

Equilibria in the spatial stochastic model of voting with party activists

Norman Schofield

Published online: 30 September 2006
© Springer-Verlag 2006

Abstract Stochastic models of elections typically indicate that all parties, in equilibrium, will adopt positions at the electoral center. Empirical analyses discussed in this paper suggest that convergence of this kind is rarely observed. Here we examine a stochastic electoral model where parties differ in their valences – the electorally perceived, non-policy “quality” of the party leader. It is assumed that valence may either be exogenous, in the sense of being an intrinsic characteristic of the leader, or may be due to the contributions of party activists, who donate time and money and thus enhance electoral support for the party. Theorem 1 shows that vote maximization depends on balancing these two opposed effects. Theorem 2 provides the necessary and sufficient conditions for convergence to the electoral mean when activist valence is zero. The paper then examines empirical electoral models for the Netherlands circa 1980 and Britain in 1979, 1992 and 1997 and shows that party divergence from the electoral mean cannot be accounted for by exogenous valence alone. The balance condition suggests that the success of the Labour party in the election of 1997 can be attributed to a combination of high exogenous valence and pro-Europe activist support.

Keywords Local Nash equilibrium · Stochastic electoral model · Valence · Activists

JEL Classification Numbers C13 · C72 · D72

N. Schofield (✉)
Center in Political Economy, Washington University in Saint Louis,
1 Brookings Drive, St Louis, MO 63130, USA
e-mail: schofield.norman@gmail.com

1 The spatial model with valence

Two very different models of political strategy suggest that parties will tend to locate themselves at the electoral center. First, stochastic models based on vote maximization suggest that the electoral origin will be a Nash equilibrium.¹ Second, models of coalition bargaining when there is no majority party indicate that a large, centrally located party, at a “core” position in the policy space, will be dominant. Such a core party can, if it chooses, form a minority government by itself and control policy outcomes.² If party leaders are aware of the fact that they can control policy from the core, then this centripetal tendency should lead parties to position themselves at the center.

Yet, contrary to this intuition, there is ample empirical evidence that party leaders do not necessarily adopt centrist positions. For example, Budge et al. (1987) and Laver and Hunt (1992), in their study of European party manifestos, found no evidence of a strong centripetal tendency. The electoral models for Italy and Israel presented in Giannetti and Sened (2004) and Schofield and Sened (2006) estimated party positions in various ways, and concluded that there was no indication of policy convergence by parties. This paper re-examines the earlier empirical analyses for the Netherlands (Schofield et al. 1998; Quinn et al. 1999; Quinn and Martin 2002) and Britain (Schofield 2004, 2005; Schofield and Sened 2005a) to determine if the non-convergence noted previously can be accounted for by a stochastic electoral model that includes “valence” (Stokes 1992).

The aforementioned empirical models have all entailed the addition of heterogeneous intercept terms for each party. One interpretation of these intercept terms is that they are valences or party biases. “Valence” refers to voters’ judgments about positively or negatively evaluated aspects of candidates, or party leaders, which cannot be ascribed to the policy choice of the party. One may conceive of the valence that a voter ascribes to a party leader as a judgement of the leader’s quality or competence. This idea of valence has been utilized in a number of recent formal models of voting.³ Theorem 2 of this paper presents necessary and sufficient conditions for convergence to the electoral origin for the stochastic model when each party’s valence is regarded as an exogenous constant, independent of the party position. The empirical section of the paper considers elections in the Netherlands in 1977 and 1981 and in Britain in 1979 and shows that the estimated parameters of the model satisfy the sufficient condition for convergence. Since there is no evidence of convergence in these elections, the conflict between theory and evidence suggests that the stochastic model should be modified to provide a better explanation of party policy choice.

¹ Adams (1999a, 2001b), Adams and Merrill (1999), Lin et al. (1999), Banks and Duggan (2005), McKelvey and Patty (2006).

² See McKelvey and Schofield (1987), Schofield et al. (1989), Laver and Schofield (1990, 1998), Sened (1995), Banks and Duggan (2000), Schofield and Sened (2006).

³ Ansolabehere and Snyder (2000), Groseclose (2001), Aragonés and Palfrey (2002, 2005).

The paper therefore considers a more general valence model based on activist support for the parties (Aldrich 1983a,b; Aldrich and McGinnis 1989; Aldrich 1995). This activist valence model presupposes that party activists donate time and other resources to their party. Such resources allow a party to present itself more effectively to the electorate, thus increasing its valence. Since activists tend to be more radical than the average voter, parties are faced with a dilemma. By accommodating the political demands of activists, a party gains resources that it can use to enhance its valence, but by adopting the radical policies demanded by activists, the party may appear too extreme and lose electoral support. The party must therefore balance the electoral effect against the activist valence effect. Theorem 1 presents the requisite balance condition between electoral and activist support. Since valence in this model is affected by activist support, it may exhibit “decreasing returns to scale” and this may induce concavity in the vote share functions of the parties. Consequently, when the concavity of activists’ valence is sufficiently pronounced then a pure strategy Nash equilibrium (PNE) of the vote maximizing game will exist. However, Theorem 1 indicates that there is no reason for this equilibrium to be one where all parties adopt centrist positions.

In some polities, activists’ valence functions will be sufficiently concave so that only one PNE will exist. Recent analyses of elections in Israel have used simulation techniques to examine the nature of these equilibria (Schofield and Sened 2005b, 2006). Computation of PNE is extremely difficult and as a first step this paper concentrates instead on conditions for existence of “local pure strategy Nash equilibria” (LNE).

Theorem 1, in the next section of this paper, presents a characterization, of LNE for the stochastic electoral activist model, in terms of the Hessians of the vote share functions of the party leaders. Throughout it is assumed that the stochastic errors have the Type I extreme value (or log Weibull) distribution, Ψ . The formal model based on Ψ parallels the empirical models based on multinomial logit (MNL) estimation (Dow and Endersby 2004). Theorem 2 specializes to the simpler case when only exogenous valence is relevant, so that the activist valence functions are zero. For the case of fixed or exogenous valence, Theorem 2 shows that the model is classified by a “convergence coefficient” which is a function of all the parameters of the model. A sufficient condition for the existence of a convergent LNE at the electoral mean is that this coefficient is bounded above by 1. When the policy space is of dimension w , then the necessary condition for existence of a PNE at the electoral mean, and thus for the validity of the “mean voter theorem” (Hinich 1977; Lin et al. 1999), is that the coefficient is bounded above by w . It is shown that the convergence coefficient is (i) an increasing function of the maximum valence difference (ii) an increasing function of the spatial parameter, β , giving the relative importance of policy difference, and (iii) an increasing function of the electoral variance of the distribution of voter preferred points.

The empirical evidence from the Netherlands in 1977 and 1981, as well as Britain in 1979, indicates that the eigenvalues of the Hessians of the vote share functions at the joint electoral origin were all negative. In other words, the joint

origin was a LNE for the stochastic model with exogeneous valence. Clearly this is compatible with the mean voter theorem. This inference does not rule out the existence of other non-convergent LNE, but no other local equilibria were found by simulation. For a two dimension stochastic model of the 1997 election in Britain, it was found that the estimated position of the Conservative Party was incompatible with the results with exogenous valence. However, this model did provide an explanation for the position of the centrist Liberal Democrat Party.

The results of Theorem 2 with activist valence are then used to explain the changes in positions of the two larger parties in Britain between the elections of 1992 and 1997. Indeed, the empirical model suggests that as the exogenous valence of the Labour Party leaders increased in the 1990s, then the party's activists became less important. This provides an explanation why the party could become more centrist on the economic axis. On the other hand, as the valences of the leaders of the Conservative Party fell in the same period, then the influence on the party of anti-Europe activists increased. This suggests why the party adopted an anti-European Union position. While these observations are particular to Britain, they appear applicable to any polity such as the US, where activist support is important (Miller and Schofield 2003; Schofield et al. 2003).

The next section presents the formal model and statement of the Theorems. Section 3 gives the empirical applications, while Sect. 4 develops the idea of the balance solution when there is exogeneous valence and two or more opposed activist groups for each party. Proofs of the two theorems are given in the Appendix.

2 Local Nash equilibrium with activists and vote maximizing parties

The electoral model presented here is an extension of the multiparty stochastic model of Lin et al. (1999), modified by inducing asymmetries in terms of valence. The basis for this extension is the extensive empirical evidence that valence is a significant component of the judgements made by voters of party leaders. There are a number of possible choices for the appropriate game form for multiparty competition. The simplest one, which is used here, is that the utility function for party j , is proportional to the vote share, V_j , of the party. With this assumption, we can examine the conditions on the parameters of the stochastic model which are necessary for the existence of a pure strategy Nash equilibrium (PNE). Because the vote share functions are differentiable, we use calculus techniques to estimate optimal positions. We can then obtain sufficient conditions for the existence of local pure strategy Nash equilibria (LNE). Clearly, any PNE will be a LNE, but not conversely. Additional conditions of concavity or quasi-concavity are sufficient to guarantee existence of PNE.

The key idea underlying the formal model is that party leaders attempt to estimate the electoral effects of party declarations, or manifestos, and choose their own positions as best responses to other party declarations, in order to maximize

their own vote share. The stochastic model essentially assumes that party leaders cannot predict vote response precisely, but can estimate an expected vote share. In the model with valence, the stochastic element is associated with the weight given by each voter, i , to the average perceived quality or valence of the party leader.

Definition 1 (The stochastic vote model $M(\lambda, \mu, \beta; \Psi)$ with activist valence) *The data of the spatial model is a distribution, $\{x_i \in X\}_{i \in N}$, of voter ideal points for the members of the electorate, N , of size n . We assume that X is a compact convex subset of Euclidean space, \mathbb{R}^w , with w finite. Each of the parties in the set $P = \{1, \dots, j, \dots, p\}$ chooses a policy, $z_j \in X$, to declare. Let $\mathbf{z} = (z_1, \dots, z_p) \in X^p$ be a typical vector of party policy positions.*

Given \mathbf{z} , each voter, i , is described by a vector

$$\mathbf{u}_i(x_i, \mathbf{z}) = (u_{i1}(x_i, z_1), \dots, u_{ip}(x_i, z_p))$$

where

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) - \beta \|x_i - z_j\|^2 + \epsilon_j = u_{ij}^*(x_i, z_j) + \epsilon_j. \tag{1}$$

Here $u_{ij}^*(x_i, z_j)$ is the observable component of utility. The term, λ_j , is the fixed or exogenous valence of agent j , while the function $\mu_j(z_j)$ is the component of valence generated by activist contributions to agent j . The term β is a positive constant, called the spatial parameter, giving the importance of policy difference defined in terms of the Euclidean norm, $\|\cdot\|$, on X . The vector $\epsilon = (\epsilon_1, \dots, \epsilon_j, \dots, \epsilon_p)$ is the stochastic error, whose multivariate cumulative distribution will be denoted by Ψ .

It is assumed that the exogenous valence vector

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_p) \text{ satisfies } \lambda_p \geq \lambda_{p-1} \geq \dots \geq \lambda_2 \geq \lambda_1.$$

Voter behavior is modeled by a probability vector. The probability that a voter i chooses party j at the vector \mathbf{z} is

$$\rho_{ij}(\mathbf{z}) = \Pr[[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)], \text{ for all } l \neq j]. \tag{2}$$

$$= \Pr[\epsilon_l - \epsilon_j < u_{ij}^*(x_i, z_j) - u_{il}^*(x_i, z_j), \text{ for all } l \neq j] \tag{3}$$

Here \Pr stands for the probability operator generated by the distribution assumption on ϵ . The expected vote share of agent j is

$$V_j(\mathbf{z}) = \frac{1}{n} \sum_{i \in N} \rho_{ij}(\mathbf{z}) \tag{4}$$

The differentiable function $V : X^p \rightarrow \mathbb{R}^p$ is called the party profile function.

The most common assumption in empirical analyses is that Ψ is the *Type I extreme value distribution* (sometimes called log Weibull). The theorems in this paper are based on this assumption.

Definition 2 (The extreme value distribution, Ψ) *The cumulative distribution, Ψ , has the closed form*

$$\Psi(h) = \exp[-\exp[-h]],$$

with probability density function

$$\psi(h) = \exp[-h] \exp[-\exp[-h]]$$

and variance $\frac{1}{6}\pi^2$.

With this assumption it follows, for each voter i , and party j , that

$$\rho_{ij}(\mathbf{z}) = \frac{\exp[u_{ij}^*(x_i, z_j)]}{\sum_{k=1}^p \exp u_{ik}^*(x_i, z_k)} \tag{5}$$

This implies that the model satisfies the independence of irrelevant alternative property (IIA) : for each individual i , and each pair, j, k , the ratio

$$\frac{\rho_{ij}(\mathbf{z})}{\rho_{ik}(\mathbf{z})}$$

is independent of a third party l .

(See Train 2003, p. 79).

In this stochastic electoral model it is assumed that each party j chooses z_j to maximize V_j , conditional on $\mathbf{z}_{-j} = (z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_p)$.

Definition 3 (Equilibrium concepts)

- (i) *A strategy vector $\mathbf{z}^* = (z_1^*, \dots, z_{j-1}^*, z_j^*, z_{j+1}^*, \dots, z_p^*) \in X^p$ is a local strict Nash equilibrium (LSNE) for the profile function $V : X^p \rightarrow \mathbb{R}^p$ iff, for each party $j \in P$, there exists a neighborhood X_j of z_j^* in X such that*

$$V_j(z_1^*, \dots, z_{j-1}^*, z_j^*, z_{j+1}^*, \dots, z_p^*) > V_j(z_1^*, \dots, z_j, \dots, z_p^*) \quad \text{for all } z_j \in X_j - \{z_j^*\}.$$

- (ii) *A strategy vector $\mathbf{z}^* = (z_1^*, \dots, z_{j-1}^*, z_j^*, z_{j+1}^*, \dots, z_p^*)$ is a local weak Nash equilibrium (LNE) iff, for each agent j , there exists a neighborhood X_j of z_j^* in X such that*

$$V_j(z_1^*, \dots, z_{j-1}^*, z_j^*, z_{j+1}^*, \dots, z_p^*) \geq V_j(z_1^*, \dots, z_j, \dots, z_p^*) \quad \text{for all } z_j \in X_j.$$

- (iii) A strategy vector $\mathbf{z}^*=(z_1^*, \dots, z_{j-1}^*, z_j^*, z_{j+1}^*, \dots, z_p^*)$ is a strict, respectively, weak, pure strategy Nash equilibrium (PSNE, respectively, PNE) iff X_j can be replaced by X in (i), (ii) respectively.
- (iv) The strategy z_j^* is termed a “local strict best response”, a “local weak best response”, a “global weak best response”, a “global strict best response”, respectively to $\mathbf{z}_{-j}^*=(z_1^*, \dots, z_{j-1}^*, z_{j+1}^*, \dots, z_p^*)$.

Obviously if \mathbf{z}^* is an LSNE or a PNE it must be an LNE, while if it is a PSNE then it must be an LSNE. We use the notion of LSNE to avoid problems with the degenerate situation when there is a zero eigenvalue to the Hessian. The weaker requirement of LNE allows us to obtain a necessary condition for \mathbf{z}^* to be a LNE and thus a PNE, without having to invoke concavity. Of particular interest is the joint mean vector

$$x^* = \frac{1}{n} \sum_{i \in N} x_i. \tag{6}$$

We first transform coordinates so that in the new coordinate system, $x^* = 0$. We shall refer to $\mathbf{z}_0 = (0, \dots, 0)$ as the *joint origin*.

Theorem 1 shows that $\mathbf{z}_0 = (0, \dots, 0)$ will generally not satisfy the first order condition for a LSNE, namely that the differential of V_j , with respect to z_j be zero. However, if the activist valence function is identically zero, so that only exogenous valence is relevant, then the first order condition will be satisfied. On the other hand, Theorem 2 shows that there are necessary and sufficient conditions for \mathbf{z}_0 to be an LSNE. A corollary of Theorem 2 gives these condition in terms of a “convergence coefficient” determined by the Hessian of party 1, with the lowest valence.

It follows from Eq. (5) that for voter i , with ideal point, x_i , the probability, $\rho_{ij}(\mathbf{z})$, that i picks j at \mathbf{z} is given by

$$\rho_{ij}(\mathbf{z}) = \left[1 + \sum_{k \neq j} \exp(f_k) \right]^{-1} \tag{7}$$

where $f_k = \lambda_k + \mu_k(z_k) - \lambda_j - \mu_j(z_j) + \beta \|x_i - z_j\|^2 - \beta \|x_i - z_k\|^2$

The Appendix uses Eq. (7) to show that the first order condition for \mathbf{z}^* to be a LSNE is that it be a *balance solution*.

Definition 4 (The balance solution for the model $M(\lambda, \mu, \beta; \Psi)$)

Let $[\rho_{ij}(\mathbf{z})] = [\rho_{ij}]$ be the matrix of voter probabilities at the vector \mathbf{z} , and let $[\alpha_{ij}] = \frac{\rho_{ij} - \rho_{ij}^2}{\sum_i (\rho_{ij} - \rho_{ij}^2)}$. The balance equation for z_j^* is given by expression

$$z_j^* = \frac{1}{2\beta} \frac{d\mu_j}{dz_j} + \sum_{i=1}^n \alpha_{ij} x_i. \tag{8}$$

The vector $\sum_i \alpha_{ij}x_i$ is called the weighted electoral mean for party j , and can be written

$$\sum_i \alpha_{ij}x_i = \frac{d\mathcal{E}_j^*}{dz_j}. \tag{9}$$

The balance equation can then be rewritten as

$$\left[\frac{d\mathcal{E}_j^*}{dz_j} - z_j^* \right] + \frac{1}{2\beta} \frac{d\mu_j}{dz_j} = 0. \tag{10}$$

The bracketed term on the left of this expression is termed the marginal electoral pull of party j and is a gradient vector pointing towards the weighted electoral mean. This weighted electoral mean is that point where the electoral pull is zero. The vector $\frac{d\mu_j}{dz_j}$ is called the marginal activist pull for party j .

If z_j^* satisfies the balance equation for all j , then call \mathbf{z}^* the balance solution.

In the case $\mu_j = 0$ for all j , then the Appendix shows that for each fixed j , all α_{ij} are identical. Thus, when there is only exogenous valence, the the balance condition gives

$$z_j^* = \frac{1}{n} \sum x_i. \tag{11}$$

By a change of coordinates we can choose $\frac{1}{n} \sum x_i = 0$. In this case, the marginal electoral pull is zero at the origin and the joint origin $\mathbf{z}_0 = (0, \dots, 0)$ satisfies the first order condition. Theorem 1 sums up the results of the Appendix.

Theorem 1 Consider the electoral model $M(\lambda, \mu, \beta; \Psi)$ based on the Type I extreme value distribution, and including both exogenous and activist valences. The first order condition for \mathbf{z}^* to be an LSNE is that it is a balance solution. If all activist valence functions are highly concave, in the sense of having negative eigenvalues of sufficiently great magnitude, then the balance solution will be a PNE.

When the valence functions $\{\mu_j\}$ are non zero, then it is the case that generically \mathbf{z}_0 cannot satisfy the first order condition. Instead the vector $\frac{d\mu_j}{dz_j}$ “points towards” the position at which the activist valence is maximized. When this marginal or gradient vector, $\frac{d\mu_j}{dz_j}$, is increased, (if activists become more willing to contribute to the party) then the equilibrium position is pulled away from the weighted electoral mean of party j , and we can say the “activist effect” for the party is increased. On the other hand if the activist valence functions are fixed, but the exogeneous valence, λ_j , is increased, or the exogenous valence

terms $\{\lambda_k : k \neq j\}$ are decreased, then the vector $\frac{d\mathcal{E}_j^*}{dz_j}$ increases in magnitude, and the equilibrium is pulled towards the weighted electoral mean. We can say the “electoral effect” is increased

The second order condition for an LSNE at z^* depends on the negative definiteness of the Hessian of the activist valence function. If the eigenvalues of these Hessians are negative at a balance solution, and of sufficient magnitude, then this will guarantee that a vector z^* that satisfies the balance condition will be a LSNE. Indeed, this condition can ensure concavity of the vote share functions, and thus of existence of a PNE.

We can use the proof technique of the Appendix to develop the necessary and sufficient condition for an LSNE when activist valence is zero.

To characterize the variation in voter preferences, we represent in a simple form the covariation matrix (or data matrix), ∇_0 , given by the distribution of voter ideal points.

Definition 5 (The electoral covariance matrix, ∇_0^*) *Let $X = \mathbb{R}^w$ be endowed with a system of coordinate axes $(1, \dots, r, s, \dots, w)$. For each coordinate axis let $\xi_r = (x_{1r}, x_{2r}, \dots, x_{nr})$ be the vector of the r th coordinates of the set of n voter ideal points. The scalar product of ξ_r and ξ_s is denoted by (ξ_r, ξ_s) .*

The symmetric $w \times w$ voter covariation matrix about the origin is denoted ∇_0 and is defined by

$$\nabla_0 = \begin{pmatrix} (\xi_1, \xi_1) & & (\xi_1, \xi_w) \\ & (\xi_r, \xi_r) & \\ (\xi_w, \xi_1) & & (\xi_w, \xi_w) \end{pmatrix}.$$

The covariance matrix ∇_0^ is defined to be $\frac{1}{n}\nabla_0$.*

We write $v_s^2 = \frac{1}{n}(\xi_s, \xi_s)$ for the electoral variance on the s th axis and

$$v^2 = \sum_{s=1}^w v_s^2 = \frac{1}{n} \sum_{s=1}^w (\xi_s, \xi_s) = \text{trace}(\nabla_0^*)$$

for the total electoral variance. The electoral covariance between the r th and s th axes is $(v_r, v_s) = \frac{1}{n}(\xi_r, \xi_s)$.

In precisely the same way, if $z \in X$, then define ∇_z^ to be the covariance matrix about z .*

In the case that all activist valence functions $\{\mu_j\}$ are identically zero, we write the electoral model as $M(\lambda, \beta; \Psi)$.

Definition 6 (The convergence coefficient of the model $M(\lambda, \beta; \Psi)$)

- (i) At the vector $\mathbf{z}_0 = (0, \dots, 0)$ the probability $\rho_{ij}(\mathbf{z}_0)$ that i votes for party, j is independent of i , and is given by

$$\rho_j = \left[1 + \sum_{k \neq j} \exp[\lambda_k - \lambda_j] \right]^{-1} \tag{12}$$

- (ii) The coefficient A_j for party j is

$$A_j = \beta(1 - 2\rho_j)$$

- (iii) The characteristic matrix for party j is

$$C_j = [2A_j \nabla_0^* - I] \tag{13}$$

where I is the w by w identity matrix.

- (iv) The convergence coefficient of the model $M(\lambda, \beta; \Psi)$ is

$$c(\lambda, \beta; \Psi) = 2\beta[1 - 2\rho_1]v^2 = 2A_1v^2. \tag{14}$$

At the vector $\mathbf{z}_0 = (0, \dots, 0)$ the probability $\rho_{ij}(\mathbf{z}_0)$ that i votes for party, j is independent of i , and is given by Definition 6(i). Obviously if all valences are identical then $\rho_1 = \frac{1}{p}$, as expected. The effect of increasing λ_j , for $j \neq 1$, is clearly to decrease ρ_1 , and therefore to increase A_1 , and thus $c(\lambda, \beta; \Psi)$.

Theorem 2 *The necessary condition for the joint origin to be a LSNE in the model $M(\lambda, \beta; \Psi)$ is that the characteristic matrix*

$$C_1 = [2A_1 \nabla_0^* - I]$$

of the party 1, with lowest valence, has negative eigenvalues.

Theorem 2 immediately gives the following Corollary (Schofield 2006).

Corollary *Consider the model $M(\lambda, \beta; \Psi)$. In the case that X is w -dimensional then the necessary condition for the joint origin to be a LNE is that $c(\lambda, \beta; \Psi) \leq w$. In the case that X is two-dimensional, a sufficient condition for the joint origin to be a LSNE is that $c(\lambda, \beta; \Psi) < 1$.*

In the following section we consider empirical models for the Netherlands and Britain. We show that the parameters of the models imply the sufficient condition for an LSNE at the joint origin was satisfied. Indeed the eigenvalues were sufficiently negative so as to imply that the joint origin was the unique PSNE. Since the parties did not appear to be positioned at the origin, we can infer either the parties were not maximizing vote share, or the activist valence functions were significant.

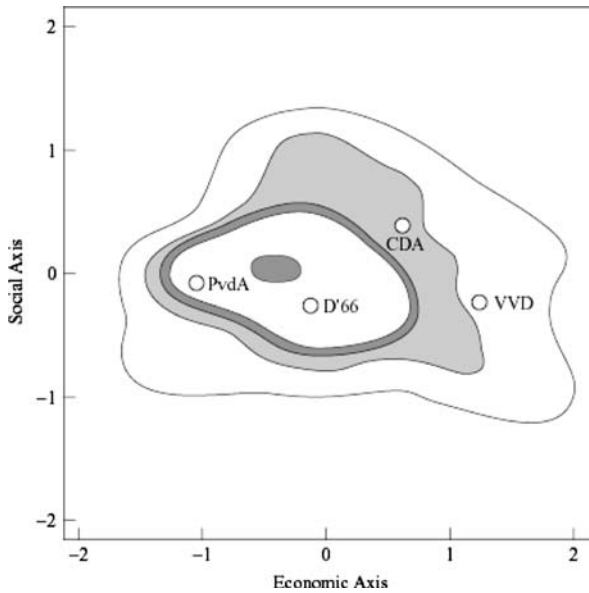


Fig. 1 Distribution of voter ideal points and party positions in the Netherlands in the 1981 Election, for a two-dimensional model, showing the highest density contours of the sample voter distribution at the 95, 75, 50, and 10% levels

3 Empirical analyses

3.1 The vote maximizing model in the Netherlands

First, we consider a multinomial logit (MNL) model for the elections of 1977 and 1981 in the Netherlands (Schofield et al. 1998; Quinn et al. 1999) using data from the middle level Elites Study (ISEIUM 1983) coupled with the Rabier and Inglehart (1981) Euro-barometer study. There are four main parties: Labour (PvdA), Christian Democratic Appeal (CDA), Liberals (VVD) and Democrats (D'66), with approximately 35%, 35%, 20% and 10% of the popular vote.

Figure 1 gives the estimated positions of the parties and the electoral distribution circa 1979.⁴

The empirical model estimated exogenous valences, which were normalized, by choosing the D'66 to have exogenous valence $\lambda_{d66} = 0$. The other valences are $\lambda_{vvd} = 1.015$, $\lambda_{cda} = 1.140$ and $\lambda_{pvda} = 1.596$. To compute the D'66 Hessian, we note that the electoral variance on the first axis is $v_1^2 = 0.658$, while on the second it is $v_2^2 = 0.289$. The covariance (v_1, v_2) is negligible.

⁴ This figure, as well as Figs. 2, 3 and 4 are taken from Schofield and Sened (2006) with permission of Cambridge University Press.

The spatial coefficient $\beta = 0.737$ for the model with exogenous valence. Thus the probability of voting for the D'66, when all parties are at the origin, is

$$\rho_{d66} = \frac{1}{1 + e^{1.596} + e^{1.403} + e^{1.015}} = 0.074.$$

Hence $A_{d66} = 0.737(0.852) = 0.627.$

$$C_{d66} = 2A_{con}\nabla_0^* - I = (1.25) \begin{pmatrix} 0.658 & 0 \\ 0 & 0.289 \end{pmatrix} - I = \begin{pmatrix} -0.18 & 0 \\ 0 & -0.64 \end{pmatrix}.$$

Thus $c(\Psi) = 1.187.$

Although the sufficient condition of the Corollary is not satisfied, the necessary condition is satisfied, and the eigenvalues can be seen to be negative. By Theorem 2, the joint origin is an LSNE for the stochastic model with exogenous valence. Quinn and Martin (2002) performed a simulation of the empirical model and showed that the positions given in Fig. 1 could not be an LSNE of the stochastic model with exogenous valence alone. This conflict between the predicted equilibrium positions of the model and the estimated positions suggest that the activists for the party played an important role in determining the party positions. We now examine this possibility in the case of Britain.

3.2 Elections in Britain in 1979 and 1992–1997

Figure 2 presents the estimated positions of of the three major parties in Britain at the election of 1979, in a two dimensional policy space.

Just as in the case of the Netherlands, the estimation used data from ISEIUM (1983) and the Rabier and Inglehart (1981) Euro-barometer study (see Quinn et al. 1999; Schofield 2005 for further details). The electoral variances were 0.605 on the first axis and 0.37 on the second, giving a total variance of 0.975. For the MNL model with exogenous valence, the spatial coefficient was $\beta = 0.27.$

The lowest valence was associated with the Conservative party, with $\lambda_{con} = -0.324.$ Since $\lambda_{lib} = 0.082, \lambda_{lab} = 0.0,$ and $\beta = 0.27,$ we find that

$$\rho_{con} = \frac{e^{-0.324}}{e^{-0.324} + e^0 + e^{0.082}} = \frac{0.723}{2.8} = 0.26.$$

Similarly, $\rho_{lib} = 0.38$ and $\rho_{lab} = 0.36.$ Thus $A_{con} = \beta(1 - 2\rho_{con}) = 0.13.$ The Hessian matrix for the Conservative Party is

$$C_{con} = 2A_{con}\nabla_0^* - I = (0.26) \begin{pmatrix} 0.605 & 0 \\ 0 & 0.370 \end{pmatrix} - I = \begin{pmatrix} -0.84 & 0 \\ 0 & -0.90 \end{pmatrix}.$$

Thus both eigenvalues are negative, and the convergence coefficient can be calculated to be 0.26. It follows that the joint origin is an LSNE. Simulation indicates that this LSNE was a PSNE, and that there existed no other equilibria.

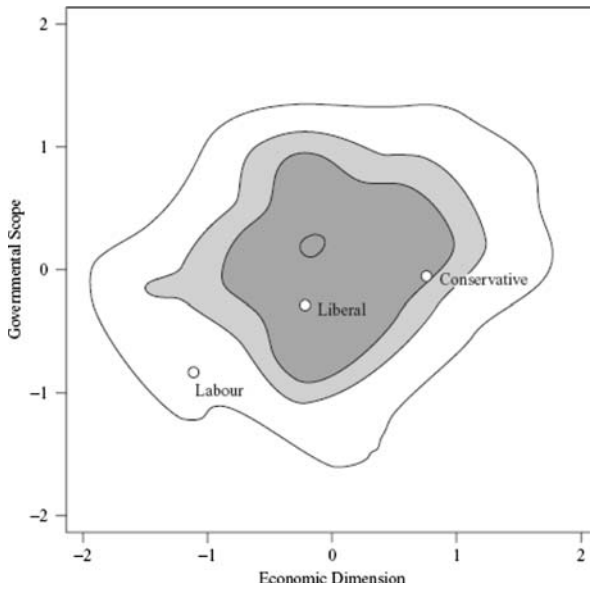


Fig. 2 Distribution of voter ideal points and party positions in Britain in the 1979 Election for a two-dimensional model, showing the highest density contours of the sample voter distribution at the 95, 75, 50, and 10% levels

This inference conflicts with the estimated positions of the parties given in Fig. 2, and suggests modification of the model to incorporate activists.

To pursue this possibility, Fig. 3 shows the estimated positions of the parties, based on a survey of Party MPs in 1997 (Schofield 2005). In addition to the Conservative Party (CONS), Labour Party (LAB) and Liberal Democrat Party (LIB) responses were obtained from Ulster Unionists (UU), Scottish Nationalists (SNP) and Plaid Cymru (PC). The first axis is economic, the second axis concerned attitudes to the European Union (pro to the “south” of the vertical axis). The electoral model with exogenous valence was estimated for the elections in 1992 and 1997. For 1992, we have $(\lambda_{con}, \lambda_{lab}, \lambda_{lib}, \beta)_{1992} = (+1.58, 0.58, 0.0, 0.56)$, so

$$\rho_{lib} = \frac{e^0}{e^0 + e^{1.58} + e^{0.58}} = \frac{1}{7.36} = 0.13.$$

For 1997, $(\lambda_{con}, \lambda_{lab}, \lambda_{lib}, \beta)_{1997} = (+1.24, 0.97, 0.0, 0.5)$ so

$$\rho_{lib} = \frac{e^0}{e^0 + e^{1.24} + e^{0.97}} = \frac{1}{7.08} = 0.14$$

$$A_{lib} = \beta(1 - 2\rho_{lib}) = 0.36.$$

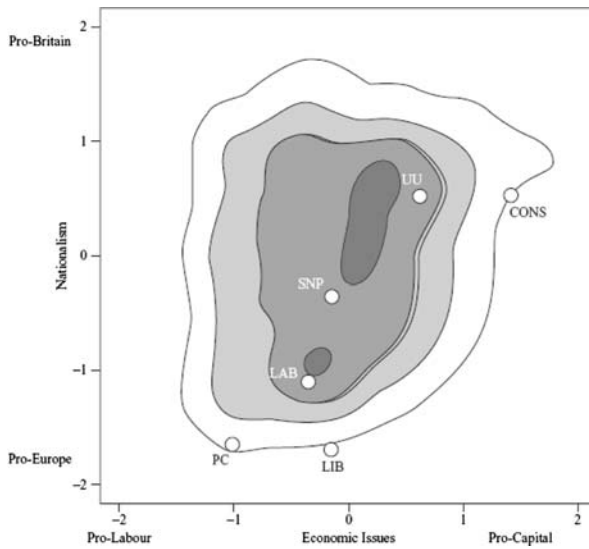


Fig. 3 Estimated party positions in the British Parliament for a two-dimensional model (based on MP survey data and the National Election Survey) showing highest density contours of the voter sample distribution at the 95, 75, 50, and 10% levels. Key: *PC* Plaid Cymru (Welsh Nationalist Party), *SNP* Scottish Nationalist Party, *UU* Ulster Unionists, *CONS* Conservative Party, *LAB* Labour Party, *LIB* Liberal Party

Since the electoral variance is 1.0 on the first economic axis and 1.5 on the European axis, we obtain

$$C_{lib} = (0.72) \begin{pmatrix} 1.0 & 0 \\ 0 & 1.5 \end{pmatrix} - I = \begin{pmatrix} -0.28 & 0 \\ 0 & +0.8 \end{pmatrix}.$$

The convergence coefficient can be calculated to be 1.8. Although the necessary condition is satisfied, the origin is clearly a saddlepoint for the Liberal Democrat Party. Note that the second “European” axis is a “principal electoral axis” exhibiting greater electoral variance. This axis is the eigenvector associated with the positive eigenvalue. Because the covariance between the two electoral axes is negligible, we can infer that, for each party, the eigenvalue of the Hessian at the origin is negative on the first or minor “economic” axis. According to the formal model with exogenous valence, all parties should have converged to the origin on this minor axis. Because the eigenvalue for the Liberal Democrat Party is positive on the second axis, we have an explanation for its position away from the origin on the Europe axis in Fig. 3. However there is no explanation for the location of the Conservative Party so far from the origin on both axes. Schofield (2005) offers a model (based on an earlier version of Theorem 1) where the falling exogenous valence of the Conservative Party leader increases the marginal importance of two opposed activist groups in the

party- one group “pro-capital” and one group “anti-Europe.” The concluding section comments on this observation.

4 Concluding remarks

The empirical analysis of the previous section showed that overall Conservative valence dropped from 1.58 in 1992 to 1.24 in 1997, while the Labour valence increased from 0.58 to 0.97. These estimated valences include both exogenous valence terms for the parties and the activist component. Nonetheless, the data presented in Clarke et al. (1995, 1997, 1998, 2004) and Seyd and Whiteley (1992, 2002) suggest that when Tony Blair took over from John Smith as leader of the Labour Party, then the exogenous valence, λ_{lab} , of the party increased up to the 1997 election. Conversely, the exogenous valence, λ_{con} , for the Conservatives fell. Since the coefficients in the equation for the electoral pull for the Conservative Party depend on $\lambda_{con} - \lambda_{lab}$, Theorem 1 implies that the effect would be to increase the marginal effect of activism for the Conservative party, thus pulling the optimal position away from the party’s weighted electoral mean. The opposite conclusion holds for the Labour Party, since increasing $\lambda_{lab} - \lambda_{con}$ has the effect of reducing the marginal activist effect.

Indeed, it is possible to include the effect of two potential activist groups for the Labour Party: one “pro-Europe,” centered at the position marked E in Figure 4 and one “pro-Labor,” marked L in the figure. The optimal Labour position will be determined by a version of the balance equation:

$$\left[\frac{d\mathcal{E}_{lab}^*}{dz_{lab}} - z_{lab}^* \right] + \frac{1}{2\beta} \left[\frac{d\mu_{lab,L}}{dz_{lab}} + \frac{d\mu_{lab,E}}{dz_{lab}} \right] = 0 \quad (15)$$

which equates the “electoral pull” against the two “activist pulls,” generated by the two different activist functions, $\mu_{lab,L}$ and $\mu_{lab,E}$. In the same way, if there are two activist groups for the Conservatives, generated by functions $\mu_{con,C}$ and $\mu_{con,B}$ centered at C and B respectively, then we obtain a balance equation:

$$\left[\frac{d\mathcal{E}_{con}^*}{dz_{con}} - z_{con}^* \right] + \frac{1}{2\beta} \left[\frac{d\mu_{con,C}}{dz_{con}} + \frac{d\mu_{con,B}}{dz_{con}} \right] = 0. \quad (16)$$

Since the electoral pull for the Conservative Party fell between the elections, the optimal position, z_{con}^* , will be one which is “closer” to the locus of points that generates the greatest activist support. This locus is where the joint marginal activist pull is zero. This locus of points can be called the “activist contract curve” for the Conservative party.

Miller and Schofield (2003) develop an activist model of this kind, where preferences of different activists on the two dimensions may accord different saliences to the policy axes. The “activist contract curves” for the two parties will be catenaries that depend on the ratios of the saliences that different activists have on the two dimensions. Figure 4 provides an illustration.

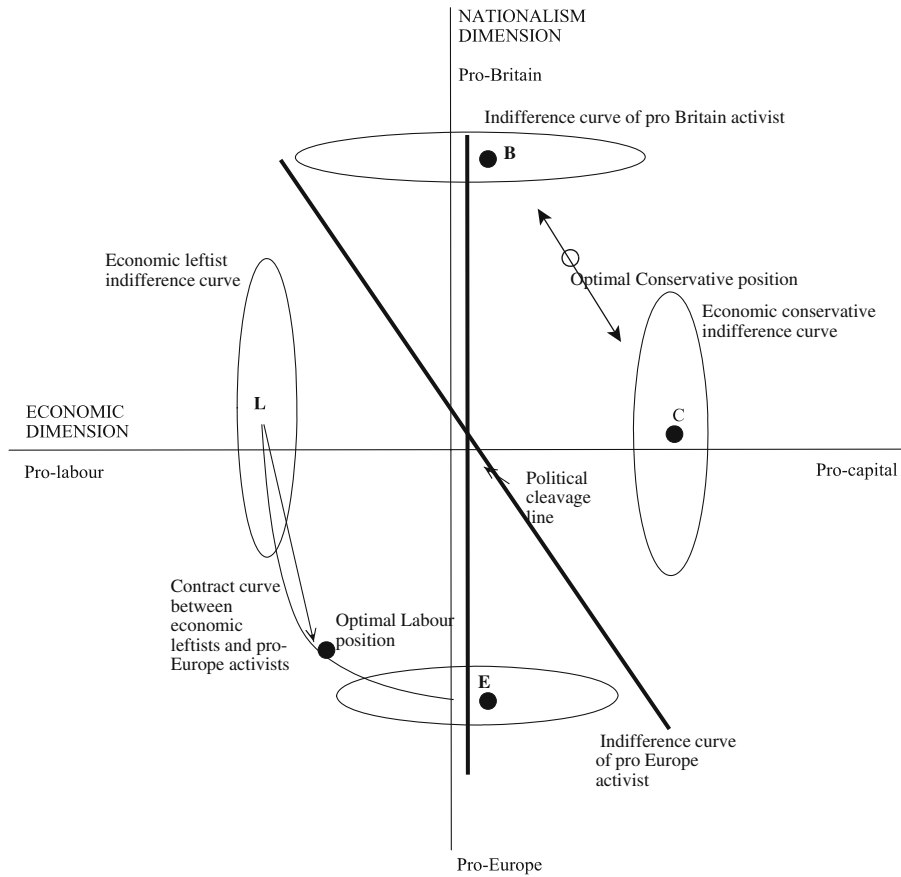


Fig. 4 Illustration of vote maximizing positions of Conservative and Labour Party leaders in a two dimensional policy space

According to Theorem 1, the reason the Labour party under Blair was able to move to a position closer to the origin between the elections of 1992 and 1997 was that his increasing valence reduced the importance of pro labour activists in the party. On the other hand, the declining valences of the Conservative Party leaders, first William Hague, and then Iain Duncan Smith, increased the importance of the marginal activist effect for the party. This appears to have the effect of obliging the party to move to the fairly extreme position shown in Fig. 3. It remains to be seen whether the new leader, David Cameron, can gain high enough valence to overcome the apparent dominant influence of anti-Europe activist sentiment in the party.

Acknowledgements This paper is based on research supported by NSF Grant SES 024173. I thank the special editor, Muhamet Yildiz, and an anonymous referee of the journal, for very helpful comments.

5 Appendix: Proof of the Theorems

5.1 Proof of Theorem 1

For the extreme value distribution Ψ we have

$$\rho_{ij}(\mathbf{z}) = \left[1 + \sum_{k \neq j} \exp(f_k) \right]^{-1}$$

where $f_k = \lambda_k + \mu_k(z_k) - \lambda_j - \mu_j(z_j) + \beta \|x_i - z_j\|^2 - \beta \|x_i - z_k\|^2$ is the comparison function used by i in evaluating party k in contrast to party j . We then obtain

$$\begin{aligned} \frac{d}{dz_j}[\rho_{ij}] &= - \left[2\beta(z_j - x_i) - \frac{d\mu_j}{dz_j} \right] \left[1 + \sum_{k \neq j} \exp(f_k) \right]^{-2} \left[\sum_k \exp(f_k) \right] \\ &= \left[2\beta(x_i - z_j) + \frac{d\mu_j}{dz_j} \right] [\rho_{ij}][1 - \rho_{ij}]. \end{aligned}$$

Thus the first order condition for maximizing V_j is

$$\begin{aligned} \sum_i \frac{d}{dz_j}[\rho_{ij}] &= \sum_i \left[2\beta(x_i - z_j) + \frac{d\mu_j}{dz_j} \right] [\rho_{ij}][1 - \rho_{ij}] = 0, \\ \text{or } \left[2\beta z_j - \frac{d\mu_j}{dz_j} \right] \sum_i [\rho_{ij}][1 - \rho_{ij}] &= \sum_i 2\beta x_i [\rho_{ij}][1 - \rho_{ij}], \\ \text{so } z_j - \frac{1}{2\beta} \frac{d\mu_j}{dz_j} &= \sum_i \alpha_{ij} x_i, \\ \text{where } \alpha_{ij} &= \frac{[\rho_{ij} - \rho_{ij}^2]}{\sum_i [\rho_{ij} - \rho_{ij}^2]}. \end{aligned} \tag{17}$$

An identical argument holds for each party j giving an equilibrium at a weighted electoral mean satisfying, for all j , the balance equation:

$$\left[\frac{d\mathcal{E}_j^*}{dz_j} - z_j^* \right] + \frac{1}{2\beta} \frac{d\mu_j}{dz_j} = 0. \tag{18}$$

where $\frac{d\mathcal{E}_j^*}{dz_j} = \sum_i \alpha_{ij} x_i$. This gives the first order condition stated in Theorem 1. Let \mathbf{z}^* be a vector satisfying the first order condition.

To examine the second order condition, note that now the Hessian of party j is given by

$$\frac{1}{n} \sum_i \frac{d^2 \rho_{ij}}{dz_j^2} \tag{19}$$

$$= \frac{1}{n} \sum_i [\rho_{ij} - \rho_{ij}^2] \left[4\beta^2 [1 - 2\rho_{ij}] [\nabla_{ij}] + \left[\frac{d^2 \mu_j}{dz_j^2} - 2\beta I \right] \right] \tag{20}$$

$$= \frac{1}{n} \left[\frac{d^2 \mu_j}{dz_j^2} - 2\beta I \right] \sum_i [\rho_{ij} - \rho_{ij}^2] + 4\beta^2 \sum_i [\rho_{ij} - \rho_{ij}^2] [1 - 2\rho_{ij}] [\nabla_{ij}^*]. \tag{21}$$

Here

$$[\nabla_{ij}] = \left[(x_i - z_j) + \frac{1}{2\beta} \frac{d\mu_j}{dz_j} \right]^T \left[(x_i - z_j) + \frac{1}{2\beta} \frac{d\mu_j}{dz_j} \right] \tag{22}$$

$$\text{and } [\nabla_{ij}^*] = \frac{1}{n} [\nabla_{ij}], \tag{23}$$

where T denotes a column vector. Since $z_j^* = \frac{1}{2\beta} \frac{d\mu_j}{dz_j} + \frac{d\varepsilon_j^*}{dz_j}$, we can regard the symmetric matrix expression in Eq. (21) involving $[\nabla_{ij}^*]$ as a measure of electoral variance taken about a weighted electoral mean. Even though this matrix term may have positive eigenvalues, if the eigenvalues of $\frac{d^2 \mu_j}{dz_j^2}$ are negative, and of sufficiently large modulus, then the Hessian will also have negative eigenvalues.

This gives a sufficient condition for existence of a LSNE at \mathbf{z}^* , and thus for a PSNE. □

5.2 Proof of Theorem 2 and the Corollary

At $\mathbf{z}^* = (0, \dots, 0)$, then by Eq. (17), we see that $\rho_{ij} = \rho_j$ is independent of i . Thus $\alpha_{ij} = \frac{1}{n}$, for all j , and so

$$z_j = \frac{1}{n} \sum_i x_i = 0$$

satisfies the first order condition. By Eq. (20), the Hessian of ρ_{i1} is

$$\frac{d^2 \rho_{i1}}{dz_1^2} = [\rho_1 - \rho_1^2] \{ 4\beta^2 [1 - 2\rho_1] [\nabla_{i1}(z_1)] - 2\beta I \}.$$

Here $[\nabla_{i1}(z_1)] = [(x_i - z_1)^T(x_i - z_1)]$ is the w by w matrix of cross product terms about the point z_1 . When $z_1 = 0$ then $\frac{1}{n} \sum_i \nabla_{i1}(z_1) = \nabla_0^*$ is the electoral covariance matrix. The Hessian of V_1 is now given by

$$\frac{1}{n} \sum_i \frac{d^2 \rho_{i1}}{dz_1^2} = [\rho_1 - \rho_1^2] \{ [1 - 2\rho_1][4\beta^2 \nabla_0^*] - 2\beta I \}.$$

By assumption $1 > \rho_1 > 0$ so $[\rho_1 - \rho_1^2] > 0$. Moreover $\beta > 0$ so the eigenvalues of V_1 will be negative iff the eigenvalues of

$$C_1 = [2\beta[1 - 2\rho_1](\nabla_0^*) - I] = [2A_1(\nabla_0^*) - I]$$

are negative. If the eigenvalues of C_1 are not negative, then $\mathbf{z}^* = (0, \dots, 0)$ cannot be a LSNE. Thus the given condition is necessary.

To prove the Corollary, note that a necessary condition for the eigenvalues of C_1 to be negative is that $\text{trace}(C_1) \leq 0$. But

$$\text{trace}(C_1) = [2A_1 \text{trace}(\nabla_0^*) - w]$$

giving the required condition $2A_1 v^2 \leq w$.

In two dimensions, if $2A_1 v^2 < 1$, then the sum of the eigenvalues of $\{C_j : j = 1, \dots, p\}$ will all be negative. Moreover, this condition is sufficient to guarantee that the determinants $\{\det(C_j); j = 1, \dots, p\}$ are all positive. Thus the stated condition is sufficient to guarantee that all eigenvalues are negative, so that $\mathbf{z}_0 = (0, 0, \dots, 0)$ is an LSNE. \square

References

- Adams J (1999a) Multiparty spatial competition with probabilistic voting. *Public Choice* 99:259–274
- Adams J (1999b) Policy divergence in multicandidate probabilistic spatial voting. *Public Choice* 100:103–122
- Adams J (2001) Party competition and responsible party government. University of Michigan Press, Ann Arbor
- Adams J, Merrill III S (1999) Modeling party strategies and policy representation in multiparty elections: why are strategies so extreme? *Am J Polit Sci* 43:765–781
- Aldrich JH (1983a) A spatial model with party activists: implications for electoral dynamics. *Public Choice* 41:63–100
- Aldrich JH (1983b) A Downsian spatial model with party activists. *Am Polit Sci Rev* 77:974–990
- Aldrich JH (1995) *Why parties?* Chicago University Press, Chicago
- Aldrich JH, McGinnis M (1989) A model of party constraints on optimal candidate locations. *Math Comput Modell* 12(4–5):437–450
- Ansolahehere S, Snyder J (2000) Valence politics and equilibrium in spatial election models. *Public Choice* 103:327–336
- Aragones E, Palfrey T (2002) Mixed equilibrium in a Downsian model with a favored candidate. *J Econ Theory* 103:131–161
- Aragones E, Palfrey T (2005) Spatial competition between two candidates of different quality: the effects of candidate ideology and private information. In: Austen-Smith D, Duggan J (eds) *Social choice and strategic decisions*. Springer, Berlin Heidelberg New York

- Banks J, Duggan J (2000) A bargaining model of collective choice. *Am Polit Sci Rev* 103:73–88
- Banks J, Duggan J (2005) The theory of probabilistic voting in the spatial model of elections. In: Austen-Smith D, Duggan J (eds) *Social choice and strategic decisions*. Springer, Berlin Heidelberg New York
- Budge I, Robertson D, Hearl D (eds) (1987) *Ideology, strategy and party change: a spatial analysis of post-war election programmes in nineteen democracies*. Cambridge University Press, Cambridge
- Clarke H, Stewart M, Whiteley P (1995) Prime Ministerial approval and governing party support: rival models reconsidered. *British J Polit Sci* 25(4):597–622
- Clarke H, Stewart M, Whiteley P (1997) Tory trends, party identification and the dynamics of Conservative support since 1992. *British J Polit Sci* 26:299–318
- Clarke H, Stewart M, Whiteley P (1998) New models for new Labour: the political economy of Labour support, January 1992–April 1997. *Am Polit Sci Rev* 92:559–575
- Clarke H, Sanders MD, Stewart M, Whiteley P (2004) *Political choice in Britain*. Oxford University Press, Oxford
- Coughlin P (1992) *Probabilistic voting theory*. Cambridge University Press, Cambridge
- Dow J, Endersby J (2004) Multinomial logit and multinomial probit: a comparison of choice models for voting research. *Elector Stud* 23:07–122
- Downs A (1957) *An economic theory of democracy*. Harper and Row, New York
- Enelow J, Hinich M (1984) *The spatial theory of voting: an introduction*. Cambridge University Press, Cambridge
- Giannetti D, Sened I (2004) Party competition and coalition formation: Italy 1994–96. *J Theor Polit* 16:483–515
- Groseclose T (2001) A model of candidate location when one candidate has a valence advantage. *Am J Polit Sci* 45:862–886
- Hinich M (1977) Equilibrium in spatial voting: the median voter result is an artifact. *J Econ Theory* 16:208–219
- ISEIUM (1983) *European elections study: European political parties' middle level elites*. Europa Institute, Mannheim
- Laver M, Hunt WB (1992) *Policy and party competition*. Routledge, New York
- Laver M, Schofield N (1990, 1998) *Multiparty governments: the politics of coalition in Europe*. Oxford University Press, Oxford, (Reprinted) University of Michigan Press, Ann Arbor
- Lin T, Enelow J, Dorussen H (1999) Equilibrium in multicandidate probabilistic spatial voting. *Public Choice* 98:59–82
- McKelvey RD, Patty J (2006) A theory of voting in large elections. *Games Econ Behav* (in press)
- McKelvey RD, Schofield N (1987) Generalized symmetry conditions at a core point. *Econometrica* 55:923–933
- Miller G, Schofield N (2003) Activists and partisan realignment in the US. *Am Polit Sci Rev* 97:245–260
- Quinn K, Martin A (2002) An integrated computational model of multiparty electoral competition. *Statist Sci* 17:405–419
- Quinn K, Martin A, Whitford A (1999) Voter choice in multiparty democracies. *Am J Polit Sci* 43:1231–1247
- Rabier J, Inglehart R (1981) Eurobarometer II April 1979. *The Year of the Child in Europe*. Inter-University Consortium for Political and Social Research, Ann Arbor
- Schofield N (1985) *Social choice and democracy*. Springer, Berlin Heidelberg New York
- Schofield N (2003) Valence competition and the spatial stochastic model. *J Theor Polit* 15:371–383
- Schofield N (2004) Equilibrium in the spatial valence model of politics. *J Theor Polit* 16:447–481
- Schofield N (2005) A valence model of political competition in Britain 1992–1997. *Elector Stud* 24:347–370
- Schofield N (2006) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. *Rev Econ Stud* (in press)
- Schofield N, Grofman B, Feld S (1989) The core and stability in spatial voting games. *Am Polit Sci Rev* 82:195–211
- Schofield N, Martin A, Quinn K, Whitford A (1998) Multiparty electoral competition in The Netherlands and Germany: a model based on multinomial probit. *Public Choice* 97:257–293

- Schofield N, Miller D, Martin A (2003) Critical elections and political realignments in the US: 1860–2000. *Polit Stud* 51:217–240
- Schofield N, Sened I (2005a) Modeling the interaction of parties, activists and voters: why is the political center so empty? *Eur J Polit Res* 44:355–390
- Schofield N, Sened I (2005b) Multiparty competition in Israel: 1988–1996. *British J Polit Sci* 36:635–663
- Schofield N, Sened I (2006) *Multiparty democracy: elections and legislative politics*. Cambridge University Press, Cambridge
- Sened I (1995) Equilibria in weighted voting games with side-payments. *J Theor Polit* 7:283–300
- Seyd P, Whiteley P (1992) *Labour's grassroots*. Clarendon Press, Oxford
- Seyd P, Whiteley P (2002) *New Labour's Grassroots*. Macmillan, Basingstoke, UK
- Stokes D (1992) Valence politics. In: Kavanagh D (ed) *Electoral politics*. Clarendon Press, Oxford
- Train K (2003) *Discrete choice methods for simulation*. Cambridge University Press, Cambridge