-334).

While the candidates' actions during the campaign contribute to their electoral fortunes, they cannot, in general, exchange payoffs simply be exchanging platforms. The party that they represent and their prior records also significantly influence the outcome of an election.

By itself, the fact that voters are predisposed in favor of one of the parties does not necessitate the elimination of the assumption of perfect sptial mobility from the formal structure of spatial models. In developing formal models of political phenomena, abstraction and simplification are the name of the game. The reasonableness of a decision to strike away particular aspects of a phenomenon can only be evaluated within the context of an investigator's objectives. Spatial analysis is designed to evaluate the desirability of various campaign strategies for a candidate. Given this goal, however, the cost of including the assumption of perfect spatial mobility in spatial models cannot be justified. By ignoring the constraints placed upon a candidate's spatial mobility, there is a substantial risk of seriously distorting any evaluation of the electoral appeal of campaign platforms.[3]

Reducing this risk within the structure of a spatial model necessitates the incorporation of the electorate's predispositions into the candidate's decision calculus. This, in turn, requires the modification of the assumption of perfect spatial mobility. But, how should it be modified? Any acceptable answer to this question must precisely specify the constraints that the electorate's predispositions place upon a candidate's spatial mobility. This goal can be achieved by altering the voter's decision calculus. In particular, by removing the constraints placed upon a voter's information, we can introduce constraints upon the candidate's mobility. If voters are not compelled to trust a candidate's promises, then advocating a particular campaign platform is not sufficient to guarantee that the candidate will be believed. When a candidate's platform differs from a voter's expectations, why should the voter completely ignore these expectations?

Any attempt to remove the constraints on a voter's state of information requires a precise determination of the information that a voter can use in deciding between the candidates. It also entails the specification of a mechanism for the modification of a voter's predispositions as he acquires additional information during the campaign. The combination of a citizen's prior expectations, his perceptions of a candidate's campaign platform, and this mechanism are sufficient to indicate the constraints a voter's predispositions place on a candidate's mobility.

The paradigm of individual voting behavior presented in the remainder of this paper provides a rule for combining all of a voter's political knowledge into an estimate of each party's future program. These estimates constitute the basis for his voting decision. By incorporating many of the essential elements of Bayesian decision theory into this paradigm, it is a framework for precisely specifying the interdependence between the voter's current electoral choice and his previous political experience.

A paradigm for dynamic models of voting behavior

During the pre-election period, a voter accumulates information about each party's platform. This information provides the basis for adjusting his beliefs about the program each party will enact if its candidates are elected. At the conclusion of the campaign, he must decide which party he will support. His vote may be essentially random. Perhaps he stumbles into the voting booth, inadvertently pressing a series of levers and then lurching out. In this instance, a citizen's vote is independent of his judgments about each party's likely program. A citizen's decision at the end of the campaign may be deliberate. His vote is directly linked to his assessment of each party's likely program. He compares the desirability of the two parties' likely future programs and votes for the one whose policies he prefers.

Only the decisions of deliberate voters are considered here. Assuming these voters are rational, their decision is derived from an ordering of the parties with respect to the attractiveness of the future programs each party is likely to implement. The decision calculus of deliberate voters can be presented as two empirically interrelated but analytically distinct phases. The first stage focuses upon the estimation of the program that the competing parties would be likely to adopt if they were to win the upcoming election. The second stage concerns the determination of the voter's preferred party.

A citizen's assessment of the likelihood that a party will enact a particular program if it wins the next election rests upon his perceptions of the policies its leaders have pursued in the past and his perceptions of its campaign platform. This judgment also depends upon his confidence that these perceptions actually are related to a party's future performance. Uncertainty surrounds any citizen's perceptions of a political party's expected program.

The political information that reaches the electorate is necessarily incomplete. Even the most zealous follower of public affairs is usually confined to the news media as a source of information. Furthermore, political parties often have strong incentives for muddying their records (Shepsle, 1972). The vagueness of a voter's perceptions of a party's record and its campaign platform contributes to his uncertainty about these policies. Moreover, his uncertainty is enhanced by the very nature of forecasting. Any attempt to predict the future using knowledge about the past is risky. Circumstances change! Perhaps the party's leaders may have deliberately misrepresented their views in order to win votes or may no longer adhere to the principles that guided their party's policies in the past. In either case, the usefulness of a voter's political information for the determination of a party's future policies is questionable.[4]

In short, any model of the relationship between a voter's political knowledge and his estimate of a party's future program must include a representation of his uncertainty that the party will enact its expected program. This criterion can be satisfied by representing a voter's estimate of a party's future program as an n-dimensional random vector (one dimension for each of the n policies that constitute its program). This joint probability density function (*jpdf*) associated with the random vector characterizes his assessment of the likelihood that a party will pursue particular courses of action.[5] Properties of this *jpdf* can be used to summarize various aspects of his expectations. For example, the mode of the distribution of this random vector represents the voter's estimate of the party's future program, while its covariance matrix can be used as the basis for an indicator of his uncertainty that the mode will be this party's future program.

Prior to the beginning of the election campaign, a citizen's estimate of a party's future program is necessarily based upon his knowledge of this party's record. His presumption is that a party's past performance indicates its future course of action. During the campaign, a voter often acquires additional information that he believes can be used in revising initial estimates of each party's future program. Since the voter's beliefs about the likely nature of a party's future program is represented by a *jpdf*, probability theory provides a mechanism for the process governing these revisions. In particular, Bayes' theorem indicates how an individual's beliefs about a phenomenon are modified by his acquisition of additional information (Kyburg, 1970, p. 74).

Specifically, let μ_{it}^{p} represent the program Party i will enact if that party wins Election t; let R_{ikt} represent Voter K's knowledge of Party i's record prior to the campaign for Election t; and let $\bar{C}_{ikt}$ represent Voter K's prior distribution of Party i's future program, denoted by $\pi'(\mu_{it}^{p} : R_{ikt})$, is defined to be his estimate of the likelihood that Party i will adopt particular programs if it wins Election t – an estimate based solely upon his knowledge of i's record. This is a probability function defined over S, the n-

dimensional issue space. Similarly, Voter K's likelihood function for Party i's campaign platform, denoted by $\pi(\bar{C}_{ikt} : \mu^p_{it}, R_{ikt})$, is a probability function defined over S. This likelihood function is defined to be Voter K's estimate of the likelihood of observing a particular $\bar{C}_{ikt}$ if Party i actually enacts program μ^p_{it} after winning Election t. This function answers the question: How likely is Party i to advocate $\bar{C}_{ikt}$ during the campaign if its candidate will enact μ^p_{it} after it wins the election? Finally, Voter K's evaluations of the likelhood that Party i will enact particular programs after winning the upcoming election based upon his knowledge of its record R_{ikt} *and* his estimate of its campaign platform $\bar{C}_{ikt}$ is defined to be his posterior distribution of Party i's future program, denoted by $\pi''(\mu^p_{it} : R_{ikt}, \bar{C}_{ikt})$.

Using this notation, Bayes's theorem clearly reveals the dependence of a citizen's estimate of a party's future program upon his knowledge of its record and its campaign platform. Formally, this relationship is given by

$$\pi''(\mu^p_{it} : R_{ikt}, \bar{C}_{ikt}) = \frac{\pi'(\mu^p_{it} : R_{ikt}) \cdot \pi(\bar{C}_{ikt} : \mu^p_{it}, R_{ikt})}{\int_\Omega \pi'(t : R_{ikt}) \cdot \pi(\bar{C}_{ikt} : t, R_{ikt})} \, d\mu(t)$$

Alternately, a voter's posterior distribution is proportional to the product of his prior distribution and his likelihood function. The conceptualization of a citizen's beliefs about a party's future program as a random vector is sufficient to provide a rationale for the modification of his initial expectations in accordance with the acquisition of additional information during the course of the campaign. The voter's posterior distribution is a complete description of his beliefs about the likely nature of a party's future program at the time of the election.

This process for modifying an individual's beliefs is dynamic. Each time the voter acquires additional relevant information, he modifies his beliefs about a party's future program by the use of Bayes' theorem. His posterior distribution at the end of one election becomes his prior distribution for the next election. Further, if a voter is more confident that a party's record indicates its future program than its campaign platform, then his final estimate of this party's policies will be 'closer' to his estimate based on its record than to his estimate based on its campaign platform. Alternately, if he places more faith in a party's promises than in its performance, then his best guess as to its future program will be closer to the party's campaign platform.

A less plausible consequence of this process is that within the framework of a Bayesian analysis, as a person gains information about an object, his beliefs about this object necessarily become more precise. This occurs because of an implicit assumption that the object remains constant over time (Barnet, 1973, p. 181). This evolution in a person's perceptions of a party's future actions is plausible as long as its program actually remains unchanged. Under these conditions, his ability to pinpoint a party's future program should increase as he obtains additional information. Over time, however, a party's program often does change. Perhaps variations in the preferences of its leaders result in the alteration of this program. Perhaps the party's electoral fortunes necessitate such a change. Whatever the reason, it is often true that the specific content of a party's policies evolves over time. If a party's program does fluctuate, then it is unreasonable to conclude that by acquiring more information, a voter must reduce his uncertainty that a party's record indicates its future policies. Nevertheless, inferences about a party's future program based on repeated use of Bayes' theorem necessitates this conclusion.

This unreasonable conclusion may be avoided by explicitly assuming that a voter realizes a party's future program may fluctuate. The precision of a voter's prior distribution provides a plausible way for incorporating this assumption into the paradigm of voting behavior presented here. This precision is defined as the voter's confidence that his estimate of a party's future program based on his knowledge of its record actually reveals the program. Consequently, if he believes that a party's program has evolved between elections, then it is reasonable to expect that his confidence that this party's record reveals its future behavior will decline. The magnitude of this adjustment depends upon his evaluation of the evidence that the future program this party is likely to enact has changed. This evidence consists of the difference between his estimates of a party's future program based on its record and its campaign platform. Small discrepancies are likely to be attributed to the vagaries of the campaign process. Substantial differences, presumably, are indicative of the possibility that its policies have changed. Further, the significance of the difference between these two estimates depends upon his confidence in these estimates. If a voter is extremely confident that a party will keep its campaign promises, then even a small difference between μ' and $\bar{c}$ probably indicates that a party's campaign program has changed. However, if he is very uncertain that a party's campaign pledges can be trusted, then even a relatively large difference between μ' and $\bar{c}$ probably does not indicate that this party's program has evolved since the last election.

Theoretically, numerous functions might be used to represent this process of adjusting $\pi''(\mu_{it}^p : R_{ikt}, \bar{C}_{ikt})$ based on an assessment of the evidence that μ_{it}^p and $\mu_{i(t-1)}^p$ are different. If a voter uses his adjusted prior distribution of a party's future program rather than his unadjusted prior distribution in estimating its future policies, then it is possible that by learning about a party's campaign platform he can actually become less certain about the future.

Perhaps the most important consequence of incorporating the assumption that voter's adjust their confidence that a party's record indicates its future program into the logical structure of this model is that its comprehensiveness is significantly enhanced. Within the context of this model, the ratio of π' to π determines a party's spatial mobility. If this ratio is approximately zero, then a party has perfect spatial mobility. Thus, the simple issue voting model which is typically associated with spatial models, can be treated as a special case of this model. If this ratio is very large, then a party has no spatial mobility. Consequently, retrospective voting models can also be treated as a special case of this model. Whenever the ratio of π' to π is extremely large, a voter's estimate of a party's future program will be based almost exclusively on his knowledge of its record.[6]

The voter's beliefs about the likely course a party will pursue after the election are completely described by his posterior distribution of this party's future program. Having established this connection between a party and the set of possible future programs, the voter must select his preferred party. This stage of the voter's decision calculus has been extensively covered elsewhere (see, for example, Riker and Ordeshook, 1973, chapter 2 or Kassouf, 1970, chapters 3-6). Briefly, the voter's determination of his most preferred party entails the specification of his utility function – a function that indicates the voter's assessment of the relative desirability of future programs. The voter then combines the information concerning his values (conveyed by his utility function) with his beliefs about a party's future program (represented by his posterior distribution for this party). This leads to an evaluation of the desirability of this party actually winning the election. A rational voter's decision-rule selects the party with the highest evaluation.

A simple Bayesian model of voting behavior

The paradigm of individual voting behavior presented in the last section is compatible with numerous models of voting behavior. To construct a particular model, it is necessary to make precise assumptions about the form of a voter's prior distribution and likelihood function for each party. It is also necessary to establish assumptions about the form of his utility function. Finally, a decision-rule must be specified. These assumptions permit the derivation of a voter's posterior distribution for each party. They also imply which party the voter should support.

To illustrate certain characteristics of individual behavior within this paradigm, it is helpful to consider one particular model in detail. This model is characterized by the assumptions that a voter's prior beliefs about a party's future program and his likelihood function for his estimate of a party's campaign platform are normally distributed. Also, the voter is assumed to know the degree of uncertainty associated with these density functions. While these are not necessarily the most intuitively plausible assumptions, given the absence of empirical evidence about the form of these distributions, this model can be used to demonstrate the general characteristics of models constructed within this paradigm of voting behavior. Under alternative assumptions, the risk of becoming needlessly inundated with mathematical detail is greater. For similar reasons, one additional assumption will be made in developing this model: the prior, likelihood and posterior distributions of a voter's judgments will be defined over a one-dimensional issue space. Mathematically, this assumption is unnecessary, but its inclusion simplifies the mathematics associated with the presentation of the model without distorting the basic features of models formulated within the paradigm.

Formally, the assumption that the citizen's estimate of a party's future program is indicated by its record and is normally distributed is given by:

Assumption 1: The probability density function (*pdf*) of Voter K's prior distribution for Party i's future program is

$$\pi'(\mu_{it}^{p} : R_{ikt}) = (\tau'_{ikt}/2)^{1/2} \exp[-\tfrac{1}{2}\tau'_{ikt}(\mu_{it}^{p} - \mu'_{ikt})^2] \quad (1)$$

where μ'_{ikt} represents the mean of his prior distribution of Party i's future program, and τ'_{ikt} represents the precision, the inverse of the variance, of this distribution.

The mean of this distribution can be interpreted as the voter's best guess as to this party's future program based solely upon his knowledge of its record. The precision of this distribution may be viewed as a measure of the voter's confidence that μ'_{ikt} actually will be Party i's future policy. As τ'_{ikt} increases, his confidence builds.

Lacking evidence about the form of the voter's prior distribution, this assumption can be viewed as a reasonable starting-point for developing a model of voting behavior.[7] The family of normal distributions permits the expression of a wide range of beliefs. First, the individual can center the distribution anywhere in the issue space by selecting a particular mean program. Moreover, by the appropriate selection of τ'_{ikt}, he can express virtually any level of confidence that the party's record actually indicates its future program.

Even with this flexibility, Assumption 1 places some restrictions on a voter's prior distribution. It requires that an individual's assessment about Party i's future policy be symmetric about μ'_{ikt}. It also implies that as the distance between a particular future policy μ_{it}^{p} and the mean of the voter's prior distribution μ'_{ikt} decreases, the citizen believes that the probability μ_{it}^{p} will be Party i's program increases. Furthermore, this assumption indicates that a voter feels that Party i's future program will almost certainly be within three standard deviations of μ'_{ikt} and will probably be within two standard deviations of μ'_{ikt}.

In this model of voting behavior, an individual's likelihood function of a party's campaign platform is also assumed to be normally distributed:

Assumption 2: The probability density function of Voter K's likelihood function for Party i is

$$\pi(\bar{C}_{ikt} : \mu^p_{it}, R_{ikt}) = (\tau_{ikt}/2)^{1/2} \exp[-\tfrac{1}{2}\tau_{ikt}(\bar{C}_{ikt} - \mu^p_{it})^2] \qquad (2)$$

where μ^p_{it} represents the mean of K's likelihood function for Party i's campaign platform and τ_{ikt} (the precision) represents his confidence that this party's platform indicates its future program.

As stated, this assumption specifies the linkage between a voter's perceptions of a party's campaign platform and its future course of action. He presumes that if his estimate were to be based solely upon the information he acquires during the campaign, then his best guess as to Party i's future program is that it will fulfill its campaign promises. Alternately, the voter's estimate of a party's campaign platform bears the same relationship to a party's future program as the correspondence between a sample mean and a population mean. The precision of the likelihood function provides a measure of his confidence that the party's campaign platform is equivalent to its ideology.

Before deriving a voter's final estimate of a party's future program, it is necessary to modify his prior distribution based on his assessment that its ideology has shifted since the last election. Lacking guidance from empirical evidence about the precise form of this adjustment mechanism, one possible function for modifying $\tau''_{i(t-1)}$ is given by:

Assumption 3: Let α_{ikt} be defined as a measure of Voter K's receptivity to the possibility that Party i's future program has changed since the last election; then Voter K's adjusted prior distribution of Party i's program if it should win Election t, denoted by $\pi^*(\mu^p_{it} : R_{ikt})$ is assumed to be normally distributed with mean μ^*_{ikt} and precision τ^*_{ikt} where

$$\mu^*_{ikt} = \mu''_{ik(t-1)} \qquad (3a)$$

and

$$\tau^*_{ikt} = \frac{\tau''_{ik(t-1)}}{1 + \alpha_{ikt}\tau_{ikt}(\mu^*_{ikt} - \bar{c}_{ikt})^2} \qquad (3b)$$

This assumption clearly states that the possibility that a party's likely course of action has changed prior to the champaign influences a voter's confidence in his estimate of its future program based upon his knowledge of this party's record. Moreover, the difference between τ^*_{ikt} and $\tau''_{ik(t-1)}$ depends upon the voter's assessment of the evidence that Party i's expected future program has changed.

A simple example covering two elections illustrates the dynamic properties of the adjustment process postulated by Assumption 3. Suppose we have three voters, Voter I, Voter II, and Voter III. After the first election, the posterior distribution for Party i of each of these voter's is $N(5, .5)$. Without the inclusion of Assumption 3 into the model this distribution would constitute the prior distribution of each voter in the next election. However, Assumption 3 requires that these posterior distributions must be modified before representing each voter's estimate of Party i's program after the next election based upon his knowledge of its record. Suppose Voter I's likelihood function for the second election is $N(0, .04)$. Moreover, he is unreceptive to the possibility that Party i's program has changed ($\alpha = .01$). According to Assumption 3, his adjusted prior distribution is $N(5, \tau^*_{ikt})$ where

$$\begin{aligned} \tau^*_{ikt} &= \frac{.5}{1 + (.01)\cdot(.04)(5 - 0)^2} \\ &= .495 \end{aligned}$$

His adjusted prior distribution for the second election is almost the same as his posterior distribution from the first election.

Next, suppose Voter II's likelihood function for Party i in the second election is $N(0, 20)$. His perceptions about its campaign platform differ from Voter I's beliefs in that he is substantially more confident that Party i will keep its campaign promises. Additionally, his receptivity to the possibility Party i's ideology has changed is the same as Voter I's level of acceptance ($\alpha = .01$). Thus, using Assumption 3, Voter II's adjusted prior distribution is $N(5, .083)$. His confidence in Party i's campaign promises, combined with the differences between μ'_{ijt} and $\bar{c}_{ikt}$, results in a substantial reduction in his confidence that Party i's record indicates its future program.

Finally, suppose that Voter III's likelihood function is the same as Voter I's. However, he is substantially more receptive to the possibility that Party i's program has varied ($\alpha = 5$). Thus, his adjusted prior distribution is $N(5, .091)$. Like Voter II, his perceptions of Party i's campaign platform result in a substantial reduction in his certainty that Party i's program will be indicated by its record. Unlike Voter II, this reduction is directly attributable to his receptivity towards the possibility that a party's program may vary.

Assumptions 1, 2, and 3 provide a basis for deriving a voter's final estimate of a party's future program, his posterior distribution. Given the information contained in these assumptions, and using Bayes' theorem, it is possible to completely describe a voter's beliefs about a party's future course of action. This description is given in the following theorem. The subscripts which denote the voter, the party, and the election will be suppressed except when they are needed to avoid confusion.

Theorem 1: For $i = (1, 2)$ and $t = (1, \ldots, m)$, let $\pi^*(\mu^p_{it} : R_{it})$ be normally distributed with mean μ^*_{it} and precision τ^*_{it} such that $-\infty < \mu^*_{it} < \infty$ and $\tau^*_{it} > 0$; and let $\pi(\bar{C}_{it} : R_{it}, \mu^p_{it})$ be normally distributed with mean μ^p_{it} and precision τ_{it} such that $-\infty < \mu^p_{it} < \infty$ and $\tau_{it} > 0$, then $\pi^{**}(\mu^p_{it} : R_{it}, \bar{c}_{it})$ is normally distributed with mean μ^{**}_{it} and precision τ^{**}_{it} where

$$\mu^{**}_{it} = \frac{\tau^*_{it}\mu^*_{it} + \tau_{it}\bar{c}_{it}}{\tau^*_{it} + \tau_{it}} \tag{5}$$

and

$$\tau^{**}_{it} = \tau^*_{it} + \tau_{it} \tag{6}$$

Expressed in this fashion, the implications of Assumption 1 and Assumption 2 for the derivation of the voter's posterior distribution are immediately evident. The voter's 'best' estimate of Party i's future program, μ^{**}, is a weighted average of μ^*, his best estimate of this party's future program based on his knowledge of its record and $\bar{c}$, his best guess about Party i's campaign platform: the components of the weights being τ^* and τ. A voter's beliefs about a party's future program prior to the start of the campaign shift in the direction of the information he acquires about this party's platform during the election campaign. The magnitude of this shift is determined by the weights associated with μ^* and $\bar{c}$. As the voter's confidence that μ^* actually will be Party i's future policy increases relative to his confidence that $\bar{c}$ indicates this course of action, the ratio of τ^* to τ increases, and the magnitude of the difference between μ^* and μ^{**} decreases.

The validity of this conclusion is illustrated in the following simple example. Let $\pi^*(\cdot) \sim N(0, 1/4)$ and let $\pi(\cdot) \sim N(10, 1/4)$, then $\pi^{**}(\cdot) \sim N(\mu^{**}, \tau^{**})$ where

$$\mu^{**} = \frac{(1/4)(0) + (1/4)(10)}{1/4 + 1/4} = 5$$

and

$$\tau^{**} = 1/4 + 1/4 = 1/2$$

Next, let $\pi^*(\cdot) \sim N(0, 1/3)$ and let $\pi(\cdot) \sim N(10, 1/6)$. In this situation μ^* and $\bar{c}$ remain unchanged, while the ratio of τ^* to τ has increased from 1 to 2. The distribution of $\pi^{**}(\cdot)$ is $N(\mu^{**}, \tau^{**})$ where

$$\mu^{**} = \frac{(1/3)(0) + (1/6)(10)}{1/3 + 1/6} = 3^1/_3$$

and

$$\tau^{**} = 1/3 + 1/6 = 1/2$$

In this case, the posterior mean shifts only 3⅓ units rather than five units. This conclusion may be interpreted as implying that the constraint placed upon a party's spatial mobility by its previous performance is determined by the ratio of τ^* to τ. As this ratio increases, its ability to shift a voter's initial estimate of its future program decreases. It will obtain perfect spatial mobility only if it convinces each voter that its campaign platform rather than its record represents its future program. This can only be accomplished when the ratio of τ^* to τ is effectively equal to zero.

A voter, having characterized his beliefs about each of the parties' future policies, still must decide which program is most attractive. This process of evaluation can be represented by first specifying his utility function which will be assumed to be single-peaked.

Assumption 4: Let $\mu^u_{kt} \in S$ represent Voter K's most preferred program at Election t, let μ_i^{**} represent the mean of Voter K's posterior distribution of Party i's expected future program, and let $U'(\cdot)$ represent Voter K's ordinal utility function defined over the issue space S, then, for all $\mu_i^{**}, \mu_j^{**} \in S$, $U'(\mu_i^{**}) > U'(\mu_j^{**})$ if and only if $\|\mu_i^{**} - \mu^u\| < \|\mu_j^{**} - \mu^u\|$ and $U'(\mu_i^{**}) = U'(\mu_j^{**})$ if and only if $\|\mu_i^{**} - \mu^u\| = \|\mu_j^{**} - \mu_u\|$.

This assumption states that Voter K prefers the party whose posterior mean is closest to his most preferred program. In other words, a voter is assumed to treat his best guess about each party's future program as if it actually will be the policy it will enact if it wins the election. Since Voter K's utility function is assumed to be a monotonically decreasing function of distance from his ideal program, the implicit assumption of voter rationality requires that the voter prefer the party whose future program is closest to his ideal point. All of these considerations can be summarized by the following decision-rule.

Decision-rule 1: If $\|\mu_i^{**} - \mu^u\| - \|\mu_j^{**} - \mu^u\| > 0$, then Voter K votes for Party j; if $\|\mu_i^{**} - \mu^u\| - \|\mu_j^{**} - \mu^u\| = 0$, then Voter K is indifferent between the two parties.

It is interesting to observe that this decision-rule, like any decision-rule which selects one particular program to represent a party's future policies, completely ignores the voter's confidence in his estimates of each party's future program. If, however, his decision-rule links each party to a set of programs, then his confidence in his best guess as to the likely nature of a party's future program can influence his decision. For example, suppose we replace Assumption 4 with the following assumption.

Assumption 4:* Let μ^u represent voter K's most preferred program at Election t, let ω represent a direct measure of the intensity of Voter K's preferences, let $U^*(\cdot)$ represent Voter K's cardinal utility function defined over S, and let $s \in S$, then

$$U^*(s) = \exp[-1/2\omega\cdot(s-\mu^u)^2] \text{ for all } s \in S .$$

If we firther assume that the voter is an expected utility maximizer, then we can derive his decision-rule:

Decision-rule 2: Given Assumptions 1-4*, Voter K's expected utility for Party i, denoted by $E[U^*(\mu_i^p)]$, is

$$E[U^*(\mu_i^p)] = [\tau_i^{**}/(\tau_i^{**}+\omega)]^{1/2} \exp[-1/2\,(\tau_i^{**}+\omega)\cdot(\mu_i^p-\mu^u)^2]$$

$$\text{for } i = 1, \ldots, n$$

Consequently, Voter K votes for Party j only if

$$E[U^*(\mu_j^p)] = \max\ \{E[U^*(\mu_i^p)]\} \text{ for } i = 1, \ldots, n$$

Basically, this assumption requires a voter to be an expected utility maximizer. It further requires his utility function to be 'bell-shaped'. While this assumption has certain drawbacks (especially its implication that $U^*(\cdot)$ is symmetric about μ^u) several reasons can be advanced for the inclusion of such a complex utility function into the model. First, since this function is bounded, it avoids many of the difficulties inherent in the use of unbounded utility functions. Second, given the assumption that Voter K's posterior distribution is normal, it is mathematically tractible: a citizen's expected utility is readily calculated. Finally, by adopting a bell-shaped utility function (and allowing it to vary over the set of real numbers) a rational voter does not necessarily vote for the party whose posterior mean is closest to his ideal point.

Given decision-rule 2, a voter can be risk averse, risk acceptant or risk neutral (see Shepsle, 1972). The relative distance between the means of the parties' posterior distributions and the individual's ideal point are not, in themselves, useful indicators of a citizen's preferences. More information is required. In particular, it is necessary to also know the precisions of these posterior distributions as well as the location of his ideal point and the variance of his utility function. This information is necessary to explain a voter's political preferences. Decision-rule 1 does not require nearly as much information about a voter's preferences in determining how he should vote. However, if a voter uses decision-rule 1 he will never vote for any party other than the party whose posterior mean is closest to his ideal point.

Conclusion

The process of model building necessarily involves abstraction and simplification. Decisions concerning the inclusion or exclusion of various aspects of the phenomenon being modeled must be based upon the theoretician's goals. If the sole purpose in formulating a model of individual voting behavior is predicting a citizen's vote, then numerous other models could be used. For example, the Six-Component Model of Voting Behavior originally presented in the *American Voter* explains a substantial proportion of the variation in individual decisions (Campbell, *et al.*, 1960, chapter 19). The 'simple' model of voting developed by Kelley and Mirer also provides accurate predictions of voters' choices (1974). These models provides at least as accurate a prediction as the Bayesian model presented in the last section.

Alternately, if the only goal is predictive efficacy, then psychological models of decision making might be preferable to the model formulated in this paper. Psychologists have developed considerably more elaborate theories of preference under uncertainty that provide superior predictions of individual behavior than simple Bayesian models (see Luce and Suppes, 1965). However, as Luce and Suppes note, 'While being elaborated as distinct and testable psychological theories, these theories of preference have begun to acquire a richness and complexity . . . that renders them largely useless as bases for economic, statistical and political theories' (p. 253). To turn a phrase, elegance and precision are lost in pursuit of the poet's pleasures.

The sole objective of a model of voting behavior that is to be incorporated into a spatial model is not predictive efficacy. Within the framework of a spatial model, a voting model is used to evaluate the desirability of alternate campaign platforms. It must indicate the number of votes a party is likely to gain – or lose – by advocating a particular policy during the election campaign. The problem with most models of voting behavior is that they cannot be used for this purpose. While the Six-Component Model does indicate the change in a party's total vote attributable to a change in the electorate's attitudes on issues, it does not provide a precise linkage between a shift in a party's policy and the electorate's attitudes. Consequently, it cannot indicate the likely change in a party's share of the total vote that is caused by its campaign platform. Numerous other models of voting behavior are vulnerable to similar criticism. The major advantages of the Bayesian models developed previously are that they are not only capable of providing a reasonable explanation of voting behavior but they also specify a precise relationship between a party's issue position and its percentage of the total vote.

Obviously, not all non-Bayesian voting models are incompatible with these two goals. Traditional rational-choice voting models do specify a connection between a party's campaign platform and its electoral support.

The primary deficiency of these voting models, however, is the restriction they place upon the information that a voter uses in estimating each party's future program. While this restriction is necessitated by the assumption of perfect spatial mobility, it increases the possibility of erroneously evaluating the support of a party will receive by advocating a particular policy.

An additional disadvantage of using a traditional rational-choice voting model in a spatial model is that it substantially reduces the range of strategic options that a party's leadership must consider during the course of the campaign.

None of these rational-choice models provide a strategist with an incentive to attempt to manipulate the relevance of his party's record in the electorate's decisions. Since, by assumption, a party's spatial mobility is unconstrained, if it wants to represent itself as a champion of consumer interests during the campaign, all it has to do is proclaim its support of these interests. Whether or not it has favored or opposed consumer interests in the past is presumed to be irrelevant to the electorate's deliberations; and, consequently, it is also irrelevant to its strategists' considerations. A second consequence of the assumption of perfect spatial mobility is that a political strategist can regard the distribution of citizens's pregerences as given. He has no incentive to initiate a costly campaign designed to alter this distribution. In other words, a party's leaders have no incentive to adopt a position of leadership. They lack motivation to attempt to convince the electorate that their stance on the issues is correct. If its leaders are only interested in winning the election, why should they initiate a massive (and expensive) propaganda campaign designed to alter public attitudes on an issue? All they have to do is bring their campaign platform into conformity with the current distribution of preferences in society.

If a Bayesian rather than a traditional rational-choice model of voting behavior is used to assess the consequences of advocating a particular campaign strategy, then a party's strategists have incentives to consider efforts designed to manipulate the relevance of their party's record in the electorate's deliberations. They also have incentives to attempt to manipulate the distribution of the electorate's preferences. The primary reason for these differences is that in a Bayesian spatial model the theorems concerning the optimal location of a party's campaign platform must be re-interpreted. In this situation, these theorems do not necessarily indicate the optimal location for this platform. Rather, they indicate the optimal location for the electorate's estimates of a party's future program. In attempting to attain this goal, a party's strategists must not only consider the content of their campaign platform but also the constraints the the electorate's perceptions of their party's record place upon their spatial mobility. If these constraints are strong enough to prevent them from accomplishing their goal, then they must attempt to manipulate these constraints in order to increase their party's chances of electoral success.

Notes

1.
Formally, symmetry may be defined in the following manner: let S_i represent candidate i's set of strategies (i = 1, 2), and let M represent the payoff function defined over $S_1 \times S_2$; then the two-person, zero-sum game $G(S_1, S_2, M)$ is symmetric if and only if $S_1 = S_2$ and for all strategy pairs $s = (s_1, s_2)$ contained in $S_1 \times S_2$,

$$M(s_1, s_2) = -M(s_2, s_1)$$

2.
Strictly speaking, McKelvey's approach in 'Some Extension . . .' (1972) does not require this assumption in its standard form. By basing his development on support functions rather than individual decision rules, he does not have to assume that the electorate has no predispositions. His results indicate optimal spatial locations. They say nothing about the relationship between a candidate's campaign platform and his location in the issue space.

3.
Some spatial theorists have confronted empirical objections to the assumption of perfect spatial mobility by implying that voter's electoral predispositions can be incorporated into the existing structure of spatial models without altering the assumption that candidates are free to manipulate their spatial location. For example, Aranson, Hinich and Ordeshook (1974, p. 136) implicitly argue that factors like age, sex and party identifications can be incorporated into a voter's decision space. A candidate can manipulate his spatial location with respect to the dimensions over which he has little or no immediate control by varying the saliency of these dimensions in the voter's calculus.

The difficulty with this approach is that if criteria over which a candidate has little or no immediate control are incorporated into a multidimensional spatial model, the assumption that candidates are free to determine the saliency of *each* dimension is incompatible with the other axioms of traditional spatial models. Specifically, this assumption ignores the zero-sum nature of electoral competition: if one of the candidates improves his position by decreasing the saliency of a dimension over which he has no immediate control, the other candidate's position is necessarily worsened. Hence the second candidate will want to increase the saliency of this dimension. Consequently, both candidates cannot possibly be free to determine the composition of those elements of the A-matrix that represent factors over which they have no immediate control. Thus, their spatial mobility is necessarily constrained! In short, in a traditional spatial model, the salience of those dimensions over which the candidate's have no immediate control must be taken as given. It cannot be assumed to be subject to manipulation by the candidates.

4.
Within the context of this paradigm, only the voter can classify a particular 'bit' of information as being politically relevant. Additionally, he is the only judge of the weight this information is given in his deliberations. More precisely, 'political information' and 'weighting' are given the status of primitive concepts in the formal system being developed. Potentially significant future modifications of this system would include formulating models of these sub-processes. For example, a candidate's personality might be directly incorporated into this paradigm by assuming that it is related to the voter's uncertainty about the candidate's future policy. As the voter's assessment of the candidate's strength and trustworthiness increases, the voter's confidence that the candidate will keep his campaign promises also increases.

5.
If a voter's estimate of a party's future program is to be defined as a random vector, then it is necessary to interpret the elements of the probability space over which this random vector is defined. Like traditional spatial models, this paradigm of voting behavior presumes that an individual's vote is determined by his comparative evaluation of the competing parties' programs. Within this context, the event space, S, is interpretable as an n-dimensional issue space. Consequently the elements of the Borel-field A, associated with S, represent different mixtures of these programs; and the *jpdf* is a rule that indicates the voter's assessment of the likelihood that different mixtures of programs will occur. The definition of a *jpdf* over the issue space requires that a voter's judgements about the likelihood a party will adopt particular programs are coherent and consistent. More specifically, coherence means that if a person believes the probability is p that a party will enact a particular program, E, then he also believes the probability that this party's program will not be E is $1-p$. Consistency implies that if he believes a party's program is more likely to be E_1 than E_2, while he also feels that E_2 is more probable than E_3, then he must believe that E_1 is more likely to be this party's program than E_3 (see Kyburg, 1970, pp. 68-71).

6.
See Gerald Kramer (1971) and Edward Tufte (1975) for examples of retrospective voting models. Morris Fiorina (1977) has also proposed a perspective on voting decisions that treats simple issue voting and retrospective voting as special cases. It might be noted that his paradigm and the paradigm formulated in this chapter are compatible. In particular, within Fionrina's conceptual framework, the perspective advanced in this paper can be viewed as an attempt to specify the logical structure of individual voting choice.

7.
Barnett (1973) argues: '[w]hen prior information consists of sparse factual measures augmented by subjective impressions, . . . a variety of different specific prior distributions may have the appropriate summary characteristics to encompass the limited information that is available. In such cases it should be quite straightforward to choose an appropriate prior distribution from the family of conjugate prior distributions' (p. 188).

References

Aranson, Peter H., Hinich, Melvin J., and Ordeshook, Peter C. 'Election Goals and Strategies: Equivalent and Nonequivalent Candidate Objectives,' *American Political Science Review* (March 1974), pp. 135-152.

Barnett, Vic. *Comparative Statistical Inference*. New York: John Wiley & Sons, Inc., 1973.

Campbell, A., Converse, P.E., Miller, W.E., and Stokes, D.E. *The American Voter*. New York: John Wiley & Sons, Inc., 1960.

Downs, Anthony. *An Economic Theory of Democracy*. New York: Harper & Row, 1956.

Fiorina, Morris P. 'An Outline for a Model of Party Choice,' *American Journal of Political Science* (August 1977), pp. 601-625.

Kelley, Stanley, Jr., and Mirer, Thad. 'The Simple Act of Voting,' *American Political Science Review* (June 1974), pp. 572-591.

Kramer, Gerald H. 'Short-term Fluctuations in U.S. Voting Behavior, 1896-1964,' *American Political Science Review* (March 1971), pp. 131-143.

Kyburg, Henry E. *Probability and Inductive Logic*. New York: The Macmillan Co., 1970.

Luce, R.D., and Suppes, R. 'Preferences, Utility, and Subjective Probability,' in *Handbook of Mathematical Psychology*, Vol. 3, edited by Luce, R.D., Bush, R.R., and Golanter, E. New York: John Wiley & Sons, Inc., 1965.
McKelvey, Richard D. 'Some Extensions and Modifications of the Spatial Models of Party Competition,' unpublished dissertation, University of Rochester, 1972.
Riker, William, and Ordeshook, Peter. *An Introduction to Positive Political Theory*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1973.
Shepsle, Kenneth A. 'The Strategy of Ambiguity: Uncertainty and Electoral Competition,' *American Political Science Review* (June 1972), pp. 555-568.
Tufte, Edward R. 'Determinants of the Outcomes of Midterm Congressional Elections,' *American Political Science Review* (September 1975), pp. 812-826.