



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Electoral Studies 24 (2005) 347–370

Electoral
Studies

www.elsevier.com/locate/electstud

A valence model of political competition in Britain: 1992–1997[☆]

Norman Schofield*

*Center in Political Economy, Campus 1208 Washington University, 1 Brookings Drive,
St. Louis, MO 63130, USA*

Abstract

Formal stochastic models of voting have concluded that vote maximizing parties will converge to the mean of the electoral distribution. Much empirical evidence has concluded that such a situation is non-generic.

This paper presents an electoral theorem for the stochastic vote model with multivariate normal disturbances, which gives necessary and sufficient conditions for the validity of the mean voter theorem. The coefficients of a multinomial conditional probit model of the 1979 election in Britain are then examined. On the basis of the estimation, the model involving socio-demographic variables and “valence”, it is concluded that vote maximizing parties would converge to the electoral origin. Since this was not observed, the formal model is extended to include the influence of activists on voter response. A survey of British MPs preferred policy positions in 1997 is used to estimate the positions of party principals, and these are compared with voter estimations of party positions. A second electoral theorem is presented which determines the optimal party location involving a trade-off between principals’ preferences and electoral response.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Electoral models; Logit and probit estimation; Mean voter theorem

[☆] This paper is based on research funded by NSF grants SES 024173, SBR 98 18562 and SBR 97-30275, involving research collaboration with Andrew Martin, Gary Miller, David Nixon, Kevin Quinn, Robert Parks, Itai Sened and Andrew Whitford.

* Tel.: +1 314 935 4774; fax: +1 314 935 4156.

E-mail address: schofld@wueconc.wustl.edu

1. Formal and empirical analyses of elections

There is an extensive literature that attempts to model voters' choices in democratic elections. These models typically include socio-demographic characteristics of voters, as well as the "distance" between voters and parties, and the voters' evaluation of the relative success of party policies and the attractiveness or "quality" of party leaders.

In contrast, formal models assume that party strategists have some understanding of this "electoral response function", and are motivated to win elections by choosing policies that maximize vote shares or some other related desideratum. Unfortunately, the complexity of the electoral response has led formal theorists to focus on relatively simple models. One class of formal models that is, in principle, compatible with empirical analysis incorporates a stochastic component to voter choice. A standard result in such models is the "mean voter theorem" which asserts that all parties, wishing to maximize vote shares, should converge to the "electoral center", the mean of the distribution of voter-preferred points (Hinich, 1977). There is, however, an extensive literature which indicates that such convergence need not occur (see Adams, 2001; Adams and Merrill, 2003; Merrill and Grofman, 1999, for example).

This disjunction between theory and observation has led to numerous attempts to amend the formal theory in some way. One obvious way is to modify the basic structure of voter model, so that it is no longer defined in terms of distance between voter and party (Merrill and Grofman, 1999). However, empirical models, dating back to Poole and Rosenthal (1984), have found the simple spatial assumption to be warranted. The empirical models discussed here all exhibit a significant Bayes' factor, or difference in log likelihood functions (Kass and Raftery, 1995) associated with the spatial structure. It would appear necessary, therefore to leave untouched this feature of the formal model.

Early formal models assumed that the only relevant aspect of the voter calculus was the distance between party and voter. Recently the idea of "valence" has been utilized to refine the formal model. Valence of a party at the time of an election can be viewed as the average perception, among the electorate, of the "quality" of the party leader (Stokes, 1992; Ansolabehere and Snyder, 2000; Aragonés and Palfrey, 2003). In both empirical and formal models, valence is assumed to be independent of the current policy choice of the party. In this formulation, valence relates to voters' judgements about negatively or positively evaluated conditions that they associate with particular parties or candidates. Indeed, it may be based on voters' perception of the past behavior of the party leader, or on average voter estimation about how the party leader is likely to perform in the future (Penn, 2003).

Empirical models usually incorporate variables defined in terms of such evaluations by voters, as well as voter response to the "attractiveness" of leaders. In the empirical models discussed here, judgements of this kind are regarded as "exogenous valence". Because the formal model attempts to characterize individual choices by a utility calculus, it is in principle difficult, if not impossible, to ascertain the relative weight that an individual assigns to valence in comparison to policy difference or class based preferences. This difficulty is dealt with by assuming that the

judgemental weight assigned by each voter to the party j is given by an expected value or intercept λ_j , say, together with a disturbance or error, ε_j . The error term is assumed to be distributed throughout the electorate. The distribution of the error vector (... ε_j ...) is assumed multivariate normal.

An argument for including valence in the formal model can be based on the well-attested correlation between the electoral satisfaction rating of party leaders in Britain, for example, and the relative success of the parties at elections (Clarke et al., 1995, 1997, 1998, 2004; Seyd and Whiteley, 2002). Empirical electoral models have found that the addition of valence terms induces high Bayes' factors, even in models that also incorporate traditional socio-demographic variables such as class, education, income, religion etc. (Quinn et al., 1999; Quinn and Martin, 2002). However, the mean voter theorem has not been demonstrated for formal models involving valence.

Previous work by Quinn et al. (1999) and Schofield et al. (1998) suggested that the mean voter theorem was, for some unknown reason, inapplicable. These studies used Eurobarometer survey data (Rabier and Inglehart, 1981) to estimate a two-dimensional electoral model, as in Fig. 1, together with a similar survey (ISEIUM, 1983) of party delegates, to obtain a multinomial conditional probit (MNP) model of the 1979 election in Britain. The exercise was to determine the linkages between the socio-demographic characteristics of voters, the difference between voter's

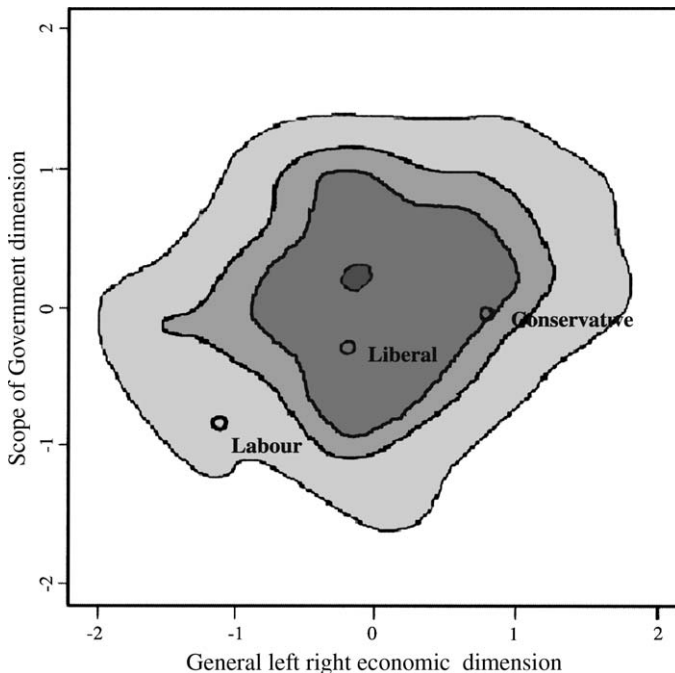


Fig. 1. Estimated party positions in the British Parliament in 1979, for a two-dimensional model (based on the European Election Study and Eurobarometer) showing the highest density plots of the voter sample, at the 95%, 75%, 50% and 10% levels.

preferences over policy and party position, and actual voter choice. They found that the superior model was a “spatial” MNP version that included valence terms for the parties, but excluded the socio-demographic characteristics, such as income, education, nature of occupation etc. The log marginal likelihood of this model was -393 , so the Bayes’ factor was $41 (= -393 + 434)$ over that of a pure socio-demographic model. Fig. 1 gives the estimated distribution of voter points (as obtained from the factor analysis of the survey). The two factors were very standard: the first a “general left right or economic dimension”, and the second a “scope of government” dimension (see Quinn et al., 1999 for details of the factor model).

The party positions were estimated by taking the median of the delegate positions on both axes. I shall refer to such an estimated party position as the position of the party “principal”. There was, of course, no guarantee that the principal’s position accurately reflected the perception in the electorate of the party position. Nonetheless, the predictive power of the model suggested that the principals’ positions closely corresponded to electoral perceptions.

It was unclear at the time whether the “mean voter theorem” was valid for this empirical model. The formal stochastic vote models that were available then and shortly later assumed that the stochastic errors were independently and identically, normally distributed (iind). Furthermore such models did not incorporate valence terms (Lin et al., 1999). Since the empirical MNP model assumed that the stochastic errors were multivariate normal, it was not possible to deduce whether or not parties could converge on the electoral center in order to increase votes. However, a simulation exercise for a similar MNP model for the 1981 election in the Netherlands did indicate that party convergence under vote maximization was likely under the conditions of the stochastic vote model (Schofield et al., 1998).

The formal result presented in the second section of this paper shows that this conclusion was correct. The First Electoral Theorem gives necessary and sufficient conditions for the validity of a version of a “mean voter theorem”, for a model involving valence and multivariate normal errors. The conditions are given in terms of a dimensionless “convergence coefficient”. For an empirical model, the coefficient is identifiable in terms of the parameters of the model and the dimension of the policy space. Using this result, it is shown that the *only* vote maximizing equilibrium for the MNP model of the 1979 election in Britain was one where all parties adopted the identical position at the mean.

This presents us with the paradox that the parties gave no indication of converging in this way.

In an attempt to develop a formal model that resolves this paradox, this paper will entertain the hypothesis that each party leader’s valence includes an exogenous component, but also depends on the behavior of activists for the party. I shall utilize and adapt an earlier model of “activist” valence due to Aldrich (Aldrich, 1983a,b; Aldrich and McGinnis, 1989; Miller and Schofield, 2003). In this model the contributions of time, effort and money by party members depend on the relationship between the party position, as chosen by the party leader, and the calculations of party members, given their preferred policy positions. In such an activist model, contributions affect the overall non-policy appeal of the party (by

allowing it greater access to the media, etc). Since valence will be, indirectly, a function of party location, the position where a party's vote share is maximized will depend on the activist contribution function. Moreover, the effect of activist effort on the party's appeal may exhibit "decreasing returns".

The mathematical consequence of this phenomenon is that the party's activist valence function will be concave. If the eigenvalues associated with this activist valence function are negative, and sufficiently great in magnitude, then the Hessian of the party's vote share function will also have negative eigenvalues at every position of the party. Under some conditions, it is possible that the vote share functions will be concave. This property is known to be sufficient for existence of a "Nash equilibrium". In this case a Nash equilibrium can be found by considering the first order conditions for vote share maximization.

The third section of this paper shows that the "first order condition" for the optimal position for each party will depend on balancing the effects of the exogenous valence of the party leader against the activist valence.

More precisely, the following can be shown to hold: other effects being equal, if the exogenous valence of the party leader falls, then the optimal party position will increase the weight afforded to the average activist preference. Conversely, if the exogenous valence of the party leader increases, then the optimal party position will decrease the weight given to activist preferences.

To examine the plausibility of this theoretical inference, a survey of the MPs of the various parties in Britain was carried out during the period of the election of 1997. As in the study for the 1979 election, the median of the estimated positions of responding party MPs was designated the "party principal position". It is argued that electoral perceptions of the party positions corresponded to the party principals' positions. The First Electoral Theorem is then used to show that these non-centrist party principals' positions could not be equilibrium in an electoral contest only involving exogenous valence. The Second Electoral Theorem, involving activist valence, is then used to show that the equilibrium of this more general vote maximizing political game will not be centrist. The final section of the paper makes a number of remarks with regard to interpreting recent British political history. Developing the model further, by estimating the activist function and determining the effect of contributions on valence, could lead to a more complex, but more realistic, method of modeling party choice.

2. A valence model of elections

In the model with valence, the stochastic element is associated with the weight given by each voter, i , to the perceived "quality" or valence of the party leader. The data of the model is a set $\{x_i\}_{i \in N}$, of voter ideal points, where each x_i lies in a Euclidean policy space X dimension w , and the electorate, N , is of size n .

Each party, j , from the set $P = \{1, \dots, j, \dots, p\}$ of exogeneously determined parties chooses a policy, $z_j \in X$ to declare to the electorate at the time of the election. Let

$\mathbf{z} = (z_1, \dots, z_p) \in X^p$ be a typical vector of party declarations. Given \mathbf{z} , each voter, i , is described by a utility vector $\mathbf{u}_i(x_i, \mathbf{z}) = (u_{i1}(x_i, z_1), \dots, u_{ip}(x_i, z_p)) \in R^p$, where

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) - \beta_j \|x_i - z_j\|^2 + \theta_j^T s_i + \varepsilon_{ij}. \quad (1)$$

Here, λ_j is the exogeneous valence or average weight representing the perceived quality of the leader of party j . The terms $\{\varepsilon_{ij}\}$ are the stochastic errors, so we can regard $\lambda_{ij} = \lambda_j + \varepsilon_{ij}$ as the quality of leader j as perceived by voter i . We shall assume that the errors $\{\varepsilon_{ij}\}$ are dependent only on j , and not on i , though it is in principle easy to extend the model so that $\{\varepsilon_{ij}\}$ may differ for i belonging to different subsets of the electorate, N . We therefore write the errors as $\{\varepsilon_j\}$. The term $\mu_j(z_j)$ is the *activist valence*, given z_j . Notice that $\mu_j(z_j)$ is assumed to be independent of i , and thus does not involve the policy differences between the voter and the party. The term β_j is a positive constant. Poole and Rosenthal (1984) assumed that the constants depended on j , but we shall assume that β_j is independent of j , so $\beta_j = \beta$ for all j . Finally, $\| \cdot \|$ is the usual Euclidean norm on \mathbf{X} . Thus $\lambda_j + \mu_j(z_j) - \beta_j \|x_i - z_j\|^2 + \varepsilon_j$ represents the perception of j by voter i . The k -vector θ_j represents the effect of k different socio-demographic parameters (such as class, domicile, education, income) on voting for agent j , while s_i is a k -vector denoting the i th individual's relevant "socio-demographic" characteristics (we use T to denote transpose, so $\theta_j^T s_i$ is a scalar).

In the formal version of the model, I assume that the errors are multivariate normal, with a non-diagonal covariance matrix, Ξ . Because only error differences are relevant, we use the notation $e(j) = [\varepsilon_p - \varepsilon_j, \dots, \varepsilon_{j+1} - \varepsilon_j, \varepsilon_{j-1} - \varepsilon_j, \dots, \varepsilon_1 - \varepsilon_j]$ for the j th error difference vector. Note that $e(j)$ is $(p-1)$ -dimensional.

The covariance matrix of $e(j)$ is denoted Ω_j , which can be obtained from Ξ by the equation $\Omega_j = F \Xi F^T$. (Here F is just a difference matrix with respect to the j th error.) Let $\text{var}(\Omega_j)$ be the sum of the terms in the matrix Ω_j . Then the set of relevant variances for the model is given by the "corrected error variance terms" $\{\kappa_j^2 = \text{var}(\Omega_j)/[p-1]^2: j = 1, \dots, p\}$. In the simple case the errors are independently and identically distributed (iid), with variance σ^2 then $\Xi = I\sigma^2$, where I is the $(p-1) \times (p-1)$ identity matrix. In this case $\text{var}(\Omega_j) = p(p-1)\sigma^2$, and $\kappa_j^2 = \kappa^2 = p\sigma^2/[p-1]$ for all j .

In empirical models based on multinomial conditional logit (MNL) estimation, it is usual to assume that the errors are iid with type I extreme value distribution (log Weibull). See Adams (2001) for discussion. Because this assumption entails an "independence of irrelevant alternatives" (IIA) assumption, which may be unwarranted, more recent empirical analyses have assumed the errors are drawn from a multivariate normal distribution. In multinomial probit (MNP) estimation, it is assumed that the errors have the general covariance matrix Ξ , and the estimation problem is to obtain the error difference matrix Ω_1 , with respect to a particular specified party. (See Alvarez and Nagler, 1998; Quinn et al., 1999; Dow and Endersby, 2004, for further discussion on the difference between logit and probit estimation.)

In substantive terms, the exogeneous valence, λ_j , in both formal and empirical models, is intended to characterize voters' judgements about positively or negatively evaluated conditions that are typically associated with party leaders. These may concern a wide range of evaluative characteristics, such as competence, compassion, integrity or honesty. In empirical estimation the average perceived qualities of the party leaders could, in principle, be obtained from electoral surveys that aim to assess the "level of satisfaction" or the expectation of future competence (Penn, 2003) associated with each party leader. Obviously, the weight given by each voter to this perceived "quality" of party j may depend on individual voter characteristics (notably education, domicile, income etc.). Quinn et al. (1999) found, by comparison of the Bayes' factors (Kass and Raftery, 1995) that including these individual characteristics in vote models for Britain and the Netherlands did little to increase the statistical validity of the estimation. As indicated in Eq. (1), the formal model, with exogeneous valence, simply assumes that the weight associated with party j is randomly distributed (and modeled by the error term ε_j) in the electoral population and has expectation equal to λ_j .

Because of the assumption of stochastic errors, individual voter behavior is modeled by a probability vector. Thus the probability that a voter i chooses a party j is

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

In formal and empirical models, this probability is derived from the assumption made on the errors. If these are multivariate normal, then Eq. (2) involves a multivariate normal integral which is quite difficult to compute analytically. For empirical models it can be estimated by Markov Chain Monte Carlo (MCMC) methods (Chib and Greenberg, 1996).

It is assumed in the formal model that party j uses some method to compute and maximize its expected vote share $V_j(\mathbf{z}) = (1/n) \sum_{i \in N} \rho_{ij}(z)$ where N is the voter population of size n .

In the formal model, the equilibrium concept generally used in is that of "pure strategy Nash equilibrium" (PNE).

Definition 1. A vector $\mathbf{z}^* \in X^p$ is a PNE if and only if, for each $j \in P$, z_j^* is a "best response" for V_j with respect to $(z_1^*, \dots, z_{j-1}^*, z_{j+1}^*, \dots, z_p^*) = \mathbf{z}_{-j}^*$.

In other words, if the other parties adopt the positions given by \mathbf{z}_{-j}^* , then z_j^* maximizes V_j .

It is an easy exercise to show, in the formal model, with exogeneous valence alone (so $\mu_j = 0$, for all j) that the first order condition for maximizing each V_j can be satisfied by choosing $z_j^* = (1/n) \sum_i x_i$, the electoral mean. New coordinates may be chosen so that the mean is 0. However, satisfying the first order condition does not guarantee that $\mathbf{z}_0^* = (0, \dots, 0)$, the joint mean position is a PNE. It is usual to assume or verify that the party vote share functions $\{V_j\}$ are concave in z_j (Banks and Duggan, 2005).

Concavity is a strong condition. It can be guaranteed if the vote share function V_j has a second differential (or Hessian) which is everywhere negative semi-definite. As

Lin et al. (1999) note, a sufficient condition for this, in the iid case, is that the standard deviation, σ , of the errors be “sufficiently” large. The results to be presented here give necessary and sufficient conditions for this Hessian condition to hold at the origin. If the necessary condition fails, then concavity of the vote share functions will also fail, and existence of PNE will be doubtful. Because the Hessian condition is local, this paper will emphasize a local variant of the equilibrium concept. Existence of this local equilibrium can be guaranteed by ensuring that the Hessian be negative semi-definite at the equilibrium vector.

Definition 2. Assume that each expected vote function V_j is twice differentiable,

- i) A vector $\mathbf{z}^* = (z_1^*, \dots, z_p^*) \in X^p$ 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 $\mathbf{z}^* = (z_1^*, \dots, z_p^*)$ 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 $V_j(z_1^*, \dots, z_j^*, \dots, z_p^*) \geq V_j(z_1^*, \dots, z_j, \dots, z_p^*)$ for all z_j in U_j . Thus, z_j^* is a weak best response to $\mathbf{z}_{-j}^* = (z_1^*, \dots, z_{j-1}^*, z_{j+1}^*, \dots, z_p^*)$ in the neighborhood, U_j . We say z_j^* is a *weak local best response*.
- iii) A vector \mathbf{z}^* is a *strict LNE* if all eigenvalues, for the Hessian of each V_j at vector \mathbf{z}^* are negative. We say that z_j^* is a *strict local best response* to \mathbf{z}_{-j}^* . Use the acronym LSNE for a strict LNE.

To compute the Hessian of the vote share functions, we formally define the “electoral variance” as follows.

Choose a system of orthogonal axes indexed by $t = 1, \dots, w$ where w is the dimension. Without loss of generality we can choose the coordinate system so that $\sum_i(x_{it}) = 0$, for $t = 1, \dots, w$. Then $x^* = (\dots(1/n)\sum_i(x_{it}), \dots)$, so that in the new coordinate system $x^* = (0, \dots, 0) \in X$. We shall write this vector more simply as 0. Let $\chi_t = (\dots x_{it} \dots) \in R^n$ be the n vector whose components are the new coordinates of the voter ideal points about the origin on the axis t . Then the voter distribution is given by the vector $(\chi_1, \dots, \chi_t, \dots, \chi_w) \in R^{nw}$. This vector encodes all the electoral data. Now define (χ_t, χ_s) to be the scalar product of the two vectors, χ_t and χ_s and consider the $w \times w$ matrix, D , whose entry in the (t, s) position is (χ_t, χ_s) . Then $v_t^2 = (1/n) \sum_i (x_{it})^2 = (1/n)(\chi_t, \chi_t) = (1/n)\|\chi_t\|^2$ is the variance of the voter ideal points about the origin on the t axis. This is simply the diagonal term in position t in the matrix $\Delta(0) = (1/n)D$. The off-diagonal term in $\Delta(0)$ is clearly the covariance between the vectors χ_t and χ_t . The matrix $\Delta(0)$ is therefore the *voter covariance matrix* about the origin. We use $v^2 = \sum_t v_t^2$ to denote the sum of these voter variance terms across all w dimensions. This is the sum of the diagonal terms in $\Delta(0)$, and is also known as the trace of the matrix $\Delta(0)$.

We use $r_{ts}^2 = |(\chi_t, \chi_t)|^2 / \|\chi_t\|^2 \|\chi_s\|^2$ to denote the correlation coefficient between the electoral vectors on the two axes, t and s .

The “principal components” of the electoral data are embodied in $\Delta(0)$, and these can be expressed in terms of the various correlation coefficients between the electoral axes. In the case of iid errors, we normalize with respect to the corrected error variance κ^2 by defining $\Delta^*(0) = \Delta(0)/\kappa^2$.

For the more general case, we define $\Delta(z_j)$ to be the covariance matrix taken about the vector z_j rather than 0, and let $\Delta_j^*(z_j) = \Delta(z_j)/\kappa_j^2$.

When only exogeneous valence is involved then the formal model is given as follows:

Definition 3. If $\mu_j = 0$, for all j , but the exogeneous valence terms are given by the vector $\lambda = (\lambda_p, \lambda_{p-1}, \dots, \lambda_1)$ and these are ranked $\lambda_p \geq \lambda_{p-1} \geq \dots \geq \lambda_1$, then the formal model with covariance error matrix Ξ is denoted $M(\beta, \lambda; \Delta^*(0), \Xi)$.

For this model define $\lambda_{av(1)} = [1/(p - 1)] \sum_{j=2}^p \lambda_j$ and let $A_1 = \lambda_{av(1)} - \lambda_1$ be the *valence difference* with respect to party 1.

To obtain the second order conditions for an equilibrium it is necessary to determine the eigenvalues of the Hessian, and these can be expressed in terms of a “convergence coefficient”.

Definition 4.

- (i) For the model $M(\beta, \lambda; \Delta^*(0), I\sigma^2)$ define the *Hessian for party 1* at \mathbf{z}_0^* to be

$$C = C(\beta, \lambda : \Delta^*(0), I\sigma^2) = 4\beta^2 A_1 [\Delta^*(0)] - 2\beta I, \tag{3}$$

where I is the $w \times w$ identity matrix.

- (ii) The *convergence coefficient* of the model is

$$c = c(\beta, \lambda : \Delta^*(0), I\sigma^2) = 2\beta A_1 [v/\kappa_1]^2 \tag{4}$$

It should be noted that the coefficient c is “dimensionless” which implies that, in empirical models, it is identifiable. Notice that the socio-demographic parameter vector θ_j and the individual socio-demographic descriptors $\{s_i\}$ are assumed not to be influenced by party positions. However, when these are included in the model, then the exogeneous valence terms will be effected. The formal results assume that these socio-demographic terms are held constant.

Schofield (2004) had proved the following **Lemma 1** and **Theorem 1** for the case with multivariate normal but identically and independently distributed errors. **Theorem 1** can be extended to the general multivariate normal model $M(\beta, \lambda; \Delta^*(0), \Xi)$, but the valence difference given in **Definition 3** has to be weighted by the error covariances. In more general case it is necessary to examine the Hessians for each vote share function, and these will involve $\Delta_j^*(0)$ rather than $\Delta^*(0)$. See also Schofield (2005 a,b).

Lemma 1. *The joint mean vector \mathbf{z}_0^* is a CNE of the model $M(\beta, \lambda: \Delta^*(0), I\sigma^2)$.*

Electoral Theorem 1. Existence of a local equilibrium at \mathbf{z}_0^* in the model

$M(\beta, \lambda: \Delta^*(0), I\sigma^2)$:

- (i) *Suppose that the policy space X is a closed bounded domain in Euclidean space. Then $\mathbf{z}_0^* = (0, \dots, 0)$ is an LSNE if and only if $C = C(\beta, \lambda: \Delta^*(0), I\sigma^2)$ has negative eigenvalues.*
- (ii) *If X is a closed bounded domain in Euclidean space of dimension w , then $\mathbf{z}_0^* = (0, \dots, 0)$ is an LSNE if $c < 1$ and is an LSNE only if $c < w$.*
- (iii) *If X is two-dimensional, then the eigenvalues of C are given by*

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

and

$$b = A \left\{ [v_t^2 + v_s^2] - [(v_t^2 - v_s^2)^2 + 4r^2 v_t^2 v_s^2]^{1/2} \right\} - 1. \tag{5}$$

In Eq. (5), $A = \beta A_1 / [\kappa_1]^2$, while v_t^2, v_s^2 are the variance terms associated with the two vectors χ_t and χ_s and r^2 is the correlation coefficient between these two electoral vectors.

As can be seen from Eq. (5), if $r^2 = 0$ then the eigenvalues will be given by $(a, b) = (2Av_t^2, 2Av_s^2)$ and these will be proportional to v_t^2, v_s^2 , respectively. With a non-zero electoral covariance, the larger eigenvalue will be associated with a “principal component” of the electoral data, determined by the eigenspace of this eigenvalue.

The necessary condition for existence of an LSNE can also be called as *the local concavity condition*. It follows immediately from Part (ii) of the theorem that $c \leq w$ is a necessary condition for \mathbf{z}_0^* to be an LNE, while $c < 1$ is also a sufficient condition for \mathbf{z}_0^* to be an LNE. When the sufficient condition fails, but the necessary condition holds, then Eq. (5), or the appropriate higher dimension analogue, can be used to compute the eigenvalues.

Theorem 1(i) is proved by showing that C , given in Eq. (3) is indeed the Hessian of the vote share function for party 1 at the vector \mathbf{z}_0^* . Moreover, if the Hessian of V_1 has negative eigenvalues at the vector \mathbf{z}_0^* , then so do V_2, \dots, V_p . The sufficient condition of Part (ii) follows by consideration of the determinant of C , while the necessary condition follows from examining the determinant of C .

Notice that in empirical models, the scale of the model is indeterminate, so the valences, themselves, cannot be identified. To identify the valence difference, it is necessary to normalize the error variances. In MNP models this can be done by choosing the variance difference term $\text{var}(\varepsilon_p - \varepsilon_1)$ to be 1.0, and defining the other

covariances in the matrix Ω_1 relative to $\text{var}(\varepsilon_p - \varepsilon_1)$. In MNL models, the error covariance is irrelevant. Instead the error variances, σ^2 , are set at $\pi^2/6 = 1.6449$. By this normalization, for the MNL model, $\kappa_1^2 = \kappa^2 = p\sigma^2/[p - 1]$. To identify β we can normalize the electoral variance, for example by choosing $v_1^2 = 1.0$. Clearly, the convergence coefficient will then be identifiable.

Theorem 1 can be used to determine conditions for existence of an LNE in MNL models, involving exogeneous valence. Although MNL models are based on the “log Weibull” extreme value distribution assumption, with variance $\sigma^2 = 1.6449$ on the errors, this distribution is “close” to the normal distribution with covariance matrix $\Xi = I\sigma^2$. Thus if we use an MNL model to estimate the various parameters, we can use **Theorem 1** to ascertain whether these parameter values are compatible with an equilibrium at the joint origin, under the assumption that the correct voting model is close to the MNL model. In this way we can address the discussion over whether convergent equilibria can exist in situations where the empirical electoral model can be assumed to be MNL (Adams, 1999a,b, 2001; Adams and Merrill, 1999, 2000; Staum, 2003).

We shall use the formal result to determine whether there exists an LNE at the origin based on the MNL model of the British election of 1979, introduced in Section 1 of this paper. As mentioned above, the voter distribution was obtained from the Rabier and Inglehart (1981) Eurobarometer Survey while the party principals’ positions were obtained by taking the two-dimensional median of estimated positions of party delegates, using data taken from the European Parties Middle Level Elites (ISEIUM, 1983). See Quinn et al. (1999) for details. The electoral variances were 0.605 on the first axis and 0.37 on the second, with negligible covariance. On the basis of an MNL model incorporating socio-demographic variables, such as religion, income and education, the estimated valence difference, between the Conservative Party on the one hand, and the average of the Liberal and Labour parties, was +0.365. The spatial coefficient, β , was 0.272. Using the value of 1.6449 for the variance of the extreme value distribution, the corrected variance, κ_1^2 , was estimated to be 2.4675, and the convergence coefficient computed to be 0.08. The eigenvalues on the two axes as derived from **Theorem 1** were calculated to be -0.97 and -0.95 , almost identical to -1.0 .

According to the definition of voter utility in Eq. (1) the socio-demographic effect on voting is not influenced by party position. It therefore follows that the theorem can be applied to demonstrate that all parties should have converged to the electoral mean. Even using the upper value of β of 0.35 from its 95% confidence interval and the upper value of $A_1 = 1.22$, this upper estimate of the coefficient was 0.34. Using these upper estimates, the two eigenvalues of the Hessian of the low valence party at the joint origin could be calculated to be -0.87 and -0.79 .

Quinn et al. (1999) also showed that the MNL model *without* socio-demographic variable had a small Bayes’ factor of 2.3 over the MNL model *with* these variables. For this model without such variables, the valence difference is naturally larger, and the coefficient could also be calculated to be $c = 0.34$, with the eigenvalues for the low valence party computed to be -0.87 and -0.79 again. Calculations for the MNP models gave very similar results.

These estimates give strong evidence that the joint origin was an LSNE, conditional on the validity of the MNL and MNP electoral models employed.

It is important to note that though the theorem merely shows that the joint origin is an LSNE, the fact that the eigenvalues were negative and of large modulus, relative to the parameters of the model, indicates that the joint origin was a PNE. Simulation of these MNL and MNP models has led to the inference that vote maximizing trajectories would lead to the joint origin, so the origin can be called an “attractor” of the vote maximizing game (Schofield et al., 1998). This implies that random learning by the parties would lead them into the electoral origin.

Obviously, there is no direct evidence that the parties were indeed located at, or close to, the principals’ positions in Fig. 1. However, using these positions, the log marginal likelihood of the MNP model was -393 . Moreover, this model was superior to MNL and MNP models based solely on socio-demographic characteristics.

The model results, taken together with the theorem, suggest four possibilities.

- (i) The party locations were, in fact, close to the principals’ positions given in Fig. 1, and these corresponded to a different LNE of a vote maximizing game that was equivalent to the formal model with exogenous valence.
- (ii) The party locations were, in fact, close to the principals’ positions given in Fig. 1. This was not an LNE of the vote maximizing game, and the parties could have increased vote shares by moving to the electoral center.
- (iii) The parties were in fact much closer to the LNE at the electoral center.
- (iv) The parties were close to those given in Fig. 1, but were maximizing vote shares with respect to a different model of electoral response.

To each of these possibilities there is a ready response:

- (i) Computation of the eigenvalues for all parties at the origin were negative and of large modulus. This implies the gradients of their vote share functions point towards the origin, even at positions far from the origin. Consequently there can exist no LNE other than the origin for the MNL model.
- (ii) It seems implausible that party principals choose their party positions on the basis of their average preferred policy, without regard to the electoral consequences. The following section will propose a model where principals choose a compromise between the preferred policy and an electorally superior one.
- (iii) Although there is no direct evidence that the positions in Fig. 1 were close to those perceived by the electorate, the high log marginal likelihoods indicate that the estimated positions are similar to actual positions. In particular, it can be inferred that both Labour and the Conservative Parties chose positions that were far from the origin.
- (iv) Although there is no guarantee that the MNL model was the correct one, the log marginal likelihood indicates that it was close to the correct one. It is unlikely that party tacticians would use electoral data to construct an inaccurate model of electoral response. It is possible that tacticians were concerned to maximize the probability of winning rather than expected vote share. However

theoretical models, particularly for two party competition, indicate that these two maximands give similar equilibria (Duggan, 2000).

In the next section we shall modify the formal model presented in this section, so as to account for non-centrist equilibrium vote maximizing positions of the parties, by utilizing the notion of activist valence.

3. Activists and elections

It is obvious that the MPs of a party can, in some sense, be viewed as activists for the party, so information on the distribution of their policy preferences will give some indication of the nature of activist preferences.

To obtain an estimate of MP positions, the British National Election survey was sent to MPs in Westminster. Approximately 60 Labour (LAB) MPs, 16 Conservative (CON) MPs, 14 Liberal Democrats (LIB), 3 Ulster Unionists (UU), and single MPs of Plaid Cymru (PC-Welsh Nationalists) and the Scottish Nationalist Party (SNP) responded.

Fig. 2 presents an estimate of the underlying electoral distribution of preferred voter positions in the two-dimensional factor space obtained from an analysis of the 1997 British National Election Survey (details of the factor model and of the estimation procedure can be found in Schofield, 2004 and Schofield and Sened, 2005). The origin (0,0), in this figure is the mean of the sample ideal points. The outer contour line in Fig. 2 contains 95% of all voter-preferred points. Each responding MP was also positioned in this factor space, and the average for each party was computed. The “left-right” axis (pro-labor against pro-capital) represents the economic factor, while the “north-south” (or nationalism) axis describes attitudes to the European Union, and is oriented so that a pro-European response is to the “south” on this second axis.

As in Section 2, we use the term “party principal” to denote that representative member of each party who has an ideal position identical to the median MP position for the party on each axis. As a first approximation we may assume that the party principal stands for an average activist for the party. These party principal positions are also marked in Fig. 2. There may of course be some bias associated with the response of these MPs, to the degree that the principal positions are not representative of the party. However, in the MP responses there was no similarity between the distribution of Labour MP positions and those of Conservative party positions. Later, I shall argue that the dispersion of MP positions within each party is indicative of the dispersion of activists within each party.

Fig. 2 makes it clear that though the Labour principal position is fairly close to the origin on the economic axis, it is very much more pro-Europe than the average voter. In contrast, the Conservative Party principal position is far more radical on both axes than the average voter. Note also that the Conservative Party principal’s position is much further from the electoral mean than the Labour Party principal’s position.

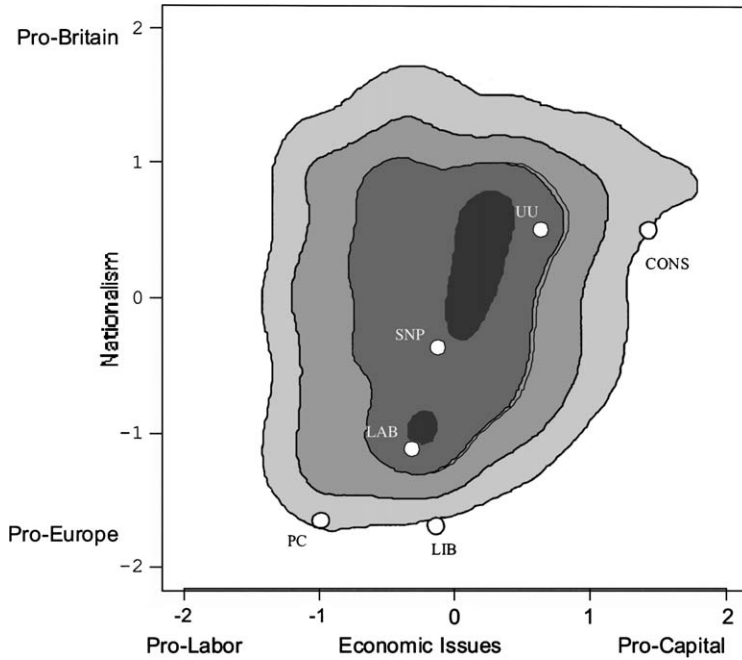


Fig. 2. Estimated party positions in the British Parliament for a two-dimensional model (based on MP survey data and a National Election Survey) showing highest density plots of the voter sample distribution at the 95%, 75%, 50% and 10% levels.

The equilibrium model of activism proposed by Aldrich (1983a,b) essentially assumed that the activists negotiated among themselves to agree to a collectively acceptable policy position. I shall adapt Aldrich’s model and suppose that the position of the party is chosen by the party principal to maximize vote share with regard to an estimated electoral response function that includes the indirect effects of activism on the party vote share.

We now examine whether a vector of positions for the three major parties (Conservative, Labour and Liberal Democrat) can be in equilibrium in a configuration that resembles the positions given in Fig. 2.

To do this, MNL models with exogenous valence, and no activist component, were estimated for the British 1992 and 1997 elections, using the British National Election Surveys for these elections. Instead of using MP responses, party positions were estimated by averaging voter perceptions on the single economic axis. The perceived positions of the three parties in 1997 were estimated to be the vector $\mathbf{z}_{97} = (z_{lab}, z_{lib}, z_{con})_{97} = (-0.20, +0.06, +1.33)$.

These positions are essentially identical to the projection of the three party principals’ positions in Fig. 2 onto the economic axis. In 1992 the estimated party positions were given by the vector $\mathbf{z}_{92} = (z_{lab}, z_{lib}, z_{con}) = (-0.64, -0.12, +1.11)$.

Fig. 3 gives the estimated electoral distribution and average perceived positions of the three parties in the one-dimensional policy space.

The estimated coefficients of the model in 1997 were $(\lambda_{con}, \lambda_{lab}, \lambda_{lib}, \beta)_{97} = (+1.24, +0.97, 0.0, 0.5)$. In 1992 these were $(\lambda_{con}, \lambda_{lab}, \lambda_{lib}, \beta)_{92} = (+1.58, +0.58, 0.0, 0.56)$.

Because the model was MNL, the variance was $\sigma^2 = 1.6449$ for the type I extreme value distribution. Since $p = 3$, the normalized variance was $\kappa^2 = p\sigma^2/[p - 1] = 2.47$. On the single economic dimension, the electoral variance for both elections was normalized to be 1.0. The average valence (of the Conservative and Labour parties) in 1997 was found to be +1.105. Since the Liberal Democrat party valence was normalized to be 0, the valence difference, for the Liberal Democrat party was given by the term $A_{97} = +1.105$. Thus the central value for the convergence coefficient was estimated to be +0.45, satisfying the sufficient condition of Theorem 1. Using the 95% confidence intervals of (0.44, 0.56) for β_{97} and (1.05, 1.25) for A_{97} the estimated confidence interval for c_{97} was (0.37, 0.56). Performing the same calculation for 1992 gave $c_{92} = +0.49$.

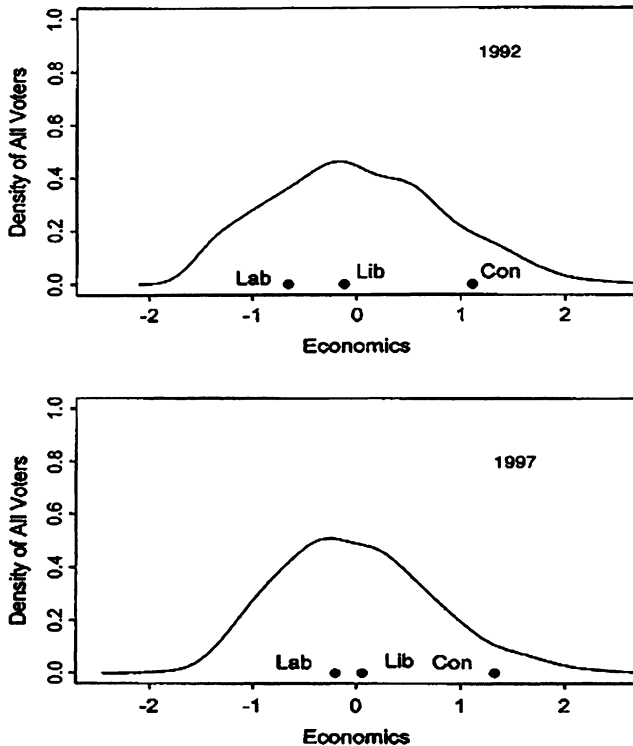


Fig. 3. Estimated electoral distributions for 1992 and 1997, together with average electorally perceived policy positions of the parties in the British Parliament, based on a one-dimensional factor model derived from the National Election Survey.

The log marginal likelihood of the 1997 MNL one-dimensional model with valence was -531 . This had a Bayes' factor of 75 over that of an MNL model without valence. This indicates valence was relevant for the estimation. A similar log marginal likelihood was obtained for 1992.

For 1997, the eigenvalue of the Liberal Democrat party Hessian of the vote share function, at the origin, when other parties are located there, was -0.55 . Correspondingly, the Conservative Party eigenvalue at the origin was -1.30 , while the Labour Party eigenvalue was -1.14 . As before, this implies that the origin is an attractor under the one-dimensional MNL model with exogenous valence. Moreover, it can then be verified that the first order condition for an equilibrium at $\mathbf{z}_{97} = (z_{lab}, z_{lib}, z_{con})_{97} = (-0.20, +0.06, +1.33)$ was not satisfied.

This MNL model did not include socio-demographic variables, such as class and education. These may have increased the predictive power of the model. However, including these variables necessarily reduces the estimated valence difference. This would have the effect of decreasing the valence difference between the Liberal Democrat party and the other parties, and the effect would be to further lower the convergence coefficient. Although the estimated exogenous valence of the Labour Party rose between 1992 and 1997, while that of the Conservative Party fell, this had no effect on the conclusion that the origin was an attractor for the one-dimensional MNL model.

We can use these conclusions about the one-dimensional model to draw inferences about the existence of an equilibrium in the two-dimensional model for the electoral data given in Fig. 2. As Eq. (5) indicates, since the covariance of the electoral distributions on the two factors of Fig. 2 is negligible, the eigenvalues on each axis will only depend on the electoral variance on that axis. Because the electoral variance on the second "European" axis is higher than on the first "economic" axis, the eigenvalue for the Liberal Democrat party on this axis, when all parties are at the origin, will be positive. Under vote maximization, this party should vacate the origin. Whether it should move "north" or "south" is indeterminate. The point denoted "LIB" in Fig. 2 is compatible with this inference. If the Liberal Democrat party moves away from the origin on the second axis, then the first order condition equilibrium for the other two parties will no longer hold. However, the negative eigenvalues on the first axis should constrain all parties to align themselves along the north-south axis, but away from the origin on the "European" axis. Indeed, we should expect the Labour and Conservative Parties to locate themselves opposite one another on the "European" axis. While the position of the Labour Party in Fig. 2 is compatible with this inference, the Conservative Party position is not.

These observations suggest that the simple vote maximization model with exogenous valence is inadequate to explain party positions in the 1997 election. We shall extend the model to include the effects of activist valence. We now use the more general form of voter utility as given in Eq. (1) where the expression includes the activist valence functions, $\{\mu_j\}$, one for each party. We denote this model by $M(\beta, \lambda, \mu; \Delta^*(0), I\sigma^2)$. Lemma 2 gives the first order condition for equilibrium in such a model (see also Schofield, 2003).

Lemma 2. *The first order condition for equilibrium in $M(\beta, \lambda, \mu; \Delta^*(0), I\sigma^2)$ is given by the condition that, for all j :*

$$z_j^* = d\mu_j^*/dz_j + \sum_i \alpha_{ij}x_i. \tag{6}$$

Here the coefficients α_{ij} are strictly increasing in λ_j and $\mu_j(z_j)$ but decreasing in $(\lambda_k; k \neq j)$. The function μ_j^* is obtained from μ_j by normalizing, by dividing by 2β .

Proof. First define $\lambda_{av(j)} = [1/(p - 1)] \sum_{k \neq j} \lambda_k$, $\mu_{av(j)}(\mathbf{z}_{-j}) = [1/(p - 1)] \sum_{k \neq j} \mu_k(z_k)$ and $A_j^*(\mathbf{z}) = (\lambda_{av(j)} - \lambda_j) + (\mu_{av(j)}(\mathbf{z}_{-j}) - \mu_j(z_j))$. These are more general analogues of the valence difference for party j , as given in Definition 4, at a party position vector, \mathbf{z} . It can be shown that the first order condition for equilibrium at \mathbf{z} has the form

$$\sum_i \alpha_i(x_i, \mathbf{z}_{-j}) [\varphi^*(A_j^*(\mathbf{z})) [2\beta(x_i - z_j) + d\mu_j/dz_j]] = 0. \tag{7}$$

Here φ^* is the probability density function of the normal distribution, while $\alpha_i(x_i, \mathbf{z}_{-j})$ are coefficients. In the simpler case without activist valence, Eq. (7) reduces to $\sum_i (x_i - z_j) = 0$, which shows that the vector \mathbf{z}_0^* is a CNE. In the more general case considered here, Eq. (7) can be written as $\sum_i \delta_{ij} [(x_i - z_j) + d\mu_j^*/dz_j] = 0$. Thus $\sum_i \delta_{ij} [z_j - d\mu_j^*/dz_j] = \sum_i \delta_{ij} x_i$. Now normalize by setting $\alpha_{ij} = \delta_{ij} / \sum_i \delta_{ij}$, so $[z_j - d\mu_j^*/dz_j] = \sum_i \alpha_{ij} x_i$. The new normalized coefficients $\{\alpha_{ij}\}$ are decreasing in A_j^* and thus increasing in λ_j and μ_j . Hence the first order condition is that each z_j^* satisfies the expression $z_j^* = d\mu_j^*/dz_j + \sum_i \alpha_{ij} x_i$, where the coefficients are as stated in the lemma. □

We shall denote the vector $\sum_i \alpha_{ij} x_i$ by dV_j^*/dz_j and call it the “(marginal) electoral pull” due to exogenous valence.

Then the first order condition can be written as

$$dV_j^*/dz_j + d\mu_j^*/dz_j - z_j^* = 0 \tag{8}$$

The first term in this expression (the “marginal electoral pull”) is a gradient vector pointing towards the “weighted electoral mean”. (This weighted electoral mean is simply that point where the electoral pull is zero.) If λ_j is exogenously increased, then this vector increases in magnitude. The vector $d\mu_j^*/dz_j$ “points towards” the position at which the total activist effort is maximized. Call this vector the “(marginal) activist pull”. When Eq. (8) is satisfied for party j , say the electoral and activist pulls are “balanced”. If the activist effort for party j is a sufficiently concave function of z_j , then the second order condition (or the negative semi-definiteness of the Hessian of the “activist pull”) will guarantee that z_j^* locally maximizes the party vote share function. If this is true for all parties then the vector \mathbf{z}^* given by the solution of Eq. (8), for all j , will be an LNE.

Electoral Theorem 2.

- (i) Consider the electoral model $M(\beta, \lambda, \mu: \Delta^*(0), I\sigma^2)$, with both exogenous valences $\{\lambda_j\}$ and activist valence functions $\{\mu_j\}$. The first order condition for \mathbf{z}^* to be an LNE, under vote maximization, 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 the greater is 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) then the nearer will z_j^* be to the activist preferred position.
- (ii) If all activist valence functions are highly concave (in the sense of having negative eigenvalues of sufficiently great magnitude) then the solution given by Eq.(8) will be a PNE.

Proof.

- (i) Follows directly from Lemma 2.
- (ii) It can be shown that the Hessian for party j at the vector \mathbf{z} has the form

$$C_j(\mathbf{z}) = [\beta^2 A_j^*(\mathbf{z})][4\Delta^*(z_j)] - 2\beta I + d^2\mu_j/dz_j^2 \quad (9)$$

Here $A_j^*(\mathbf{z})$ is an expression involving position vectors and valences, as given in the proof of Lemma 2. The term $\Delta^*(z_j)$ is the voter variance matrix taken about the point z_j , normalized by dividing by the corrected error variance, κ^2 . Obviously if the eigenvalues of the second order term $+d^2\mu_j/dz_j^2$ are negative and of large modulus, then the Hessian $C_j(\mathbf{z})$ at a CNE will also be negative definite. Indeed, if μ_j is sufficiently concave, then the Hessian will be negative definite everywhere. Thus if the vote share functions for all j are “sufficiently” concave then all Hessians $C_j(\mathbf{z})$, at the CNE, \mathbf{z} will be negative definite. This is a sufficient condition for \mathbf{z} to be a PNE. \square

The model implicit in Theorem 2 considers the intrinsic trade-off between the motivation of activists to adopt a “radical” position, and the necessity that the party appeals to the typical voter. The divergence between the electoral mean and the principals’ positions for the two parties in Fig. 2 suggests that the activist or principal’s influence over perceived party position was much stronger for the Conservative Party than for the Labour party in the elections under consideration. Theorem 2 also allows for the existence of more than a single activist group, or principal, within each party.

Modeling intra-party bargaining between the party principals appears necessary because of the wide variation in the estimated preferred points of the MP’s of each party.

Approximately six of the 16 surveyed Conservative MP positions were highly opposed to Europe, and had positions “north” of the 95% electoral contour line,

given in Fig. 2. In contrast, 16 of the 60 surveyed Labour MP positions were very pro-Europe and lay “south” of the 95% electoral contour line, while six of the Labour positions were weakly anti-Europe, to the extent of lying north of the zero-position on this axis. Almost all the Liberal Democrat MP positions were strongly pro-Europe. The single members of Plaid Cymru and the Scottish Nationalist Party could easily have been members of Labour. The three surveyed members of the Ulster Unionist Party could also have been members of the Conservative Party. Obviously enough, the simplified model implied by Fig. 2 (in which there is a single set of coherent activist groups, represented by their principals) is inadequate for determining the full effect of intra-party bargaining. Implicitly, the statement of Theorem 2 refers to the situation when there is only one activist group, or principal. The variation in MP preferred points suggests that the model be extended to include the effect of multiple potential activist groups for each party. This extension can be done relatively easily.

Consider a “pro-British” principal (with ideal point at the position marked B in Fig. 4) and “pro-Capital” principal (marked C in the figure). The optimal Conservative position will be determined by a version of Eq. (8) which equates the “electoral pull” against the “activist pulls” generated by the two principals. The MNL model estimates that λ_{con} fell from 1.58 to 1.24 between 1992 and 1997. Thus the optimal position z_{con}^* , will be one where z_{con}^* is “closer” to the locus of points where the marginal activist pull is zero (i.e., where $d\mu_{con}/dz_{con} = 0$). This locus of points can be called the “activist contract curve” for the Conservative party. To generate this contract curve we may suppose that the utility function of the economic principal at the point (x, y) has the “ellipsoidal” form

$$u(x, y) = -[(x - s_1)^2/a^2] - [(y - t_1)^2/b^2] \tag{10}$$

Here (s_1, t_1) are the coordinates of the pro-capital point, C, and $a < b$, to indicate that the principal accords greater salience to the first axis. Similarly, suppose that the utility function of the anti-Europe principal has the form

$$u(x, y) = -[(x - s_2)^2/e^2] - [(y - t_2)^2/f^2] \tag{11}$$

Again (s_2, t_2) are the coordinates of the pro-British point, B, with $f < e$ to indicate that this principal affords greater salience to the second axis. The contract curve between these two principals is then given by the catenary equation

$$(y - t_1)/(x - s_1) = S(y - t_2)/(x - s_2) \tag{12}$$

In this equation the coefficient $S = (be/af)^2$, and it is this coefficient that determines the curvature of the locus of points satisfying the equation $d\mu_{con}/dz_{con} = 0$. Thus, the greater the salience ratio, (be/af) , and the lower the exogeneous valence λ_{con} , the further will the optimal Conservative position be from the weighted electoral mean.

A similar analysis can be performed for the Labour party. The “activist contract curve” given in Fig. 4, for Labour, is the locus of points satisfying the equation $d\mu_{lab}/dz_{lab} = 0$. This curve represents the balance of power between Labour

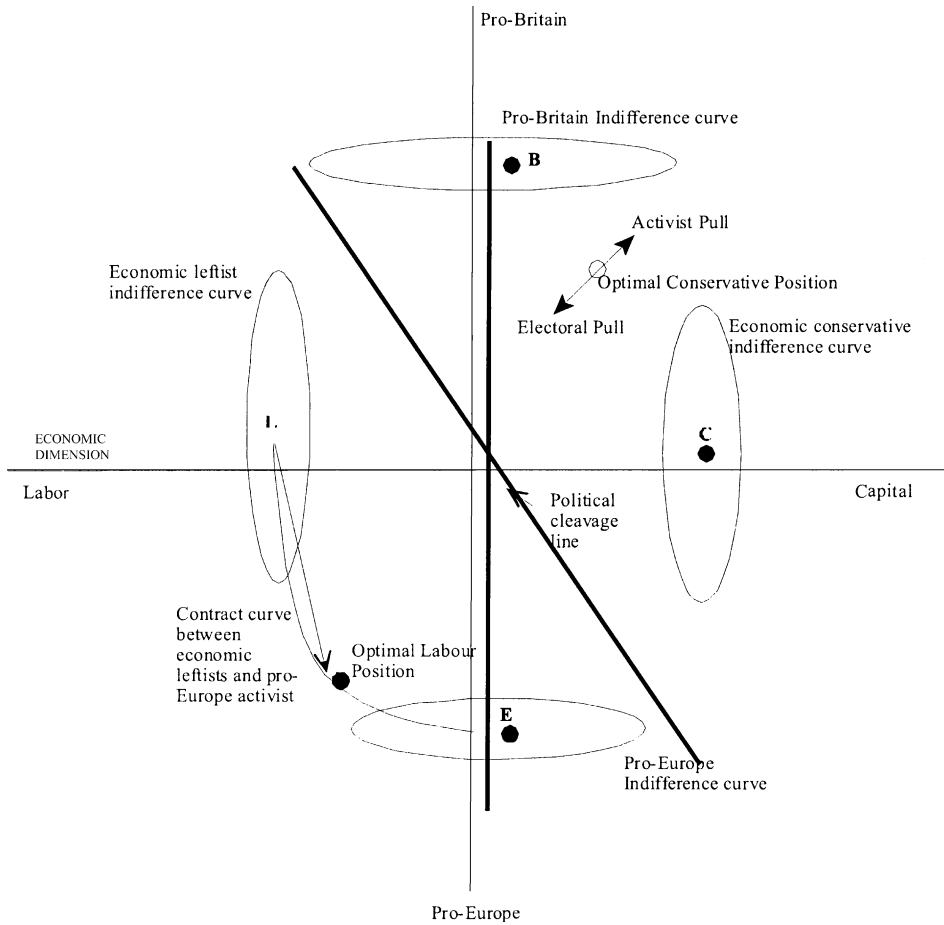


Fig. 4. Illustration of vote maximizing positions of Conservative and Labour Party Leaders in a two-dimensional policy space.

supporters most interested in economic issues (centered at L in the figure) and those more interested in Europe (centered at E).

Thus the optimal positions for the two parties will be at appropriate points on the locus between the respective “activist contract curves” and a point “near” the origin (where the electoral pull is zero). The “political cleavage” line in Fig. 4 is a representation of the electoral dividing line if there were only the two parties in the election. (See Miller and Schofield, 2003, where such a figure is used to discuss US politics.)

As Theorem 2 implies, if the relative exogenous valence for a party leader falls, then the optimal party position will approach the activist contract curve. Moreover, the optimal position on this contract curve will depend on the relative intensity, and the saliences of political preferences of the activists of each party. For example, if

grass roots “pro-British” Conservative party activists have intense preferences on this dimension, then this feature will be reflected in the activist contract curve and thus in the optimal Conservative position.

For the Labour party, it seems clear that two effects are apparent. Blair’s high exogenous valence in 1997 generated an optimal Labour party position that was closer to the electoral center than the optimal position of the Conservative party. Moreover, this affected the balance between pro-Labour or “old left” activists in the party, and “new Labour” activists, concerned to modernize the party through a European style “social democratic” perspective.

The logic of the model is that if the exogenous valence or perceived quality of one of the party leaders rises, in contrast to the exogenous valence of the opposing leader, then the optimal trade-off between activist preferred policies and the electorally advantageous position will be moved towards a more centrist position.

4. Discussion and conclusion

The model presented here may be relevant for understanding the changes in Labour party policy positions after John Smith became leader in 1992, and even more so, after Tony Blair took over in 1994. The model suggests that the low exogenous valence of the Labour leader, Neil Kinnock, prior to 1992, implied that Labour’s optimal electoral position would be dominated by the more radical party activists (Seyd and Whiteley, 1992). John Smith’s greater exogenous valence, by weakening the importance of the Left-wing activists, could induce a more “centrist” party position. Blair’s even greater exogenous valence meant that the influence of the Labour Party economic activists was sufficiently weakened, so that many of the traditional platforms of the Labour party could be abandoned.

As Seyd and Whiteley (2002) have observed, Blair also transformed activist support by efforts to bring new members into the party. Fig. 2 suggests that there is a high density domain of pro-Europe voters who are economically centrist, and who are candidates for “New Labour” activists. By changing the relative weight attached to these “New Labour” activists (whose average preferred position is denoted E in Fig. 4) the electorally optimal position for Labour became one which was much more centrist on the economic axis. It is compatible with this logic that the distribution of Labour MP positions also shifted somewhat, so that the overall average MP position also became more centrist on the economic axis.

The exogenous valence or perceived quality of the Conservative party leaders (in comparison first to John Smith and then to Tony Blair) appears to have declined over time. The MNL estimate, discussed previously, for the 1992 and 1997 elections, gave some evidence, albeit weak, that the overall Conservative valence dropped from 1.58 in 1992 to 1.24 in 1997, while the Labour valence increased from 0.58 to 0.96. These estimated valences include both the exogenous terms of the party leaders and effects due to the activist valence functions. Nonetheless, the data presented in Clarke et al. (1998, 2004) suggest that the Labour exogenous valence (λ_{lab}), due to Blair, rose in this period. Conversely, the relative exogenous term, λ_{con} for the

Conservative leader fell. Since the coefficients (in the equation for the electoral pull) for the Conservative party depend on $(\lambda_{\text{con}} - \lambda_{\text{lab}})$, these must all fall in this period. This has the effect of increasing the marginal effect of activism for the Conservative party.

Consequently, the optimal Conservative party position would move further away from the electoral center. The model proposed here suggests that the drop in the Conservative Party vote, from 41.9% in 1992 to 31.4% in 1997 followed from the combined effects of a decrease in the Conservative Party leader exogeneous valence, a rise in that of the Labour leader's, together with the increased influence of Conservative Party activists over the policy position adopted by the party. These observations appear compatible with Figs. 2 and 3.

The model suggests that the decreasing valence of the previous Conservative Party leader, Iain Duncan Smith, would increase the power of the two opposed principals of the party. Eventually this led to a leadership contest between the various sections of the party, and on October 30, 2003, Michael Howard became the new party leader.

In the above discussion, I have focused on the competition between the two more prominent parties and have not attempted to explain the low overall valence of the Liberal Democrat party. However, if Blair's exogeneous valence were to fall as a result of the international situation, as suggested by the local election results of June 2004, then the Liberal Democrat party might well move to a centrist position on the European axis as well, in order to make further electoral gains

The purpose in introducing the notion of "activist valence" has been to explore the possibility that the relationship between the party and the potential party activists will be affected by the exogeneous popularity of the leader. Party leaders can either exploit changes in their valence, or become victims of such changes. The theoretical framework can be used to provide an explanation for the seemingly irrational behavior of the Labour party during the period of Conservative government from 1979 until about 1992, and the similarly counter-intuitive Conservative Party policy choices over the last two elections.

Acknowledgements

I thank Martin Battle for research assistance, and Alexandra Shankster for assistance in preparing the figures. The anonymous referees made some very helpful comments. I would also like to express my appreciation for support from Washington University and from the Fulbright Foundation for the opportunity provided by my tenure of the Fulbright Distinguished Chair at Humboldt University, Berlin in 2002–2003.

References

- Adams, J., 1999a. Multiparty spatial competition with probabilistic voting. *Public Choice* 99, 259–274.
- Adams, J., 1999b. Policy divergence in multicandidate spatial voting. *Public Choice* 100, 103–122.

- Adams, J., 2001. *Party Competition and Responsible Party Government*. University of Michigan Press, Ann Arbor.
- Adams, J., Merrill III, S., 1999. Modeling party strategies and policy representation in multiparty elections: why are strategies so extreme? *American Journal of Political Science* 43, 765–791.
- Adams, J., Merrill III, S., 2000. Spatial models of candidate competition and the 1998 French presidential election: are presidential candidates vote-maximizers? *Journal of Politics* 62, 729–756.
- Adams, J., Merrill III, S., 2003. Centrifugal incentives in multicandidate elections. *Journal of Theoretical Politics* 14, 275–300.
- Aldrich, J., 1983a. A spatial model with party activists: implications for electoral dynamics. *Public Choice* 41, 63–100.
- Aldrich, J., 1983b. A Downsian spatial model with party activists. *American Political Science Review* 77, 974–990.
- Aldrich, J., McGinnis, M., 1989. A model of party constraints on optimal candidate positions. *Mathematical and Computer Modelling* 12, 437–450.
- Alvarez, M., Nagler, J., 1998. When politics and models collide: estimating models of multi-candidate elections. *American Journal of Political Science* 42, 55–96.
- Ansolabehere, S., Snyder, J., 2000. Valence politics and equilibrium in spatial election models. *Public Choice* 103, 327–336.
- Aragones, E., Palfrey T., 2003. *Spatial competition between two candidates of different quality: the effects of candidate ideology and private information*. Typescript, California Institute of Technology.
- Banks, J., Duggan, J., 2005. The theory of probabilistic voting in the spatial model of elections: the theory of office-motivated candidates. In: Austen-Smith, D., Duggan, J. (Eds.), *Social Choice and Strategic Behavior*. Springer, Heidelberg, pp. 1–41.
- British Election Study, 1992. National cross-section survey dataset. ESRC Data Archive, Essex University.
- British Election Study, 1997. National cross-section survey dataset. ESRC Data Archive, Essex University.
- Chib, S., Greenberg, E., 1996. Markov chain Monte Carlo simulation methods in econometrics. *Econometric Theory* 12, 409–431.
- Clarke, H., Stewart, M., Whiteley, P., 1995. Prime ministerial approval and governing party support: rival models reconsidered. *British Journal of Political Science* 25, 597–622.
- Clarke, H., Stewart, M., Whiteley, P., 1997. Tory trends, party identification and the dynamics of Conservative support since 1992. *British Journal of Political Science* 26, 299–318.
- Clarke, H., Stewart, M., Whiteley, P., 1998. New models for new Labour: the political economy of Labour support, January 1992–April 1997. *American Political Science Review* 92, 559–575.
- Clarke, H., Sanders, D., Stewart, M., Whiteley, P., 2004. *Political Choice in Britain*. Oxford University Press, Oxford.
- Dow, J.K., Endersby, J.W., 2004. Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral Studies* 23, 107–122.
- Duggan, J., 2000. *Equilibrium equivalence under expected plurality and probability of winning maximization*. Typescript, University of Rochester.
- Hinich, M., 1977. Equilibrium in spatial voting: the median voter result is an artifact. *Journal of Economic Theory* 16, 208–219.
- ISEIUM, 1983. *European Elections Study: European Parties' Middle Level Elites*. Europa Institut, Mannheim.
- Kass, R., Raftery, R., 1995. Bayes' factors. *Journal of The American Statistical Association* 90, 773–795.
- Lin, T.-M., Enelow, J., Dorussen, H., 1999. Equilibrium in multicandidate probabilistic spatial voting. *Public Choice* 98, 59–82.
- Merrill III, S., Grofman, B., 1999. *A Unified Theory of Voting*. Cambridge University Press, Cambridge.
- Miller, G., Schofield, N., 2003. Activists and partisan realignment. *American Political Science Review* 97, 245–260.
- Penn, E., 2003. *A model of farsighted voting*. Typescript, Carnegie Mellon University.
- Poole, K., Rosenthal, H., 1984. US presidential elections 1968–1980: a spatial analysis. *American Journal of Political Science* 28, 283–312.

- Quinn, K., Martin, A., Whitford, A., 1999. Voter choice in multiparty democracies. *American Journal of Political Science* 43, 1231–1247.
- Quinn, K., Martin, A., 2002. An integrated computational model of multiparty electoral competition. *Statistical Science* 17, 405–419.
- Rabier, J.-R., Inglehart, R., 1981. Eurobarometer II April 1979, The Year of the Child. Inter-University Consortium for Political and Social Research, Ann Arbor, MI.
- Schofield, N., 2003. Valence competition and the spatial stochastic model. *Journal of Theoretical Politics* 15, 371–383.
- Schofield, N., 2004. Equilibrium in the spatial valence model of politics. *Journal of Theoretical Politics* 16, 447–481.
- Schofield, N., 2005a. Local political equilibria. In: Austen-Smith, D., Duggan, J. (Eds.), *Social Choice and Strategic Decisions*. Springer, Heidelberg, pp. 43–77.
- Schofield, N., 2005b. Social choice and elections. In: Alt, J., Aldridge, J., Lupia, A. (Eds.), *A Positive Change in Political Science*. University of Michigan Press, Ann Arbor.
- Schofield, N., Martin, A., Quinn, K., Whitford, A., 1998. Multiparty competition in the Netherlands and Germany: a model based on multinomial probit. *Public Choice* 97, 257–293.
- Schofield, N., Sened, I., 2005. Modelling the interaction of parties, activists and voters. *European Journal of Political Research* 39, in press.
- Seyd, P., Whiteley, P., 1992. *Labour's Grassroots*. Clarendon Press, Oxford.
- Seyd, P., Whiteley, P., 2002. *New Labour's Grassroots*. Macmillan, Basingstoke, UK.
- Staum, J., 2003. Ideological platforms and probabilistic voting. Typescript, Cornell University.
- Stokes, D., 1992. Valence politics. In: Kavanagh, D. (Ed.), *Electoral Politics*. Clarendon Press, Oxford, pp. 141–162.