CHAPTER 2

How Voters Decide

The Components of the Unified Theory of Voting

2.1. Introduction

As discussed in Chapter 1, behavioral researchers and spatial modelers have quite different perspectives on how voters decide. Spatial modelers typically assume that voters are entirely (or at least chiefly) motivated by the policies that the competing parties or candidates present in the current campaign, and, in their empirical applications, policy factors are the chief (if not the only) influences that are incorporated into their models. Behavioral researchers, by contrast, emphasize that voter choices are affected by a variety of considerations besides the parties' current policies, including voters' party identifications and sociodemographic characteristics, their perceptions of economic conditions, and retrospective evaluations of incumbent performance.

Furthermore, while spatial modelers typically posit that voters employ a *proximity metric* to evaluate parties' policies – that is, that voters prefer parties whose positions are close to their own positions along salient policy dimensions – some research, both theoretical and empirical (see Grofman 1985; Fiorina 1994, 1996; Lacy and Paolino 1998, 2001; Merrill and Grofman 1999; Kedar 2002; Adams, Bishin, and Dow 2004; Lewis-Beck and Nadeau 2004), suggests that this behavior is modified by voters' realization that parties/candidates will probably not be able to implement the full extent of the policies that they advocate. Hence, a proximity model with *discounting* may be appropriate. As we argue beginning in Chapter 3, both of these perspectives contribute to our understanding of party policy strategies and the nature of mass–elite policy linkages.

In this chapter, we compare the spatial-voting model to the behaviorist's perspective on voting, developing the various components of the unified multivariate voting specification that we use in subsequent chapters to analyze elections in the four countries we study.

2.2. The Policy-Only Model

2.2.1. Deterministic Policy Voting

Although spatial modelers and behavioral researchers have profoundly different perspectives about how voters decide, the two camps agree that policies matter, at least to some voters. In order to introduce the way in which the policies of candidates and parties enter the voter's decision calculus in both behavioral and spatial models, we begin with a simple model in which policy debates revolve around a single overarching issue dimension. For simplicity, we speak about candidates, not parties, although the model applies equally well to party-centered elections.

We assume that voters and candidates are located at ideal points on some continuum, representing their preferred positions along a single issue or ideological dimension. While the measured range of variation along this continuum may, in principle, be either finite or infinite, we will normally employ a finite interval from 1 to 7 (or a similar finite interval) in order to conform to the options presented to respondents in many national election studies. The 1–7 Left–Right scales that we use to place the voters, parties, and candidates in Figures 2.1 and 2.2 later in this chapter are typical of such intervals. Normally, the value 1 on such a scale represents a strongly liberal or leftist position on the issue, whereas the value 7 represents a strongly conservative or rightist position, with intermediate values representing intermediate positions.

Given the voter and candidate positions, the voter's utility (evaluation) of the candidate may employ this information in various ways. Spatial modelers typically posit that voters evaluate candidates according to their proximity along the policy interval, so that left-wing voters prefer left-wing candidates, centrist voters prefer centrist candidates, and so on. The most common way to implement this definition is via *quadratic proximity utility*, specified as the negative of the squared distance between the voter i at location x_i on the scale and a candidate k at location s_k:

$$V_i(k) = -(x_i - s_k)^2, \tag{2.1}$$

where $V_i(k)$ represents voter i's utility for candidate k's position.[1] The negative sign is used so that utility declines with distance. This proximity model has been the cornerstone of most empirical analyses of voting behavior (see Markus and Converse 1979; Page and Jones 1979; Alvarez 1997; Dow 2001).

Incorporating multiple policy dimensions. So far, we have restricted the development to a single issue. More realistically, however, voters arrive at preferences on the basis of more than one issue. Suppose that there are J issues, that the voter's position on the jth issue is x_{ij}, and that the candidate's position on the jth issue is s_{kj}. Multidimensional utility functions are obtained by summing the corresponding one-dimensional utilities over the various issue dimensions. Thus the voter's utility for the candidate's policies becomes

$$V_i(k) = -\sum_j a_j (x_{ij} - s_{kj})^2, \tag{2.2}$$

where a_j denotes the policy-salience parameter for the jth issue.

2.2.2. *Unmeasured Voter Motivations: Probabilistic Voting Models*

Many spatial modelers employ the specifications given in equations 2.1–2.2; that is, they assume that voters evaluate the candidates entirely on the basis of their positions along important policy dimensions. This implies that voters

[1] An alternative to quadratic utility is *linear utility*, in which the utility of a voter i for candidate k is given by $V_i(k) = |x_i - s_k|$. It is difficult to distinguish empirically whether quadratic or linear utility (or some other function of distance) better represents voters' relative evaluations of candidates. There is evidence that linear utility gives a better fit to thermometer scores interpreted as utilities, such as those solicited from respondents in the American National Election Studies and other national election studies (Westholm 1997: 876; Merrill and Grofman 1999: 173–5; Berinski and Lewis 2001). Inferring that the utility scale itself is linear from such evidence is, however, problematic, because both the policy scales from which distance is measured and the thermometer scores are constrained to specified finite intervals (typically 1–7 or 1–10 for the policy scales and 0–100 for the thermometer scales), whereas the utility scale need not be so constrained. On the other hand, quadratic utility is more convenient mathematically for most applications and has frequently been the preferred choice (Erikson and Romero 1990; Alvarez and Nagler 1995). We will use quadratic utility in our analyses unless otherwise stated. In our empirical applications, the policy optima that we determine for quadratic and linear losses are fairly similar (see, for instance, Table 1 in Merrill and Adams 2001), so that our substantive conclusions do not depend on the utility function we use.

invariably support the candidate who best reflects their policy beliefs – that is, that voters employ a *deterministic policy-only model*. While this assumption simplifies theorizing about candidate strategies, empirical voting studies find that unobservable factors – which may stem from voters' evaluations of the candidates' personal qualities, retrospective evaluations of incumbent performance, or from policy motivations that are not accurately measured in voter surveys – can motivate voters to support a particular candidate even when some rival candidate better reflects these voters' measured policy beliefs (Campbell et al. 1960; Alvarez 1997; Schofield et al. 1998). These factors – unobserved both by researchers and, in general, by the candidates themselves – are best modeled as random variables. Accordingly, we extend the utility formula given in equation 2.2 by incorporating a voter-specific random utility term X_{ik}, which represents unmeasured components of the voter i's utility for candidate k. Thus the formula for a random utility model becomes

$$U_i(k) = V_i(k) + X_{ik} = -\sum_j a_j(x_{ij} - s_{kj})^2 + X_{ik}. \tag{2.3}$$

Note that we have retained the notation $V_i(k)$ to refer to the deterministic component of the random utility model – here, the voter's evaluation of the candidate's policies. The utility of the full random utility model is denoted by $U_i(k)$. The random term X_{ik} in equation 2.3 renders the voter's choice indeterminate from measured policy components alone, so that this choice is probabilistic from the candidates' (and the analyst's) perspectives.

We label the specification given in equation 2.3, in which voters employ the proximity metric and policies are the only measured influence on the vote, the (probabilistic) *policy-only model*. Because probabilistic models will be our default, for simplicity we omit the term "probabilistic" in specifying this and similar models. According to this model, the voter supports the candidate who maximizes his combination of measured policy-related utilities and his unmeasured utilities. In this model, the vote probability depends on the distribution of X_{ik}. One plausible assumption, which we employ in subsequent chapters, is that the values of X_{ik} are generated independently from a type I extreme-value distribution – the assumption that characterizes the *conditional logit* (CL) model, which has been used extensively by behavioral researchers in empirical voting studies (Endersby and Galatas 1997; Adams and Merrill 1999a,b) and by spatial modelers (Coughlin 1992;

Adams 1999a,b).[2] Given a random-utility model for a K-candidate election, a CL model specifies that the probability that voter i chooses candidate k is proportional to the exponential of the deterministic component of utility and that the sum of these probabilities over candidates is one. Thus, the probability that a voter i votes for candidate k is given by

$$P_{ik} = \frac{\exp[V_i(k)]}{\sum_{j=1}^{K} \exp[V_i(j)]}, \tag{2.4}$$

where $V_i\,(k)$ is defined by equation 2.2. It follows that in an election in which n voters cast votes, the expected vote share for candidate k is given by

$$EV(k) = \frac{1}{n}\sum_{i=1}^{n} P_{ik}. \tag{2.5}$$

It is the straightforward nature of formula 2.4 for the choice probabilities – a formula that follows logically from the definition of utility in the model specification (see Train 1986) – that makes the conditional logit model so convenient to use.

In Appendix 2.1, we discuss alternatives to the conditional logit model – in particular, the generalized extreme value (GEV) model and the multinomial probit (MNP) model – and argue that the greater specificity obtained in these more complicated models usually does not outweigh the disadvantages inherent in their complexity.

2.3. Nonpolicy Factors: The Unified Model

2.3.1. A Unified Specification That Incorporates Measured Nonpolicy Variables

Probabilistic policy-voting models have been employed extensively by spatial modelers in their theoretical work on both two-candidate elections (Enelow and Hinich 1984; Coughlin 1992; Hinich and Munger 1994) and multicandidate elections (Lin, Enelow, and Dorussen 1999; Adams 1999a,b;

[2] The cumulative distribution function for a type I extreme value distribution is $F(x) = \exp[-\exp(-x)]$. The distribution has a variance of $\pi^2/6$ and a mean equal to Euler's constant. We note that this distribution is similar to the normal in shape but has the advantage that the choice probabilities have a simple, tractable form. See Alvarez and Nagler (1998a) for a comparison of conditional logit and related models.

https://doi.org/10.1017/CBO9780511614433.003 Published online by Cambridge University Press

Schofield 2002). Yet empirical voting studies suggest that without further modifications, these models are only marginally more realistic than the deterministic policy-voting model of equations 2.1–2.2. The reason is that voters' candidate preferences are also influenced by a number of non-policy-related factors that are measured in voter surveys and that are arguably observable by the candidates during the election campaign. Some of these factors, such as partisanship, sociodemographic traits (race, gender, class, income, education, and the like), and retrospective evaluations of party performance, are voter-specific. Others, such as voters' impressions of the candidates along evaluative dimensions of character such as integrity and leadership ability, vary not only across voters but also across candidates. These nonpolicy attributes, commonly available in national election studies, are often referred to as *measured* nonpolicy variables, as opposed to random variables (which may be either non-policy-based or policy-based).

Distinguishing between the effects of policies and measured nonpolicy factors is by no means easy, and there are ongoing scholarly controversies about the relative influence of policy-related and nonpolicy factors on the vote. In particular, the electoral impact of party identification is a subject of debate (see Appendix 2.2 for a summary of this issue). For our purposes, however, it is important only that both policy and nonpolicy factors do influence voter choice. We are concerned in this book with the joint effects of both types of factors and less concerned with their relative degree of contribution.

Central to our analysis is the claim that both policy and nonpolicy motivations influence voter choice. We illustrate the effects of liberal–conservative ideology and partisanship in the 1988 American National Election Study (ANES).[3] Figure 2.1 plots the proportion of validated voters in the 1988 ANES who voted for the Democratic candidate (Michael Dukakis) versus voters' liberal–conservative self-placement. The separate plots for Democratic and Republican identifiers show that self-identified Democrats were significantly more likely to vote for the Democratic candidate than self-identified Republicans, regardless of the voter's ideological location.[4]

[3] The 1988 election is the primary American election that we study in this book because it is the most recent presidential election for which validated voting is available in the ANES study. The latter is important because of the central use we make of voter turnout decisions in our model of U.S. party competition. However, in Chapter 8 we present additional empirical applications to the 1980, 1984, 1996, and 2000 U.S. presidential elections.

[4] This plausibly occurs in part because party identity is used as a cue or proxy for party policy; alternatively, partisan voters may be persuaded to adjust their policy preferences in response to the appeals of party elites (see Appendix 2.2). But our objective in this book is to show

https://doi.org/10.1017/CBO9780511614453.003 Published online by Cambridge University Press

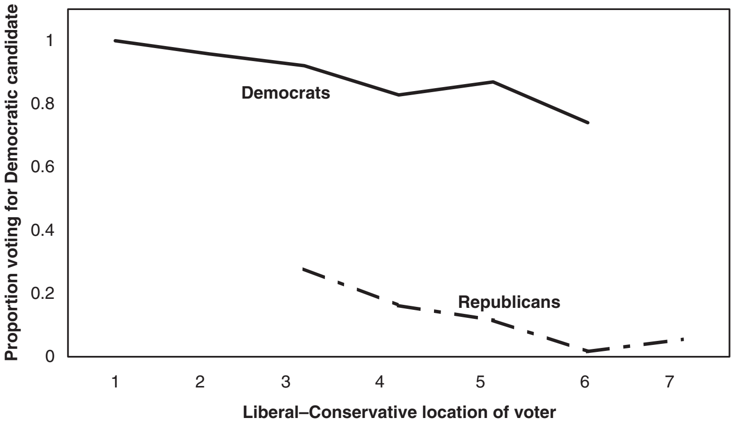

Figure 2.1. Proportion of validated voters voting for the Democratic candidate, stratified by party identification and liberal–conservative locations: United States, 1988. ($N = 701$)
Note: Strong and weak Democrats and others leaning toward the Democrats were classified as Democrats (0, 1, and 2 on the ANES party scale); strong and weak Republicans and others leaning toward the Republicans were classified as Republicans (4, 5, and 6 on the ANES party scale).

Moreover, the downward slope of each of the plots shows that self-identified conservative voters were less likely to vote Democratic, regardless of party identification, than liberal voters were. These relationships suggest that policy (ideology) and partisanship together have greater influences on voter choice than either factor alone.

Second, our analysis depends on the fact that nonpolicy factors, such as party identification, are correlated with voters' policy positions. This is illustrated for the 1988 ANES in Figure 2.2, which shows that although the two sets of partisans are spread over wide, overlapping ranges, self-identified Democrats are, not surprisingly, significantly more leftist than self-identified Republicans. This separation of partisan distributions is typical, and is even more distinct in European electorates than in the United States (see Figures 3.3 and 6.1A in later chapters).

We note that the contribution of each nonpolicy factor to voter utility – whether this contribution is large or small – may be modeled by introducing

that policy and nonpolicy factors together have certain effects on party positioning and that neither should be omitted from a fully specified model; it is less important, for our purposes, to determine the exact mix of these effects or the reciprocal influences between them.

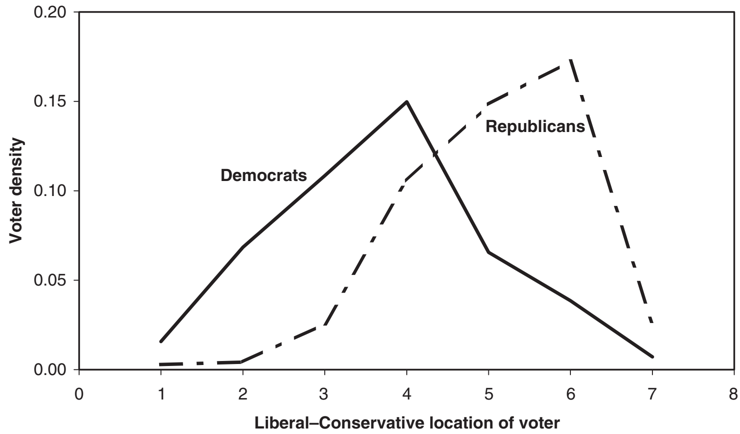

Figure 2.2. Distribution of self-identified Democrats and Republicans by liberal–conservative location: United States, 1988. (N = 701; validated voters only)
Note: Strong and weak Democrats and others leaning toward the Democrats were classified as Democrats (0, 1, and 2 on the ANES party scale); strong and weak Republicans and others leaning toward the Republicans were classified as Republicans (4, 5, and 6 on the ANES party scale).

a variable t_{ik}, which specifies a characteristic of voter i in relation to candidate k. For example, for the partisanship variable, t_{ik} might equal 1 if voter i identifies with candidate k's party and zero otherwise. This variable is then multiplied by a salience parameter b_k, which is estimated from the data and which represents the importance or contribution of the corresponding variable in determining voter utility. In order to model the influence of several nonpolicy factors – partisanship, sociodemographic characteristics, retrospective evaluations of incumbent performance, and so on – we introduce a vector $\mathbf{t}_{ik}$ of nonpolicy variables with coordinates t_{ikl} for the kth candidate and the lth nonpolicy variable, and a vector $\mathbf{b}_k$ of parameters with coordinates b_{kl}.[5] Combining policy factors and nonpolicy factors (both measured and unmeasured) in the same model, we obtain

$$U_i(k) = V_i(k) + X_{ik} = -\sum_j a_j(x_{ij} - s_{kj})^2 + \mathbf{b}_k\mathbf{t}_{ik} + X_{ik}, \quad (2.6)$$

[5] In general, the nonpolicy contribution to voter utility is given by

$$\mathbf{b}_k\mathbf{t}_{ik} = \sum_l b_{kl}t_{ikl}.$$

where a_j is the policy-salience parameter for the jth issue to be estimated from the data, and, as before, $V_i(k)$ denotes the deterministic component of utility.

Note that in equation 2.6, $V_i(k)$ now represents *all* measured elements – both policy-related and non-policy-related – of the voter's utility. We label equation 2.6, which represents the behavioral researcher's typical specification for analyzing voting, the *unified model* (see Markus and Converse 1979; Page and Jones 1979; Erikson and Romero 1990; Alvarez and Nagler 1995; Dow 1999). However, we caution the reader that this singular designation does not imply that there is a single, universally accepted explanation for how voters decide! Rather, behavioral researchers have advanced many alternative theories of voting, which can be captured using differing specifications of equation 2.6. That is, one can capture alternative behavioral voting models by varying the set of explanatory variables included on the right-hand side of equation 2.6 as well as the variable parameters. We elaborate on this point in the following discussion.

2.4. The Unified Discounting Model

Heretofore, we have assumed that in a Downsian or proximity model, voters take their notion of candidate/party location at face value – that is, they evaluate the parties as if they will implement the policies they are perceived to advocate. Downs himself (1957: 39) notes, however, that "[the voter] cannot merely compare platforms; instead he must estimate in his own mind what the parties would actually do were they in power."

There are several reasons why we may expect that parties or candidates may be unable to implement the full extent of the policies that they advocate. The effect will be most pronounced in nations in which power is shared between an executive and a legislature, which are frequently controlled by different parties. Notable examples are the United States, in which either or both houses of Congress may be controlled by other than the president's party, and France, in which a similar cohabitation between different parties may occur. Empirical work by Lacy and Paolino (1998, 2001) documents that U.S. voters in both presidential and gubernatorial elections distinguish between the policies that candidates propose and the policy outcomes that voters expect from each candidate in government, and furthermore, that voters' proximities to the expected position of the government under each candidate influences the vote more strongly than proximity to the candidates'

https://doi.org/10.1017/CBO9780511614453.003 Published online by Cambridge University Press

policy proposals.[6] In a multiparty parliamentary system, as in Norway, the governing party frequently holds a minority of seats or is part of a formal coalition. In either case, compromise is required to implement policies. Empirical work by Merrill and Grofman (1999) and Kedar (2002) supports the hypothesis that voters in multiparty parliamentary democracies account for these expected policy compromises when casting their votes.[7] Even in a country such as Britain, where normally a single party holds a majority of seats in Parliament, the government does not have totally free rein because of the constraints of shifting public opinion (see Kedar 2002).

Grofman (1985) suggests the following implementation of the discounting idea: Voters perceive a status quo point in the spatial model and judge that the candidates'/parties' capacity to move that status quo to the location that they appear to advocate must be discounted. For example, suppose that, on a 1–7 Left–Right scale, the status quo is at 5; party *L* is perceived to advocate a position at 1, while party *R* appears to advocate a position at 6 (see Figure 2.3A). If voters discount both parties' ability to move the status quo to their respective advocated locations by, say, 50 percent, then they would expect party *L* to implement policy at 3.0 and *R* to implement policy at 5.5. Note that under deterministic voting, such discounting would change the vote choice of a voter located at 4 on the scale. Because 4 is closer to 6 than to 1, such a voter who does not discount would choose party *R*. However, the discounted location of *L*, namely 3.0, is closer to this voter than the discounted location of *R* (5.5), so a discounting voter would choose party *L*.

Hence we might expect that optimal party strategies will be affected if discounting takes place. Because, relative to the advocated policies, discounting will tend to draw the expected policies to be implemented toward the center, we would expect that voters would vote for parties that advocate relatively extreme policies in an attempt to obtain implemented policies that they favor. Accordingly, we expect that parties – anticipating such discounting – may

[6] Fiorina's (1994, 1996) work on divided government, which posits that citizens cast their votes based on policy expectations rather than candidate positions, is a well-known version of this thesis (see also Alesina and Rosenthal 1995; Adams, Bishin, and Dow 2004; Lewis-Beck and Nadeau 2004).

[7] McCuen and Morton (2002) report experimental support for this proposition. In addition, we note that Iversen's (1994a,b) Representational Policy Leadership Model – in which voters prefer parties presenting policies that are similar to but more extreme than the voters' own beliefs – is consistent with the logic that voters project likely government policy outputs when casting their ballots (see also Adams and Merrill 1999a,b, 2000).

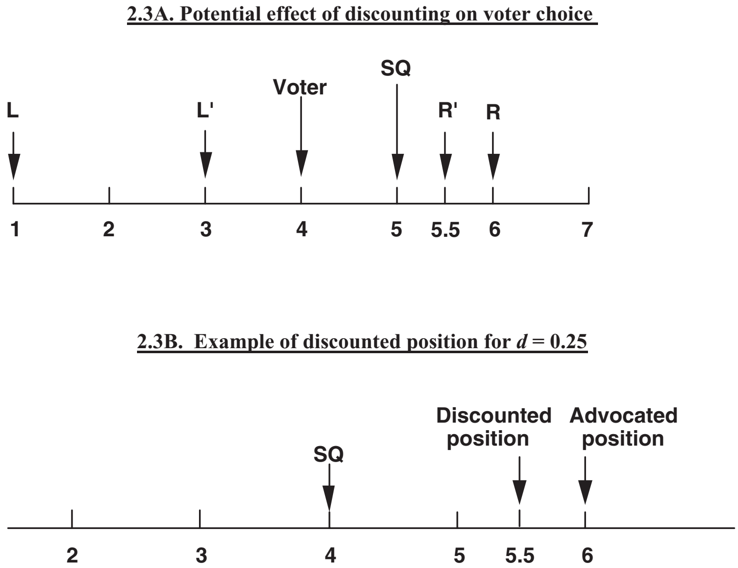

Figure 2.3. Candidate location under the discounting model.
Note: For 2.3A, advocated positions are **L** and **R**, discounted positions are **L**′ and **R**′.

advocate policies that are more extreme than those they would advocate if discounting did not occur.

In order to implement the Grofman discounting model, we denote by d_k the discounting factor for party k, which, for simplicity, we will assume is the same for all voters and for each issue dimension. The factors d_k all lie between 0 and 1. When we say, for example, $d_k = 0.25$, we mean that the voters discount candidate k's capacity to draw the status quo to her advocated position by 25 percent.[8] In effect, voters multiply the distance from the status quo to the candidate's advocated position by one minus the discount factor – that is, by 0.75. Thus if the status quo is at 4.0 and the candidate advocates a policy at 6.0, the voters anticipate implementation at 5.5 (see Figure 2.3B). In general, if SQ denotes the status quo point, the voters anticipate that candidate k, who advocates policy at s_k, will implement policy at location $SQ + (1 - d_k)(s_k - SQ)$. If $d_k = 0$, there is no discounting; if $d_k = 1$, then there is complete discounting and the status quo will be implemented.

[8] Note that the discounting factor d_k as defined here is equal to $(1 - d_k)$, as defined in Merrill and Grofman (1999).

https://doi.org/10.1017/CBO9780511614453.003 Published online by Cambridge University Press

There is a simple relationship between the level of discounting, specified by the value of d, and the policy position that must be advocated by a candidate if voters are to anticipate policy implementation at a specific position. For example, if $d = 0.5$, $SQ = 4$, and voters are to expect policy implementation at 5.0, then the candidate must advocate policy at 6.0 – that is, at $4.0 + (5.0 - 4.0)/(1 - d) = 6.0$. Thus if voters discount candidates' capacities to implement the policies they advocate, the candidates can be expected to advocate policies that are more extreme than those the voters actually want – in fact, more extreme by approximately the factor $1/(1 - d)$. For example, if discounting is 25 percent, policy advocacy can be expected to be about 33 percent more extreme; if discounting is 50 percent, policy advocacy can be expected to be about 100 percent more extreme, and so on.

Thus, under the Grofman discounting model, the voter's utility function given in equation 2.6 is replaced by

$$U_i(k) = V_i(k) + X_{ik} = \sum_j a_j \left[x_{ij} - \left(SQ + (1 - d_k)(s_{kj} - SQ)\right)\right]^2 + \mathbf{b}_k \mathbf{t}_{ik} + X_{ik} \qquad (2.7)$$

where, again, a_j is the policy-salience parameter for the jth issue, $V_i(k)$ is the deterministic component of utility, and $\mathbf{b}_k \mathbf{t}_{ik}$ represents the measured nonpolicy utility components. We label the specification given in equation 2.7 the *unified discounting model* of voting. Appendix 2.3 presents a discussion of the relationship between our discounting formulation and the directional voting model proposed by Rabinowitz and Macdonald (1989). Appendix 6.1 presents a similar discussion for the outcome-oriented model of Kedar (2002).

2.5. Discussion

In this chapter, we have presented a range of increasingly complex models so as to be able to incorporate a wide range of motivations in voter decision making. We began with a spatial model based on one or more policy dimensions and then incorporated both measured and unmeasured (random) components representing nonpolicy factors affecting voting decisions. Finally, we returned to policy considerations and explored alternative voter metrics (direct proximity, discounted proximity) for evaluating candidates' associations with policies.

https://doi.org/10.1017/CBO9780511614453.003 Published online by Cambridge University Press

A question that has arisen in connection with empirically assessing the relative contributions of the proximity and discounting (and directional) utilities is whether to use voter-specific placements of candidates or mean placements when computing the distance between voters and candidates (see Rabinowitz and Macdonald 1989; Merrill and Grofman 1999). We believe that conceptually idiosyncratic placement is preferable because, unlike mean placement, it is known to each voter. For the voter-choice models used in this volume – as opposed to regression models using a thermometer scale as dependent variable – use of mean or voter-specific placements yields very similar parameter estimates and equilibrium positions. Hence, for simplicity, we report results based on mean placements throughout the book.

These models of voting – the policy-only model, the unified model, and the unified discounting model – will be used throughout the book to assess candidates' and parties' responses to voter behavior in historical elections in four countries exemplifying a range of electoral rules and circumstances. In particular, for elections in each of these four countries, we will determine the strategies (i.e., policy positions) that parties might choose to maximize their vote shares, and, in order to assess the *responsible party model*, we will compare each party's vote-maximizing location with the mean location of its supporters. The effects of nonvoting will be investigated primarily in the context of American elections – in which this phenomenon is particularly prominent – although the effect of abstention will also be analyzed for Britain. Because there are only two major parties in the United States, abstention is crucial to our arguments for party divergence in America. For these reasons, technical discussion of the extension of the model to accommodate the effects of nonvoting on party positioning will be delayed until Chapter 7.

In addition to their responses to voters, however, parties have their own motivations concerning policy. In Chapters 11 and 12, we will see how these latter motivations, coupled with uncertainty about the expected vote, lead to significant modifications of the parties' optimal strategies. We begin the process of investigating alternative models in Chapter 3 via illustrative arguments about the ways in which different degrees of partisan voting affect candidates' policy strategies.

https://doi.org/10.1017/CBO9780511614453.003 Published online by Cambridge University Press