PA

The Neglected Role and Variability of Party Intercepts in the Spatial Valence Approach

Ingrid Mauerer[ID]

Department of Political Science, LMU Munich, Oettingenstr. 67, D-80538 Munich, Germany.
Email: ingridmauerer@web.de

Abstract

Empirical applications of the spatial theory of elections typically rely on the discrete choice framework to arrive at probabilistic voting models. Whereas in the classic model voter choice is solely a function of spatial proximity, neo-Downsian models also incorporate voter-specific nonpolicy attributes, which are represented by sociodemographic characteristics. One prominent line of such probabilistic models, Schofield's Valence Model, additionally includes party valences into voter utility functions. The model rests on the estimated party intercepts to measure the valence advantages empirically. The party intercepts are ordered based on their values, and then this valence ranking is used further to predict equilibrium locations. The paper demonstrates that this measurement strategy does not provide unique results in fully specified models due to central properties of discrete choice models and the specific nature of party intercepts in these models. Drawing on a simple example based on mass election surveys from Germany, it is shown that the valence ranking, the crucial factor to investigate how valence differences affect the nature of spatial competition, is highly sensitive to arbitrary coding decisions. As a consequence, it is impossible to represent valence with the constants and to infer something substantial from the resulting valence ranking.

Keywords: probabilistic voting, spatial competition, valence, party intercepts, discrete choice models

1 Introduction

Spatial models of voting and electoral competition (Downs 1957; Davis, Hinich, and Ordeshook 1970; Enelow and Hinich 1990) have become tremendously important and powerful to understand the interplay between voter preferences and spatial party strategies. In the search for a better understanding of how citizens form their preferences to arrive at an electoral decision and how parties or candidates[1] strategically adopt policy platforms in response to the electorate, a huge literature emerged (see Dewan and Shepsle 2011; Jackson 2014, for reviews). The still growing 'neo-Downsian agenda' (Grofman 2004, 40) spans many different subfields, addressing various aspects such as: Are voters solely driven by policy preferences, or do they have nonpolicy motivations? What role play nonpolicy characteristics of parties, considered as valence, and how do these dimensions of competition interact? And how do all these factors affect optimal party locations? The discrete choice framework seems to provide the promising modeling approach to translate the theory into empirical models that allow addressing these and many more questions in real-world elections.

Based on random utility maximization, discrete choice models rely on an underlying microfoundation that can be directly linked to the logic of probabilistic voting (e.g., Enelow and Hinich 1989; Coughlin 1992; Adams 1999; Lin, Enelow, and Dorussen 1999; Thurner 2000). Discrete choice models based on a logit or probit link function have become well-established tools in the rational choice modeling of voting behavior and party competition. Especially the usage of

---

1 Throughout the article, I employ the language of party competition, which equally applies to candidate competition.

conditional logit models (McFadden 1974) has become common practice.[2] Although differing in the operational model they employ and how underlying latent policy dimensions are constructed, empirical applications share one common approach: they rest on the model parameters to understand both voter choice and party policy strategies. Based on survey data, they first estimate voter choice models. Then, the empirically estimated parameters are used further on to calculate equilibrium constellations and to explore their nature, conditions, and characteristics. Work in this genre produced comprehensive findings on equilibria in spatial competition (e.g., Adams and Merrill III 2000, 2009; Merrill III and Adams 2001; Adams *et al.* 2013).

While in the traditional spatial framework voter choice is solely a function of spatial proximity on one single dimension, neo-Downsian models are multidimensional and incorporate both spatial proximity and voters' nonpolicy motivations. The latter is typically represented by sociodemographic voter attributes to account for differences in voting behavior among the electorate. Several probabilistic voting models also include the parties' nonspatial qualities, considered as party valences, into voter utility functions. One prominent and influential line of such probabilistic models, Schofield's valence model (e.g., Schofield 2003, 2004; Schofield and Sened 2005a,b, 2006; Schofield and Miller 2007), relies on the estimated party intercepts in discrete choice models to measure the valence qualities and advantages of parties empirically. These party-specific constants are assumed to capture the average weight voters attach to the parties' valence qualities. The respective estimates are ordered based on their absolute values[3] to assess how quality differences affect the nature of spatial competition.

I show here that using this valence ranking, the key factor in Schofield's valence model, further on to calculate and analyze equilibrium strategies does not allow us to generate generalizable conclusions about the impact of valence on parties' spatial strategies. Under the discrete choice framework, only differences in utility matter and the absolute values are irrelevant. Due to this central property and the specific nature of party intercepts within this framework, the valence ranking is highly sensitive to the coding of voter attributes. Depending on arbitrary coding decisions, the valence ranking differs, as changes in the coding lead to changes in absolute values. As a consequence, it is impossible to represent valence with the constants and to learn something substantial from the resulting valence ranking. I recommend that researchers do not rely on party intercepts, which reflect the relative average effect of unmeasured sources of utility—that might or might not be nonspatial party qualities—but to make the valence qualities measurable instead and incorporate them into voter utility functions by observable components.

2 Neo-Downsian Spatial Voting Models

Neo-Downsian models usually operate in a probabilistic framework where both observed and unobserved factors determine voter utility. The random component is assumed to be independent of spatial factors and is formulated in various ways in the literature (see, e.g., Enelow and Hinich 1989; Coughlin 1992; Burden 1997; Adams 1999; Lin, Enelow, and Dorussen 1999). Another central aspect and major model extension relates to the importance of voters' nonpolicy motivations in addition to policy in the voting calculus (e.g., Adams and Merrill III 1999a,b, 2000; Thurner 2000; Merrill III and Adams 2001; Adams, Merrill III, and Grofman 2005). By accounting for nonspatial motivations, a behavioral perspective was incorporated into the spatial voting model in which psychological or sociological factors also influence voter choice.

---

2 Conditional logit applications to analyze both party-centered and candidate-centered elections in many different contexts include Alvarez and Nagler (1995, 1998); Endersby and Galatas (1998); Lacy and Burden (1999); Adams and Merrill III (1999a,b, 2000, 2009); Quinn, Martin, and Whitford (1999); Alvarez, Nagler, and Bowler (2000); Thurner (2000); Merrill III and Adams (2001); Kedar (2005); Sanders *et al.* (2011); Mauerer, Thurner, and Debus (2015); Mauerer *et al.* (2015); Stoetzer and Zittlau (2015); Mauerer (2016); and Ansolabehere and Puy (2018). Numerous other work could be included in this list.

3 An absolute value is understood as a value that is not compared to another value. It does not refer to the unsigned value and can, therefore, take positive or negative values.

The nonpolicy considerations enter empirical models by measured variables that represent the voters' socioeconomic characteristics, class, or religious ties, etc. The incorporation of such voter attributes allows accounting for segment-specific evaluations of parties.

Several probabilistic voting models posit that voters also evaluate and value the nonspatial qualities of parties. In these models, party competition is characterized by an additional dimension, namely valence. The concept of valence, going back to Stokes (1963, 1992) and his fundamental critique of the Downsian framework, received much scholarly attention in both formal theoretical and empirical work. In the electoral research literature, valence is very diversely understood, resulting in many different debates and subfields with varied definitions, subcategories, and empirical measurement or operational strategies (for an overview, see Green and Jennings 2017, and references therein). Within the spatial voting literature, valence is very broadly defined and generally conceptualized as a nonpolicy-related characteristic of a party (or a party leader) that distinguishes parties from one another. The literature proposes several spatial models that connect the two dimensions of competition to study their interplay, yielding a whole class of models in which a valence term enters voters' utility functions. The formal work produced a lot of theoretical results regarding equilibrium constellations and properties of spatial competition (e.g., Enelow and Hinich 1982; Ansolabehere and Snyder Jr. 2000; Groseclose 2001; Aragonès and Palfrey 2002; Ashworth and de Mesquita 2009; Serra 2010). This prominent stream of research demonstrates that party quality differences affect the nature of political competition (for a recent review of formal models, see Evrenk 2019).

As diverse as the understanding of the concept is in the different theoretical debates, are the ways how it is quantified and modeled. Empirical probabilistic voting models typically assume the party valence component to be unrelated to spatial factors. The main difference in how valence is operationalized in the empirical work is whether the party quality term enters the systematic or the unobserved component of voter utility functions when using discrete choice models to arrive at an operational model. Some studies model it directly by observed variables using survey or expert data designed to capture valence qualities, so that the valence term enters utility functions as measured components (Stone, Maisel, and Maestas 2004; Stone and Simas 2010; Adams *et al.* 2011; Buttice and Stone 2012; Franchino and Zucchini 2015). Other works consider valence as a dichotomous concept consisting of measured and unmeasured attributes (Adams, Merrill III, and Grofman 2005; Adams and Merrill III 2009; Adams *et al.* 2013).

One prominent line of probabilistic models, which became an influential strategy for incorporating valence considerations into empirical models, is Schofield's spatial valence model of politics. The model and extensions thereof were applied to many different countries, such as the United States (e.g., Schofield, Miller, and Martin 2003; Schofield and Miller 2007; Gallego and Schofield 2016), Britain (e.g., Schofield 2004, 2005; Schofield, Gallego, and Jeon 2011), Germany (e.g., Kurella and Pappi 2015; Schofield and Kurella 2015), Israel (e.g., Schofield 2004; Schofield and Sened 2005b), or Russia (Schofield and Zakharov 2010). To measure the valence advantages empirically, the model exclusively relies on unobserved information by using the party intercepts in discrete choice models. The estimated party-specific constants are assumed to reflect "the average perception, among the electorate, of the 'quality' of the party leader" (e.g., Schofield 2005, 348). These valence characteristics are defined as a nonpolicy-related term λ_{ij} that is "interpreted as the weight that voter i gives to the competence of candidate j" (e.g., Schofield 2004, 448).[4] The valence terms are ranked based on the absolute constant estimates so that $\lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_1$, where p denotes the number of parties and 1 the party with the lowest valence. The average valence of the parties other than the lowest ranked is given by $\lambda_{av(1)} = [1/(p-1)] \sum_{j=2} \lambda_j$, and

---

4   Note that in the following λ_{ij} is not subscripted by i anymore: "In empirical estimations, it is usual to assume that the term λ_{ij} comprises an expected 'exogenous' term, λ_j, together with a 'stochastic' error term, ϵ_j." (Schofield 2004, 448).

$\Lambda = \lambda_{av\,(1)} - \lambda_1$ defines the valence difference for the lowest valence party. The valence ranking is then the crucial factor to investigate how valence differences influence parties' spatial strategies. Whereas the formal model does not account for sociodemographic voter characteristics (e.g., Schofield and Sened 2006, 33), which are "regarded as *sociodemographic valences*, generated by common perceptions of the parties by different societal subgroups" (Schofield and Zakharov 2010, 179), empirical models incorporate them in order to avoid that the valence terms are exaggerated.

However, empirical applications of the formal model do not yield unique results in fully specified voter models (i.e., when both voters' spatial and nonpolicy motivations are incorporated into voter utility functions) due to central properties of discrete choice models and the specific role party intercepts play in these models. The next section summarizes these fundamental properties to identify the problem.

3 Spatial Voting Models within the Discrete Choice Framework

Within the spatial voting framework, we assume voter $i \in \{1, \ldots, n\}$ faces a choice among J parties and each party $j \in \{1, \ldots, J\}$ provides a specific amount of utility, denoted by U_{ij}. Following the principle of utility maximization, voter i chooses party j if and only if: $U_{ij} > U_{ir}, \forall j \neq r$. Probabilistic voting models incorporate a random element in the voter's calculus so that both observed and unobserved factors determine voter utility. This rationale is in accord with discrete choice theory and can be related to the concept of random utility maximization (Manski 1977), which is the most widely used derivation of discrete choice models (see, e.g., Louviere, Hensher, and Swait 2009; Train 2009; Hensher, Rose, and Greene 2015). Under this approach, a probabilistic choice mechanism is introduced so that the true utilities of the alternatives are random variables. The randomness is caused by the researcher's lack of full information about the decision process, i.e., by the nonobservability of all relevant decision criteria due to unobserved characteristics of the alternatives and individuals, measurement or specification errors.

The limitation on the side of the researcher is captured by dividing the overall utility U_{ij} into two parts of the utility expression so that $U_{ij} = V_{ij} + \epsilon_{ij}$. The first part represents the utility contributions that are measured, and the second one contains the utility sources we cannot observe. Relating this perspective to voter choice and after rearranging, the probability of voter i to decide for party j equals the probability that the difference in the unobserved factors induced by any other party r compared to party j is below the difference in the observed factors of these two parties: $Pr(\epsilon_{ir} - \epsilon_{ij} < V_{ij} - V_{ir}), \forall j \neq r$.

As the choice probabilities are only a function of differences, the utilities derived from discrete choice models are based on ordinal utility theory. Under this framework, utilities express the preference ordering of parties, and the absolute values are irrelevant. Ordinal utilities, therefore, indicate relative utility differences which are maintained by order-preserving changes in the absolute values. The fact that the choice probabilities only depend on utility differences has major implications for model specification and identification, parametrization, and parameter interpretation. Before we discuss these issues in more detail, let us consider the components that typically enter voter utility functions in neo-Downsian spatial models.

I specify a simple model that is multidimensional and incorporates both a policy component and nonspatial factors into the choice rule. The spatial term contains voter-party proximity measures. For simplicity, I assume a city-block voter policy metric based on individual perceptions of party positions. Let $z_{ijk}, k \in \{1, \ldots, K\}$, represent the proximity between the voters' self-placements x_{ik} and the voter-specific party placements p_{ijk} on each issue k so that $z_{ijk} = -|x_{ik} - p_{ijk}|$.[5] Sociodemographic variables typically represent voters' nonpolicy motivations.

---

5 The points I discuss here hold regardless of how we construct the policy space and what functional form the voter policy metric takes. That is, the elements in z_{ij} can be altered without a qualitative change of major points, as the corresponding parameters are unaffected.

Let s_{im}, $m \in \{1, \ldots, M\}$, contain these voter characteristics. For $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, J\}$, define V_{ij} as linear predictors that accumulate the voter utility functions in scalar quantities:

$$V_{ij} = \beta_{j0} + \sum_{k=1}^{K} z_{ijk}\alpha_k + \sum_{m=1}^{M} s_{im}\beta_{jm} = \beta_{j0} + z_{ij}^T\alpha + s_i^T\beta_j. \tag{1}$$

The parameters $\beta_{10}, \ldots, \beta_{J0}$ are intercepts specific to each party j.[6] These constants play a central role in the utility expression of discrete choice models. They capture the average effect of unobserved sources of utility associated with each party j.[7] Thus, the constant terms represent the average role of all factors that are central to the electoral choice but are not measurable, i.e., cannot be related to observable party or voter attributes. $\alpha^T = (\alpha_1, \ldots, \alpha_K)$ is the parameter vector related to voter-party proximity measures z_{ij} on K different policy dimensions. These coefficients indicate the weight voters attach to the policy issues.[8] $\beta_j^T = (\beta_{j1}, \ldots, \beta_{jM})$ is the coefficient vector related to the M-dimensional individual-specific covariate vector s_i. These parameters are specific to each party j and indicate segment-specific evaluations of parties. To derive at the choice probabilities of the conditional logit model (McFadden 1974), the working horse in empirical applications, we assume a logit link function that connects the V_{ij} with the choice probabilities, yielding

$$\pi_{ij} = Pr(Y = j | z_{ij}, s_i) = \frac{\exp(\beta_{j0} + z_{ij}^T\alpha + s_i^T\beta_j)}{\sum_{r=1}^{J} \exp(\beta_{r0} + z_{ij}^T\alpha + s_i^T\beta_r)}, \tag{2}$$

where π_{ij} denotes the J-dimensional vector of choice probabilities and $Y \in \{1, \ldots, J\}$ the j-categorical, probabilistic response variable.

Since only differences in utility are relevant, and the absolute utility values are meaningless in discrete choice models, particular parameters in Equation (2) need special consideration. Here it is important to note that two different types of predictors represent the spatial and nonspatial components of the voter utility functions: variables that are specific to parties and voters (voter-party proximity measures z_{ij}), and variables that are specific to voters only (voter attributes s_i). This distinction is fundamental as only those parameters can be estimated freely that relate to variables which capture differences across parties. As the voter attributes s_i are invariant across parties, the level of utility has to be set to an arbitrary value for at least one pair of the parameters $\beta_j^T = (\beta_{j1}, \ldots, \beta_{jM})$ to make the remaining ones estimable. The same applies to the constants $\beta_{10}, \ldots, \beta_{J0}$. As a result, Equation (2) refers to the conditional logit model in its generic form where at most $J - 1$ intercepts and $(J - 1) \times M$ parameters for the voter attributes are identifiable. The standard side constraint to identify the model is to select one party l whose constant and coefficients related to voter attributes are normalized by setting $\beta_{l0} = 0$ and $\beta_l^T = (0, \ldots, 0)$. The choice probabilities for $j \in \{1, \ldots, J - 1\}$ and for the reference party l are given by

$$Pr(Y = j | z_{ij}, s_i) = \frac{\exp(V_{ij})}{\exp(V_{il}) + \sum_{r=1}^{J-1} \exp(V_{ir})}; \quad Pr(Y = l | z_{ij}, s_i) = \frac{\exp(V_{il})}{\exp(V_{il}) + \sum_{r=1}^{J-1} \exp(V_{ir})}. \tag{3}$$

---

6 To demonstrate the interplay between party intercepts and voter attributes, the party intercepts are denoted by β_{j0} and not λ_j as in Schofield's valence model.

7 Note that ϵ_{ij} has a zero mean when party intercepts enter the utility functions.

8 For simplicity, I assume that the issue dimensions are separable and independent. For the relaxation of this assumption, see Stoetzer and Zittlau (2015). I also constrain the parameters α to be the same for all J parties, i.e., $\alpha_1 = \cdots = \alpha_j := \alpha$. For the relaxation of this assumption, see Mauerer, Thurner, and Debus (2015), Mauerer (2016), Thurner (2000).

Since $\beta_{I0} = 0$ and $\beta_I^T = (0, \ldots, 0)$, the log odds for $j \in \{1, \ldots, J-1\}$, giving the utility differences associated with party j as compared to the reference party I, simplify to

$$\log\left(\frac{Pr(Y = j | \mathbf{z}_{ij}, \mathbf{s}_i)}{Pr(Y = I | \mathbf{z}_{ij}, \mathbf{s}_i)}\right) = V_{ij} - V_{il} = \beta_{j0} + \mathbf{s}_i^T \boldsymbol{\beta}_j + (\mathbf{z}_{ij} - \mathbf{z}_{il})^T \boldsymbol{\alpha}. \tag{4}$$

The relative utility difference between each of the parties remains the same by order-preserving changes in absolute values. Changes in the coding of voter attributes $\mathbf{s}_i$ induce changes in the absolute parameter values for the constants. Depending on how we code voter attributes to account for differences in the decision-making process among members of particular social segments, which is an arbitrary choice we make, the absolute constant values differ. For instance, assume $\mathbf{s}_i$ only contains a dummy-coded variable:

$$V_j - V_I = \begin{cases} \beta_{j0} + \beta_j + (\mathbf{z}_{ij} - \mathbf{z}_{il})^T \boldsymbol{\alpha} & \text{if } s = 1, \\ \tilde{\beta}_{j0} + (\mathbf{z}_{ij} - \mathbf{z}_{il})^T \boldsymbol{\alpha} & \text{if } s = 0. \end{cases} \tag{5}$$

When we solely reverse the coding of one single voter attribute, all constants are affected. The respective parameters adjust to leaving the utility differences and therefore, the corresponding choice probabilities unchanged. By contrast, the parameters related to the policy component are unaffected, as the party-specific variables $\mathbf{z}_{ij}$ define differences between parties.

4 Empirical Illustration

For illustration purposes, I use survey data on voter choices over the five major parties in the 1998 German parliamentary election: the Christian Democrats (CDU), the Social Democratic Party (SPD), the Liberal Party (FDP), the Greens, and the Left Party (PDS). The 1998 German national election study (Falter, Gabriel, and Rattinger 2012) includes three policy issues and the ideological Left–Right scale on which respondents were asked to locate themselves and the parties. These placements are used to construct the voter-party proximity variables $\mathbf{z}_{ij}$. As voter-specific nonpolicy attributes $\mathbf{s}_i$, six standard sociodemographics were selected: working class, union membership, religious denomination, age, gender, and region.[9]

I begin by illustrating the important role the party intercepts play in discrete choice models. The estimates do not only capture the average effects of unobserved sources of utility, but they also reproduce the observed vote shares in the sample within the conditional logit framework. The parameters should be interpreted with that in mind. To demonstrate that, I fitted constant-only models which do not contain any voter choice predictors but only use the information on the vote shares in the data to estimate the choice probabilities. The deterministic part of utility associated with each party j simplifies to $V_j = \hat{\beta}_{j0}$. Since no covariates enter the models, the utility functions and the choice probabilities $\hat{\pi}_j$ are not subscripted by voter i and are therefore identical for each voter in the sample. Table 1 compares the resulting utility functions and choice probabilities. The first column gives the vote shares as observed in the sample.[10] The remaining columns report the party intercept estimates where each time the constant associated with one of the parties is normalized to zero to ensure model identification. The values give the utility functions as compared to the respective reference party. The last column reports the estimated vote shares. Since the models based on differing reference parties contain the same information, all $J - 1$ combinations of the constant terms are functionally equivalent and yield the observed vote shares

---

9 For operationalization details, see Section A in the Supplementary Materials. All models were estimated with the statistical software R. For Dataverse replication materials, see Mauerer (2019).

10 Note that the sample vote shares displayed in Table 1 do not perfectly match the vote shares in the population (see Section A in the Supplementary Materials). However, this does not change the points I wish to make.

https://doi.org/10.1017/pan.2019.41 Published online by Cambridge University Press

Table 1. Intercept-only model with different reference parties as normalization.

	Sample	$V_j = \hat{\beta}_{j0}$										**Estimated**
	Vote Shares π_j	**CDU**		**SPD**		**FDP**		**Greens**		**Left**		**Vote Shares** $\hat{\pi}_j$
CDU	0.30	—		−0.42	(0.09)	1.95	(0.19)	1.04	(0.13)	1.32	(0.15)	0.30
SPD	0.46	0.42	(0.09)	—		2.37	(0.19)	1.46	(0.13)	1.74	(0.14)	0.46
FDP	0.04	−1.95	(0.19)	−2.37	(0.19)	—		−0.91	(0.21)	−0.63	(0.22)	0.04
Greens	0.11	−1.04	(0.13)	−1.46	(0.13)	0.91	(0.21)	—		0.28	(0.17)	0.11
Left	0.08	−1.32	(0.15)	−1.74	(0.14)	0.63	(0.22)	−0.28	(0.17)	—		0.08

Source: 1998 German national election study (Falter, Gabriel, and Rattinger 2012).
Note: Entries are to be read columnwise. Numbers in parentheses show standard errors.
$N = 715$, $df = 4$, AIC = 1864.86, BIC = 1883.15, Log $L = -928.43$.

in the sample. When defining the SPD as the reference party, for instance, and plugging the vote shares for the CDU and the SPD into Equation (4), one obtains the log odds. These give the relative difference in utilities that equals the intercept associated with the CDU:

$$\log\left(\frac{Pr(Y = \text{CDU})}{Pr(Y = \text{SPD})}\right) = \log\left(\frac{0.30}{0.46}\right) = -0.42 = V_C = \hat{\beta}_{C0}.$$

The utility difference between these two parties and any party pairs are preserved regardless of what party is defined as the reference because the relative utility differences are maintained by order-preserving changes in the absolute values.[11]

Next, I specify a voter choice model that adds substantive information so that the relative differences in unobserved utilities between the parties should decrease significantly. In accord with Equation (1), the model is fully specified and contains four spatial voter-party proximities (immigration, European unification, nuclear energy, Left–Right) and six voter attributes. The model uses the CDU as the reference party and is based on 32 degrees of freedom: 4 issue weights, $5-1$ intercepts, and $(5-1) \times 6$ parameters related to voter attributes. The maximum likelihood point estimates and associated standard errors for the policy component are as follows: $\hat{\boldsymbol{\alpha}}^T = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3, \hat{\alpha}_4) = (0.09, 0.20, 0.31, 0.39)$; $\hat{\boldsymbol{\sigma}}_{\boldsymbol{\alpha}}^T = (\hat{\sigma}_1, \hat{\sigma}_2, \hat{\sigma}_3, \hat{\sigma}_4) = (0.05, 0.06, 0.04, 0.03)$. The party intercept estimates are $\hat{\boldsymbol{\beta}}_0^T = (\hat{\beta}_{S0}, \hat{\beta}_{F0}, \hat{\beta}_{G0}, \hat{\beta}_{L0}) = (-0.06, -1.73, -2.44, -0.37)$; $\hat{\boldsymbol{\sigma}}_{\boldsymbol{\beta}_0}^T = (0.28, 0.57, 0.58, 0.38)$.

When we rank the constant terms from p to 1, where p denotes the number of parties and 1 the lowest valence party, as proposed in Schofield's valence model by $\hat{\boldsymbol{\beta}}_0^T = (\hat{\beta}_{[p]0}, \ldots, \hat{\beta}_{[1]0}) = (0, -0.06, -0.37 - 1.73, -2.44)$, we get the ordering $(\hat{\beta}_{C0}, \hat{\beta}_{S0}, \hat{\beta}_{L0}, \hat{\beta}_{F0}, \hat{\beta}_{G0})$.[12] The average valence of the parties other than the lowest ranked $\Lambda = 1.90$. According to this framework, the result suggests that the CDU is the party with the highest valence, followed by the SPD, whereas the Greens are the lowest valence party.[13] The valence ranking reflects the one empirical applications of the valence model typically obtain: the major parties have a high valence and parties gaining small vote shares have low valences. Under the discrete choice framework, such a result, however, indicates that our empirical voter choice models are good at capturing utility differences between large parties but not between large and small parties.

Next, I demonstrate that the valence ranking is highly sensitive to the coding of voter attributes. Figure 1 illustrates how arbitrary coding decisions affect the estimates for the party intercepts and the valence ranking. Each time I reversed the coding of one of four dummy-coded covariates and re-estimated the models. From the left panel in Figure 1, we observe that the signs of all estimates related to these covariates are reversed, but the sizes remain the same. The coefficients for all other voter attributes and the voter-party issue proximities, which are not displayed (see Section D in the Supplementary Materials for estimation tables), are unaffected by the reversed coding, but the constants are not. Since only utility differences matter, the party intercepts adjust to the sign change to reproduce these differences and to maintain the respective choice probabilities.[14]
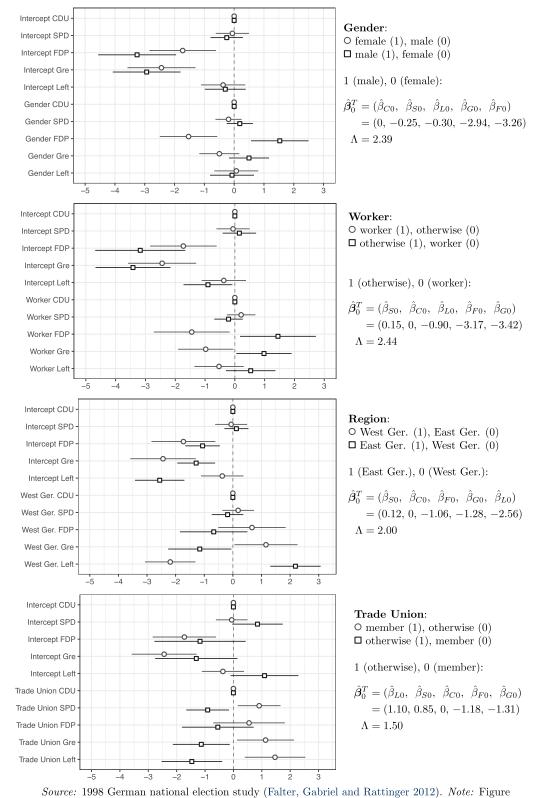
As shown in the right panel of Figure 1, the reversed coding of the standard variable gender would lead us to conclude that the FDP is the lowest valence party instead of the Greens, what we

---

11 CDU reference: $\Delta(V_{CDU} - V_{SPD}) = 0 - 0.42 = -0.42$; FDP reference: $\Delta(V_{CDU} - V_{SPD}) = 1.95 - 2.37 = -0.42$; Greens reference: $\Delta(V_{CDU} - V_{SPD}) = 1.04 - 1.46 = -0.42$; Left reference: $\Delta(V_{CDU} - V_{SPD}) = 1.32 - 1.74 = -0.42$.

12 Note that the intercept for the reference party is not an actual value because it is not estimated. However, since such values are always included in the valence ordering in empirical applications, I proceed in the same way.

13 Again, this ordering is preserved no matter what party is selected as reference (see Table A2 in Section C in the Supplementary Materials), as all types of normalization reproduce the same utility differences.

14 For instance, assume a person that is not a member of a trade union, non-Catholic, middle-aged, based in former East Germany with zero voter-party proximities so that only the constants and the parameter related to the variable gender remain in the linear predictors. When we insert the corresponding values of the models reported in Table A5 (Section D in the Supplementary Materials) into Equation (3), we get identical choice probabilities for females, for instance, to vote for the SPD: $\exp(-0.06 - 0.19)/(1 + \exp(-0.06 - 0.19) + \exp(-1.73 - 1.53) + \exp(-2.44 - 0.50) + \exp(-0.37 - 0.07)) = 0.30$; $\exp(-0.25)/(1 + \exp(-0.25) + \exp(-3.26) + \exp(-2.94) + \exp(-0.30)) = 0.30$.
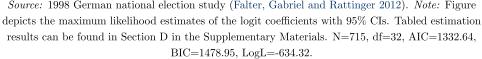

Figure 1. Impact of reverse coding of voter attributes on valence (CDU as reference party).

would infer under the original coding. When we reverse the coding of the variable worker, the SPD is the highest valence party instead of the CDU. The reversed coding of the region variable leads to the Left as the lowest valence party, which was ranked in the middle under the original coding. The most dramatic change in valence ranking is observed when we recode the variable trade union: the Left moves to the highest valence party, ranking before the major parties SPD and CDU.

As a consequence, when we rely on the estimated party intercepts to measure the valence advantages of parties empirically, a single change in the coding of voter attributes affects the results and their substantial interpretation. Depending on how we code voter attributes to account for differences in the decision-making process among members of particular social segments, the ordering of valence parties differs. Using the valence ranking further on to calculate and analyze equilibrium positions will produce different conclusions we would draw from the model. To illustrate this, let us consider how the convergence coefficient[15] is defined in the valence model:

$$c = \frac{2(1 - 2\rho_1)\text{trace}(\tilde{\alpha}\nabla_0^*\tilde{\alpha})}{\frac{1}{K}(\hat{\alpha}_1 + \hat{\alpha}_2 + \cdots + \hat{\alpha}_K)}. \tag{6}$$

$(\hat{\alpha}_1, \ldots, \hat{\alpha}_K)$ are the estimated issue weights for the K different policy dimensions, and $\tilde{\alpha}$ is a $K \times K$ diagonal matrix of issue weights. $\nabla_0^* = (1/n)\nabla_0$ is the $K \times K$ electoral covariance matrix of n voters' ideal points. The smallest valence party comes into play by the term ρ_1, which defines the probability that the voter chooses the lowest valence party when all parties are at the joint electoral mean (e.g., Schofield and Zakharov 2010, 180):

$$\rho_1 = \left[1 + \sum_{j \neq 1} \exp(\beta_{j0} - \beta_{10})\right]^{-1}. \tag{7}$$

As $\mathbf{z}_{ij}^T = (0, \ldots, 0)$ when all parties are located at the joint electoral mean, the vote shares are assumed to be independent of voter i's ideal points. Since the policy component drops out of the equation, the choice probability of the party with the smallest valence is only a function of differences in absolute constant estimates. As demonstrated above, depending on how voter attributes are coded, a different lowest valence party can result. That is, depending on arbitrary coding decisions each time another party would be considered in the calculation of the convergence coefficient.

The intercept values are also sensitive to the variables included in the model. When we delete or add covariates, the intercept estimates, and therefore the resulting valence ranking can change as well. The variable selection affects the information contained in the observed part of utility, and consequently, the unobserved utility sources represented by the constants. This equally applies to unstandardized and standardized variables.

One might argue that the solution to the variability of party intercepts might be to specify policy-only models as the voter-party proximities $\mathbf{z}_{ij}$ are unaffected by the coding of voter attributes. However, by neglecting the importance of segment-specific evaluations of parties, we run the risk of missing crucial parts of how voters form their preferences to arrive at an electoral choice. A model comparison based on performance measures (Log-Likelihood, AIC, BIC) shows

---

15 To ensure consistency, I stick to the notation introduced above. Note that Equation (6), which I slightly adjusted to assure uniqueness of the elements in c, gives the convergence coefficient for the multidimensional model (see Schofield and Kurella 2015, 298).

that the full model clearly outperforms the policy-only model: Voter attributes are essential for the specification of a realistic model of voting and party competition in mass elections.[16]

It is also not a remedy to look at predicted probabilities (e.g., by calculating choice probabilities of a particular combination of voter attributes or the 'average voter'). Under this empirical measurement strategy, no variable enters the model that explicitly measures valence, and that can be increased to examine how a variable change affects the choice probabilities. Since such computations condition the predicted probabilities on voter attributes and spatial proximities, they would not help us to understand how much weight or importance voters attach to valence considerations—broadly understood as a positive nonspatial characteristic of a party. That is, a concept that relates first and foremost to parties, not voters even though the valence perceptions might vary across voters.

5   Discussion

The preceding empirical illustration highlights the analytical challenge in studying valence and the need for alternative ways to measure valence empirically. It raises the question of how we can deal with the concept within the spatial voting framework. What is the exact meaning and definition of valence? That is the most important question to be addressed first, as conceptual work must lay the foundation for its empirical analysis. In his critique of the Downsian voting framework, Stokes (1963, 363) defines valence issues as "those that merely involve the linking of the parties with some condition that is positively or negatively valued by the electorate". Within the spatial voting paradigm, valence is very broadly understood and is generally defined as any positive nonpolicy-related attribute of a party (or party leader) that voters value. But, what are desirable qualities, traits, or skills that form an additional base—next to spatial considerations—on which voters evaluate how attractive the different parties are to them? Within the spatial voting literature, a very promising conceptualization of valence is employed in Stone, Maisel, and Maestas (2004), Stone and Simas (2010) and Adams *et al.* (2011). This work proposes to differentiate between two different types or dimensions of valence: character or personal qualities and strategic qualities. Character qualities refer to attributes that define abilities and skills such as competence, job performance, leadership, integrity, trustworthiness, or diligence. By contrast, strategic qualities are instrumental in nature and comprise, for instance, name recognition, campaign or fundraising abilities. Voters are expected to respond first and foremost to character valence-based characteristics because strategic qualities are ones that parties might possess or not enjoy, but voters not intrinsically value in parties (or party leaders), and therefore might not take into account (see Stone and Simas 2010; Stone 2017).

The next concern is how to translate the theoretical concept into empirically quantifiable units. I believe it is worth to consider how the main competing model to the spatial framework deals with this task, the 'valence politics model of party choice' (e.g., Clarke *et al.* 2004, 2009, 2011; Sanders *et al.* 2011; Whiteley *et al.* 2013). Core components of the model, which directly resulted from Stokes's critique on the spatial voting framework, are party leader images, problem-solving capacities, or party identification. This strand of literature arrives at empirical measures of valence by relying on mass survey responses to questions such as who is best able to handle the most important issue a country is facing, feelings about party leaders (e.g., competence, responsiveness, and trust), or performance evaluations. Candidate trait measures might best capture how valence is understood in Schofield's valence model. However, by relying on performance evaluations, one might argue that such considerations are contrary to the

---

16 Contrasting the Log-Likelihood of the constant-only with the full model, we observe a large reduction from -928.43 (AIC = 1864.86, BIC = 1883.15) to -634.32 (AIC = 1332.64, BIC = 1478.95), whereas the policy-only model delivers only a value of -722.08 (AIC = 1460.17, BIC = 1496.75). The estimation table of the policy-only model can be found in Section B in the Supplementary Materials.

prospective logic and the concept of expected utility within the spatial voting paradigm. An additional difficulty arises when we operate in a multiparty competition environment where such valence measures are necessary for several parties. Another critical challenge is how to empirically separate valence factors from spatial considerations which might be correlated (see, e.g., Feldman and Conover 1983; Granberg and Brown 1992). Spatial and nonspatial considerations might not be separated by relying on survey data so that valence judgments independent from individual perceptions are necessary (see, e.g., Stone 2017).

6 Concluding Remarks

Based on a discussion of central properties of discrete choice models and the specific nature of the estimates related to party intercepts within this framework, my objective here was to demonstrate why and how empirical applications of Schofield's formal valence model, which is based on the ordering the party intercepts, do not yield unique results in fully specified voter models. Arbitrary changes in the coding of voter attributes affect the valence ranking and therefore, the conclusions we would draw from the model. It is impossible to depict valence with the constants and to ascertain something substantial from ordering absolute constant estimates. I recommend that researchers do not rely on party intercepts, but to make the valence qualities of parties (or party leaders) measurable instead and include them into voter utility expressions by observed variables. Future research should seek to quantify valence in a way that is in accord with the decision-theoretical logic of spatial voting.

Data availability

Supplementary materials for this article are available on the *Political Analysis* website. For Dataverse replication materials, see Mauerer (2019).

Supplementary material

For supplementary material accompanying this paper, please visit
https://doi.org/10.1017/pan.2019.41.

References

Adams, J. 1999. "Multiparty Spatial Competition with Probabilistic Voting." *Public Choice* 99(3):259–274.
Adams, J., L. Ezrow, S. Merrill III, and Z. Somer-Topcu. 2013. "Does Collective Responsibility for Performance Alter Party Strategies? Policy-Seeking Parties in Proportional Systems." *British Journal of Political Science* 43(1):1–23.
Adams, J., and S. Merrill III. 1999a. "Modeling Party Strategies and Policy Representation in Multiparty Elections: Why Are Strategies so Extreme?" *American Journal of Political Science* 43(3):765–791.
Adams, J., and S. Merrill III. 1999b. "Party Policy Equilibrium for Alternative Spatial Voting Models: An Application to the Norwegian Storting." *European Journal of Political Research* 36(2):235–255.
Adams, J., and S. Merrill III. 2000. "Spatial Models of Candidate Competition and the 1988 French Presidential Election: Are Presidential Candidates Vote-Maximizers?" *The Journal of Politics* 62(3):729–756.
Adams, J., and S. Merrill III. 2009. "Policy-Seeking Parties in a Parliamentary Democracy with Proportional Representation: A Valence-Uncertainty Model." *British Journal of Political Science* 39(3):539–558.
Adams, J., S. Merrill III, and B. Grofman. 2005. *A Unified Theory of Party Competition: A Cross-national Analysis Integrating Spatial and Behavioral Factors*. New York: Cambridge University Press.

Adams, J., S. Merrill III, E. N. Simas, and W. J. Stone. 2011. "When Candidates Value Good Character: A Spatial Model with Applications to Congressional Elections." *The Journal of Politics* 73(1):17–30.

Alvarez, R. M., and J. Nagler. 1995. "Economics, Issues and the Perot Candidacy: Voter Choice in the 1992 Presidential Election." *American Journal of Political Science* 39(3):714–744.

Alvarez, R. M., and J. Nagler. 1998. "When Politics and Models Collide: Estimating Models of Multiparty Elections." *American Journal of Political Science* 42(1):55–96.

Alvarez, R. M., J. Nagler, and S. Bowler. 2000. "Issues, Economics, and the Dynamics of Multiparty Elections: The British 1987 General Election." *American Political Science Review* 94(1):131–149.

Ansolabehere, S., and J. M. Snyder Jr. 2000. "Valence Politics and Equilibrium in Spatial Election Models." *Public Choice* 103(3):327–336.

Ansolabehere, S., and M. S. Puy. 2018. "Measuring Issue-Salience in Voters' Preferences." *Electoral Studies* 51:103–114.

Aragonès, E., and T. R. Palfrey. 2002. "Mixed Equilibrium in a Downsian Model with a Favored Candidate." *Journal of Economic Theory* 103(1):131–161.

Ashworth, S., and E. B. de Mesquita. 2009. "Elections with Platform and Valence Competition." *Games and Economic Behavior* 67(1):191–216.

Burden, B. C. 1997. "Deterministic and Probabilistic Voting Models." *American Journal of Political Science* 41(4):1150–1169.

Buttice, M. K., and W. J. Stone. 2012. "Candidates Matter: Policy and Quality Differences in Congressional Elections." *The Journal of Politics* 74(3):870–887.

Clarke, H. D., D. Sanders, M. C. Stewart, and P. Whiteley. 2004. *Political Choice in Britain*. Oxford: Oxford University Press.

Clarke, H. D., D. Sanders, M. C. Stewart, and P. Whiteley. 2009. *Performance Politics and the British Voter*. New York: Cambridge University Press.

Clarke, H. D., D. Sanders, M. C. Stewart, and P. Whiteley. 2011. "Valence Politics and Electoral Choice in Britain, 2010." *Journal of Elections, Public Opinion and Parties* 21(2):237–253.

Coughlin, P. J. 1992. *Probabilistic Voting Theory*. New York: Cambridge University Press.

Davis, O. A., M. J. Hinich, and P. C. Ordeshook. 1970. "An Expository Development of a Mathematical Model of the Electoral Process." *The American Political Science Review* 64(2):426–448.

Dewan, T., and K. A. Shepsle. 2011. "Political Economy Models of Elections." *Annual Review of Political Science* 14(1):311–330.

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

Endersby, J. W., and S. E. Galatas. 1998. "British Parties and Spatial Competition: Dimensions of Party Evaluation in the 1992 Election." *Public Choice* 97(3):363–382.

Enelow, J. M., and M. J. Hinich. 1982. "Nonspatial Candidate Characteristics and Electoral Competition." *The Journal of Politics* 44(1):115–130.

Enelow, J. M., and M. J. Hinich. 1989. "A General Probabilistic Spatial Theory of Elections." *Public Choice* 61(2):101–113.

Enelow, J. M., and M. J. Hinich. 1990. *Advances in the Spatial Theory of Voting*. Cambridge: Cambridge University Press.

Evrenk, H. 2019. "Valence Politics." In *The Oxford Handbook of Public Choice*, edited by R. D. Congleton, B. Grofman, and S. Voigt, 266–291. New York: Oxford University Press.

Falter, J. W., O. W. Gabriel, and H. Rattinger. 2012 "Political Attitudes, Political Participation and Voter Conduct in United Germany 1998." doi:10.4232/1.11461, ZA3066, Data file Version 3.0.0. Cologne, Germany: GESIS Data Archive.

Feldman, S., and P. J. Conover. 1983. "Candidates, Issues and Voters: The Role of Inference in Political Perception." *The Journal of Politics* 45(4):810–839.

Franchino, F., and F. Zucchini. 2015. "Voting in a Multi-dimensional Space: A Conjoint Analysis Employing Valence and Ideology Attributes of Candidates." *Political Science Research and Methods* 3(2):221–241.

Gallego, M., and N. Schofield. 2016. "Do Parties Converge to the Electoral Mean in All Political Systems?" *Journal of Theoretical Politics* 28(2):288–330.

Granberg, D., and T. A. Brown. 1992. "The Perception of Ideological Distance." *The Western Political Quarterly* 45(3):727–750.

Green, J., and W. Jennings. 2017. "Valence." In *The SAGE Handbook of Electoral Behavior*, edited by K. Arzheimer, J. Evans, and M. S. Lewis-Beck, 538–560. Los Angeles, CA: Sage.

Grofman, B. 2004. "Downs and Two-Party Convergence." *Annual Review of Political Science* 7(1):25–46.

Groseclose, T. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science* 45(4):862–886.

Hensher, D. A., J. M. Rose, and W. H. Greene. 2015. *Applied Choice Analysis*. 2nd edn. Cambridge: Cambridge University Press.

Jackson, J. E. 2014. "Location, Location, Location: The Davis-Hinich Model of Electoral Competition." *Public Choice* 159(1):197–218.

Kedar, O. 2005. "How Diffusion of Power in Parliaments Affects Voter Choice." *Political Analysis* 13(4):410–429.

Kurella, A., and F. U. Pappi. 2015. "Combining Ideological and Policy Distances with Valence for a Model of Party Competition in Germany 2009." *Journal of Theoretical Politics* 27(1):86–107.

Lacy, D., and B. C. Burden. 1999. "The Vote-Stealing and Turnout Effects of Ross Perot in the 1992 U.S. Presidential Election." *American Journal of Political Science* 43(1):233–255.

Lin, T., J. M. Enelow, and H. Dorussen. 1999. "Equilibrium in Multicandidate Probabilistic Spatial Voting." *Public Choice* 98(1):59–82.

Louviere, J. J., D. A. Hensher, and J. D. Swait. 2009. *Stated Choice Methods: Analysis and Applications*. Cambridge: Cambridge University Press.

Manski, C. F. 1977. "The Structure of Random Utility Models." *Theory and Decision* 8(3):229–254.

Mauerer, I. 2016. "A Party-Varying Model of Issue Voting. A Cross-National Study." Doctoral Dissertation, University of Munich (LMU), Germany.

Mauerer, I. 2019. "Replication Data for: The Neglected Role and Variability of Party Intercepts in the Spatial Valence Approach." https://doi.org/10.7910/DVN/PIV1KX, Harvard Dataverse, V1.

Mauerer, I., P. W. Thurner, and M. Debus. 2015. "Under Which Conditions Do Parties Attract Voters' Reactions to Issues? Party-Varying Issue Voting in German Elections 1987–2009." *West European Politics* 38(6):1251–1273.

Mauerer, I., W. Pößnecker, P. W. Thurner, and G. Tutz. 2015. "Modeling Electoral Choices in Multiparty Systems with High-Dimensional Data: A Regularized Selection of Parameters Using the Lasso Approach." *Journal of Choice Modelling* 16(3):23–42.

McFadden, D. L. 1974. "Conditional Logit Analysis of Qualitative Choice Behaviour." In *Frontiers in Econometrics*, edited by P. Zarembka, 105–142. New York: Academic Press.

Merrill, S. III, and J. Adams. 2001. "Computing Nash Equilibria in Probabilistic, Multiparty Spatial Models with Nonpolicy Components." *Political Analysis* 9(4):347–361.

Quinn, K. M., A. D. Martin, and A. B. Whitford. 1999. "Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models." *American Journal of Political Science* 43(4):1231–1247.

Sanders, D., H. D. Clarke, M. C. Stewart, and P. Whiteley. 2011. "Downs, Stokes and the Dynamics of Electoral Choice." *British Journal of Political Science* 41(2):287–314.

Schofield, N. 2003. "Valence Competition in the Spatial Stochastic Model." *Journal of Theoretical Politics* 15(4):371–383.

Schofield, N. 2004. "Equilibrium in the Spatial 'Valence' Model of Politics." *Journal of Theoretical Politics* 16(4):447–481.

Schofield, N. 2005. "A Valence Model of Political Competition in Britain: 1992–1997." *Electoral Studies* 24(3):347–370.

Schofield, N., and A.-S. Kurella. 2015. "Party Activists in the 2009 German Federal Elections." In *The Political Economy of Governance: Institutions, Political Performance and Elections*, edited by N. Schofield and G. Caballero, 293–311. Cham: Springer International Publishing.

Schofield, N., and A. Zakharov. 2010. "A Stochastic Model of the 2007 Russian Duma Election." *Public Choice* 142(1):177–194.

Schofield, N., and G. Miller. 2007. "Elections and Activist Coalitions in the United States." *American Journal of Political Science* 51(3):518–531.

Schofield, N., G. Miller, and A. D. Martin. 2003. "Critical Elections and Political Realignments in the USA: 1860–2000." *Political Studies* 51(2):217–240.

Schofield, N., and I. Sened. 2005a. "Modeling the Interaction of Parties, Activists and Voters: Why Is the Political Center So Empty?" *European Journal of Political Research* 44(3):355–390.

Schofield, N., and I. Sened. 2005b. "Multiparty Competition in Israel, 1988–96." *British Journal of Political Science* 35(4):635–663.

Schofield, N., and I. Sened. 2006. "Multiparty Democracy. Elections and Legislative Politics." In *Political Economy of Institutions and Decisions*, New York: Cambridge University Press.

Schofield, N., M. Gallego, and J. S. Jeon. 2011. "Leaders, Voters and Activists in the Elections in Great Britain 2005 and 2010." *Electoral Studies* 30(3):484–496.

Serra, G. 2010. "Polarization of What? A Model of Elections with Endogenous Valence." *The Journal of Politics* 72(2):426–437.

Stoetzer, L. F., and S. Zittlau. 2015. "Multidimensional Spatial Voting with Non-separable Preferences." *Political Analysis* 23(3):415–428.

Stokes, D. E. 1963. "Spatial Models of Party Competition." *The American Political Science Review* 57(2):368–377.

Stokes, D. E. 1992. "Valence Politics." In *Electoral Politics*, edited by D. Kavanagh, 141–164. Oxford: Clarendon Press.

Stone, W. J. 2017. *Candidates and Voters: Ideology, Valence, and Representation in U.S. Elections*. Cambridge: Cambridge University Press.

Stone, W. J., and E. N. Simas. 2010. "Candidate Valence and Ideological Positions in U.S. House Elections." *American Journal of Political Science* 54(2):371–388.

Stone, W. J., L. S. Maisel, and C. D. Maestas. 2004. "Quality Counts: Extending the Strategic Politician Model of Incumbent Deterrence." *American Journal of Political Science* 48(3):479–495.

Thurner, P. W. 2000. "The Empirical Application of the Spatial Theory of Voting in Multiparty Systems with Random Utility Models." *Electoral Studies* 19(4):493–517.

Train, K. E. 2009. *Discrete Choice Methods with Simulation*. 2nd edn. New York: Cambridge University Press.

Whiteley, P., H. D. Clarke, D. Sanders, and M. C. Stewart. 2013. *Affluence, Austerity and Electoral Change in Britain*. Cambridge: Cambridge University Press.