Journal of Theoretical Politics 16(4): 447–481
DOI: 10.1177/0951629804046150

EQUILIBRIUM IN THE SPATIAL 'VALENCE' MODEL OF POLITICS

Norman Schofield

ABSTRACT

It has been a standard result of the stochastic, or probabilistic, spatial model of voting that vote maximizing candidates, or parties, will converge to the electoral mean (the origin). This conclusion has appeared to be contradicted by empirical studies.

Here, a more general stochastic model, incorporating 'exogeneous' valence, is constructed. Contrary to the standard result, it is shown in Theorem 1 of this paper that a potentially severe domain constraint (determined by the electoral and stochastic variance, valence as well as the dimension of the space) is *necessary* for the existence of equilibrium at the electoral mean. A more stringent condition, independent of the dimension of the space, is shown to be sufficient. An empirical study of Israel for 1992 shows that the necessary condition failed. This suggests that, in proportional electoral systems, a pure strategy equilibrium will almost always fail to exist at the electoral mean. Instead, in both the formal and empirical models, each party positions itself along a major electoral axis in a way which is determined by the valence terms.

A second empirical analysis for Britain for the elections of 1992 and 1997 shows that, in fact, the necessary and sufficient condition for the validity of the 'mean voter theorem' was satisfied, under the assumption of unidimensionality of the policy space. Indeed the low valence party, the Liberal Democrat Party, did appear to locate at the electoral center. However, the high valence parties, Labour and the Conservatives, did not. This suggests that, in polities based on plurality rule, valence is a function of activist support rather than a purely exogenous factor. Theorem 2 shows, as in Britain, that exogeneous and activist valence produce opposite effects.

KEY WORDS • elections • stochastic model • valence

Earlier versions of this paper were presented at the European Public Choice Meeting, Aarhus, Denmark, April 2003, and the World Congress of the International Economic Association, Lisbon July, 2002. I thank my colleagues Randy Calvert, John Carey, Gary Miller and Olga Shvetsova, as well as seminar participants at CERGE (Prague), School of Economics (Warsaw), Trinity College (Dublin) and Nuffield College (Oxford). I would like to express my appreciation for support from the Fulbright Foundation, from Humboldt University and from Washington University in St Louis, as well as from the NSF under grant SES 0241732. Joint work with Andrew Martin, David Nixon, Kevin Quinn and Itai Sened formed the basis for the empirical analyses discussed in Section 3 of this paper. In addition, Dganit Ofek and Martin Battle provided invaluable research assistance and Lexi Shankster kindly typed the manuscript. John Duggan and an anonymous referee of the journal made a number of very helpful comments.

1. Introduction

Formal spatial models of voting have been available for many decades and have been, to some degree, the inspiration for empirical analyses of elections in the US (Poole and Rosenthal, 1984) and, more recently, in the context of multiparty competition under proportional electoral rules (Merrill and Grofman, 1999; Adams and Merrill, 1999; Adams, 2001). It has become apparent, however, that conclusions derived from at least one class of formal models are contradicted by some of the empirical evidence.

This paper will focus on the 'stochastic' formal voting model and will offer a variation of the model which appears compatible with empirical analyses of Israel and Britain. The conclusion of the standard stochastic model is the 'mean voter theorem' – that 'vote-maximizing' parties or candidates will converge to a 'pure strategy Nash equilibrium'(PSNE) located at the center of the electoral distribution (Hinich, 1977; Lin et al., 1999). The empirical evidence offered here is that the standard stochastic model has to be modified to include non-policy preferences of the electorate. In particular, it is necessary to add to each voter's evaluation of candidate, j, say, a *valence* term (Stokes, 1992). This term, labeled λ_{ij}, can be interpreted as the weight that voter i gives to the competence of candidate j. In empirical estimations, it is usual to assume that the term λ_{ij} comprises an expected 'exogeneous' term, λ_j, together with a 'stochastic' error term, ε_j. The error vector, ε, is drawn from a multivariate normal or log-Weibull extreme value distribution (Poole and Rosenthal, 1984).

Recent formal analyses of the non-stochastic voting model with valence (but limited to the case with two candidates) suggest that the candidates will not converge to the electoral center (Ansolabehere and Snyder, 2000; Groseclose, 2001).The purpose of this paper is to extend the work of Groseclose (2001) on the stochastic model with valence to the case with an arbitrary number of candidates and dimensions. We shall obtain necessary and sufficient conditions for convergence to the electoral mean in this general model. The first theorem of this paper asserts that the formal stochastic model is classified by a number called the 'convergence coefficient', c, defined in terms of all the parameters of the model. The parameters include: the vector of valences, $\{\lambda_p, \ldots, \lambda_1\}$, the variance of the stochastic errors, σ^2, the 'variance' of the voter ideal points, v^2, the spatial coefficient, β, the number, p, of political parties, and the dimension, w, of the 'policy' space, X.

Suppose the exogeneous valences are ranked $\lambda_p \geq \ldots \lambda_1$. Define

$$\lambda_{\mathrm{av}(1)} = [1/(p-1)] \sum_{j=2} \lambda_j$$

where this is the average valence of the parties other than the lowest ranked. Define the valence difference for the lowest valence party to be $\Lambda =$

$\lambda_{av(1)} - \lambda_1$. This valence difference is the key to the analysis. Obviously if all valences are equal then $\Lambda = 0$.

Define the 'convergence coefficient', c, of the model to be

$$c = 2\beta\Lambda(p-1)\nu^2/[p\sigma^2]$$

The main theorem shows that if c is bounded above by 1, then a centrist equilibrium will exist. If the policy space has dimension w and if c exceeds w, then such a centrist equilibrium is impossible. It is inferred, therefore, that the vector of vote-maximizing party positions will depend in a subtle fashion on the actual distribution of voter-preferred positions.

An empirical analysis of Israel (developing earlier work by Schofield et al., 1998) for 1992 shows that the estimated 'non-convergent' party positions are indeed close to the vote-maximizing positions derived from a multinomial logit (MNL) model. To interpret this analysis, the notion of a 'local Nash equilibrium' (LNE) is introduced. This concept depends on the idea that a vector of party positions is in 'local' equilibrium if no party may effect a unilateral, but small, change in position so as to increase its expected vote share. The formal analysis is then used to obtain conditions for existence of a LNE at the mean voter position.

In the case of Israel for the election of 1992, the empirical MNL analysis in two policy dimensions obtained high estimates for the valence terms (λ_{Likud} and λ_{Labor}) for the two major parties, Likud and Labor, and a very low valence term for one of the small parties, Shas. This gave a maximum valence difference of $\Lambda = 5.08$ and a value for the convergence coefficient of 9.72, well above the critical value of 2.0. Theorem 1 asserts that the electoral origin cannot be an equilibrium for the low valence parties such as Shas. Instead, this party should, in equilibrium, adopt a policy position far from the electoral center. This is confirmed by the empirical analysis. Moreover, parties with large valence (such as Likud and Labor) will not adopt identical positions at the electoral center. Instead, their equilibrium positions will depend on the differences in their valences and on the distribution of voter ideal points.

Whether or not a vector of party positions is a LNE can be determined by examining first- and second-order conditions on the differentials of the expected vote functions. Typically, these conditions can be met by more than one vector of party positions. However, once the empirical model has been developed, it is possible to determine the full set of LNE by simulation. These simulations have been carried out for a number of polities (Schofield et al., 1998; Schofield and Sened, 2004a,b) and have found no evidence of convergence. Since the empirical analysis mandates the inclusion of valence terms, the formal anlysis presented here gives an explanation as to why convergence is generally not observed in polities using proportional electoral methods.

Since a Nash equilibrium in party position is characterized by 'global' rather than 'local' unilateral moves, the set of LNE must contain the set of Nash equilibria. Thus, the results reported here, for the non-existence of local equilibria at the mean, also imply the non-existence of Nash equilibria. One situation where the two equilibrium concepts are identical is when the expected vote functions are, in fact, 'concave'. The first Electoral Theorem suggests that this condition will typically fail at the electoral mean, when the model is based on exogeneous valence alone. However, concavity can be expected to hold when the valence terms for the parties are not based simply on the exogenous valence of party leaders but are also derived from the behavior of party activists, who contribute time and money to their party. Schofield (2003a) argues that concavity is plausible in this case, because activism will naturally exhibit decreasing returns to scale.

An empirical analysis of the British elections of 1992 and 1997, based on a single economic dimension, shows that the convergence coefficient was below the critical value of 1 (details of the empirical analysis can be found in Schofield [2004a]). As a consequence, convergence to the electoral mean would be predicted for all parties, on the basis of the formal stochastic model involving exogeneous valence. The empirical evidence indicated that the low valence party, the Liberal Democrats, did adopt a centrist position, as predicted. However, the two high valence parties, Conservative and Labor, did not adopt (and were certainly not seen by the electorate to have adopted) centrist positions.To provide an explanation for this non-convergence, the formal model was extended to include activist valence. This model is then used to interpret the vote-maximizing positions of the parties as a balance between the centripetal tendency due to exogeneous valence and the centrifugal tendency induced by activist valence. The second theorem of this paper asserts that if the exogeneous valence of a party leader is increased, then the marginal effect of activist valence falls. The theorem is used to account for the move somewhat to the electoral center by the Labour party, under the leadership of Tony Blair, in the lead up to the 1997 election. Previous results on the activist model (Schofield, 2003a) also imply that if the activist valence function of one of the parties is sufficiently concave, then it will, in equilibrium, adopt a non-centrist position. This is used to account for the location of the Conservative Party at a policy position far from the electoral center.

A number of recent papers have presented empirical analyses of Britain, Germany, and The Netherlands (Alvarez and Nagler, 1998; Schofield et al., 1998; Quinn et al., 1999) similar in manner to the earlier work of Poole and Rosenthal (1984) but using multinomial probit (MNP) rather than multinomial conditional logit (MNL) techniques. Although the formal analysis presented here assumes independent and identically distributed stochastic errors and is, therefore, compatible with MNL, it is, in principle, possible to extend it to a more general model based on a general multivariate normal stochastic model. This extension would be compatible with MNP. Consequently, these MNP models could be used as the basis for simulation analysis to determine if the estimated party locations are indeed LNE or Nash equilibria of the appropriate vote-maximizing model.

It would then be possible to 'endogenize' party positions as a function of exogenous valence terms and endogenous activist support. Moreover, it should be possible to incorporate the effect of the electoral system on transforming vote shares to seat shares. By this means it would be possible to extend the earlier work by Rae (1971) and Riker (1982, 1986) on the relationship between electoral laws and party configurations.

2. The Stochastic Vote Model

In the voting model, it is generally assumed that political choice is described by a policy space, $\mathbf{X}$, of dimension w. As usual,we require $\mathbf{X}$ to be a compact convex subset of Euclidean space.The society is N of size n. Every voter, i, in N, is characterized by a bliss point, x_i, in $\mathbf{X}$, and the voters political preferences are given by a 'utility function' that is a monotonically decreasing function of the distance of the final outcome from x_i. There is a set P, of size p, consisting of candidates or parties. We shall use the term 'agent' for a member of P. Agents offer differing policy positions in $\mathbf{X}$. Each voter evaluates these policy positions and votes for one of the agents. Since the total vote for each agent depends on the whole vector of agent positions, it is usual to assume that the final vector of agent positions will be in 'Nash equilibrium', so that each agent maximizes some evaluation function, conditional on the other agents' positions. In the simplest 'stochastic' model, it is typically assumed that each agent's evaluation function is that agent's 'expected' vote share. The notion of 'expected' share is utilized because it is acknowledged that there is a degree of uncertainty in each voter's response. Consequently, from the point of view of the agents, when the vector of positions is given by $\mathbf{z} = (z_1, \ldots, z_p)$, where each z_j is a point in $\mathbf{X}$, then each voter, i, is characterized by a vector of probabilities $\rho_i(\mathbf{z}) = (\rho_{i1}, \ldots, \rho_{ip})(\mathbf{z})$. In the standard formal model (without valence), these probabilities are derived in the following fashion. Given the vector $\mathbf{z}$, each voter, i, in N, possesses a 'utility' vector

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

where

$$u_{ij}(x_i z_j) = -\beta||x_i - z_j||^2 + \varepsilon_j \tag{1}$$

The point x_i is i's ideal point and $||x_i - z_j||$ is the Euclidean distance between x_i and z_j. The coefficient, β, is always assumed to be positive. The error vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_p)$ has a multinomial normal distribution with mean 0 and covariance matrix Θ. A more restrictive assumption is that all $\{\varepsilon_j\}$ are independent and identically normally distributed (iind) with mean 0 and variance σ^2, so $\Theta = \sigma^2 I$, where I is the p by p identity matrix.

The probability, $\rho_{ij}(\mathbf{z})$, that i picks j is given by

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

The expected vote share of party j is $V_j(\mathbf{z}) = (1/n)\sum_{i \in N} \rho_{ij}(\mathbf{z})$.

In this standard model, it is assumed that each agent, j, adopts a position z_j so as to maximize $V_j(\mathbf{z})$, subject to the positions of the other agents. We shall consider a variety of models in this paper, based on different forms of Equation 2 and, therefore, different specifications of the expected vote shares. For each such model, we consider the existence of pure strategy Nash Equilibrium.

Definition 1: A *Pure Strategy Nash Equilibrium* (PNE), under expected vote maximization, is a vector $\mathbf{z}^* = (z_1^*, \ldots, z_p^*)$, where $z_j^* \in \mathbf{X}$, and for each $j \in P$,

$$V_j(z_1^*, \ldots, z_j^*, \ldots, z_p^*) \geq (z_1^*, \ldots, z_j, \ldots, z_p^*) \tag{3}$$

for every $z_j \in \mathbf{X}$.

It is possible to consider Nash equilibrium under different evaluation functions. For example, in two-party competition ($P = \{1, 2\}$), it is appropriate to consider plurality maximization, where, for example, party 1 maximizes $V_1(\mathbf{z}) - V_2(\mathbf{z})$. This is essentially the same as the assumption that each party maximizes its probability of winning (Duggan, 2000). The conditions under which Nash equilibria can be shown to exist are well known (Banks and Duggan, 2004). The usual assumptions are that $\mathbf{X}$ is a compact, convex subset of Euclidean space, of dimension w, and that each V_j is a continuous function of all variables. A crucial sufficient condition for existence of PNE is that each V_j is concave in the strategy of player j.

Definition 2. The function V_j is concave in z_j if and only if, for any real α,

$$V_j(z_1, \ldots, \alpha z_j + (1-\alpha)z_j', \ldots) \geq \alpha V_j(z_1, \ldots, z_j, \ldots) + (1-\alpha)V_j(z_1, \ldots, z_j', \ldots) \tag{4}$$

The function is *strictly concave* if this weak inequality is replaced with a strict inequality.

For example, Banks and Duggan (2004), in their examination of two-party competition, show that (with concavity) there exists a convergent Nash equilibrium (z_1^*, z_2^*) under plurality maximization where $z_1^* = z_2^* = x^*$ and $x^* = (1/n)\sum_{i \in N} x_i$. Call this point x^*, the 'electoral mean'. Lin et al. (1999) extended the earlier work of Hinich (1977), Enelow and Hinich (1984) and Coughlin (1992) by using calculus techniques to determine conditions for the existence of Nash equilibrium at the electoral mean in the general case with $p \geq 2$ and iind errors. Because the assumptions of the standard model imply that each V_j is not only continuous but differentiable, it is possible to examine the Hessian of V_j at the mean and to show that as long as the variance, σ^2, of the error terms is sufficiently large, then V_j can be expected to be concave, leading to Nash equilibrium, at the joint mean position $\mathbf{z}^* = (x^*, \ldots, x^*)$.

The following example shows why the constraint involving σ occurs.

Example 1. Consider the standard model (without valence), with two voters $\{1, 2\}$ at $x_1 = -1$ and $x_2 = +1$, and two parties $\{1, 2\}$ at z_1, z_2. The probability that voter i picks party 1 over 2 is

$$\text{Prob}[-\beta(x_i - z_1)^2 + \varepsilon_1 + \beta(x_i - z_2)^2 - \varepsilon_2 > 0] \tag{5}$$

Let $g_i = -\beta(x_i - z_1)^2 + \beta(x_i - z_2)^2$. Call g_i the 'comparison function' of i at (z_1, z_2). Then the previous probability can be written $\Phi(g_i)$, where $\Phi(\cdot)$ is the cumulative normal distribution for the variate $\varepsilon_1 - \varepsilon_2$, with variance $2\sigma^2$ and expectation 0.

Suppose now that $z_2^* = 0$ (the mean of the distribution of bliss points). We seek to determine whether $z_1 = 0$ comprises a best response to $z_2^* = 0$, thus making up a component of a Nash equilibrium. To do so we seek to determine the first- and second-order conditions on V_1 (i.e. we assume z_2 fixed and compute derivatives with respect to z_1). Without loss of generality, we may assume that $V_1 = \Phi(g_1) + \Phi(g_2)$. It is easy to show that

$$dV_1/dz_1 = \phi(g_1)\, dg_1/dz_1 + \phi(g_2)\, dg_2/dz_1$$

where $\phi(g_i)$ is the value of the normal probability density function, ϕ, at g_i. Since $z_1 = z_2$, we see $g_1 = g_2$ and so the first-order condition is

$$dg_1/dz_1 + dg_2/dz_1 = 0 \quad \text{or} \quad z_1^* = (1/2)(x_1 + x_2) = 0$$

(It should be obvious that the first-order condition, in the more general case with n voters, is satisfied at $z_1^* = x^*$.)

The Appendix computes the second differential (the Hessian) of V_1. Equation A7 of the Appendix can be used to show that the Hessian of $\Phi(g_i)$ is given by

$$d^2\Phi(g_i)/dz_1^2 = \phi(g_i)[d^2 g_i/dz_1^2 - [g_i/2\sigma^2](dg/dz_1)^2] \tag{6}$$

where $\phi(g_i)$ denotes the value of the probability density function (pdf) associated with Φ. Clearly, if $z_1 = 0$, then $g_i = 0$ for $i = 1, 2$. Moreover, $d^2 g_i/dz_1^2 = -2\beta$ is negative definite. Consequently, V_1 has a negative definite Hessian at $z_1 = 0$. Although V_1 has a local maximum at $z_1 = 0$, this does not imply that (0, 0) is a Nash equilibrium. Indeed, the vector (0, 0) is what we shall call a Local Nash Equilibrium or LNE (see Definition 3). The sufficient condition for $z_1 = 0$ to be a best response to $z_2 = 0$, when z_1 can lie in the domain $[-1, +1]$, is that V_1 is a concave function in z_1 (for z_2 fixed at 0), in the domain $z_1 \in [-1, +1]$. A standard result in convex analysis is that a necessary and sufficient condition for the concavity of a twice differentiable function on some domain is that its Hessian be negative semi-definite on the domain (see, for example, Schofield, 2003b: 189). To see whether this is so, evaluate the Hessian of $\Phi(g_2)$ at $z_1 = -1$.

Using Equation 6 and $g_2 = -\beta(2)^2 + \beta(1)^2 = -3\beta$; we obtain

$$d^2\Phi(g_2)/dz_1^2 = \phi(-3\beta)[-2\beta + [3\beta/2\sigma^2](4\beta)^2]$$

This is negative only if $\sigma^2 > 12\beta^2$. Obviously, if $\sigma < \beta\sqrt{12}$, then the Hessian of $\Phi(g_2)$ at $z_1 = -1$ will be positive definite. Moreover, for σ 'sufficiently' small, the Hessian of V_1 at $z_1 = -1$ will also be positive definite. In this example, therefore, for some value of σ, with $\sigma < \beta\sqrt{12}$, the expected vote function, V_1 will not be concave on $[-1, +1]$. Concavity is, of course, a strong sufficient condition for existence of a PNE and its failure does not imply that the joint mean fails to be a PNE. As the example illustrates, concavity imposes a relationship between the variance of the voter bliss points and the error variance. The more general point is that even when the joint mean is a LNE, it need not be a PNE.

In Example 1, calculus techniques were used to determine whether a vector $\mathbf{z} = (z_1, \ldots, z_p)$ satisfied the first- and second-order conditions for a PNE, with respect to the expected vote functions. Since these two conditions, at **z,** are necessary but not sufficient for PNE, we introduce appropriate variants of the Nash equilibrium concept, based on these two local conditions.

Definition 3. Assume each expected vote function V_j is twice differentiable.

(i) A vector $\mathbf{z}^* = (z_1^*, \ldots, z_p^*)$ is a *critical Nash equilibrium* (CNE), under expected vote maximization, if, for each $j \in P$, the vector $\mathbf{z}^*$ satisfies the first-order condition $dV_j/dz_j = 0$ at $z_j = z_j^*$ (under the restriction that z_k^*, for all $k \neq j$, are kept fixed).
(ii) A vector $\mathbf{z}^* = (z_1^*, \ldots, z_p^*)$ is a *local Nash equilibrium* (LNE) if $\mathbf{z}^*$ is a CNE and, in addition, for each $j \in P$, there exists an open neighborhood

U_j of z_j^* in X such that $V_j(z_1^*, \ldots, z_j^*, \ldots, z_p^*) \geq V_j(z_1^*, \ldots, z_j, \ldots, z_p^*)$ for all z_j in U_j. Thus, z_j^* is a weak best response to

$$z_j^* = (z_1^*, \ldots, z_{j-1}^*, z_{j+1}^*, \ldots z_p^*)$$

in the neighborhood, U_j. We say z_j^* is a weak local best response.

(iii) A vector $\mathbf{z}^*$ is a strict LNE, written LSNE, if and only if it is an LNE and, for all j, all the eigenvalues of the Hessian of V_j are negative. We say z_j^* is a strict local best response.

From standard results in calculus, if $\mathbf{z}^*$ is an LSNE, then it is an LNE. Moreover, a necessary condition for $\mathbf{z}^*$ to be an LNE is that all the relevant Hessians be negative semi-definite. As in Example 1, any PNE must be a LNE. Moreover, if $\mathbf{z}^*$ is an LNE and each V_j is a concave function in z_j for all z_j in $\mathbf{X}$ (and for all values of $(z_1, \ldots, z_p)$), then $\mathbf{z}^*$ is also a PNE. There are weaker conditions on V_j (such as quasi-concavity) sufficient to guarantee that $\mathbf{z}^*$ is a PNE. However, using calculus and requiring that the Hessian be everywhere negative semi-definite is a sufficient condition for existence of PNE in the case of differentiable evaluation functions. Because this condition is unlikely to be satisfied, we focus instead on local properties of the Hessian. Finally, because of difficulties that arise when eigenvalues are zero, we use the notion of LSNE. Since any PNE must be an LNE, necessary conditions for the existence of LSNE can be used to determine the necessary conditions for PNE. In the empirical analyses detailed later, the local conditions can be readily tested by simulation, while concavity is extremely difficult to verify.

For reasons mentioned in the introduction, and discussed in the next section, it is appropriate to extend the standard model to include *valence*.

Definition 4: *The probabilistic model with exogeneous valence* is defined as follows:

(i) Given the vector $\mathbf{z}$ of party positions, the utility vector $\boldsymbol{u}_i(x_i, \mathbf{z})$ of voter i has jth component

$$u_{ij}(x_i, z_j) = -\beta||x_i - z_j||^2 + \lambda_{ij}$$

where λ_{ij} is drawn from a normal distribution with expected value λ_j. This value is termed the 'exogeneous' valence' of agent j. The stochastic variation in the exogeneous valence of agent j is denoted ε_j, with expected value 0 and variance σ^2. The stochastic vector $\{\varepsilon_1, \ldots, \varepsilon_p\}$ is assumed to be multivariate normal with expectation $(0, \ldots, 0)$ and variance covariance matrix Θ with diagonal variance terms $\{\sigma_j^2\}$.

(ii) The valence terms are given by the vector $\lambda = (\lambda_p, \lambda_{p-1}, \ldots, \lambda_1)$ and the indices are chosen so that $\lambda_p \geq \lambda_{p-1} \geq \ldots \geq \lambda_1$.

(iii) The probability $\rho_{ij}(\mathbf{z})$ that i picks j (at $\mathbf{z}$) is

$$\text{Prob}[u_{ij}(x_i, z_j) > u_{ij}(x_i, z_l) \text{ for all } l \neq j]$$

(iv) The vote share of agent j is $V_j(\mathbf{z}) = (1/n)\sum_{i \in N} \rho_{ij}(\mathbf{z})$.

The main difficulty with analysis of the model with $p \geq 3$ is that vote shares involve a multivariate normal integral, whose variables are the correlated error differences. However, the Appendix shows that it is possible to perform a transformation so that the variates become independent. This allows us to show that the joint mean vector is a CNE. However, to show this vector is an LNE, it is necessary to consider the Hessian of the vote-share functions. We focus on LSNE since this allows us to ignore degenerate situations with eigenvalues that are zero. The necessary and sufficient conditions for LSNE turn on what we call the *voter covariance matrix* and the other parameters of the model.

The voter covariance matrix

We first choose a system of orthogonal axes indexed by $t = 1, \ldots w$ where w is the dimension. Without loss of generality, we can choose the coordinate system so that $\Sigma_i x_{it} = 0$, for $t = 1, \ldots w$ We use x^* to denote the vector mean $x^* = (\ldots, (1/n)\Sigma_i(x_{it}), \ldots)$, so that in the new coordinate system $x^* = (0, \ldots 0)$. We shall write this vector more simply as 0. Now let $\chi_t = (\ldots x_{it} \ldots)$ be the nth vector whose components are the new coordinates of the voter ideal points in dimension t. Then the voter distribution is given by the vector $(\chi_1, \ldots, \chi_t, \ldots, \chi_w)$. Define (χ_t, χ_s) to be the scalar product of the two vectors, χ_t and χ_s and consider the w by w matrix, D, whose entry in the (t, s) position is (χ_t, χ_s). Then $v_t^2 = (1/n)\Sigma_i(x_{ij})^2) = (1/n)(\chi_t, \chi_t) = (1/n)||\chi_t||^2$ is the variance of the voter ideal points about the origin on the t axis and this is simply the diagonal term in position t in the matrix $\Delta = (1/n)D$. We shall call Δ the *voter covariance matrix*.

Let $v^2 = \Sigma_t v_t^2$ be the total of these voter variance terms in all w dimensions. This is the sum of the diagonal terms in Δ, also known as the trace of the matrix Δ.

We shall show that the probabilistic voting model is characterized by a number called the 'convergence coefficient'.

DEFINITION 5:

(i) When the errors are identically and independently normally distributed (iind) with variance covariance matrix, $I\sigma^2$, where I is the p by p identity matrix, then the probabilistic vote model with valence is denoted

$M(\beta, \lambda : \sigma^2, \Delta)$. In the more general case with variance/covariance matrix, Θ, the model is denoted $M(\beta, \lambda : \Theta, \Delta)$.

(ii) Given the vector $\lambda = (\lambda_p, \lambda_{p-1}, \ldots \lambda_1)$ of exogeneous valences, define

$$\lambda_{\text{av}(1)} = [1/(p-1)] \sum_{j=2} \lambda_j.$$

Let $\Lambda = \lambda_{\text{av}(1)} - \lambda_1$ be the *valence difference* for agent 1.

(iii) For the model $M(\beta, \lambda : \sigma^2, \Delta)$, define the *convergence coefficient* denoted

$$c = c(\beta, \lambda : \sigma^2, \Delta), \text{ by } c(\beta, \lambda : \sigma^2, \Delta) = 2\beta\Lambda v^2\{(p-1)/[p\sigma^2]\}. \quad (7)$$

We now seek conditions under which the *joint origin* $\mathbf{z}_0^* = (0, \ldots, 0) \in X^p$ will be an LSNE for the model $M(\beta, \lambda : \sigma^2, \Delta)$. Obviously a necessary condition is that $\mathbf{z}_0^*$ is a CNE.

LEMMA: *The joint origin*, $\mathbf{z}_0^*$, *is a CNE of the model* $M(\beta, \lambda : \sigma^2, \Delta)$.

The second-order conditions can be expressed in terms of the convergence coefficient.

ELECTORAL THEOREM 1 (*Existence of a local equilibrium at the joint mean in the model* $M(\beta, \lambda : \sigma^2, \Delta)$): *Suppose that the policy space X is a closed bounded domain in Euclidean space of dimension w. Then,* $\mathbf{z}_0^* = (0, \ldots, 0)$ *is an LSNE if* $c(\beta, \lambda : \sigma^2, \Delta) < 1$ *and is an LSNE only if* $c(\beta, \lambda : \sigma^2, \Delta) < w$.

The proof of the lemma and Theorem 1 are given in the Appendix. Note that an LSNE is an LNE so the given sufficient condition of Theorem 1 is also *sufficient* for $\mathbf{z}_0^*$ to be an LNE. The proof of the Appendix also shows that $w \geq c$ is a *necessary* condition for $\mathbf{z}_0^*$ to be an LNE. In essence, the theorem implies that the dimensionless coefficient c is a measure of the degree of the centrifugal tendency in the electoral system, in the sense that the larger c is, then the more likely it is that there will be divergence of party positions from the origin. The coefficient thus provides a means of classifying elections on the basis of empirically determined parameters. The proof technique of the Appendix also shows that in the more general model $M(\beta, \lambda : \Theta, \Delta)$ it is possible to define a coefficient $c(\beta, \lambda : \Theta, \Delta)$ which classifies $M(\beta, \lambda : \Theta, \Delta)$ in exactly the same way. As observed in the Appendix, the convergence coefficient is dimensionless. What this implies in empirical applications is that the coefficient can be identified. Note, in particular, that the valences can only be defined up to a constant term, so that only valence differences have any significance. In fact, only product terms like $\beta\Lambda$ can be identified. Furthermore, the electoral variances and co-variances can only be identified with respect to the stochastic variances. This means that the ratios v_t^2/σ^2 can be identified and indeed the matrix $(1/\sigma^2)\Delta$ is identifiable. For convenience, the error variance σ^2 is normalized, depending on the model. With this normalization β, Λ and Δ can be identified.

To indicate how the proof is obtained, define the constant

$$A = \beta\Lambda\{(p-1)/[p\sigma^2]\}.$$

The Appendix shows that the Hessian of the vote share function V_1 for the lowest valence agent at $\mathbf{z}_0^* = (0, \ldots, 0)$ is given by the matrix $C = 4A\Delta - 2nI$, where I is the w by w identity matrix. The necessary condition is that the trace of C is negative, and this gives the condition $2Av^2 < w$. Substituting the definition of c in this inequality gives the necessary condition $c < w$. When this condition fails, then agent 1 can increase vote share by vacating the origin, immediately implying that $\mathbf{z}_0^*$ can be neither an LSNE nor an LNE. The sufficient condition follows from considering the determinant of C. When the condition $2Av^2 < 1$ is satisfied, then the Hessian for V_1, and indeed for $V_2, \ldots, V_p$ are all negative definite and so $\mathbf{z}_0^*$ is indeed an LSNE. In the case $w = 1$, the necessary and sufficient conditions are essentially identical.

We can illustrate the theorem in this simple case by considering a variant of Example 1 but with exogeneous valence.

Example 2. Using the expression for the Hessian given in Example 1, we can compute the Hessian for V_1 at $z_1 = 0$ when $z_2 = 0$. Clearly, the comparison functions, g_1, g_2 for voters 1 and 2 are both $\lambda_1 - \lambda_2$. Moreover,

$$dg_i/dx_i = 2\beta(x_i - z_1) = -2\beta \text{ at } i = 1 \text{ and } = +2\beta \text{ at } i = 2$$

Thus, the Hessian of V_1 is

$$\phi(\lambda_1 - \lambda_2)\{[(\lambda_2 - \lambda_1)/2\sigma^2]([-2\beta]^2 + [+2\beta]^2) - 4\beta\}$$

(This follows because of the symmetry between the two voters.) Now the probability density function ϕ is positive in value. Because party 2 has a strictly higher exogeneous valence than party 1, it therefore follows that the single necessary and sufficient condition for an LSNE at the origin is $\beta(\lambda_2 - \lambda_1)/\sigma^2 < 1$. This follows directly from Theorem 1, because $w = 1$, $p = 2$, $\lambda_{\text{av}(1)} = \lambda_2$, and $v^2 = 1$. Clearly, there exist values of $\lambda_1, \lambda_2, \beta$, and σ such that $c > 1$, so that the Hessian of V_1 at $z_1 = 0$ is positive definite. In this case, the position $z_1 = 0$ is a local 'worst' response by 1 to $z_2 = 0$. Since the Hessian of V_2 involves $(\lambda_1 - \lambda_2)$, it is obvious that $z_2 = 0$ is a best response by 2 to $z_1 = 0$. When the necessary condition is violated, $(0, 0)$ cannot be a PNE. Since 1 will adopt a position $z_1 \neq 0$, it follows that an LSNE will generally exist but not at the electoral mean. Indeed symmetry suggests there will be two LSNE $\{(z_1^a, z_2^a), (z_1^b, z_2^b),\}$ say. At any such LSNE, the points z_1^a and z_2^a will be on opposite sides of the origin with $|z_2^a| < |z_1^a|$. Indeed, the greater the asymmetry between λ_1 and λ_2, the closer z_2^a will be to the origin and the further z_1^a will be from the origin. (See also Groseclose [2001] for somewhat similar observations in a one-dimensional model.)

Even when the second-order condition at the origin is satisfied, the constraint sufficient for concavity and thus existence of a PNE will typically fail, as the third example illustrates.

Example 3. As before, let $z_2 = 0$, and let z_1 range over the domain $[-1, 1]$. For $z_1 = 0$ to be a best response in this domain, we seek a condition for concavity of V_1 on this domain. Suppose again that $\lambda_2 > \lambda_1$, and examine $z_1 = -1$. The comparison function g_2 of voter 2 is now $\lambda_1 - \beta(x_2 - z_1)^2 - \lambda_2 + \beta(x_2 - z_2)^2 = \lambda_1 - \lambda_2 - 3\beta$.

The Hessian at $z_1 = -1$ is now $\phi(\lambda_1 - \lambda_2 - 3\beta)\{-2\beta + [(3\beta + \lambda_2 - \lambda_1)/2\sigma^2][4\beta]^2\}$.

The condition for the Hessian to be negative definite at $z_1 = -1$ then becomes $\sigma^2 > 4(3\beta + \lambda_2 - \lambda_1)\beta$. Since $\lambda_2 > \lambda_1$, this condition is more severe than the one obtained in Example 1. □

This example indicates that even in one dimension, a very severe constraint involving the variance of the voter ideal points, the variance σ^2, β and Λ is required for existence of a PNE at the origin. When the weaker necessary condition $w \geq c(\beta, \lambda : \sigma^2, \Delta)$ fails, then some of the eigenvalues of the Hessian matrix for agent 1 must be positive. This implies that there exists a lower dimensional space (the eigenspace) within which the origin is a local minimum of V_1. The empirical analyses discussed in the next section suggest that the much more restrictive sufficient condition, $c(\beta, \lambda : \sigma^2, \Delta) < 1$, will generally fail. Of course, the failure of the sufficient condition need not necessarily imply that the origin fails to be a LSNE or LNE. It is, however, possible to compute the eigenvalues, using the information on the parameters of the model, voter variances $\{||\chi_t||^2 : t = 1, \ldots w\}$ and covariances as encoded in Δ. When the Hessian does have positive eigenvalues, then it implies that if an LNE $(z_1^*, \ldots, z_p^*)$ exists, it will be one where z_p^* lies closer to the origin than z_1^*. Consequently, agents with low valence will, in equilibrium, adopt a position nearer to the electoral periphery.

The simplest case is of course when $w = 1$, so the necessary and sufficient condition for an LSNE is simply that $c(\beta, \lambda : \sigma^2, \Delta) < 1$. We can apply the theorem to empirical situations where the agents are political parties, and examine the dependence of equilibria on dimension. In two or more dimensions, the difference between necessary and sufficient conditions leads to the three possible scenarios:

(i) When there is statistical evidence that the convergence coefficient greatly exceeds w, then we can infer that at least one of the eigenvalues is strictly positive. (In w dimensions, there will be w eigenvalues and the relative size of these will determine optimal positions for the parties.) For the formal model, an LSNE at the origin is impossible. In estimated models, an LSNE at the origin is highly unlikely. Typically, it will be the case that at least one low valence party j (for which the term $\lambda(_{\text{av}(j)} - \lambda_j)$ is large) can be expected to vacate the origin and adopt a position at the electoral periphery. Moreover, there will be many distinct LSNE, particularly when there are many low valence parties. This 'centrifugal' tendency is readily apparent for the case of recent elections in Israel (Schofield and Sened, 2004a). For the empirical analysis for Israel for 1992, presented in the next section, the estimated value of $c(\beta, \lambda : \sigma^2, \Delta)$, for a two-dimensional model, implies that the joint origin could not be an LNE and, thus, could not be a PNE. Simulation based on this estimated model showed that there did indeed exist numerous LSNE, none of which were at the joint origin. In particular, the LSNE positions were determined by the large positive eigenvalue associated with a particular axis of the policy space.

In this case of Israel in 1992, the product moment correlation coefficient of the voter ideal points, obtained from a survey, was fairly high as can be seen from an examination of the estimated distribution of voter ideal points in Figure 1. It is obvious from this figure that there is a major axis, aligned at approximately 40 degrees to the security axis and a minor axis, perpendicular to the major axis. Using these two principal components of the voter covariance matrix Δ as the axes facilitates the calculation of the Hessian eigenvalues. Since the variance of these data points on the major axis greatly exceeds the variance on the minor axis, it turns out that the eigenvalue on major axis is large and positive, while the eigenvalue on the minor axis is small and positive (i.e. the origin is a minimum and cannot be an LSNE for low valence parties).

As a consequence, vote-maximizing parties will move away from the origin along this major axis. This is precisely the phenomenon observed in simulation of the model (Schofield and Sened, 2004a).

(ii) A second possibility is that $c(\beta, \lambda : \sigma^2, \Delta)$ is estimated to lie in the range $[1, w]$. Empirical analysis, based on two dimensions, for elections in The Netherlands in 1977 and 1981 suggests that neither extreme divergence (as in Israel) nor convergence to the mean is expected in this case. Instead parties locate themselves at a moderate distance from the center, essentially in a circular configuration. This phenomenon is particularly pronounced

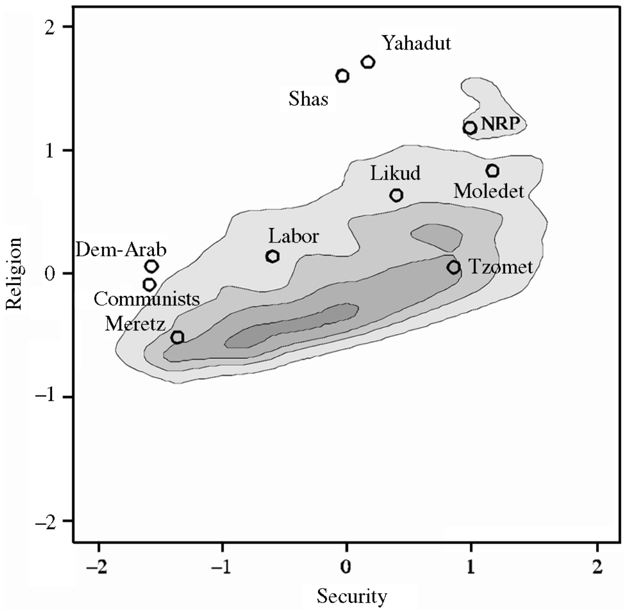

Figure 1. Estimated Party Positions in the Israeli Knesset (Based on Survey Data for 1992 from Arian and Shamir, 1995) Showing Highest Density Plots of the Voter Sample Distribution at the 95, 75, 50 and 10 percent levels) (Reprinted from Schofield [2002] with permission from North Holland)

when the variance on the axes are approximately equal and the covariance terms in D are close to 0 (Schofield and Sened 2004b).

(iii) The third possibility is that $c(\beta, \lambda : \sigma^2, \Delta)$ is clearly bounded above by 1. In this case, the eigenvalues of the Hessians will be strictly negative. Theorem 1 predicts that there will be one LSNE under which all parties converge to the origin. The empirical analysis presented in the next section suggests that this was the case for the British elections of 1992 and 1997, when only one (economic) dimension was considered. However, only one party, the small Liberal Democrat Party, clearly adopted the centrist position. When the model was extended to two dimensions (incorporating attitudes to Europe), it was found that the eigenvalue for this low valence party on the second axis was positive (but relatively small in magnitude). Theorem 1 predicts that vote-maximizing would lead that party away from the origin. Indeed, it was estimated that the Liberal Democrat Party adopted a non-centrist position on this axis. Because the covariance between the two axes was very close to zero, the eigenvalues of the Hessians on each axis are independent. Consequently, all eigenvalues for each of the parties on the economic axis were negative. All three parties should have adopted centrist positions on this axis. Although the Labour Party was perceived to be at such a position, the Conservative Party was not.

To account for divergence from the center by this high valence party, an activist model of valence was adopted. This model, introduced in recent papers (Schofield, 2003a; Miller and Schofield, 2003; Schofield et al., 2003) assumes that valence is affected by activist support. The interpretation of this is that λ_j is a function of z_j, rather than being endogenously determined. If λ_j is a concave function of z_j, then a unique PNE could indeed exist. In particular, if the (negative) eigenvalues of the Hessians of the activist valence functions have suffiently large modulii, then concavity of the vote share functions can be assured.

The next section compares the two differing cases of Israel and Britain as illustrations of cases (i) and (iii).

3. Empirical Analyses of Israel and Britain

Table 1 presents election results for the Israeli Knesset for 1988 to 2003, while Figure 1 presents an estimate of the various party positions at the election of 1992. Using the same methodology as Schofield et al. (1998), factor analysis of survey data obtained by Arian and Shamir (1995) was utilized to obtain a two-dimensional policy space. The left–right axis describes voter attitudes towards the PLO (with left being supportive). The north–south axis denoted attitudes to the importance of the Jewish faith in Israel (north being supportive). Party positions were estimated by subjecting the party manifestos to analysis based on the survey questions (see Schofield and Sened, 2004a, for further details). The outer contour line in Figure 1 contains 95 percent of the voter ideal points and the origin of the figure is the electoral mean of the estimated distribution of sample ideal points. A multinomial logit (MNL) model was used to estimate voter response and a value of $\beta = 1.25$ was obtained for the spatial coefficient. As in earlier empirical studies (Quinn et al., 1999; Quinn and Martin, 2002), sociodemographic characteristics of voters were utilized so that the valence terms were not exaggerated. Analysis of the Bayes' factors was used to determine whether the model with or without exogeneous valence was statistically superior. The log marginal likelihood of the model with valence was -834 and the Bayes' factor (Kass and Raftery, 1995) for the valence against the non-valence model was 50. This implies that the valence model was statistically superior. Because only valence differences can be estimated, it was necessary to normalize by choosing one party to have valence zero. The small party Maretz was chosen. Although ten parties obtained vote shares above the 1.5 percent cutoff, the estimation was carried out for just seven parties.

In 1992 the highest valence party was Likud with $\lambda_{Likud} = 2.73$, followed by Labor ($\lambda_{Labor} = 0.91$), Meretz ($\lambda_{Meretz} = 0$), Molodet ($\lambda_{Molodet} = -0.36$), NRP ($\lambda_{NRP} = -0.43$) and Shas ($\lambda_{Shas} = -4.66$). These valence terms are in some correspondence to the seat shares as given in Table 1. Now use the suffix 1 to denote the lowest valence party, Shas.

In the two-dimensional Figure 1, distance is normalized with respect to $\sigma^2 = 1.65$. Simulation of the model found that LNE party positions were not at the electoral mean (0,0). Calculation gives $\lambda_{av(1)} - \lambda_1 = (0.42 + 4.66) = 5.08$.

Table 1. Elections in Israel

	Party	Knesset Seats				
		1988	1992	1996	1999	2003
Left	Labor	39	44	34	28	19
	Meretz	–	12	9	10	6
	Others	9	–	–	–	3
	Democrat Arab	1	2	4	5	2
	Communists	4	3	5	3	3
	Balad	–	–	–	2	3
Subtotal		*53*	*61*	*52*	*48*	*36*
Center	Olim	–	–	7	6	2[a]
	Third Way	–	–	4	–	–
	Center	–	–	–	6	–
	Shinui	2	–	–	6	15
Subtotal		*2*		*11*	*18*	*15*[a]
Right	Likud	40	32	30	19	38[b]
	Gesher	–	–	2	–	–
	Tsomet	2	8	–	–	–
	Yisrael Beiteinu	–	–	–	4	7
Subtotal		*42*	*40*	*43*	*35*	*47*
Religious	Shas	6	6	10	17	11
	Yahadut	7	4	4	5	5
	NRP	5	6	9	5	6
	Moledet	2	3	2	4	–
	Techiya	3	–	–	–	–
Subtotal		*23*	*19*	*25*	*31*	*22*
TOTAL		*120*	*120*	*120*	*120*	*120*

[a] Olim joined Likud to form one party immediately after the last election.
[b] Likud with Olim has 40 seats.

The results of Theorem 1 indicate why convergence is not to be expected. The convergence coefficient is given by

$$c = c(\beta, \lambda : \sigma^2, \Delta) = 2\beta v^2\{(p-1)/[p\sigma^2]\}\{\lambda_{\text{av}(1)} - \lambda_1\}.$$

We may choose new orthogonal coordinate axes, labeled 1and 2, along which estimated voter variances on these axes are of the order of 1.2 and 0.235 so that the total voter variance is 1.435. For example, we may take the first major axis to run through the Molodet/Meretz positions in Figure 1. Then the convergence coefficient is approximately 9.72, well above the necessary limit of 2.0. Using the method of calculation outlined in the Appendix, the product of the two eigenvalues for Shas can be calculated to be 1.86, while the sum is $c - 2 = 7.72$. Thus, the eigenvalues can be estimated to be 7.47 and +0.25. If all parties were located at the electoral mean, then the vote-share function for Shas would increase very rapidly away from the origin along the major axis. (This is because this axis is the eigenvector for the large positive eigenvalue of 7.47.) Because the second eigenvalue is small and positive, the optimal position for Shas will lie on (or close to) the major axis. Whether the party should move up or down, this axis is indeterminate. As soon as Shas moves away from the origin, then so should the other parties, simply in order to gain votes. Simulation of the MNL vote model found that equilibrium locations of all parties were strung along the major axis, as predicted by the theorem (Schofield and Sened, 2004a).

The basic distribution assumptions of the MNL model and the formal stochastic vote model are not identical. The empirical MNL model assumes that the voter probabilities satisfy the following 'independence property': that for any voter i and parties j, k then the ratio

$$\rho_{ij}/\rho_{ik} = [\exp(-\beta\delta_{ij}^2)]/[\exp(-\beta\delta_{ik}^2)]$$

(see Adams, 2001). This 'independence of irrelevant alternatives' property is not satisfied by the formal stochastic vote model. The empirical model based on the assumption that the errors are multivariate normal is termed multinomial probit (MNP). Thus, the theoretical underpinning of the MNP model and the formal model are very close. Estimates of parameters on the basis of the MNL and MNP models tend to be similar. (See Quinn et al. [1999] and Dow and Endersby [2004] for a comparison of the empirical MNL and MNP models.) This suggests that the formal model as developed here can be used as a benchmark for the empirical MNL model. Moreover, the technique for proving Theorem 1 can be used for a more general model based on the multivariate normal distribution, which is the basis for multinomial probit estimation (Schofield et al., 1998).

The difficulty of policy decisions in Israel has meant that the exogeneous valences of the parties in government have waxed and waned, expanding or contracting the electoral 'sphere of influence' of each party. As Table 1 indicates, the changes in seat shares (and vote shares in this highly proportional electoral system) were matched by these fluctuating valences. This suggests that while equilibrium policy positions do change over time, the changes are driven by exogenous transformations in valence. The situation in Israel is, of course, made even more complex by the ability of small parties to enter the political arena relatively easily (because of the low threshold in this proportional electoral system). This has led to a degree of coalitional instability as the major parties have been obliged to seek support from small parties, such as Shas or, more recently, Shinui and the NRP (Schofield and Sened, 2002).

In contrast to the situation where the convergence coefficient is large, the results of an analysis of the British election suggest that convergence will be predicted if a single dimension is considered. Figure 2 presents estimates of party positions (for Labour [Lab], the Conservative Party [Con] and the Liberal Democrat Party [Lib]) together with an estimate of the distribution of voter ideal points for Britain in the two elections of 1992 and 1997, using a single (economic) dimension (Schofield, 2004a).

Party positions were derived from average voter perceptions, as obtained from the British National Election Survey. Using an MNL empirical model,

Table 2. Sample and Estimation Data for Britain 1992–1997

1992	Sample (%)	Coefficients	Confidence Interval	Correct Prediction (%)
UK without Scotland		$\beta = 0.56$	[0.50,0.63]	50.0
Conservative	49.3	$\lambda = 1.58$	[1.38,1.75]	62.9
Labour	32.4	$\lambda = 0.58$	[0.40,0.76]	47.4
Liberals	18.2	$\lambda = 0.00$		19.6
Scotland		$\beta = 0.50$	[0.31,0.67]	35.6
Conservative	30.2	$\lambda = 1.68$	[1.18,2.21]	48.6
Labour	31.4	$\lambda = 0.91$	[0.38,1.51]	37.2
SNP	25.9	$\lambda = 0.77$	[0.26,1.30]	29.5
Liberals	12.3	$\lambda = 0.00$		12.8
1997				
UK without Scotland		$\beta = 0.50$	[0.44,0.56]	45.7
Conservative	31.9	$\lambda = 1.24$	[1.03,1.44]	45.2
Labour	49.6	$\lambda = 0.97$	[0.85,1.07]	56.1
Liberals	18.5	$\lambda = 0.00$		19.0
Scotland		$\beta = 0.50$	[0.40,0.64]	40.5
Conservative	14.2	$\lambda = 0.92$	[0.58,1.24]	33.7
Labour	52.7	$\lambda = 1.33$	[1.10,1.57]	56.3
SNP	20.4	$\lambda = 0.42$	[0.16,0.72]	21.5
Liberals	12.5	$\lambda = 0.00$		12.8

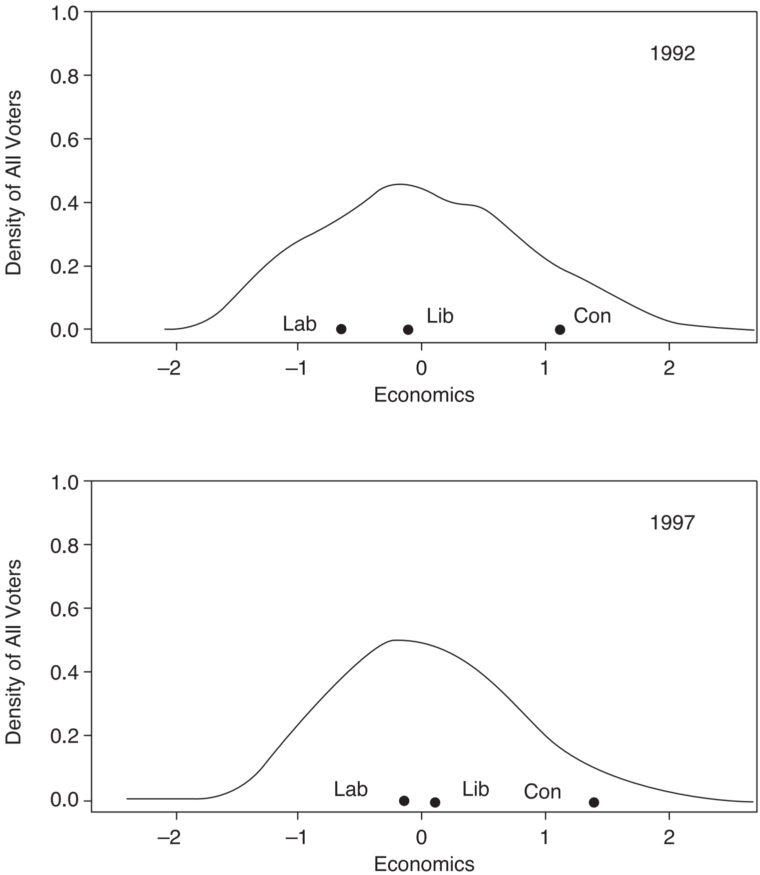

Figure 2. Estimated Party Positions in the British Parliament in 1992 and 1997 for a One-dimensional Model (based on the British National Election Survey and Voter Perceptions)

the estimated valence coefficients for 1997 were $\lambda_{\text{Labor}} = 0.96$, $\lambda_{\text{Cons}} = 1.24$ and $\lambda_{\text{Lib}} = 0.0$ while the spatial coefficient, β, was estimated to be 0.50. Table 2 presents the estimation results for the 1992 and 1997 elections. Because the issue of Scottish Nationalism was relevant in Scotland, estimations were performed for Britain without Scotland and separately for Scotland. The empirical model was highly successful (approximately 50 percent) at predicting individual voter response (See also Quinn et al. [1999] for a similar empirical analysis of earlier British elections). For Britain without Scotland in 1997, the estimated value of

$$\lambda_{\text{av(lib)}} = (1/2)(\lambda_{\text{Cons}} + \lambda_{\text{Labor}}) = 1.11$$

was obtained, while $\beta v^2 = 0.5$. We can calculate the convergence coefficient of Theorem 1 to be 0.34. Even using the upper estimated bound of $\lambda_{\text{Cons}} = +1.44$ and $+\lambda_{\text{Labor}} = 1.07$ and $\beta = 0.56$, we obtain the upper estimated bound for the coefficient (at the 95 percent confidence level) of 1.0. Thus, on the basis of the formal model, we can assert with a high degree of certainty that the low valence party, the Liberal Democrats, can be located at an LSNE at the origin if all other parties also locate there. The same argument can be made for the Labour and Conservative Parties. Thus, under the assumptions of exogeneous valence, vote maximization, and unidimensionality, a version of the 'mean voter theorem' should have been valid for the British election of 1997 (and indeed for 1992). Although Figure 2 indicates that a position close to the center was adopted (or seen to be adopted) by the Liberal Democrats in 1992 and 1997, this was not so for the two other parties (the Labour Party and, even more obviously, the Conservative Party) in 1992. Although the Labour approached closer to the center in 1997 than in 1992, the Conservatives were perceived to become more extreme over the same period.

To test whether this disjunction between theory and empirical analysis was the result of the assumption of unidimensionality, the empirical model was extended to include a second dimension, based on attitudes to the European Union.

Figure 3 presents the result of this extended model based on a two-dimensional factor analysis of voter response for the election of 1997. The outer contour contains 95 percent of all estimated voter bliss points and, again, the origin (0,0) is the electoral sample mean. The 'positions' of the parties were estimated by taking the mean, for each party, of the estimated bliss points of a sample of party MPs (obtained from written responses to the British National Election Survey). These estimated mean MP party positions need not, of course, be identical to the party positions (as perceived by the electorate). Indeed, there may be some element of bias in the MP response. In the next section of the paper, we shall examine a model in which activists for the party affect the choice of party position to be declared to the electorate. I shall interpret the mean MP position obtained for each party as a proxy for the preferred position of party activists. This point I shall call the 'principal's' position for the party. Figure 3 indicates the positions for the principals of the various parties.

Figure 3 suggests that the only centrist principal's position was that of the single respondent of the Scottish National Party (labeled SNP). It is clear that the principals' party positions for the Conservatives (CONS), Ulster Unionists (UU), Plaid Cymru (PC) and Liberal Democrats (LIB) were far removed from the center on the second axis, while the positions of the Labour and Liberal principals were centrist only on the economic axis.

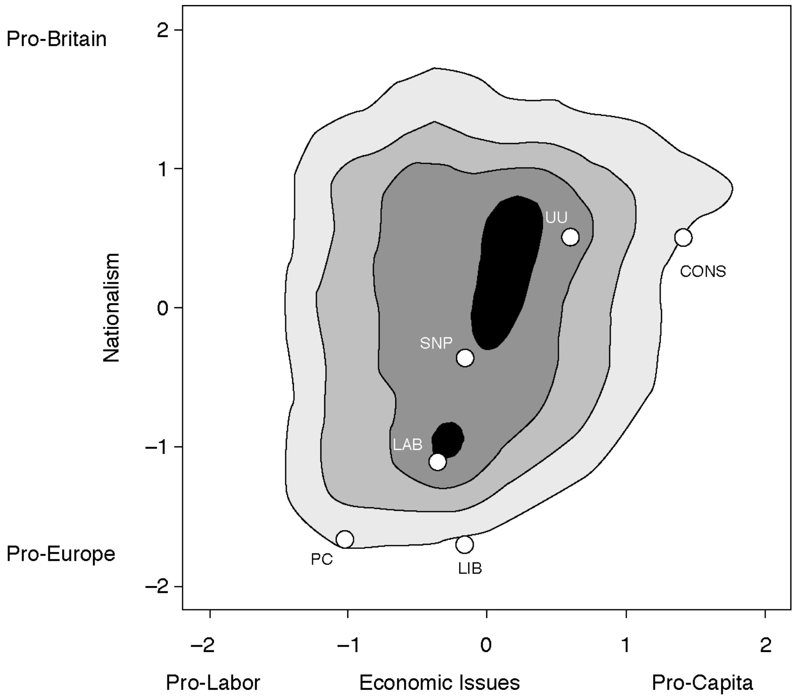

Figure 3. Estimated Party Positions in the British Parliament in 1997, for a Two-dimensional Model (Based on MP Survey Data and a National Election Survey) Showing Highest Density Plots of the Voter Sample Distribution at the 95, 75, 50 and 10 percent levels

We can repeat the calculation of the convergence coefficient in the two-dimensional case. For purposes of illustration, we may assume that the covariance between the electoral data on the economic axis and on the Nationalism axis is zero: that is, the electoral data on the two axes are uncorrelated. This allows us to compute coefficients c_1 and c_2 separately for each axis. The separate necessary condition is that each is bounded above by 1. We can immediately conclude that the Liberal Democrat eigenvalue on the economic axis is negative (just as in the one-dimensional case). However, the electoral variance on the Nationalism axis is much greater than on the economic axis. This leads to the conclusion that the coefficient $c_2 = c_2(\beta, \lambda : \sigma^2, \Delta)$ on this Britain/Europe axis exceeds 1. Calculation shows that the Liberal Democrat eigenvalue on this axis is positive. Thus the origin will be a saddlepoint for this party. We can infer that the Liberal Democrat Party moved 'south' away from the origin, in the direction of the gradient associated with this eigenvalue. Although I have emphasized that the principals' positions need not coincide with the positions of the parties themselves or with the positions of the leaders, it is remarkable that the estimated principals' positions in Figure 3, when projected onto the economic axis, coincide with the electorally perceived positions given in Figure 2. This suggests what I have termed the principals' positions are, on average, accurately perceived by the electorate. Another interpretation is that activists for each of the parties have a considerable impact on the perceived position of the party. The analysis just conducted suggests that the relationship between principal's position and party position is particularly simple for the Liberal Democratic Party. As we have seen the Liberal Democrat eigenvalue on the Europe axis at the origin is positive, so the party could just have readily moved 'north' as 'south' to maximize vote share. The move 'south' was consistent with the principal's position.

In contrast, if we assume that valence is exogeneous and accurately reflected in the estimation, then this eigenvalue analysis suggests that the Conservative Party should have adopted a centrist position on the economic axis and a position on the Nationalism axis closer to the center than the Liberal Democrats. This contradiction leads us to consider an activist model.

4. The Activist Valence Model

The disjunction observed in the previous section between theory and empirical analysis suggests that the valence model, given in Definition 4, should be modified.

For example, Adams and Merrill (1999) and Merrill and Grofman (1999) have proposed theoretical electoral models based on directional voting and these do not lead to convergence theorems. However, the success of the empirical MNL model for Britain, in predicting voter response, suggests that the directional model cannot be used to account for non-convergence in these elections.

An earlier model based on 'activist' valence (Aldrich, 1983a,b; Aldrich and McGinnis, 1989) has proved useful in providing an explanation for the non-convergence of candidates in US presidential elections (as noted earlier by Poole and Rosenthal [1984]; see also Miller and Schofield [2003] and Schofield et al. [2003]).The 'exogeneous valence' model, discussed earlier, can be readily modified to accommodate 'activist valence'. With activism, the valence term is not due to exogenous factors alone but is a function of the willingness of party activists to support their party with time and money. Consequently, activist valence depends on the interaction of activist preferences and party location. One formal difference between the proposed activist valence model and the exogeneous valence model discussed in the previous section is that 'activist valence' can be assumed to be a concave function of party position. As shown in Schofield (2003a), this will more readily guarantee that PNE will exist. Indeed it is possible to obtain a Lagrangian by which to determine, theoretically, the location of PNE when there are multiple potential activist groups. This activist model is most useful when there are two political dimensions and we can use the two-dimensional estimation of the previous section to study intraparty bargaining.

We now modify the definition of voter utility, given in Definition 4 and assume instead that each voter, i, is described by a utility vector

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

where

$$u_{ij}(x_i, z_i) = \lambda_j + \mu_j(z_j) - \beta||x_i - z_j||^2 + \varepsilon_j \tag{8}$$

As before, λ_j is the expected 'exogenous' valence of party j, while $\mu_j(z_j)$ is the 'activist' valence, β is the positive spatial coefficient and ε_j is the error term. Note, in particular, that $\mu_j(x_j)$ is a function of z_j. I shall use the expression $M(\beta, \lambda, \mu : \sigma^2, \Delta)$ to denote both formal and empirical versions of the model when exogeneous and activist valence are incorporated in the model.

Schofield (2003a) has shown that the first-order condition for an LSNE in this valence model is that the equilibrium position for each party j satisfies the condition that

$$z_j^* = d\mu_j^*/dz_j + \sum_i \alpha_{ij} x_i \tag{9}$$

In this equation, the coefficients $\{\alpha_{ij}\}$ depend on the values of the functions $\{\mu_j\}$ at $\mathbf{z}$. Most importantly, the coefficients depend on the exogeneous valance differences $\{\lambda_j - \lambda_k\}$. The function μ_j^* is renormalized and satisfies $\mu_j^* = (1/2\beta)\mu_j$. Let us denote the vector $\sum_i \alpha_{ij} x_i$ by dV_j^*/dz_j and call it the 'marginal electoral pull' of party j, due to valence. The first order condition can then be written

$$dV_j^*/dz_j + d\mu_j^*/dz_j - z_j^* = 0 \tag{10}$$

The first term in this expression is the 'marginal electoral pull', namely a gradient vector pointing towards a weighted electoral mean. As $(\lambda_j - \lambda_k)$ is exogeneously increased, this vector increases in magnitude. The second term, $d\mu_j^*/dz_j$, is the 'marginal activist pull'. This vector points towards that point or locus of points where the activist contributions of the party are maximized. The simplest case to study is when there are only two candidates or parties. In this case, the following theorem follows from Schofield (2003a).

Electoral Theorem 2 (*The Stochastic Model with exogeneous and activist valences*): *Consider a formal or empirical vote maximization model*

$M(\beta, \lambda, \mu : \sigma^2, \Delta)$. *The first-order condition for* $\mathbf{z}^*$ *to be a local equilibrium is that, for each j, the electoral and activist pulls must be balanced. Other things being equal, the position* z_j^* *will be closer to a weighted electoral mean for larger values of the exogenous valence* λ_j. *Conversely, if the activist valence function* μ_j *is increased (due to the greater willingness of activists to contribute to the party),* z_j^* *will be nearer to the activist preferred position. If all activist valence functions are highly concave, in the sense of having negative eigenvalues of sufficiently great magnitude, then the solution given by the first-order condition will be a PSNE.* □

We can use this theorem to provide an account of the strategies adopted by the two high valence parties in Britain. Since the exogeneous valence model used in the previous section provides a reasonable account of the Liberal Democrat Party position, we exclude it from the analysis. First, the individual response data for the Conservative MP cohort shows that they generally adopted quite extreme positions, particularly on the second Nationalism dimension (pro-Britain or anti-Europe). The fact that the Conservatives lost the 1997 election (with 31.4 percent of the vote to 44.4 percent for Labour), even though they had quite a high exogeneous valence, suggests that 'anti-Europe' Conservative Party activists dominated the Conservative Party position in 1997. Although the mean Conservative MP position was 'pro-capital' on the economic axis and 'anti-Europe' on the second axis, there were a number of 'moderates' on both axes. The Conservative MP responses suggest that there were, in fact, two distinct groups of MPs in the party. For purposes of illustration, we may suppose that there are two activist groups (or two principals) for the party, one at the position denoted C (for pro-capital) in Figure 4, and one at position B (for pro-Britain). The optimal Conservative position will be determined by a version of Equation 10 which equates the 'electoral pull' against the two 'activist pulls' of the principals at C and B. Since the exogeneous valence term for the Conservative Party fell between 1992 and 1997, the party's 'electoral pull' also fell between the elections. The optimal position, z^*_{cons}, will be one where z^*_{cons} is 'closer' to the locus of points where the marginal activist pull is zero (i.e. where $d\mu_{\text{cons}}/dz_{\text{cons}} = 0$). This locus of points we can call the 'activist contract curve' for the Conservative Party and we may identify it as the set of policy points satisfying the first-order condition for maximizing activist contributions to the party. As Figure 4 indicates, the features of the curve will depend on assumptions about the preferences of the party activists. Miller and Schofield (2003) suggest that activists exhibit different saliences along different axes (as indicated by the 'ellipsoidal' indifference curves in Figure 4).

For the Conservative Party leader, the fall in exogeneous valence forces the party to acquiesce to the demands of activists, just to gain as many votes as

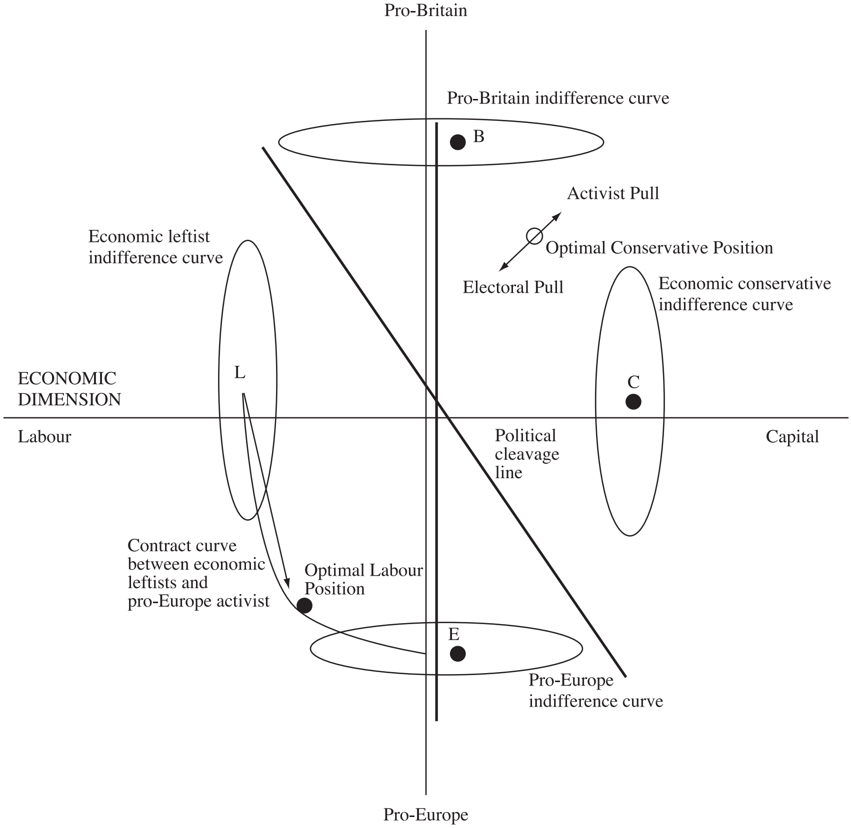

Figure 4. Illustration of Vote-maximizing Positions of Conservative and Labour Party Leaders in a Two-dimensional Policy Space

possible. Until last year, the low exogeneous valence of the party leader, Iain Duncan Smith, brought about an activist contest between the moderates, led by Michael Portillo and Kenneth Clarke, against the anti-Europe activists, and Duncan Smith, himself. This led eventually to the selection of a new party leader, Michael Howard, who appears, at present, to have a higher valence.

The opposite effect may have been apparent in the case of the Labour Party in the period from 1983 to 1997. Since the exogeneous valences of Labour leaders were low, prior to Blair's selection as leader in 1994, the activists dominated and Labour's electoral vote share was in the range 28–35 percent. However, the replacement of Thatcher by John Major as Conservative leader in 1990 and then of Kinnock by John Smith as Labour leader in 1992, led to an increase in Labour's exogeneous valence. The increase in the Labour exogeneous valence, due to Tony Blair's leadership after 1994, further weakened the importance of the Labour party activist valence.

Figure 4 gives a schematic representation of the 'activist' contract curve for Labour. In this case, the two opposed Labour principals are denoted L (for pro-labour) and E (for pro Europe). Blair's high exogeneous valence had two effects. First, it allowed him to overcome grass-roots Labour 'radicalism' by bringing new middle-class pro-Europe activists into the party. Second, this strengthened the electoral pull, so that the party was able to adopt a much more centrist electoral manifesto for the 1997 election (Seyd and Whiteley, 2002).

This more general model, incorporating both popularity and activist valence, suggests that it was because Blair came to be seen as an attractive candidate for the position of Prime Minister that he was able to move the party to a more centrist position. From this perspective, the causality runs from valence to centrism to electoral victory. The present danger for the Labour party is two-fold. Blair's current position on Iraq has lowered his exogeneous valence and this has led to a grass-roots and MP revolt against the party leadership. Second, as Figure 3 illustrates, there is some degree of anti-European Union sentiment in the electorate, while there is much pro-Europe support among Labour MPs. If Blair's valence falls, then the party could face increased conflict between the two principal activist groups in the party. Thus, weak exogeneous valences for both major parties may thus lead to activist conflict, increased partisanship and uncertain elections.

5. Conclusion

In their interpretation of their analysis of the US Presidential elections of 1968–80, Poole and Rosenthal (1984: 288) commented that their results were 'at variance with the simple spatial theories which [hold] that candidates should converge to the center of the [electoral] distribution'. The purpose of this paper has been to show that the apparent contradiction between empirical analysis and the convergence theorem of the probabilistic vote model has a relatively simple resolution. Since empirical analyses demand that valence terms be included, it is necessary to examine more carefully the second-order conditions for equilibrium, when valence is involved.

For the empirical model of Israel, it is clear that convergence to the electoral mean cannot be expected. Indeed, since the valence differences between the parties are significant, low valence parties will have vote functions that are minimized at the origin.

Because most low valence parties must retreat from the electoral center, the first-order condition for equilibrium, with high valence parties at the origin, will not be satisfied. More generally, the failure of concavity of the vote-share functions in multiparty electoral systems suggests that the Nash equilibrium concept is not the appropriate equilibrium concept. Instead, LNE can be determined through simulation, once the parameter values of the model have been empirically obtained. It may be the case that, in contrast to the convergence result, there will exist multiple LNE. Simulation of the model, at least in the case of Israel, suggests that most low valence parties will adopt quite radical positions. Moreover, the model also suggests that if a party suffers from falling valence over time, then its optimal equilibrium strategy will be to adopt a more radical position. An obvious conjecture is that this phenomenon will be pronounced in all multiparty polities based on proportional electoral systems. Because of the necessity of coalition government, the result may be political instability (Laver and Schofield, 1990).

The estimated parameters of a one-dimensional empirical model with exogeneous valence for Britain, for the elections of 1992 and 1997, suggested that the convergence condition for LNE, at the joint mean position, was satisfied. Although the perceived Labour position was closer to the origin in 1997 than in 1992, the Conservative position appeared to become more radical. This phenomenon is compatible with a more general valence model, where there is a second component of valence which is affected by activist support. Analysis of this model (Schofield, 2003a) suggests that as the 'exogeneous valence' of the party leader falls, then the effect of activists becomes more pronounced. Empirical results, based on two-dimensional factor analysis (involving a second 'European axis) for Britain, were compatible with this inference.

The complexity of the set of possible LNE in the exogeneous valence model parallels the rich variety of party configurations found in many polities using proportional electoral methods (Schofield and Sened, 2004b). For plurality electoral systems, such as Britain and the US, the activist valence model appears to provide a useful framework with which to study the inherent conflict between electoral support and activist motivation.

Appendix: Proof of the Lemma and Theorem 1

The argument depends on a consideration of the vote share function, $V_1(\mathbf{z})$, for the lowest valence party. However, this function involves a multivariate integral whose variates are correlated. We therefore make a linear transformation so that the covariance matrix of the new random variables are independent. This allows us to construct a vote-share function, denoted $V_{1p}(\mathbf{z})$ with which we can examine the properties of $V_1(\mathbf{z})$: the function $V_{1p}(\mathbf{z})$ is a univariate integral and models a specific contest between the lowest valence agent, 1, and an imaginary agent whose valence is the average of the valences of the other $(p-1)$ agents. First- and second-order conditions on this function then allow us to infer the same conditions on $V_1(\mathbf{z})$.

Let i be any voter.The probability that i picks 1 is

$$\rho_{i1}(\mathbf{z}) = \text{Prob}[\lambda_1 - \beta||x_i - z_1||^2 + \varepsilon_1 > \lambda_k - \beta||x_i - z_k||^2 + \varepsilon_k : \text{ for all } k \neq 1] \quad \text{(A1)}$$

This expression can be written as $\Phi(g_{i1}(\mathbf{z}))$, where

$$g_{i1}(\mathbf{z}) = [[\lambda_1 - \beta||x_i - z_1||^2 - \lambda_2 + \beta||x_i - z_2||^2], \ldots, [\lambda_1 - \beta||x_i - z_1||^2 - \lambda_p + \beta||x_i - z_p||^2]] \quad \text{(A2)}$$

Here $g_{i1}(\mathbf{z})$ is a $(p-1)$-dimensional variable, while Φ is the cumulative probability function (cpf) of the $(p-1)$-dimensional variate $\text{e}(1) = [\ldots \varepsilon_j - \varepsilon_1 \ldots : j = 2, \ldots, p]$. Here we use the notation $\Phi(d) = \text{Prob}[\text{e}(1) < d]$, where d is a $(p-1)$ vector. The variance–covariance matrix, Ξ, of this variate is a $(p-1)$ by $(p-1)$ matrix which can be written $\Xi = \sigma^2 FF^t$, where F^t stands for the transpose of the $(p-1)$ by p difference matrix. It is obvious that Ξ is only diagonal in the case $p = 2$, in which case the variance is $2\sigma^2$. In the more general case with p greater than or equal to 3, the matrix Ξ has diagonal terms $2\sigma^2$ and off-diagonal terms σ^2. Consequently, the components of e(1) are correlated and this makes it difficult to compute $\rho_{i1}(\mathbf{z})$. However, since e(1) is multivariate normal, by the assumption on the errors, we can utilize a matrix transformation, B, chosen so that the transformed variate $u(1) = B(\text{e}(1))$ has covariance matrix $B\Xi B^t = \sigma^2 I$ (again B^t stands for the transpose of B, while I is the identity matrix). Because the transformed variates are independent, we may now write

$$\Phi(g_{i1}(\mathbf{z})) = \Phi(B_1 g_{i1}(\mathbf{z}))\Phi(B_2(g_{i1}(\mathbf{z})))\Phi(B_{p-1}(g_{i1}(\mathbf{z}))) \quad \text{(A3)}$$

In this expression each linear transformation, B_j, is represented by the jth row of B, while the integral $\Phi(B_j(g_{i1}(\mathbf{z})))$ is associated with the jth component of $u(1)$. Thus, each $B_j(g_{i1}(\mathbf{z}))$ gives the the upper bound of the integral. Because of the number of degrees of freedom associated with this transformation, we can choose each vector B_j, for $j = 1, \ldots, p-2$, so that its coordinates sum to zero. This has the consequence that the bounds, $B_j(g_{i1}(\mathbf{z}))$,for $j = 1, \ldots, p-2$, are independent of z_1. Moreover, we can also choose B_{p-1} so that all its coordinates are identical. Note also that such a transformation can be found in the more general case where the errors are correlated. Now write

$$h_{i1p}(\mathbf{z})) = [\lambda_1 - \beta||x_i - z_1||^2] + (1/(p-1)\sum_j [(-\lambda_j + \beta||x_i - z_j||)^2]$$

Then the last component of Equation A3 can be written $\Phi^*(h_{i1}p(\mathbf{z}))$ where we use Φ^* to denote the cpf of the variate $u = (1/(p-1)\sum_{j=2}(\varepsilon_j - \varepsilon_1)$.

It is easy to see that the variance of u is $[1/(p-1)^2][(p-1)^2 + (p-1)]\sigma^2 = p/(p-1)\sigma^2$. We can now prove the lemma.

Proof of Lemma. From Equation A3,

$$V_1(\mathbf{z}) = (1/n)\sum_i \Phi(g_{i1}(\mathbf{z}))$$
$$= (1/n)\sum_i [\Phi(B_1(g_{i1}(\mathbf{z})))\ldots\Phi(B_{p-2}(g_{i1}(\mathbf{z})))\ldots\Phi^*(h_{i1p}(\mathbf{z}))] \quad \text{(A4)}$$

The terms involving B_1, B_2, B_{p-2} are all independent of z_1 but they do depend on $\{x_i\}$ and on $\mathbf{z}_{-i} = (z_2, \ldots, z_p)$. We seek to show that $z_1 = 0$ is a local best response to $\mathbf{z}_{-1} = (0, \ldots, 0)$. At $\mathbf{z}_{-1} = (0, \ldots, 0)$, all the terms involving $B_1(g_{i1}(\mathbf{z})), \ldots, B_2(g_{i1}(\mathbf{z})), \ldots, B_{p-2}(g_{i1}(\mathbf{z}))$ are independent of x_i but they do depend on $\{\lambda_j\}$. We may write α for the product of these first $(p-2)$ terms and, for purposes of differentiation, treat it as a positive constant. Thus,

$$V_1(\mathbf{z}) = V_1(z_1, \mathbf{z}_{-1}) = (\alpha/n)\sum_i \Phi^*(h_{i1p}(\mathbf{z}))$$

Differentiating this expression gives the first-order condition

$$\sum_i \alpha[\varphi^*(h_{i1p}(\mathbf{z}))\, dh_{i1p}(\mathbf{z})\, dz_1] = 0 \quad \text{(A5)}$$

Here $\varphi^*(h_{i1p}(\mathbf{z}))$ is the value of the probability density function(pdf) of the variate u at $h_{i1p}(\mathbf{z})$, given $\mathbf{z}_{-1} = (0, \ldots, 0)$. This is positive. Moreover, at $z_1 = 0$, this term is independent of x_i. Finally, $dh_{i1p}/dz_1 = -2\beta(z_1 - x_i)$ so $\sum_i dh_{i1p}/dz_1$ only involves terms in $\beta\sum_i(z_1 - x_i)$. Thus, Equation A5 reduces to the condition that $z_1 = (1/n)\sum_i x_i$. But we have normalized coordinates so that $\sum_i x_i = 0$.

Thus, $z_1 = 0$ solves the first-order condition for maximizing $V_1(\mathbf{z})$ subject to the constraint that $\mathbf{z}_{-i} = (0, \ldots, 0)$. But the same argument can be carried out for each vote function $V_j(\mathbf{z}), j = 1, \ldots, p$, proving that $\mathbf{z}^* = (0, \ldots, 0)$ is a CNE. □

To determine whether $\mathbf{z}^* = (0, \ldots, 0)$ is an LNE requires differentiating the LHS of Equation A5. Because Equation A1 involves a multivariate integral, this cannot readily be computed. However, after the transformation, B, we can define $V_{1p}((z_1, \mathbf{z}_{-1}) = (1/n)\sum_i \Phi^*(h_{i1p}(\mathbf{z}))$. At $\mathbf{z}_{-1} = (0, \ldots, 0)$, we can write $V_1(z_1, \mathbf{z}_{-1}) = \alpha V_{1p}(z_1, \mathbf{z}_{-1})$, where α is independent of z_1. Thus, we can use $V_{1p}(z_1, \mathbf{z}_{-1})$ as a proxy for $V_1(z_1, \mathbf{z}_{-1})$.

Proof of Theorem 1. The Hessian of $V_{1p}(z_1, \mathbf{z}_{-1})$ is given by

$$H_1(z) = (1/n)\sum_i H_{i1}(z) \quad \text{(A6)}$$

Here $H_{i1}(z)$ denotes the Hessian of $\Phi^*(h_{i1p}(\mathbf{z}))$. Because this is a univariate integral, it can be shown to be given by

$$H_{i1}(z) = \phi(h_{i1p}(z)[-(p-1)[D_{i1}]h_{i1p}(z)/p\sigma^2 + [d^2h_{i1p}/dz_1^2]] \quad \text{(A7)}$$

The term $[D_{i1}]$ is a $w \times w$ symmetric matrix involving the differentials dh_{i1p}/dz_1, while $[d^2h_{i1p}/dz_1^2] = -2\beta I$ is the negative-definite Hessian of h_{i1p}. (Here we use I to denote the w by w identity matrix.) At the joint origin $(z_1, \mathbf{z}_{-1}) = (0, \ldots, 0)$, it is clear that

$$h_{i1p}((z_1, \mathbf{z}_{-1}) = \lambda_1 - [1/(p-1)] \sum_{j=2} \lambda_j$$

Now the matrix $D = \sum_i [D_{i1}]$ is a quadratic form, whose diagonal entry in the (t, t) cell is $\sum_i x_{it}^2$. As in the preamble to Theorem 1, let $\chi_t = (\ldots x_{it} \ldots)$ be the nth vector of the t coordinates of the bliss points. The diagonal terms in D are of the form $||\chi_t||^2$, while the off-diagonal terms in the (t, s) cell are scalar products (χ_t, χ_s). Note that, for each coordinate axis, $v_t^2 = (1/n)\sum_i x_{it}^2 = (1/n)||\chi_t||^2$ is the electoral variance on the t coordinate axis, while $v^2 = \sum_t v_t^2$ is the total electoral variance. We now use $\lambda_{\text{av}(1)} = [1/(p-1)]\sum_{j=2} \lambda_j$ as given in Definition 5.

From Equation A7, the condition that all eigenvalues of the Hessian of V_{1p} be negative gives the same requirement on the matrix

$$C = \beta\{(p-1)/[p\sigma^2]\}\{\lambda_{\text{av}(1)} - \lambda_1\}[4D] - 2nI$$

Writing $\Delta = (1/n)D$ for the *voter covariance matrix* and taking $\Lambda = (\lambda_{\text{av}(1)} - \lambda_1)$ gives the Hessian matrix

$$C(\beta, \lambda : \sigma^2, \Delta) = 2\beta\Lambda\{(p-1)/[p\sigma^2]\}\Delta - I$$

We now need to obtain the conditions under which the matrix C has negative eigenvalues. The argument is easiest in the two-dimensional case but can be pursued in any dimension. If we can show that the determinant, det[C], of C is positive and the trace, trace [C] of C is strictly negative, then both eigenvalues will be strictly negative.

Now let $A = \beta\Lambda\{(p-1)/[p\sigma^2]\}$, so $(1/2n)C$ has diagonal terms $(2A/n)||\chi_t||^2 - 1$ and off-diagonal terms $(2A/n)(\chi_t, \chi_s)$. Thus, $\det(H_1)$ can be determined from

$$(1/2n)^2 \det[C] = (2A/n)^2)[||\chi_s||^2||\chi_t||^2 - (\chi_t, \chi_s)^2] + 1 - (2A/n)[||\chi_t||^2 + ||\chi_s||^2] \quad \text{(A8)}$$

The first term is positive by the triangle inequality. Thus det[C] is strictly positive if

$$(2A/n)[||\chi_t||^2 + ||\chi_s||^2] < 1$$

Now $v_t^2 = (1/n)\sum_i x_{it}^2 = (1/n)||\chi_t||^2$ and $v^2 = [v_t^2 + v_s^2]$, so this condition is

$$2\beta\Lambda(p-1)v^2 < p\sigma^2 \quad \text{(A9)}$$

Moreover,

$$(1/2n)\,\text{trace}([C]) = (2A/n)[||\chi_t||^2 + ||\chi_t||^2] - 2$$

and this is negative if

$$2\beta\Lambda(p-1)v^2 < 2p\sigma^2 \quad \text{(A10)}$$

Obviously Equation A9 implies Equation A10. By Definition 5, $c = c(\beta, \lambda : \sigma^2, \Delta) = 2\beta\Lambda v^2(p-1)/[p\sigma^2]$. Thus, the condition $c < 1$ is *sufficient,*while the condition $c < 2$ necessary for C to have negative eigenvalues. Indeed,we have shown that the two eigenvalues a, b of $(1/2n)C$ satisfy the condition $a + b = c - 2$. Moreover, $ab = (1/2n)^2 \det[C]$ and this is positive if $c < 1$. Obviously when Equation A10 is satisfied then both a, b are strictly negative, so this condition is sufficient for the LSNE condition to be satisfied for agent 1. When Equation A9 fails, it is still possible for the eigenvalues to be negative. However, if Equation A10 fails, then both of the eigenvalues cannot be negative.

We can proceed in the same way for agent 2. Let $\lambda_{av(2)} = [1/(p-1)][\lambda_1 + \sum_{j=3}\lambda_j]$. Since $\lambda_2 \geq \lambda_1$, it follows that $(\lambda_{av(1)} - \lambda_1) \geq \lambda_{av(2)} - \lambda_2)$. Therefore, if $c < 1$, then the determinant of the Hessian for agent 2 will also have negative eigenvalues.

Thus, $z_2 = 0$ will be a local best response to $\mathbf{z}_{-2} = (0, \ldots, 0)$. For agents $j = 3, \ldots, p$ we can proceed in exactly the same way to show that the condition $c < 1$ is sufficient for $z_j = 0$ to be a local best response to $\mathbf{z}_{-j} = (0, \ldots, 0)$. Thus, the condition $c < 1$ is sufficient for an LSNE at the origin.

Suppose now that Equation A10 fails, so that $2\beta\Lambda(p-1)v^2 \geq 2p\sigma^2$. This is identical to the condition that $c \geq 2$.Then $a + b \geq 0$, so at least one eigenvalue, say a, must be non-negative. Consequently, one of the eigenvalues of the Hessian of $V_I(\mathbf{z})$ must be non-negative at the joint origin, so $z_1 = 0$ cannot be a strict local best response to $\mathbf{z}_{-1} = (0, \ldots, 0)$.

Thus, $\mathbf{z}^* = (0, \ldots, 0)$ cannot be an LSNE. This proves the condition $c < 2$ is necessary in two dimensions. Indeed, it is easy to see that the condition $2 \geq c$ is also necessary for $\mathbf{z}^*$ to be an LNE. The higher dimensional case follows in precisely the same fashion, by considering the determinant and trace of the w by w matrix C. □

Discussion

We can illustrate the theorem by considering the one- and two-dimensional cases:

(i) Consider the simplest case with $p = 2$ and $w = 1$. Then, from Equation A7, the Hessian is given by

$$H_1(\mathbf{z}) = (1/n)\sum_i H_{i1}(\mathbf{z})$$

$$= (1/n)\sum_i \phi(h_{i1p}(\mathbf{z})\{(p-1)[D_{i1}][-h_{i1p}(\mathbf{z})]/p\sigma^2 + [d^2h_{i1p}/dz_1^2]\}$$

Now $[D_{i1}] = 4(\beta x_i)^2)$ so $(1/n)\sum_i[D_{i1}] = 4\beta^2v^2$ where v^2 is the total electoral variance. At $\mathbf{z} = 0$, we see $h_{i1p}(\mathbf{z}) = (\lambda_1 - \lambda_2)$. Moreover, $[d^2h_{i1p}/dz_1^2] = -2\beta$. Finally, since ϕ is a pdf, its value is strictly positive. Then requiring that the Hessian be strictly negative gives the condition that

$$2\beta^2v^2(\lambda_2 - \lambda_1)/\sigma^2 < 2\beta \quad \text{or} \quad \beta(\lambda_2 - \lambda_1)v^2 < \sigma^2$$

Clearly this is necessary and sufficient for the Hessian of V_{1p} to be negative at $z_1 = 0$.

(ii) Computation of eigenvalues when $w = 2$. We have called the number $c = 2\beta\Lambda v^2\{(p-1)/[p\sigma^2]\}$ the 'convergence coefficient'. Notice that in empirical applications both the product $\beta\Lambda$ as well as the ratio $[v/\sigma]^2$ can be identified, so that c is certainly identifiable. When this coefficient is strictly less than 1, then we may say that 'strict local concavity' is satisfied at the origin. We can be sure that both eigenvalues are negative and the origin will be an LSNE. Conversely, as we shown, in the two-dimensional case, when the coefficient is greater than 2, then $z_1 = 0$ cannot be a strict best response to $(0, \ldots, 0)$ and the origin cannot be a strict LNE.

In the range $[1, 2]$, both eigenvalues may be negative. We can estimate the eigenvalues by using the product moment correlation coefficient, r^2, between the vectors χ_t, χ_s representing the voter ideal points on the two axes. By definition,

$$||\chi_s||^2||\chi_t||^2 - (\chi_t, \chi_s)^2 = (1 - r^2)||\chi_s||^2||\chi_t||^2$$

Using this we can solve the quadratic expression to obtain a, b. Thus, if we let $ab = 1 - c + y$ and $a + b = c - 2$, we can calculate the eigenvalues to be

$$a = (1/2)[c - 2 + [c^2 - 4y]^{1/2}] = A\{[v_t^2 + v_s^2] + [[v_t^2 - v_s^2]^2 + 4r^2v_t^2v_s^2]^{1/2}\} - 1$$

$$b = (1/2)[c - 2 - [c^2 - 4y]^{1/2} = A\{[v_t^2 + v_s^2] - [[v_t^2 - v_s^2]^2 + 4r^2v_t^2v_s^2]^{1/2}\} - 1 \qquad \text{(A11)}$$

Obviously for r^2 close to 1, the two vectors will be highly correlated and y will be small.Then b will be close to -1,while a will be close to $2A[v_t^2 + v_s^2] - 1 = c - 1$. As an illustration, in the Israel example for 1992, $c = 9.72$, while $a = 7.47$ and $b = +0.25$. The large positive eigenvalue is associated with a particular eigenspace, the major principal component of the electoral distribution.

When the two vectors, χ_t, χ_s are uncorrelated, so that r^2 is close to 0, then as Equation A11 shows, the term $+[c^2 - 4y]$ in the expression for a involves the electoral variance difference $[v_t^2 - v_s^2]$. Indeed, in this case, $a = 2Av_t^2 - 1$ and $b = 2Av_s^2 - 1$. When this difference term is small, then the two eigenvalues will be almost identical. When the difference between these two variance terms is sufficiently large, then the term in y will be large, so that the eigenvalue on the high variance axis will be positive but the eigenvalue on the second axis may be negative.

The one-dimensional case is the clearest because the local concavity condition is both necessary and sufficient.

The analysis can be extended to the general case of w dimensions but conditions for obtaining negative eigenvalues follow in an analogous fashion.

Note, for the case studied by Lin et al. (1999), with all $\lambda_j = 0$, that the local concavity condition is always satisfied. It follows that the joint origin can always be assured of being an LSNE. In the case with non-zero valences, however, if $\beta\Lambda$ is large or if σ is 'small' relative to v, so that c exceeds w, then the origin cannot be an LSNE. The two examples that we have just considered show that there are two separate situations: (i) if the product moment correlation is close to 1, then one axis is degenerate and the origin may be a saddle; and (ii) if the product moment correlation is close to 0, then the origin will either be a maximum or a minimum depending on the relative size of the electoral variances on the two axes.

Extension to the case of multivariate normal errors

When the errors have a non-diagonal covariance matrix Ω, then instead of seeking a solution to the matrix equation $B\Xi B^t = \sigma^2 I$ where $\Xi = \sigma^2 FF^t$, we must find a solution to the equation $BF\Omega(BF)^t = G$ where G is a diagonal matrix and B has the same properties as before. The matrix B will of course now depend on the error covariances but it still can be found because of the number of degrees of freedom associated with this matrix equation.

REFERENCES

Adams, James (2001) *Party Competition and Responsible Party Government*. Ann Arbor: University of Michigan Press.

Adams, James and Samuel Merrill III (1999) 'Modeling Party Strategies and Policy Representation in Multiparty Elections: Why are Strategies so Extreme?', *American Journal of Political Science* 43(3): 765–91.

Aldrich, John (1983a) 'A Spatial Model with Party Activists: Implications for Electoral Dynamics', *Public Choice* 41: 63–100.

Aldrich, John (1983b) 'A Downsian Spatial Model with Party Activists', *American Political Science Review* 77(4): 974–90.

Aldrich, John and Michael McGinnis (1989) 'A Model of Party Constraints on Optimal Candidate Positions', *Mathematical and Computer Modelling* 12: 437–50.

Alvarez, Michael and Jonathan Nagler (1998) 'When Politics and Models Collide: Estimating Models of Multi-candidate Elections', *American Journal of Political Science* 42(1): 55–96.

Ansolabehere, Stephen and James Snyder (2000) 'Valence Politics and Equilibrium in Spatial Election Models', *Public Choice* 103: 327–36.

Arian, Asher and Michael Shamir (1995) *The Election in Israel: 1992*. Albany: SUNY Press.

Banks, Jeff and John Duggan (2004) 'Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-Motivated Candidates', in David Austen Smith and John Duggan (eds) *Social Choice and Strategic Decisions:Essays in Honor of Jeffrey S. Banks*, pp. 1–41. Heidelberg: Springer.

Coughlin, Peter (1992) *Probabilistic Voting Theory*. Cambridge: Cambridge University Press.

Dow, Jay and James Endersby (2004) 'Multinomial Logit and Mutinomial Probit: A comparison of Choice Models for Voting Research', *Electoral Studies* 23(1): 107–22.

Duggan, John (2000) 'Equilibrium Equivalence under Expected Plurality and Probability of Winning Maximization',Typescript, Wallis Institute, Rochester University.

Enelow, James and Melvin Hinich (1984) *The Spatial Theory of Voting*. Cambridge: Cambridge University Press.

Groseclose, Timothy (2001) 'A Model of Candidate Location when one Candidate has Valence Advantage', *American Journal of Political Science* 45(3): 862–86.

Hinich, Melvin (1977) 'Equilibrium in Spatial Voting: The Median Voter Result is an Artifact', *Journal of Economic Theory* 16(2): 208–19.

Kass, Robert and Adrian Raftery (1995) 'Bayes' Factors', *Journal of the American Statistical Association* 90(3): 773–95.

Laver, Michael and Norman Schofield (1990) *Multiparty Government:The Politics of Coalition in Europe*. Oxford: Oxford University Press and Ann Arbor: University of Michigan Press.

Lin, Tse-Min, James Enelow and Han Dorussen (1999) 'Equilibrium in Multicandidate Probabilistic Spatial Voting', *Public Choice* 98: 59–82.

Merrill III, Samuel and Bernard Grofman (1999) *A Unified Theory of Voting*. Cambridge: Cambridge University Press.

Miller, Gary and Norman Schofield (2003) 'Activists and Partisan Realignment in the US', *American Political Science Review* 97(2): 245–60.

Poole, Keith and Howard Rosenthal (1984) 'US Presidential Elections 1968–1980: A Spatial Analysis', *American Journal of Political Science* 28(2): 283–312.

Quinn, Kevin, Andrew Martin and Andrew Whitford (1999) 'Voter Choice in Multiparty Democracies', *American Journal of Political Science* 43(4): 1231–47.

Quinn, Kevin and Andrew Martin (2002) 'An Integrated Computational Model of Multiparty Electoral Competition', *Statistical Sciences* 17(3): 405–19.

Rae, Douglas (1971) *The Political Consequences of Electoral Laws*. New Haven, CT: Yale University Press.

Riker, William (1982) 'The Two-party System and Duverger's Law:An Essay on the History of Political science', *American Political Science Review* 82(3): 753–66.

Riker, William (1986) 'Duverger's Law Revisited', in Bernard Grofman and Arend Lijphart (eds) *Electoral Laws and their Consequences*. New York: Agathon.

Schofield, Norman (2002) 'Representative Government as Social Choice', in Kenneth Arrow, Amartya Sen and Kotaro Suzumura (eds) *The Handbook of Social Choice and Welfare*, pp. 425–55. New York: North Holland.

Schofield, Norman (2003a) 'Valence Competition in the Spatial Stochastic Model', *Journal of Theoretical Politics* 15(4): 371–83.

Schofield, Norman (2003b) *Mathematical Methods in Economics and Social Choice*. Heidelberg: Springer.

Schofield, Norman (2004a) 'A Valence Model of Political Competition in Britain: 1992–97',Typescript, Washington University in St Louis.

Schofield, Norman (2004b) 'Local Political Equilibria', in David Austen-Smith and John Duggan (eds) *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, pp. 43–77. Heidelberg: Springer.

Schofield, Norman, Andrew Martin, Kevin Quinn and Andrew Whitford (1998) 'Multiparty Competition in The Netherlands and Germany: A Model based on Multinomial Probit', *Public Choice* 97: 257–93.

Schofield, Norman, Gary Miller and Andrew Martin (2003) 'Critical Elections and Political Realignment in the US: 1860–2000', *Political Studies* 51: 217–40.

Schofield, Norman, Itai Sened and David Nixon (1998) 'Nash Equilibrium in Multiparty Competition with "Stochastic" Voters', *Annals of Operations Research* 84: 3–27.

Schofield, Norman and Itai Sened (2002) 'Local Nash Equilibrium in Multiparty Politics', *Annals of Operations Research* 109: 193–210.

Schofield, Norman and Itai Sened (2004a) 'Multiparty Competition in Israel: 1988–96', *British Journal of Political Science*.

Schofield, Norman and Itai Sened (2004b) 'Modelling the Interaction of Parties, Activists and Voters: Why is the Political Centre so Empty', *European Journal of Political Research*.

Seyd, Patrick and Whiteley, Paul (2002) *New Labour's Grassroots*. Basingstoke: Macmillan.

Stokes, Donald (1992) 'Valence Politics', in Denis Kavanagh (ed.) *Electoral Politics*, pp. 141–62. Oxford: Clarendon Press.

NORMAN SCHOFIELD is the Director of the Center in Political Economy at Washington University in St Louis. His recent publications include *Mathematical Methods in Economics and Social Choice* (Springer, 2003) and he is currently completing a book on *Multiparty Parliaments*. In 2002–2003, Professor Schofield was the Fulbright Distinguished Professor of American Studies at Humboldt University, Berlin. ADDRESS: Center in Political Economy, Washington University in St. Louis, 1 Brookings Drive, St Louis, MO 63130, USA [email: schofld@wueconc.wustl.edu].