Local Political Equilibria *

Norman Schofield

Washington University at St. Louis schofld@wueconc.wustl.edu

Summary. This article uses the notion of a "Local Nash Equilibrium" (LNE) to model a vote maximizing political game that incorporates valence (the electorally perceived quality of the political leaders). Formal stochastic voting models without valence typically conclude that all political agents (parties or candidates) will converge towards the electoral mean (the origin of the policy space). The main theorem presented here obtains the necessary and sufficient conditions for the validity of the "mean voter theorem" when valence is involved. Since a pure strategy Nash equilibrium (PNE), if it exists, must be a LNE the failure of the necessary condition for an LNE at the origin also implies that PNE cannot be at the origin. To further account for the non-convergent location of parties, the model is extended to include activist valence (the effect on party popularity due to the efforts of activist groups). These results suggest that it is very unlikely that Local Nash equilibria will be located at the electoral center. The theoretical conclusions appear to be borne out by empirical evidence from a number of countries. Genericity arguments demonstrate that LNE will exist for almost all parameters, when the policy space is compact, convex, without any restriction on the variance of the voter ideal points or on the party valence functions.

1 Introduction

Many of the decisions of a society depend on choices of elected representatives, over social and property rights, taxation, government regulation, etc. These

* This article is based on research supported by NSF grant SES 0241732. I would also like to acknowledge the debt that I owe to Jeff Banks, for inspiration and motivation in my attempt to outline an equilibrium theory of politics. I am very grateful to Lexi Shankster for preparing the figures, and to Ben Klemens for his help in dealing with Scientific Word. John Duggan very kindly went through earlier versions of the paper and pointed out problems with the proof of the Theorem. Gaetano Antinolfi showed me how to use Mathematica to check some calculations.

representatives, or political agents, are comparable to the entrepreneurs of an economic system. How these political decisions are made, and the relationship between the decisions and the nature of economic equilibrium, are part of the discipline of political economy. To understand this relationship it is necessary to have an equilibrium theory of political choice that is comparable to the equilibrium theory of economics (Austen-Smith and Banks [8, 9]). In principle, it would then be possible to create a more general theory of political economy. It has, however, proved difficult to construct such a political theory. The general perspective I adopt is that it worthwhile attempting to use the very abstract ideas of topology that proved so useful in the earlier demonstration of the existence of economic equilibria by Dierker [18] and Smale [48]. These studies did not use the apparatus of convex analysis and fixed point theory as usual in economics, but rather the notions of critical and local equilibria taken from the qualitative theory of dynamical systems. While these equilibrium concepts are weaker than the usual "global" concepts, it is possible to prove the "generic" existence of local economic equilibria under very weak assumptions. In the last main section of this article, I shall offer a way of using the same tools to demonstrate generic existence of local political equilibria. A development of this technique would then provide a way of studying the political economy in a qualitative fashion. However, before discussing political equilibria, it is necessary to have a good idea about what is entailed in such a notion. The usual concept of equilibrium is the "global" one of Nash equilibrium (see Nash [31]), but proof of existence is extremely difficult because of the necessity of requiring "convexity", in some form, of the preferences of political agents (see Schofield [38]). For example, even in a very simple model, where it is assumed that representatives adopt positions to maximize votes, there is no obvious reason why the convexity assumption should be satisfied (see Schofield and Parks [42]). As an illustration of the applicability of the concept of local political equilibrium, this article will examine this model in detail, for the general case in which the number of political agents, and the dimension of the space of political decisions, are not arbitrarily restricted.

More specifically, we shall model the choices of the leaders of political parties as equilibria to a vote maximizing electoral game. Since the work of Downs [19], such models have tended to lead to the inference that representatives will adopt a position at the electoral mean (see Banks and Duggan [11] and Hinich [23] or near the electoral median (Banks, Duggan and Le Breton [13, 14] and McKelvey [27]). However, this conclusion seems to contradict extensive evidence that parties do not converge in this way. In electoral systems based on proportional representation (PR), there may be a large number of parties, located at very different non-centrist positions (see Schofield and Sened [43, 44]). In electoral systems based on plurality (or first past the post), either there is no party near the center, as in the U.S. (Miller [29], Schofield et al. [41], Poole and Rosenthal [32]), or, as in Britain, the centrist party is not a likely candidate for government (Schofield [36]).

The model presented here attempts to account for non-centrist choices of parties by extending the stochastic model of voting due to Lin, Enelow and Dorussen [26] to include " valence"(Stokes [49]). Valence here is assumed to have two different sources. It may either be an exogeneous effect due to the differences in the average perceived quality of the party leaders (see Ansolabehere and Snyder [6], Aragones and Palfrey [7], Groseclose [22], Schofield [36]), or more generally, valence can be endogeneous, due to the indirect effect of party leader choices on the willingness of activists to support the party by fund raising, *et cetera*. Such activist support allows a party leader to increase the electoral perception of the valence of the party , through the use of the media and by other means (see Aldrich [1, 2] and Aldrich and McGinnis [3]).

The chapter first models the situation where each agent (whether a candidate or party leader), j, is characterized by an average level of (exogenous) perceived competence (or quality), λ_j. That is to say, a typical voter, i, when making comparisons between agents j and k will compare their policies, z_j and z_k, and also their valences λ_j and λ_k. We assume that each agent adopts a policy position z_j so as to maximize the agent's own share of the popular vote. Because the agents do not know precisely what weight each voter gives to the perceived quality of the agent, the agent uses the "expected" vote share as a utility function.

This version is an extension of the standard "probabilistic" or stochastic vote model (Banks and Duggan [11] and Coughlin [17]). The addition of exogenous valence changes the result known as the "mean voter theorem."Although the first-order condition for vote maximization is satisfied when all agents adopt the mean voter position, it may be the case that one of the agents with a low valence will, in fact, minimize vote share at the mean. Consequently, the second-order (or Hessian) condition has to be considered.

A standard condition, or assumption, that is usually made for proof of existence of pure strategy Nash equilibrium (PNE) is "concavity" of the payoff function (in this case the vote share). Concavity is equivalent to the requirement that the Hessian, for the vote share function of each agent, be negative semi-definite on the entire strategy domain. Because this condition is unlikely to be met, I consider a weaker equilibrium concept, that of "local strict pure strategy Nash equilibrium" (LSNE). A vector of strategies is a LSNE whenever the appropriate Hessian for each agent is negative definite at that strategy vector. I shall refer to this condition as strict local concavity. Since it is generically the case that smooth functions do not exhibit degenerate critical points (with at least one zero eigenvalue of the Hessian) we may assert that, generically, a PNE, if it exists, is a LSNE. The converse, of course, is not true, since a LSNE may not be a PNE.

Theorem 1 gives the necessary and sufficient conditions for strict local concavity at the joint electoral mean position. These conditions can be expressed as upper bounds for a single convergence coefficient which is defined in terms of the parameters of the model, including the number of parties, the maximum valence difference, the dimension of the policy space, the stochastic variance σ^2, together with the variance of the set of voter ideal points. Empirical studies of elections in the Netherlands (Schofield et al. [40]) and in Israel (Schofield and Sened [44]) have obtained values for these parameters such that the necessary condition is violated. Consequently, there is no reason to expect convergence by parties to the electoral mean in these polities.

In the case when the policy space is unidimensional, then a corollary of Theorem 1 gives a readily computed necessary and sufficient condition for existence of a LSNE at the voter mean. This chapter uses related empirical work by Schofield [36], based on a unidimensional model, to show that this condition was satisfied in elections in Britain in 1992 and 1997. It may not prove surprising that the mean voter theorem is invalid for multiparty competition. In the U.S. for example, candidates must win primaries in order to compete in presidential elections, and it may be the case that a policy position which wins a primary is much more radical than a position which wins the presidential election. However, political parties in Britain do not have to face primaries. Moreover, the fact that there are two principal parties in Britain, competing under an electoral system based on plurality rule, implies that maximizing vote share is almost identical to maximizing the probability of winning the election (Duggan [20]). According to the empirical analysis, under the assumption of exogenous valence, all parties would have maximized their vote share by adopting an identical position at the voter mean in these two elections. The evidence offered in Schofield [36] and supported by Alvarez [5] is that the British parties were perceived by the electorate to have very different positions on the single economic policy axis.

To account for this commonly perceived divergence of party position, the model of British elections was developed in two directions. Firstly, the empirical analysis was extended to include a second dimension involving nationalism, in particular attitudes towards the European Union. The increase in the number of dimensions, and in the electoral variance in this policy space, implied that the necessary condition for strict local concavity at the electoral mean was violated. Although this explains the non-centrist position of the low valence Liberal Democrat Party, the analysis does not explain the location of the high valence Conservative Party, perceived to be far from the center on both axes. In an attempt to account for the extreme divergence of this party, the simple voting model with exogenous valence was extended to include activist support (Aldrich and McGinnis [3]). The activist model assumes that by contributing time and support to a preferred candidate, activists enhance overall voter support. Such activist support for a candidate is a function of the party or candidate position. Moreover, activist contributions to a party can be expected to exhibit decreasing marginal returns. It is plausible, therefore, that each party's activist valence function will be *concave* in the party strategy.

The most general model that I consider is one where both exogenous valence and activist valence are included. It is shown that the first-order condition for the optimal party position involves a trade-off or balance between what I shall call the marginal "electoral pull" and the marginal "activist pull." In other words, the electoral gradient vector and the activist gradient vector must be equal in magnitude, but in opposite directions, at an equilibrium position for each party. Another way of interpreting this result is that when exogenous valence falls, for whatever reason, then activist valence becomes more significant in determining the optimal position of the party. Moreover, if we assume that activist valence is indeed a concave function of strategy, then it is possible to determine conditions on the eigenvalues of the Hessian of each of the parties' activist vote-share functions, sufficient to guarantee that PNE do indeed exist.

The purpose of this article is to present the argument that, in the general spatial electoral model considered here, existence of PNE need not be guaranteed. However, transversality arguments, based on differentiabilty, can be used to show that LSNE will typically exist. Since the set of LSNE generically includes the set of PNE, it is possible to use simulation techniques in actual empirical situations to determine the set of LSNE, and by further analysis determine whether any of these LSNE are PNE. To determine whether PNE do indeed exist, it appears necessary to model, in a more complete fashion, the effect of activists on political support.

The notion of LSNE has not generally been used in modeling political phenomena. It is much more common when existence of PNE cannot be guaranteed, to use the notion of mixed strategy Nash equilibria (MNE). There seem to be two reasons to reject the applicability of the notion of MNE. First of all, as Banks, Duggan and LeBreton [13, 14] demonstrate, in two-party competition, the support of MNE will belong to a subset of the uncovered set (McKelvey [27]). It is conjectured that the uncovered set will tend to be small and centrally located with regard to the electoral distribution; see for example Schofield [35]. Consequently, models based on the MNE concept suggest that two party competition may lead to party locations that are "close" to the electoral center. In general, there is no empirical evidence for such a conclusion. A second objection is that party leaders are usually obliged to make policy pronouncements in the form of manifestos, etc. Empirical analyses with current techniques require that parties be precisely located and the stochastic models of electoral behavior, based on specific estimates of party position, typically give statistically significant empirical models of voter behavior (see Alvarez and Nagler [5], Poole and Rosenthal [32], Schofield [36] and Schofield et al. [41]). It would appear, therefore, that formal models, using the theoretically appealing notion of MNE, cannot be readily tested by empirical analysis.

The next section of the article introduces the stochastic model, and proves Theorem 1, and its two corollaries, on existence of electoral equilibrium at the electoral mean. Section 3 offers some examples based on the theorem. Section

4 extends the model to include activist valence and presents a preliminary result on optimal party strategy in this more complex game. The empirical example from Britain is given to illustrate the balancing of marginal "electoral pull" and "activist pull". A short Section 5 attempts to address possible shortcomings of the model, but argues for the usefulness of the notion of local Nash equilibrium. Since there are many possible alternatives to the assumption of vote share maximization, Section 6 considers the more general question of generic existence of local Nash equilibria when agent payoffs are differentiable. Section 7 concludes.

2 Nash Equilibria in the Voter Model

The situation we consider is that of a collection $P = \{1, \ldots, j, \ldots, p\}$ of political agents (whether candidates or parties). Each agent, j, chooses a policy z_j in a set X. Let $\mathbf{z} = (z_1, \ldots, z_p) \in X^p$ denote a strategy vector for the set of agents. The game form $h : X^p \to W$ maps from the set of strategy vectors to a space, W, of outcomes, on which the jth agent has a utility function $U_j : W \to \mathbb{R}$. The game is $\{U_j^h : X^p \to \mathbb{R}\}_{j\in P}$ or $U^h : X^p \to \mathbb{R}^p$, where $U_j^h(\mathbf{z}) = U_j(h(\mathbf{z}))$.

Definition 1. *A pure strategy Nash equilibrium (PNE) for the game $\{U_j^h\}_P$ is a vector $\mathbf{z}* \in X^p$ with the property that, for each $j \in P$, there exists no $z_j \in X$ such that*

$$U_j^h(z_1^*, \ldots, z_{j-1}^*, z_j, z_{j+1}^*, \ldots, z_p^*) > U_j^h(z_1^*, \ldots, z_{j-1}^*, z_j^*, z_{j+1}^*, \ldots, z_p^*).$$

A more general notion, that of mixed strategy Nash equilibrium (MNE) is similar, but considers strategies for each agent in a space of lotteries, or mixtures, defined over X. In the case where X is a compact, convex subset of a topological vector space, there are well known properties of U^h sufficient to gurantee existence (Austen-Smith and Banks [8]). Some of these focus on the properties of the underlying preferences induced by U^h on X. These are based on the Fan [21] Theorem. For example, quasi concavity and continuity of U_j^h are sufficient for existence of PNE. A function U is *quasi-concave* if

$$U(ax + (1-a)y) \geq \min[U(x), U(y)] \text{ for all } x, y \in W \text{ and } a \in [0, 1].$$

Concavity is a stronger property that also suffices. The function U is *concave* if

$$U(ax + (1-a)y) \geq aU(x) + (1-a)U(y), \text{ for every real } a.$$

In the topological category where X is a topological space, a "weaker"equilibrium concept is local Nash equilibrium.

Definition 2.

(i) A local pure strategy Nash equilibrium (LNE) for the game $\{U_h^j\}_P$ *is a vector* $\mathbf{z}^* \in X^p$ *with the property that, for each* $j \in P$*, there exists an open neighborhood* W_j *of* z_j^* *in* X*, such that, for no* z_j *in* W_j *is it the case that*

$$U_j^h(z_1^*, \ldots, z_{j-1}^*, z_j, z_{j+1}^*, \ldots, z_p^*) > U_j^h(z_1^*, \ldots, z_{j-1}^*, z_j^*, z_{j+1}^*, \ldots, z_p^*).$$

(ii) An LNE, $\mathbf{z}^*$*, is locally isolated if there exists a neighborhood* W *of* $\mathbf{z}^*$ *in* X^p *such that* $\mathbf{z}^*$ *is the unique LNE in* W*.*

In the differentiable category, where X has a differentiable structure and $\{U_j^h : X^p \to \mathbb{R}\}$ are smooth functions, then it is natural to use the weaker equilibrium concept of critical Nash equilibrium (CNE). We give the definition when X is a vector space. However, the definition is also applicable to the general case when X has a differential structure; that is, when X is a smooth manifold (see Hirsch [24]).

Definition 3. *Suppose that* X *is a compact topological vector space of dimension* w *with smooth boundary. Let* $U^h : X^p \to \mathbb{R}^p$ *be* C^1*-differentiable. Then* $\mathbf{z}^* \in X^p$ *is a critical Nash equilibrium (CNE) iff the first order vector equation* $\frac{\partial U_j^h}{\partial z_j}(\mathbf{z}^*) = 0$ *is satisfied for all* $j \in P$*.*

In this category, if all U_j^h are C^2-differentiable, then analysis of the second order Hessian conditions at $\mathbf{z}^*$ can be used to determine if the CNE $\mathbf{z}^*$ is a LNE. Excluding boundary situations, every LNE must be a CNE, and the Hessian of each $U_j^h : X^p \to \mathbb{R}$ must be negative semi-definite (with all eigenvalues non-positive) at $\mathbf{z}^*$, with respect to z_j. It is therefore natural to say that a strategy vector $\mathbf{z}^*$ is local strict Nash equilibrium, or LSNE, if and only if it is a CNE the and the Hessian of each U_j^h at $\mathbf{z}$ is negative definite. Consequently, if the eigenvalues of all Hessians are negative at a CNE, $\mathbf{z}^*$, then this vector will be a LSNE and therefore a LNE. Since every PNE must be a LNE, this can be used to determine existence of PNE. Since we focus on LSNE, we essentially ignore the case where one of the eigenvalues is zero.. In this situation, the determinant of the Hessian is zero, so that the CNE is degenerate. As we shall see, this corresponds to a knife edge property of the parameters of the model. More abstractly, the degeneracy of a critical point is a non-generic property in the space of C^2differentiable utility profiles, when this is endowed with the C^2-topology (see Hirsch [24]). It is obvious that a PNE can be characterized by a locally "flat" domain in utility space, so that such a PNE is not a LSNE. However, the generic situation is when a PNE is strict, in the sense that the payoff for each agent j at z_j^* is strictly greater than at any different z_j.Such a strict PNE is clearly also a LSNE.Thus the necessary condition for a LSNE is generically a necessary condition for existence of a PNE.

To examine LSNE, we first seek the condition for $\mathbf{z}^*$ to be a CNE, and then fix $\mathbf{z}^*_{-j} = (z^*_1, \ldots, z^*_{j-1}, z^*_{j+1}, \ldots)$. We then determine the conditions under which the Hessian of the j's vote share function is negative definite at z^*_j and then examine z_j in the jth strategy set. If the induced utility function for j is concave on this strategy set (with $\mathbf{z}^*_{-j}$ fixed) then the Nash equilibrium property holds for j at $\mathbf{z}^*$. Reiteration for each j gives a method of determining, at least generically, whether PNE exist. I shall use the term " strict local concavity" for the Hessian condition for $\mathbf{z}^*$ to be a LSNE. This technique allows us to determine those constraints on the parameters of the model which are necessary and sufficient for strict local concavity.This allows us to obtain a necessary condition for LNE , and thus for PNE.

The stochastic electoral model was originally developed for two-agent competition. In this case, it is natural to suppose that, for each agent j, $U^h_j(\mathbf{z}) = V_j(\mathbf{z}) - V_k(\mathbf{z})$ where k is the opposing agent to j, and V_j is j's "expected vote share function". Implicitly, such a model assumes a game form $h : X^2 \rightarrow W = [0,1]^2$, with $h(\mathbf{z}) = (V_j(\mathbf{z}), V_k(\mathbf{z}))$ where the prize for each agent j is the plurality over j's opponent. Banks and Duggan [11] give an extensive discussion of this symmetric zero-sum case. In addition to compactness and convexity of X, and joint continuity of $(V_j,\, V_k)$ they assumed a further property, "aggregate strict concavity"(a property on voter utility functions), and showed that then there would exist a unique symmetric PNE $(z^*_j,\, z^*_k)$ with $z^*_j{=}z^*_k$ equal to the mean of the voter ideal points.

One purpose of this article is to present an extension of the multiparty stochastic model of Lin, Enelow and Dorussen [26] by inducing asymmetries due to differing valences between the parties. As we shall see, this changes the "mean voter theorem", so as to bring its conclusions more into line with the empirical evidence as regards non-convergence. There are a number of possible choices for the appropriate game form for multiparty competition (see Schofield and Sened [43]). The simplest one, which is used here, is based on the assumption that the prize for agent j is proportional to V_j. With this assumption, we can examine the conditions on the parameters of the stochastic model which are necessary and sufficient for strict local concavity and thus for existence of strict LSNE. From this we can infer the necessary condition for existence of PNE.

The key idea underlying the formal model is that party leaders attempt to estimate the electoral effects of party declarations, or manifestos, and choose their own positions as best responses to other party declarations, in order to maximize their own vote share. The stochastic model essentially assumes that party leaders cannot predict vote response precisely. In the model with "exogenous" valence, the stochastic element is associated with the weight given by each voter, i, to the average perceived quality or valence of the party leader.

The data of the model is a set $\{x_i \in X\}_{i \in N}$ of "voter ideal points" for the members of the electorate, N, of size n. As usual we assume that X is a compact convex subset of Euclidean space, $\mathbb{R}^w$, with w finite.

Each party in the set $P = \{1, \ldots, j, \ldots, p\}$ chooses a policy, $z_j \in X$, to declare. Let $\mathbf{z} = (z_1, \ldots, z_p) \in X^p$ be a typical vector of party policy positions. Given $\mathbf{z}$, each voter, i, is described by a vector $\mathbf{u}_i(x_i, \mathbf{z}) = (u_{i1}(x_i, z_1), \ldots, u_{ip}(x_i, z_p))$, where

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

Here, λ_j is the "exogenous" valence of party j, $\mu_j(z_j)$ is the "activist" valence, β is a positive constant and $|| \cdot ||$ is the usual Euclidean norm on X. The terms $\{\epsilon_j\}$ are the stochastic components, assumed to be independently and identically normally distributed (iind), with zero expectation, each with standard deviation σ. The independence assumption is used in the estimations discussed in Section 4, and in the proof of Theorem 1. However, we show that analogous results can be obtained, if the more general assumption is made that the stochastic errors are multivariate normal with general variance/covariance matrix, θ.

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

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

Here Pr stands for the probability operator. The expected vote share of party j is

$$V_j(\mathbf{z}) = (1/n) \sum_{i \in N} \rho_{ij}(\mathbf{z}).$$

In the vote models it is assumed that each agent j chooses z_j to maximize V_j, conditional on $\mathbf{z}_{-j} = (z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_p)$.

The first result will focus on "exogenous" valence and assume that $\mu_j \equiv 0$ and that the exogeneous valences are ranked $\lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1$. I denote this model by M.

In this model it is natural to regard λ_j as the "average" weight given by a member of the electorate to the perceived competence or quality of candidate j. The "weight" will in fact vary throughout the electorate, in a way which is described by the normal distribution. Because of the differentiability of the cumulative normal distribution, the individual probability functions $\{\rho_{ij}\}$ are C^2-differentiable in the strategies $\{z_j\}$. Thus, the vote share functions will also be C^2-differentiable. Let $x^* = (1/n)\Sigma_i x_i$. Then the mean voter theorem for the stochastic model asserts that the "joint mean vector" $\mathbf{z}_0^* = (x^*, \ldots, x^*)$ is a PNE. Lin, Enelow and Dorussen [26] used C^2-differentiability of the expected vote share functions, in the situation with zero valence, to show that the validity of the theorem depended on the condition that σ^2 was "sufficiently large."

To determine whether this result can be applied to empirical cases, where the parameters of the model can be estimated, I shall obtain the precise necessary and sufficient conditions for existence of a LSNE of the form $(x^*, \ldots, x^*)$. Of course, even if the sufficient condition for $\mathbf{z}_0^* = (x^*, \ldots, x^*)$ to be an LSNE is satisfied, this does not imply that it is a PNE. However, the necessary condition for $\mathbf{z}_0^*$ to be a LSNE immediately gives a necessary condition for $\mathbf{z}_0^*$ to be a PNE. I shall argue that, in general, this necessary condition will be violated.

First, we may transform coordinates so that in the new coordinates, $x^* = 0$, and I shall often refer to $\mathbf{z}_0^*$ as the joint origin.

The voter variance covariance matrix: To characterize the variation in voter preferences, we must represent in a simple form the variance covariance matrix (or data matrix, ∇) of the voter ideal points. Let X be endowed with a system of coordinate axes $(1, \ldots, t, s, \ldots, w)$. For each coordinate axis let $\xi_t = (x_{1t}, x_{2t}, \ldots, x_{nt})$ be the vector of the t^{th} coordinates of the set of n voter ideal points. We use (ξ_s, ξ_t) to denote scalar product.

The symmetric $w \times w$ voter variance/covariance data matrix ∇ is then defined to be

$$\nabla = \begin{pmatrix} (\xi_1,\xi_1) & & & \cdots & (\xi_t,\xi_1) & (\xi_s,\xi_1) & \cdots & \\ & (\xi_2,\xi_2) & (\xi_3,\xi_2) & & & & & \\ & (\xi_2,\xi_3) & (\xi_3,\xi_3) & & & & & \\ \vdots & & & \ddots & & & & \\ (\xi_1,\xi_t) & & & & (\xi_t,\xi_t) & (\xi_s,\xi_t) & & \\ (\xi_1,\xi_s) & & & & (\xi_t,\xi_s) & (\xi_s,\xi_s) & & \\ \vdots & & & & & & \ddots & \\ & & & & & & & (\xi_w,\xi_w) \end{pmatrix}.$$

We next define the *convergence coefficient* of the model M.

Definition 4.

(i) For each agent j, define $\lambda_{av(j)} = \frac{1}{p-1}\sum_{k \in P-\{j\}} \lambda_k$.

(ii) Define the coefficient A_j for the contest of agent j against the competing agents to be

$$A_j = \frac{(p-1)\beta}{p\sigma^2}(\lambda_{av(j)} - \lambda_j).$$

(iii) The Hessian matrix C_j associated with agent j is defined to be $C_j = (2A_j/n)\nabla - I$, where I is the $w \times w$ identity matrix.

(iv) The "convergence coefficient" of the model M is given by $c(M) = 2A_1 v^2$ where $v^2 = (\frac{1}{n})\sum_{t=1}^{w}(\xi_t, \xi_t)$ is the total variance of the voter ideal points.

We now state the main result on the model M, with arbitrary p and w.

Theorem 1. *The necessary and sufficient condition for $\mathbf{z}_0^*$ to be a LSNE is that the eigenvalues of the Hessian matrix C_1 be all negative.*

The proof of Theorem 1, below, depends on considering the first and second order conditions at $\mathbf{z}_0^*$ for each vote share function. The first order condition is obtained by setting $dV_j/dz_j = 0$ (where we use this notation for full differentiation, keeping $z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_p$ constant).

Because the voter probabilities involve correlated errors, we first transform the problem so that these terms are uncorrelated. This allows us to show that $\mathbf{z}_0^*$ satisfies the first order condition. The second order condition is that the Hessian d^2V_j/dz_j^2 be negative definite at the joint origin. If this holds for all j at $\mathbf{z}_0^*$, then $\mathbf{z}_0^*$ is a LSNE. However, we need only examine this condition for the vote function V_1 for the lowest valence party. As we shall show, this condition on the Hessian of V_1 is equivalent to the condition on C_1, and if it holds for V_1, then the Hessians for $V_2, \ldots, V_p$ are all negative definite at $\mathbf{z}_0^*$. As usual conditions on C_1 for the eigenvalues to be negative depend on the trace, trace(C_1), and determinant, det(C_1), of C_1. These turn on the value of A_1 and on the electoral variance/covariance matrix, ∇. Using the determinant of C_1, we can show that $1 > 2A_1v^2$ is a sufficient condition for the eigenvalues to be negative. In terms of the "convergence coefficient" this is $1 > c(M)$. In a policy space of dimension w, the necessary condition on C_1, therefore, on the Hessian of V_1 is that $w \geq 2A_1v^2$. This condition is obtained from examining the trace of C_1 and gives $w \geq c(M)$. Conversely, if the necessary condition for V_1 fails, then $\mathbf{z}_0^*$ can be a neither a LNE nor a LSNE.

Ceteris paribus, a LNE at the joint origin is "less likely" the greater are the parameters p, β, $\lambda_{av(1)} - \lambda_1, v^2$, and is "more likely " the greater is the stochastic variance.To illustrate, in the case with p=2, the expression for A_1 is just $\frac{\beta}{2\sigma^2}(\lambda_2 - \lambda_1)$. Thus, in the very simplest case, with two parties and one dimension the convergence coefficient of the model M is

$$c(M) = \frac{\beta}{\sigma^2}(\lambda_2 - \lambda_1)v^2.$$

As Corollary 1 shows, the single eigenvalue is $c(M) - 1$, so the necessary condition for an LNE is $c(M) \leq 1$.

Note that the case $\lambda_p = \lambda_1$ was studied by Lin, Enelow and Dorussen [26]. In that case, the convergence coefficient $c(M)$ is zero so the joint origin, $\mathbf{z}_0^*$, is an LSNE. However the examples presented below to illustrate the theorem suggest that even when the joint origin is an LSNE, the concavity condition will not be satisfied. This suggests that PNE are unlikely to exist at the origin.

The proof of the theorem is presented in two parts. In Part A we show that the multivariate integral problem can be reduced to a univariate problem by suitable choice of a matrix transformation. In Part B, we examine the first and second order conditions.

Proof of Theorem 1.

Part A: Choice of orthonormal variates. Because the vote functions involve correlated variates, we show first how to construct a transformation of the variates so as to make them orthonormal. This allows us to show that the competition between an agent j and the other agents can be represented as a contest between j and a composite competitor. We then use the "convergence coefficient", $c(M)$, as given in Definition 4, to show that $c(M)$ determines the eigenvalues of the Hessians, and thus classifies the model M in a crucial sense.

From the above definitions, the probability, $\rho_{i1}(\mathbf{z})$, that voter i chooses agent 1 is given by

$$\Pr\left[[\lambda_1 - \beta||x_i - z_1||^2 + \epsilon_1 > \lambda_j - \beta||x_i - z_j||^2 + \epsilon_j], \text{ for all } j \neq 1\right]$$

$$= \Pr\left[[\lambda_1 - \beta||x_i - z_1||^2 - \lambda_j + \beta||x_i - z_j||^2 > \epsilon_j - \epsilon_1], \text{ for all } j \neq 1\right].$$

Now let $e_1 = (\epsilon_2 - \epsilon_1, \epsilon_3 - \epsilon_1, \ldots, \epsilon_p - \epsilon_1)$ be the $(p-1)$ dimensional variate. It is obvious that e_1 has the mutivariate normal distribution with covariance matrix Σ. Unfortunately the components of e_1 are correlated, so that Σ has off-diagonal terms. Indeed it is easy to see, in the case $p = 4$, for example, that

$$\Sigma = \sigma^2 \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

To show this, note that $e_1 = F(\epsilon)$ where ϵ is the error vector, and

$$F = \begin{pmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

Thus $\Sigma = \sigma^2 F \cdot F^T$, where T denotes transpose. Because the components of e_1 are correlated, the expression for $\rho_{i1}(\mathbf{z})$ cannot be readily differentiated. However we may make a transformation to new orthogonal variates. Consider a transformation matrix B_1 of rank $(p-1)$ with $y_1 = B_1(e_1)$. A standard result is that the random vector y_1 has the multivariate normal distribution with covariance matrix $(B_1)\Sigma(B_1)^T$. Now consider the solution to the matrix equation

$$(B_1)\Sigma(B_1)^T = I\sigma^2,$$

where I is the $(p-1) \times (p-1)$ identity matrix. It can be readily shown that a solution

$$B_1 = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,j} & b_{1,p-1} \\ . & . & . & . \\ . & . & b_{k,j} & . \\ b_{p-1,1} & . & b_{p-1,j} & b_{p-1,p-1} \end{pmatrix}$$

can be found such that $\sum_{j=1}^{p-1} b_{k,j} = 0$ for each $k = 1, \ldots, p-2$, and such that $b_{p-1,j} = b_{p-1,p-1}$ for $j = 1, \ldots, p-2$. These additional $2(p-2)$ restrictions can be satisfied because the symmetry of the covariance matrix, Σ, gives $\frac{(p-1)(p-2)}{2}$ degrees of freedom on the choice of entries of B_1.

We have shown that this transformation exists in the simple case that the original components of ϵ were independent and identically normally distributed (iind). Suppose more generally that these errors were mutivariate normal with an arbitrary covariance matrix θ. Then it is easy to show that the required matrix B_1 satisfies the matrix equation $(B_1F)\theta(B_1F)^T = G$ where G is an appropriate diagonal matrix. In principle there is no difficulty in finding B_1 in this more general case.

For the purpose of exposition, this chapter will focus on the simpler,iind case. Consider the transformed $(p-1)$ dimensional variate

$$y_1 = \sum_{j=2}^{p} b_{k,j}(\epsilon_j - \epsilon_1).$$

Because the components of y_1 are independent, we may write

$$\rho_{i1}(\mathbf{z}) = \Phi_1(g_{i1,1}(\mathbf{z})) \cdots \Phi_k(g_{i1,k}(\mathbf{z})) \cdot \Phi_{p-1}(g_{i1,p-1}(\mathbf{z})),$$

where $\Phi_k(d)$ stands for the univariate cumulative normal distribution (cnd) with appropriate variance, up to the value d. The values $\{g_{i1,k}(\mathbf{z})\}$ are the upper bounds of these univariate integrals. It follows directly from the definition of B_1 that

$$g_{i1,k}(\mathbf{z}) = \sum_{j=1}^{p-1} b_{k,j} \left[\lambda_1 - \beta||x_i - z_1||^2 - \lambda_{j+1} + \beta||x_i - z_{j+1}||^2\right].$$

Since $\sum_{j=1}^{p-1} b_{k,j} = 0$ for $k = 1, \ldots, p-2$, we see that this term is independent of z_1. However it is a function of

$$\sum_{j=1}^{p-1} b_{k,j}[-\lambda_{j+1} + \beta||x_i - z_{j+1}||^2].$$

Now consider

$$g_{i1,p-1}(\mathbf{z}) = \sum_{j=1} b_{p-1,j}[\lambda_1 - \beta||x_i - z_1||^2 - \lambda_{j+1} + \beta||x_i - z_{j+1}||^2].$$

It is easy to see that the variate $\sum_{j=2}(\epsilon_j - \epsilon_1)$ has variance

$$[(p-1) + (p-1)^2]\sigma^2 = p(p-1)\sigma^2.$$

Thus the coefficients $\{b_{p-1,j}\}$ must be $\frac{1}{\sqrt{p(p-1)}}$. Without loss of generality we may define the $(p-1)$ dimensional variate $y_{1,p-1}$ by

$$\begin{aligned} y_{1,p-1} &= \frac{1}{p-1}\sum_{j=2}(\epsilon_j - \epsilon_1) \\ &= -\epsilon_1 + \frac{1}{p-1}\sum_{j=2}\epsilon_j. \end{aligned}$$

Note that this new variate has variance $\frac{p}{p-1}\sigma^2$. The relevant upper bound on the integral is then

$$h_{i1,p-1}(\mathbf{z}) = \frac{1}{p-1}\sum_{j=2}^{p}[\lambda_1 - \beta||x_i - z_1||^2 - \lambda_j + \beta||x_i - z_j||^2].$$

Now for each $k = 1, \dots, p$ define

$$\lambda_{av(k)} = \frac{1}{p-1}\sum_{j\in P-\{k\}}\lambda_j.$$

It then follows that

$$h_{i1,p-1}(\mathbf{z}) = [\lambda_1 - \lambda_{av(1)} - \beta||x_i - z_1||^2] + \frac{1}{p-1}\sum_{j=2}^{p}\beta||x_i - z_j||^2.$$

As a consequence we may write

$$\rho_{i1}(\mathbf{z}) = \Phi_1(g_{i1,1}(\mathbf{z}))\cdots\Phi_k(g_{i1,k}(\mathbf{z}))\cdot\Phi_{p-2}(g_{i1,p-2}(\mathbf{z}))\cdot\Phi_{p-1}(h_{i1,p-1}(\mathbf{z})).$$

The terms $\Phi_k(g_{i1,k}(\mathbf{z}))$, for $k = 1, \dots, p-2$, are independent of z_1. The $(p-1)$th term is given by the univariate cnd for the variate $y_{1,p-1}$ with variance $\frac{p}{p-1}\sigma^2$. Now define, for any agent j, the $(p-1)$ vector $\mathbf{z}_{-j} = (z_1, \dots, z_{j-1}, z_{j+1}, \dots)$. It then follows by an identical argument to the above, that we may write

$$\rho_{ij}(\mathbf{z}) = \Phi_1(g_{ij,1}(\mathbf{z}_{-j}))\dots\Phi_k(g_{ij,k}(\mathbf{z}_{-j}))\cdot\Phi_{p-1}(h_{ij,p-1}(\mathbf{z})).$$

Again, the first $(p-2)$ terms are independent of z_j, while the expression

$$\Phi_{p-1}(h_{ij,p-1}(\mathbf{z}))$$

involves the term $[\lambda_j - \lambda_{av(j)} - \beta||x_i - z_j||^2]$. In essence, the electoral competition of agent j against the other agents can be regarded as a contest between j and a composite opponent with the average exogeneous valence $\lambda_{av(j)}$. The relevant stochastic variance associated with this contest is $\frac{p}{p-1}\sigma^2$.

Thus, the coefficient $A_j = \frac{(p-1)\beta}{p\sigma^2}(\lambda_{av(j)} - \lambda_j)$, previously defined, characterizes the contest of party j against the opposing parties. Having shown Part A we may now proceed to the second part of the proof of the theorem.

Part B: Equilibrium conditions.

We first seek the first order condition. Consider the lowest valence party, 1, and any voter i. From Part A, we have seen that the probability that voter i picks agent 1 over the other agents, at the joint strategy $\mathbf{z}$, is given by

$$\rho_{i1}(\mathbf{z}) \quad = \Phi_1(g_{i1,1}(\mathbf{z}_{-1})) \cdots \Phi_k(g_{i1,k}(\mathbf{z}_{-1})) \cdot \Phi_{p-1}(h_{i1,p-1}(\mathbf{z})).$$

Now $V_1(\mathbf{z}) = \frac{1}{n}\Sigma_i \rho_{i1}(\mathbf{z})$. To obtain the condition for z_1 to be a local strict best response to $\mathbf{z}_{-1} = (0, \ldots, 0)$, observe that we may write

$$\rho_{i1}(\mathbf{z}) = a_{i1}(x_i, \mathbf{z}_{-1}) \cdot \Phi_{p-1}(h_{i1,p-1}(x_i, \mathbf{z})).$$

Because of the definition of B_1, the first term in this expression is independent of z_1, and we may treat it as a constant. Notice however that this term is a function of x_i and $\mathbf{z}_{-1}$. Taking the first differential gives the first order condition on $V_1(\mathbf{z})$ as

$$\Sigma_{i=1}^{n} a_{i1}(x_i, \mathbf{z}_{-1}) \phi(h_{i1,p-1}(\mathbf{z})) dh_{i1,p-1}/dz_1 = 0,$$

where $\phi(h_{i1,p-1})$ is the normal probability density function for the variate $y_{1,p-1}$ at the value $h_{i1,p-1}$. Now,

$$\frac{dh_{i1,p-1}}{dz_1} = -2\beta(z_1 - x_i).$$

Consider the vector $\mathbf{z}^*_{-1} = (0, \ldots, 0)$, where $\sum x_i = 0$. By the definition of B_1, the term $a_{i1}(x_i, \mathbf{z}^*_{-1})$ is now independent of x_i, and can be written as a_1. Note that a_1 is positive. To test whether $z_1 = 0$ satisfies the first order condition, let $\mathbf{z}^*_0 = (0, \ldots, 0)$.

Then, for each i,

$$h_{i1,p-1}(\mathbf{z}^*_0) = g_1 = \lambda_1 - \lambda_{av(1)}$$

is independent of x_i, and thus of i. Hence, the first order condition for a CNE with respect to agent 1 is

$$\phi(g_1) \sum_i (z_1 - x_i) = 0.$$

The solution to this equation is $z_1 = \frac{1}{n} \sum x_i$. But $\sum x_i = 0$, by the choice of coordinates. Thus the solution is $z_1 = 0$. Obviously, an identical argument can be carried out for each agent j, by noting that $g_j = \lambda_j - \lambda_{av(j)}$. Thus $\mathbf{z}^*_0 = (0, \ldots, 0)$ solves the first order condition.

The second order condition on the Hessian $H_1 = d^2V_1/dz_1^2$ is that this be negative definite. With $\mathbf{z}^*_{-1} = (0,\ldots,0)$ we may write $h_{i1,p-1}(\mathbf{z}) = h_{i1,p-1}(\mathbf{z}_{-1},z_1) = g_{i1}(z_1)$. Computation of H_1 shows that it is given by

$$H_1(\mathbf{z}_{-1},z_1) \tag{1}$$
$$= \frac{a_1}{n}\sum_i H_{i1}(z_1) \tag{2}$$
$$= \frac{a_1}{n}\sum_i \left([\phi(g_{i1}(z_1))]\left(\frac{p-1}{p}\left[\frac{-g_{i1}(z_1)[\nabla_{i1}(z_1)]}{\sigma^2}\right] + \left[\frac{d^2 g_{i1}(z_1)}{dz_1^2}\right]\right)\right).$$

Without loss of generality we may normalize this expression by dividing by a_1 . The term $[\nabla_{i1}(z_1)]$ is a $w\times w$ symmetric matrix involving the differentials dg_{i1}/dz_1, whose entries are $4\beta^2(x_{is}-z_{1s})(x_{it}-z_{1t})$ in the (s,t) position. The term $[d^2g_{i1}/dz_1^2]$ is the negative definite Hessian of g_{i1p} (namely, $-2\beta I$, where I is the identity matrix). Thus, $\sum_{i\in N}\nabla_{i1}(0) = 4\beta^2\nabla$, where ∇ is the voter covariance matrix, given above. As we have just seen, $g_{i1}(0) = g_1 = \lambda_1 - \lambda_{av(1)}$ is independent of x_i. We thus obtain

$$\begin{aligned} H_1(\mathbf{z}_0^*) &= \frac{2\beta}{n}\phi(\lambda_1-\lambda_{av(1)})\left\{\left[\frac{2(p-1)\beta}{p\sigma^2}\right]\left[\lambda_{av(1)}-\lambda_1\right]\nabla - nI\right\} \\ &= 2\beta\phi(\lambda_1-\lambda_{av(1)})\left[\frac{2A_1}{n}\nabla - I\right]. \end{aligned}$$

Obviously, the first term involving $\phi(\lambda_1-\lambda_{av(1)})$ in this expression is positive. So the condition that $z_1 = 0$ constitutes a local best response to $\mathbf{z}_{-1} = (0,\ldots,0)$ depends on the eigenvalues of $C_1 = [\frac{2A_1}{n}\nabla - I]$. If C_1 has negative eigenvalues, then the same argument can be repeated for each agent $j = 2,\ldots,p$. To see this note that the above argument immediately shows that the Hessian for j is

$$H_j(\mathbf{z}_0^*) = 2\beta\phi(\lambda_j-\lambda_{av(j)})\left[\frac{2A_j}{n}\nabla - I\right].$$

Thus the Hessian for j has negative eigenvalues if and only if C_j is negative definite. Because the valences are ranked $\lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1$, it follows that $A_1 \geq A_2 \geq \cdots \geq A_p$. This implies that $\text{trace}(C_1) \geq \text{trace}(C_2) \geq \cdots \geq \text{trace}(C_p)$ and $\det(C_1) \geq \det(C_2) \geq \cdots \geq \det(C_p)$. Thus if C_1 has negative eigenvalues then so do $C_2,\ldots,C_p$, and this implies that $z_1 = z_2 = \cdots = z_p = 0$ will all be mutual local best responses. This shows that the stated condition is sufficient for $\mathbf{z}_0^* = (0,0,\ldots,0)$ to be an LSNE. Obviously, if C_1 does not have negative eigenvalues, then $\mathbf{z}_0^*$ cannot be a LSNE. This completes the proof of the theorem.

Note also that if the valences are not all identical then $0 > A_{j_0}$ for some $j_0 > 1$, thus guaranteeing that C_j has negative eigenvalues whenever $j \geq j_0$.

On the other hand, if all valences are identical then C_1 must have negative eigenvalues.

Although Theorem 1 is stated in terms of LSNE, it clearly gives analogous conditions for $\mathbf{z}_0^* = (0, \ldots, 0)$ to be an LNE. Thus if $\mathbf{z}_0^*$ is an LNE then the eigenvalues of C_1 must be non-positive, while negative eigenvalues imply that $\mathbf{z}_0^*$ must be an LNE. It is also worth remarking that with a general covariance matrix, θ, for the errors, we can use the above procedure, based on the matrix transformation $(B_1F)\theta(B_1F)^T = G$, to obtain a Hessian matrix $C_1(\theta)$ whose eigenvalues give conditions for a LSNE at the joint origin.

To illustrate this theorem, we can give corollaries for the one and two dimensional cases.

Corollary 1. *If X is one dimensional, and there are two parties, then the necessary condition for $\mathbf{z}_0^*$ to be a LNE is that $\frac{\beta}{\sigma^2}(\lambda_2 - \lambda_1)v^2 \leq 1$. The sufficient condition for the joint origin to be a LSNE is that this inequality holds strictly.*

Proof. Since we assume $p = 2$, then $\lambda_{av(1)} = \lambda_2$. Just as in the above, $g_{i12}(0) = \lambda_1 - \lambda_2$ is the same, for all i. Moreover,

$$\nabla_{i1} = (dg_{i1}/dz_1)^2 = 4x_i^2\beta^2.$$

Thus,

$$H_1(0) = \phi(\lambda_1 - \lambda_2)\frac{2\beta}{n}\sum_i \left[\frac{(\lambda_2 - \lambda_1)(x_i^2\beta)}{\sigma^2} - 1\right]. \tag{3}$$

Now $v^2 = \frac{1}{n}\sum_i x_i^2$ is the total electoral variance. Clearly, $H_1(0)$ is non-positive if and only if $\sigma^2 \geq \beta(\lambda_2 - \lambda_1)v^2$. If this condition fails, then $z_1 = 0$ is a local strict minimum of the vote share function, V_1, given $z_2 = 0$, so that $\mathbf{z}_0^*$ cannot be an LNE. This gives the necessary condition. On the other hand, if $\sigma^2 > \beta(\lambda_2 - \lambda_1)v^2$, then $z_1 = 0$ gives a local strict maximum of V_1. The sufficient condition for V_2 is $\sigma^2 > \beta(\lambda_1 - \lambda_2)v^2$. Since $\lambda_2 \geq \lambda_1$, this condition must hold. Consequently the strict inequality gives a sufficient condition for $\mathbf{z}_0^*$ to be a LSNE.

Notice that as $\sigma^2 \to 0$, then it becomes impossible for $\mathbf{z}_0^*$ to be an LSNE if the valences differ. (See also Groseclose [22] for a one-dimensional model involving valence.)

Corollary 2. *Assume X is two dimensional. Then $\mathbf{z}_0^*$ is a LSNE if $c(M)$ is strictly less than one, and is a LNE only if $c(M) \leq 2$.*

Proof. The condition that both eigenvalues be negative is equivalent to the condition that $\det(C_1)$ is positive and $\text{trace}(C_1)$ is negative. Now

$$\begin{aligned}\det(C_1) = {} & (2A_1/n)^2\left[(\xi_1, \xi_1)\cdot(\xi_2, \xi_2) - (\xi_1, \xi_2)^2\right] \\ & +1 - (2A_1/n)\left[(\xi_1, \xi_1) + (\xi_2, \xi_2)\right].\end{aligned}$$

The first bracket $\left[(\xi_1,\xi_1)\cdot(\xi_2,\xi_2)-(\xi_1,\xi_2)^2\right]$ is an "inverse correlation coefficient" associated with the covariance matrix of the voter distribution. By the triangle inequality, this term must be non-negative. Thus $\det(C_1)$ is positive if

$$1>(2A_1/n)[(\xi_1,\xi_1)+(\xi_2,\xi_2)],\text{ or } 1>\frac{2(p-1)}{p\sigma^2}\beta(\lambda_{av(1)}-\lambda_1)v^2. \quad (4)$$

This gives the sufficient condition that $1>c(M)$ for a LSNE at the joint origin, $\mathbf{z}_0^*$.

The necessary condition for $\mathbf{z}_0^*$ to be an LNE is that the eigenvalues be non-positive. Since trace(C_1) equals the sum of the eigenvalues we can use the fact that trace$(C_1)=(2A_1/n)[(\xi_1,\xi_1)+(\xi_2,\xi_2)]-2$, to obtain the necessary condition

$$\frac{2(p-1)}{p\sigma^2}\beta(\lambda_{av(1)}-\lambda_1)v^2\leq 2. \quad (5)$$

Thus $c(M)\leq 2$ gives the necessary condition, completing the proof.

If the first inequality (given in (4)) fails, but the second (given in (5)) is satisfied, then the eigenvalues may still be non-positive, and can be explicitly computed in terms of the model parameters and data. If the second condition fails then obviously at least one of the eigenvalues must be strictly positive, and so $\mathbf{z}_0^*$ cannot be an LNE. The condition $c(M)=2$ includes the non-generic case where both eigenvalues are zero.

In the two dimensional case there is one situation where computation of eigenvalues is particularly easy. If the covariance (ξ_1,ξ_2) of the electoral data on the two axes is (close to) zero, then the voter covariance matrix is (approximately) diagonal and the two policy dimensions can be treated separately. In this case we obtain two separate necessary conditions $\frac{2(p-1)\beta}{p\sigma^2}(\lambda_p-\lambda_1)v_t^2\leq 1$, for both $t=1,2$, for $\mathbf{z}_0^*$ to be a LNE. If $v_t^2\leq v_s^2$ then we obtain essentially identical necessary and sufficient conditions in terms of v_s^2. It is obvious that trace(C_1) involves the dimension, w, so we obtain the necessary condition $c(M)\leq w$ for the joint origin to be a LNE in this more general case.

In the examples below we deal with the one dimensional case first, and then give a two dimensional illustration of the result.

3 Examples

Example 1. Consider the model, M, but with zero valence, in one dimension, with two voters $\{1,2\}$ at $x_1=-1$ and $x_2=+1$, and two parties $\{1,2\}$ at z_1, z_2. As observed above, with $p=2$, it is not necessary to perform the transformation to orthogonalise the variables. The probability voter i picks party 1 over 2 is

$$\Pr\left[-\beta(x_i - z_1)^2 + \epsilon_1 + \beta(x_i - z_2)^2 - \epsilon_2 > 0\right].$$

We can write $g_i = -\beta(x_i - z_1)^2 + \beta(x_i - z_2)^2$ for the "comparison function" for voter i in comparing party 1 with party 2 when the parties are at (z_1, z_2). This probability is $\Phi(g_i)$, where Φ is the cumulative normal distribution for the variate $\epsilon_2 - \epsilon_1$, with variance $2\sigma^2$, and expectation 0.

Suppose now that $z_2^* = 0$ (the mean of the voter ideal points). We can proceed just as above to test whether $z_1 = 0$ comprises a best response to $z_2^* = 0$, thus making up a component of a Nash equilibrium.. Now $V_1 = \frac{1}{2}[\Phi(g_1) + \Phi(g_2)]$, so $dV_1/dz_1 = \frac{1}{2}[\phi(g_1)dg_1/dz_1 + \phi(g_2)dg_2/dz_1]$, 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 the first order condition is $dg_1/dz_1 + dg_2/dz_1 = 0$ or $z_1^* = (1/2)(x_1 + x_2) = 0$. As before, $z_1^* = x^*$.

From (2),

$$\frac{d^2\Phi(g_i)}{dz_1^2} = \phi(g_i)\left[\frac{d^2 g_i}{dz_1^2} - \frac{g_i}{2\sigma^2}\left(\frac{dg_i}{dz_1}\right)^2\right].$$

Clearly, if $z_1 = 0$, then $g_i = 0$ for $i = 1, 2$. Moreover, $d^2g_i/dz_1^2 = -2\beta$ 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. (Obviously, the vector $(0, 0)$ is what we have termed a LSNE.) 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 $z_1 \in [-1, +1]$. We use the standard result in convex analysis that a necessary and sufficient condition for concavity of a differentiable function on some domain is that its Hessian be negative semi-definite on the domain. To see whether this is so, we evaluate the Hessian of $\Phi(g_2)$ at $z_1 = -1$.

We obtain $g_2 = -\beta(2)^2 + \beta(1)^2 = -3\beta$. So,

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

Similarly $d^2\Phi(g_1)/dz_1^2 = \phi(\beta)[-2\beta]$, which is clearly negative. The first expression is negative only if $\sigma > 12\beta^2$. If $\sigma < \beta\sqrt{12}$, then the Hessian of $\Phi(g_2)$ at $z_1 = -1$ will be positive. Thus, for σ "sufficiently" small, the Hessian of V_1 at $z_1 = -1$ will also be positive. In this example, therefore, for some value of σ, with $\sigma < 12\beta^2$, the expected vote function, V_1 will not be concave on $[-1, +1]$. Although concavity may fail, the weaker condition of quasi-concavity appears to be satisfied, so that the origin is a local "attractor" on the domain.

As the example should make clear, the requirement of concavity imposes a relationship between the variance of the voter ideal points and the stochastic variance. As expressed in Theorem 1, if the stochastic variance is given, then the constraint is imposed on the variance of the set of bliss points. Alternatively, as in the example, if the bliss points are given then the constraint is imposed on the stochastic variance.

Example 2. We now modify Example 1, by introducing valence terms λ_2, λ_1 with $\lambda_2 > \lambda_1$. We first use Corollary 1 to determine the conditions for $(0,0)$ to be a LNE. Now $dg_i/dz_1 = 2\beta x_i = \pm 2\beta$ depending on whether $i = 1$ or 2. By (3), $H_1(0) = \frac{1}{2}\phi(\lambda_1 - \lambda_p)\sum_i [((\lambda_p - \lambda_1)/2\sigma^2)(2\beta x_i)^2 - 2\beta]$, so the necessary and sufficient condition is $(\lambda_p - \lambda_1)\beta \leq \sigma^2$, as stated in Corollary 1. (In this case the electoral variance is 1.) Suppose this condition is satisfied, so that $(0,0)$ is an LNE. Then we can obtain a condition for concavity so as to ensure that $(0,0)$ is a PNE.

Let $z_2 = 0$. A sufficient condition for $z_1 = 0$ to be a best response in the domain $[-1,+1]$ is that the Hessian of V_1 be non-positive on this domain. Consider $z_1 = -1$. The comparison functions g_2 and g_1 are now

$$\begin{aligned} g_2 &= \lambda_1 - \beta(x_2 - z_1)^2 - \lambda_2 + \beta(x_2 - z_2)^2 \\ &= \lambda_1 - \lambda_2 - 3\beta \end{aligned}$$

$$\begin{aligned} g_1 &= \lambda_1 - \beta(x_1 - z_1)^2 - \lambda_2 + \beta(x_1 - z_2)^2 \\ &= \lambda_1 - \lambda_2 + \beta. \end{aligned}$$

The Hessian at $z_1 = -1$ is then

$$\phi(\lambda_1 - \lambda_2 - 3\beta)\left\{-2\beta + \left[(3\beta + \lambda_2 - \lambda_1)/2\sigma^2\right][4\beta]^2\right\} + \phi(\lambda_1 - \lambda_2 + \beta)\{-2\beta\}.$$

While the second term is clearly negative, the first term may be positive. A sufficient condition for the first term to be negative at $z_1 = -1$ is then $\sigma^2 > 4(3\beta + \lambda_2 - \lambda_1)\beta$. Since $\lambda_2 > \lambda_1$ this condition is more severe than the one obtained in Example 1. Obviously a sufficient condition for the Hessian of V_1 to be negative at $z_1 = -1$ is much more restrictive than the one obtained in Example 1. Moreover, even when the condition for $(0,0)$ to be a LNE is satisfied, there is no formal reason to expect $(0,0)$ to be a PNE. Results in Groseclose [22] suggest that even this simple model it is difficult to determine even if PNE exist. Symmetry suggests however that if the origin is not an LSNE, there will be two symmetrically located LSNE, with the high valence party nearer the origin.

Example 3. To illustrate the computation of eigenvalues in two dimensions, consider a situation where the valences for 1, 2, 3 are ranked with $\lambda_1 < \lambda_2 < \lambda_3$, and three voters, labeled $\{1,2,3\}$, have the ideal points in $\mathbb{R}^2$ given by $(0,1)$, $(-\sqrt{3/4}, -\frac{1}{2})$ and $(\sqrt{3/4}, -\frac{1}{2})$. With all agents at the origin, the comparison function g_i for each voter is given by $\lambda_1 - \frac{1}{2}(\lambda_2 + \lambda_3)$, in comparing agent 1 against agent 2 and 3. So the Hessian for agent 1 involves the matrix

$$H_1 = \Sigma_i \phi(g_i)\left[\frac{-g_i(p-1)}{p\sigma^2}[\nabla_i]\beta - 2I\right].$$

Here ∇_i is the 2×2 matrix generated by the gradients of g_i at x_i. Summing the terms $[\nabla_i]$ gives 4∇, where ∇ is the voter variance/covariance data matrix

$$\nabla = \begin{pmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}.$$

As before we let $\xi_1 = (-\sqrt{\frac{3}{4}}, 0, \sqrt{\frac{3}{4}})$ and $\xi_2 = (-\frac{1}{2}, 1, -\frac{1}{2})$ be the vectors of voter ideal points on the first and second coordinates. The the diagonal terms in ∇ are given by the quadratic forms $(\xi_1, \xi_1) = (\xi_2, \xi_2) = \frac{3}{2}$ and the off diagonal terms are given by the scalar product $(\xi_1, \xi_2) = 0$. Thus the covariance between ξ_1 and ξ_2 is zero. The Hessian for agent 1 is given by

$$\frac{2(p-1)\left[\frac{1}{2}[\lambda_2 + \lambda_3] - \lambda_1\right]}{p\sigma^2}[\nabla]\beta - 3I.$$

A necessary condition for this matrix to have negative eigenvalues is that its trace be negative. If the trace is zero or positive, then one of the eigenvalues must be positive (or both eigenvalues are zero). Now the variance of the voter ideal points on the two axes is given by $v^2 = \frac{1}{3}[(\xi_1, \xi_1) + (\xi_2, \xi_2)] = \frac{1}{3}[v_1^2 + v_2^2] = 1$. Because of the zero covariance, the two eigenvalues of the Hessian must be equal. Thus the necessary condition that the the eigenvalue be negative gives the same condition on the trace

$$\frac{\beta(p-1)[\lambda_2 + \lambda_3 - 2\lambda_1]}{p\sigma^2} - 2.$$

Because of the symmetry, this condition is identical to the one obtained from the determinant. If the trace is positive, then each eigenvalue of the Hessian will be positive at the origin. Thus, if σ^2 is sufficiently small then $z_1 = 0$ cannot be a local best response and so the origin cannot be an LSNE. Note that due to the symmetry of the voter ideal points, if the above condition fails, then agent 1 can increase vote share by moving in any direction away from the origin. In the non symmetric case, where ∇ is non-diagonal, when the trace condition fails then one eigenvalue will exceed the other, and agent 1 should move away from the origin in the "major" eigenspace associated with this eigenvalue. This eigenspace will depend on ∇, and in particular, on the variances v_1^2 and v_2^2. Again, symmetry suggests that, there are two different cases, depending on the degree of covariance, and on the variances v_1^2 and v_2^2 . In the case where there is a major axis, and one eigenvalue is positive, while the other is negative, then, in equilibrium, all three agents will be positioned on this major axis. In the second case, where both eigenvalues are positive, then the agents will move away from the origin in these eigenspaces. In both cases, the highest valence agent will be positioned nearest the origin. Obviously, by symmetry, there will be many LSNE.

4 Activist Valence

We now consider a more general model, $M(\mu)$, involving both exogenous valence $\{\lambda_j\}$ and activist valence $\{\mu_j\}$. To keep the analysis relatively simple we shall focus below on the competition between two parties, called j and p. We can then apply the results to the case of Britain, where there are two principal parties.

As in the proof of Theorem 1, the first order condition for agent j is

$$\sum_i \phi(g_{ij}(\mathbf{z}))\frac{dg_{ij}}{dz_j} = 0.$$

However, now

$$g_{ij}(\mathbf{z}) = \lambda_j + \mu_j(z_j) - \beta||x_i - z_j||^2 - \lambda_p - \mu_p(z_p) + \beta||x_i - z_p||^2.$$

Hence the first order solution for agent j is

$$z_j^* = \frac{1}{2\beta}\frac{d\mu_i}{dz_j} + \sum_i \alpha_{ij}x.$$

In this equation, the coefficients $\{\alpha_{ij}\}$ involve $\mathbf{z}$ and the exogenous valence terms $\{\lambda_j\}$. Moreover, each α_{ij} is strictly increasing in λ_j and decreasing in λ_p. Let us 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

$$\frac{dV_j^*}{dz_j} + \frac{1}{2\beta}\frac{d\mu_j}{dz_j} - z_j^* = 0. \tag{6}$$

Say the electoral pull and activist pull are "balanced" if this equation is satisfied.

The first term in this expression (the "marginal or gradient 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.) As λ_j is exogenously increased, this vector increases in magnitude. The vector $d\mu_j/dz_j$ "points towards" the position at which the total of activist "contributions" is maximized. We may term this vector the "(marginal or gradient) activist pull." Moreover, if the activist function is "sufficiently concave" (with negative eigenvalues of large modulus), then the second order condition (the negative definiteness of the Hessian of the "activist pull") will guarantee that the vector $\mathbf{z}^*$ given by the solution of the system of equations given by (6), for all j, will be an LSNE. This can be seen by examining the Hessian as in (2). The following theorem states these conclusions (see Schofield [37] for a proof).

Theorem 2. *Consider a vote maximization model, $M(\mu)$, with both exogenous popularity valences $\{\lambda_j\}$ and activist valences $\{\mu_j\}$. The first order condition for $\mathbf{z}^*$ to be an 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*

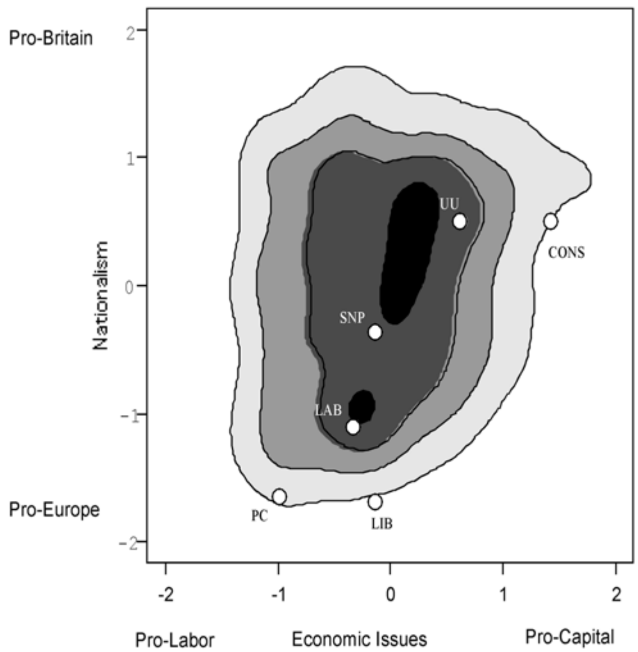

Fig. 1. Estimated party positions in the British Parliament in 1997 (based on MP survey data and a National Election Survey), with the highest density plots of the voter sample distribution at 95%, 75%, 50% and 10% levels.

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.*

If all activist valence functions are sufficiently concave (in the sense of having negative eigenvalues of sufficiently great magnitude) then the solution given by (6) will be a PNE.

Example 4. Fig. 1 presents data on the voter distribution, in a two-dimensional policy space derived from an empirical model for Britain (see Schofield [36] for further details). In this figure, the positions of the Conservative Party (CONS), Labour (LAB), Liberal Democrats (LIB), Plaid Cymru (PC), Ulster Unionists (UU) and Scottish Nationalists (SNP) were estimated from MP surveys. The two dimensional policy space was estimated from factor analysis of an electoral sample survey. The left right axis is the usual economic axis, while the north south axis represents attitudes to the European Union (with south being pro-Europe, and north pro-Britain.) Fig. 1 shows the estimated density function of the distribution of voter ideal points. The origin in this

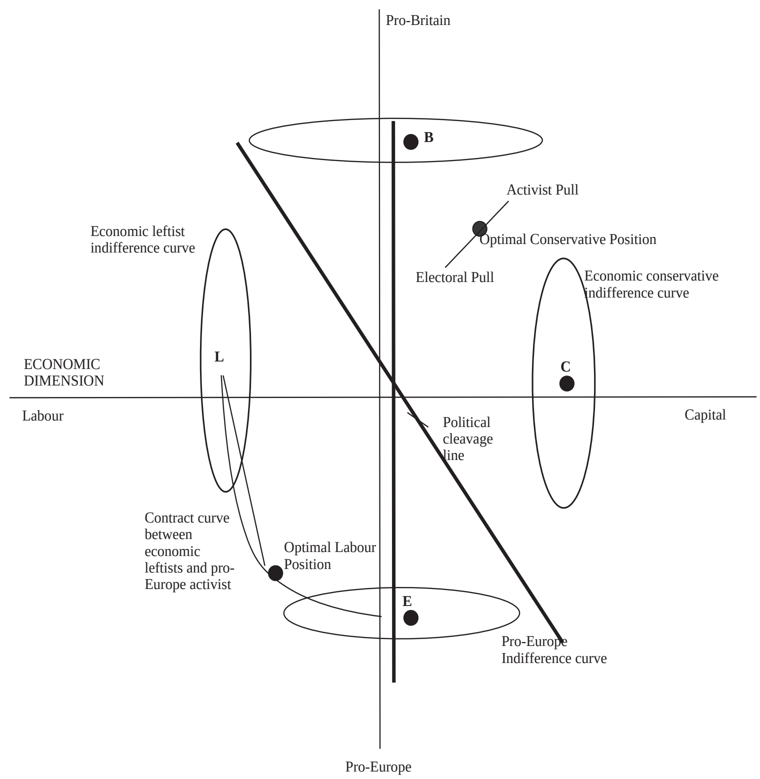

Fig. 2. Illustration of vote maximizing positions of Conservative and Labour party leaders in a two-dimensional policy space.

figure gives the mean of the voter distribution. It is obvious from the figure that the covariance (ξ_1, ξ_2) term in the voter matrix ∇ is close to zero. We can therefore estimate the eigenvalues of the Hessians on these two axes separately. Using the single economic axis alone, the empirical estimate of the exogenous valence of the Conservative Party dropped from 1.58 at the election of 1992 to 1.24 at the 1997 election, while the Labour valence increased from 0.58 to 0.97 over the same period. The Liberal Democrat Party was given the label 1, and its valence was set at zero for both elections. Because the scale of the figure is indeterminate, we are at liberty to use the standard deviation as the unit of measurement. For 1997, the spatial coefficient, β, was estimated to be 0.50, while the average valence difference $\lambda_{av(1)}$ was estimated to be 1.105. The ratio $\frac{v^2}{\sigma^2}$ of the electoral variance (on the economic axis) to the stochastic variance was 2/3. Using the equation

$$A_1 = \frac{2(p-1)\beta}{3p}(\lambda_{av(1)} - \lambda_1)$$

and considering only competition between three parties (p=3) we find $A_1 = 0.25$, so the estimated convergence coefficient of the model is 0.5. Therefore the necessary and sufficient conditions for an LSNE at the joint origin were satisfied, at least when only one axis of policy is relevant. Indeed, computation of the vote share functions indicated that, in one dimension, all the gradients of the party vote share functions pointed towards the origin. In other words, the origin (in one dimension) was an attractor of the electoral model. We can infer that all parties should have converged to this electoral mean.

This conclusion changes, however, when the second policy axis, involving attitudes to Europe, is included. As Fig. 1 indicates, the electoral variance v^2 increases to approximately 2, and we can estimate that the sufficient condition for a LSNE at the origin is not satisfied. When the eigenvalues for the Liberal Democratic Party are computed, this party's eigenvalue at the origin on the European axis is positive. As Fig. 1 suggests, the Liberal Democratic party did indeed vacate the electoral origin, and adopted a pro-European position. We may infer that such a position was also more acceptable to its party activists than the electoral mean. Once the Liberal Democrat Party is no longer positioned at the origin, the results presented in Theorem 1 cannot be used. However, the high relative valences of the Conservative and Labour parties would suggest that their local equilibrium positions would be nearer to the origin than the Liberal Democrat Party. Simulation of vote maximizing models by Schofield and Sened [44] suggests that parties with high exogeneous valence tend to adopt positions nearer the origin than parties with low valence.

I conjecture that the Conservative and Labour Parties did adopt positions close to those marked CONS and LAB, respectively, in Fig. 1. A basis for this conjecture can be found in Theorem 2, which develops the idea that the high valences for these two parties were due partly to exogenous valence and partly to valence derived from activist support. Analysis of survey data suggests that the Labour exogenous valence (λ_{LAB} due to Blair) rose in this period (see Clarke, Stewart and Whitely [16]). Conversely, the relative exogenous term, λ_{CON}, for the Conservatives fell. Since the coefficients (in the equation for the electoral pull) for the Conservative party depend on $(\lambda_{CON} - \lambda_{LAB})$, these must all fall in this period. This has the effect of increasing the marginal effect of activism.

Indeed, it is possible to include the effect of two potential activist groups for the Conservative Party — one "pro-British" and one "pro-Capital." The optimal Conservative position will be determined by a version of (6) which equates the "electoral pull" against the two "activist pulls." Since the electoral pull fell between the elections, 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 I shall refer to as the "activist contract curve" for the Conservative party.

Note that in Fig. 2, the indifference curves of representative activists for the parties are described by ellipses. This is meant to indicate that preferences of different activists on the two dimensions may accord different saliences to the policy axes. The "activist contract curve" given in the figure, for Labour say, is the locus of points satisfying the equation $d\mu_{LAB}/dz_{LAB} = 0$. This curve represents the balance of power between Labour supporters most interested in economic issues and those more interested in Europe. 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). As relative exogenous valence for a party 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 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 gave an optimal position closer to the electoral center than the optimal position of the Conservative party. Moreover, this affected the balance between "old left" activists in the party, and "new Labour" activists, concerned to modernize the party through a European style "social democratic" perspective. This conclusion is compatible with Blair's successful attempts to bring "New Labour" members into the party (see Seyd and Whiteley [46]).

5 Comments on the Electoral Model

A number of objections can be raised against the model presented here. I believe it worthwhile addressing these.

Voter rationality. Why is it appropriate to build a stochastic element into the model? Does the model suppose that voters randomize across their choices?

The purpose of the model is to determine how political agents choose their positions, given their information about electoral response. Opinion poll data will give approximate vote shares for each agent. Each agent can build a rough model of electorate response to party declarations. This is proxied by the econometric models of electoral response. These models are stochastic, and have proven quite successful at estimating voter response. Various assumptions are made about the distribution of the error terms, but the model based on the multivariate normal distribution is the most general (see Alvarez et al. [4], [5], Poole and Rosenthal [32], Quinn et al. [33], Schofield et al. [40, 44, 45]).

Agent rationality. Why should political agents (party leaders or candidates) attempt to maximize expected vote share?

In two party competition, maximizing expected plurality or probability of winning is an obvious alternative (see Duggan [20]). In this case, empirical estimates of the variance of the vote share functions are generally low, so probability of winning is a close proxy to expected vote share.

In multiparty competition, the objections are more serious. A party on the electoral periphery with a large vote share is unlikely to be asked to join a government coalition. If party leaders care about policy, then vote share alone is not an accurate measure of payoff. Modeling a party's effect on final policy outcome has proved very difficult, since it is necessary to combine a model of post-election bargaining (Banks and Duggan [12]) with pre-election maneuvering (Schofield and Sened [43]).

Although the simple model of vote maximization presented here is likely inadequate for multiparty situations, it is offered as a first step at classification of electoral systems. It suggests that convergence to the electoral mean is unlikely, and that the notion of LSNE is compatible with the very different political configurations that have been found to occur.

Activists. An early model of activists by Aldrich [1, 2] and Aldrich and McGinnis [3] assumed that activists controlled candidate policies. In contrast, the activist model outlined here is based on the assumption that different potential activist groups compete with one another to influence candidates or parties. Moreover, activists are likely to prefer policy positions far from the electoral origin. Theorem 2 asserts that this poses a complex optimization problem for party leaders. If a leader's exogenous valence falls, then activists become more important for the party. Thus a party led by a low valence politician may be forced away from the electoral origin.

The effect of activists on valence. It seems most natural in modeling plurality electoral systems to assume activists affect valence. In such systems, a small advantage in vote shares results in a large seat advantage. Activist effort and contribution then becomes crutial in obtaing votes. I have not attempted to model the activist calculus, since there are many possible ways to do this. It is plausible that activist motivations play a predominant role in the two-party system of the U.S. (see Miller and Schofield [29] and Schofield et al. [41]).

Local Nash equilibria. In empirical studies the LSNE can be determined by simulation. These local equilibria are not identical to actual policy positions of parties, but they can be used to distinguish between parties that principally interested in maximizing vote shares, and those that are policy motivated (see Schofield and Sened [44]). It is clear from the simulation that exogenous valence terms have considerable impact on equilibrium positions in multiparty systems based on proportional representation. Since these valence terms fluctuate because of exogeneous shocks, such shocks will induce small changes in local equilibria. Indeed, the local equilibrium concept is compatible with small policy changes by parties as they attempt to adapt to the changing electoral environment. A possible conjecture is that such shocks will induce only small changes in party configurations, when the electoral system is based on proportional representation. In contrast, under plurality rule, exogeneous shocks may increase the importance of activists for some parties, thus inducing large policy changes.

6 Extensions of the Model and Generic Existence of LSNE

One feature of the "mean voter theorem" is that each agent's optimum strategy is independent of the strategies of other agents, and is determined only by the distribution of voter ideal points. This has been regarded as an attractive feature of the spatial model. However, as Theorem 1 indicates, there are conditions under which the joint electoral origin will not be a LSNE, and definitely not be a PNE. As Example 4 indicates, a key element of the necessary condition for convergence is the ratio of the variance of the voter ideal points to the stochastic variance. When the electoral variance is high, and there are many parties with differing exogeneous valences, as in Israel, then simulation of the vote maximizing model has found multiple non-convergent LSNE (see Schofield and Sened [44]). For these non-convergent LSNE, agent equilibrium strategies are "strongly interdependent", in the sense that each party's equilibrium position is sensitive to the positions of all other parties. Simulation, under the vote maximizing assumption, shows that higher valence parties adopt positions nearer the electoral center than low valence parties. Determining the nature of these LSNE by analysis, rather than by simulation, is extremely difficult. Because convexity properties, like quasi-concavity of the utility functions, are not be satisfied, it is difficult, if not impossible to determine if PNE exist. For example, failure of the second order, Hessian condition at the origin immediately implies that the usual sufficient conditions for existence of PNE will not be satisfied. For these reasons we now consider the question of existence of LSNE under more general conditions that those discussed above.

Firstly, in the proof of Theorem 1, it was assumed that the errors were independent and identically normally distributed (iind). However, estimation of such voter models has found it necessary to adopt the more general hypothesis that the error structure is multivariate normal (that is, allowing for non-zero covariance terms in the error covariance matrix θ). See Alvarez and Nagler [4], Quinn et al. [33] and Schofield et al. [40]. The proof procedure of Theorem 1 does carry through in this case. In particular, it will still be possible to find a solution to the orthonormal matrix equation $(B_1F)\theta(B_1F)^T = G$, but the solution will depend on the error covariances. The solution will then lead to the determination of a convergence coefficient, $c(\theta)$. Because of the symmetry induced at the origin, the computation of $c(\theta)$ is relatively simple. The problem arises when the induced necessary condition for an LSNE at the origin is not satisfied. In this case, the analytical computation of the first and second order equations for equilibria away from the origin will be extremely difficult. The equilibria can generally be found by simulation.

The first order equations are matrix equations involving n smooth functions $\{g_i\}_N$. Since there are p agents, each choosing a strategy in X (of dimension w) we obtain wp equations. In the case of a general covariance matrix, these wp equations will involve interaction terms induced by the error covariance. Nonetheless, transversality theory (Austen-Smith and Banks [9], Banks [10] and Saari [34]) can be used to show that these wp equations are "generically" independent. Since wp is the dimension of the joint strategy space, the solution is generically of dimension 0. Consequently, the first order equations can be solved, even for general valence functions, as long as these functions, and the multivariate cumulative probability distribution of the errors are differentiable. Thus, CNE will generically exist. Indeed, it can be shown that one of these CNE will be a LSNE. A similar argument can be carried out even when the payoff functions of the agents are defined not simply by their expected vote shares, but also by agent policy preferences.

To outline a proof of this assertion, we need to introduce the idea of a tangent bundle (see Hirsch [24]). At $z \in X^p$, the tangent space is $T_z(X^p)$ and the tangent bundle is defined by $T(X^p) = \cup_z T_z(X^p)$. We assume all utility profiles in the set $E = \{U^h : X^p \to \mathbb{R}^p\}$ are C^2 differentiable. The differential $dU_j^h(z)$ at z can be regarded as a linear map from $\mathbb{R}^{wp}$ to $\mathbb{R}$ so $dU_j^h : X^p \to \text{Lin}(\mathbb{R}^{wp}, \mathbb{R})$. Since dU_j^h is C^1 differentiable, this map is continuous. Moreover, there is a C^2-topology on the set, E, under which two profiles are close, if their components are close as linear maps, and all their Hessians are close as bilinear maps, at every $z \in X^p$ (see Smale [48]). The differential dU_j^h can also be projected onto $T_z(X)$. Then this projection $DU_j^h : X^p \to \text{Lin}(\mathbb{R}^w, \mathbb{R})$ can be identified with the gradient of U_j^h in X, when z_k (for $k \neq j$) are held fixed. That is to say, $DU_j^h(z)$ can be regarded as an element of a subtangent space $T_z(X) \subset T_z(X^p)$, where X corresponds to the jth strategy space. This uses the fact that $T_z(X^p) = T_z(X) \times \cdots \times T_z(X)$. We use the idea of a conic field generated by U^h. Let $D(z) = (\text{Con}U^h)(z)$ be the convex hull in $T_z(X^p)$ of the set of vectors $\{DU_j^h(z) : j \in P\}$. Then $\text{Con}U^h : X^p \to TX^p$ is a generalized vector field over X^p. Its image in T_zX^p is convex and the field is continuous as a correspondence. The Selection Theorem of Michael [28] can be deployed to construct a continuous selection $m : X^p \to TX^p$ of D. That is to say, at every point $z \in X^p$, $m(z) \in D(z) \subset T_zX^p$ (see Schofield [35]).

Note that by this construction each $DU_j^h(z)$ lies in a different tangent space. Thus, $0 \in \text{Con}U^h(z)$ if and only if $DU_j^h(z) = 0$ for every $j \in P$.

We shall now use standard transversality theory to show CNE generally exist, and are “locally isolated.” We say a property of points in X^p is *locally isolated* if it holds for 0-dimensional submanifolds of X^p. This is simply a more general version of the definition given with regard to local equilibria. Say a property K is C^2-generic in E if the set $E^1 = \{U^h \in E : U^h \text{ has property } K\}$ is open dense in the C^2-topology on E.

Definition 5. *Let K_1 be the property that there exists a locally isolated LSNE for the profile U^h and let K_2 be the property that a CNE exists and is locally isolated.*

We seek to show K_1 is C^2-generic. We shall prove this by two lemmas.

Lemma 1. *K_2 is C^2-generic.*

Proof. Given U^h, let $T_j(U^h)$ be the set in the jth strategy space such that $\frac{dU_h^j}{dz_j} = 0$. By the inverse function theorem, $T_j(U^h)$ is generically a smooth submanifold of X^p of dimension $(p-1)\dim(X)$, i.e., of codimension $\dim(X)$. But then $\cap_j T_j(U^h)$ is of codimension $p \cdot \dim(X)$. This is precisely that the property K_2 is C^2- generic.

Definition 6. *A profile U^h satisfies the boundary condition if for any z in the boundary, ∂X^P, the differential $(DU_j^h(z), \ldots, DU_j^h(z))$ points inward.*

This boundary condition simply means that if $\delta(z)$ is the normal to the boundary at z, then $DU_j^h(z)(\delta(z)) > 0$, for all j.

Lemma 2. *If U^h satisfies the boundary condition and if U^h exhibits a locally isolated CNE, then there exists a LSNE.*

Proof. We have defined the conic field $D = \text{Con}(U^h)X^p \to TX^p$ in the preliminary to this section. The conic field D admits a selection $m : X^p \to TX^p$, that is a continuous function such that $m(z) \in \text{Con}(dU^h(z))$. Moreover, $m(z^*) = 0$ if and only if $DU_j^h(z^*) = 0$, for all $j \in P$. Clearly, $0 \in \text{Con}(dU^h(z))$ iff $DU_j^h(z) = 0$ for all $j \in P$. The selection is a vector field on X^p. Moreover, m can be selected to be a gradient vector field on X^p which satisfies the boundary condition, i.e., $m(x)(\delta(x)) > 0$ for all x on the boundary of X^p. Clearly, $m(z^*) = 0$ iff z^* is a CNE. Because of the Morse inequalities (Milnor [30]), m must exhibit a stable equilibrium, z^*, say.

Let $m_j(z)$ be the projection of m at z onto the jth tangent space. Because z^* is a stable equilibrium, there exists a neighborhood Y of z^* with the following property: if $z \in Y$ then for each $j \in P$, $m_j(z)$ “points” towards z^*. In the vector space context this means that $m(z)(z^* - z) > 0$. Let Y_j be the projection of Y onto the jth strategy space. Then $DU_j(z)$ points towards z_j^*

for all $z_j \in Y_j$. The Cartesian product of $\{Y_j\}$ gives a neighborhood of z^* satisfying all second order conditions. Clearly, z^* is an attractor and therefore a LSNE.

(Note that this lemma is proved here for X compact and convex, but it is valid for a gradient field on a space with non-zero Euler characteristic; see Brown [15].)

Theorem 3. *Existence of LSNE is generic in the topological subspace* $E' = \{U^h \in E : U^h$ *satisfies the boundary condition*$\}$.

Proof. By Lemma 1, there exists an open dense subset E'' on which K_2 is satisfied. But then $E' \cap E''$ gives an open dense subspace of E. For every profile in this set, the procedure of Lemma 2 gives a LSNE.

Now let $E^* = \{u : X^{n+p} \to \mathbb{R}^{np}\}$ denote the set of all electoral systems with n voters and p agents. In this formulation the valence functions are treated as parameters. We can also regard the covariance matrix θ, and therefore the error variances $\{\sigma_1^2, \ldots, \sigma_p^2\}$, as fixed, since these are essentially scale factors. We may think of a specific u as the political institution. Once voter ideal points in X^n, and agent strategies in X^p, are specified, then the $n \times p$ array of voter utilities is defined by u. Then $u \in E^*$ and the set of voter ideal points $x \in X^n$ leads to the political game $U^h : \ X^p \to \mathbb{R}^p$.The vector $x \in X^n$ may then be thought of an electoral map $x : E^* \to E$. The electoral models that we have considered here possess the property that this electoral map is continuous. By Theorem 3, existence of LSNE is generic in E' within the space E. Assuming the electoral map x is continuous, the inverse image of an open set in E' is open in the co-image of x. Suppose x is also proper: that is, if Y is open in E^*, then its image is open in E'. In this case, the inverse image of a dense set is dense. This suggests the following conjecture.

Conjecture 1. Existence of LSNE is a generic property in the space of political institutions.

The results on transversality theory used in Theorem 3 can be found in Hirsch [24]. Dierker [18] and Smale [47, 48] discuss the Morse inequalities. A review of this material can be found in Schofield [38].

A minor point concerns the nature of the topology. Theorem 3 uses topologies on utilities. Equilibrium concepts in both politics and economics should be based on preferences. Schofield [35] has proposed a C^1-topology on preferences, which can be used when the utility representation of the preferences are differentiable. This suggests that the Conjecture can be expressed as a genericity result for preferences.

7 Conclusion

In this exposition, I have assumed that the policy space, X, is fixed. In fact, a possible extension is where there is a map $\psi : X \to Z$, where Z is the full economic commodity space. Voters have economic preferences on Z, and these can, in theory, induce preferences on $\psi^{-1}(Z) \subset X$. These preferences then define voter ideal points in X. There are certain subtleties of such a model which were explored by Konishi [25]. The principal difficulty with Konishi's model was in locating appropriate political equilibria in X. This paper has presented one way of inferring existence of LSNE in X^P, given electoral data about voter ideal points. The results of Banks and Duggan [11], for example, can then be used to infer the political outcome in X.

Note, however, that there are a number of further theoretical difficulties. Firstly, the map ψ may be multi-valued. Given the structure of the economy, there will be many possible Pareto incomparable local economic equilibria. Each of these can arise from a single political decision in X. However, it is likely that ψ will be locally single-valued. That is, if the economy is initially at a particular state, π, in Z, then voters may compute back to X to determine how changes in political decisions will affect economic outcomes in a neighborhood of π within Z. Similarly, political decision-making will be locally single-valued. That is, for a given economic and political situation, the local political equilibrium will be determined by the particular basin of attraction within which the status quo is located.

Although this article has focused on party leaders who attempt to maximize vote share, the general model can, in principle, be extended to include more general candidate or leader motivations, as long as these can be assumed to be C^2-differentiable functions.

References

[1] Aldrich, J. (1983) A spatial model with party activists: implications for electoral dynamics. *Public Choice*, 41: 63–100.

[2] Aldrich, J. (1983) A Downsian spatial model with party activists. *American Political Science Review*, 77: 974–990.

[3] Aldrich, J. and M. McGinnis (1989) A model of party constraints on optimal candidate positions. *Mathematical and Computer Modelling*, 42: 437–450.

[4] Alvarez, M. R. and J. Nagler (1998) When politics and models collide: estimating models of multicandidate elections. *American Journal of Political Science*, 42: 55–96.

[5] Alvarez, M. R., J. Nagler and S. Bowler (2000) Issues, economics and the dynamics of multiparty elections: the British 1987 general election. *American Political Science Review*, 94: 131–150.

[6] Ansolabehere, S. and J. M. Snyder (2000) Valence politics and equilibrium in spatial election models. *Public Choice*, 103: 327–336.

[7] Aragones, E and T. Palfrey (2004) Spatial competition between two candidates of different quality: the effects of candidate ideology and private information. This volume.

[8] Austen-Smith, D. and J. S. Banks (1998) Social choice theory, game theory and positive political theory. In N. Polsby (ed) *Annual Review of Political Science* vol I: 259–287.

[9] Austen-Smith, D. and J. S. Banks (1999) *Positive Political Theory I: Collective Preference* Ann Arbor: The University of Michigan Press.

[10] Banks, J. S. (1995) Singularity theory and core existence in the spatial model. *Journal of Mathematical Economics*, 24: 523–536.

[11] Banks, J. S. and J. Duggan (2004) The theory of probabilistic voting in the spatial model of elections: the theory of office-motivated candidates. This volume.

[12] Banks, J. S. and J. Duggan (2000) A bargaining model of collective choice. *American Political Science Review*, 94: 73–88.

[13] Banks, J. S., J. Duggan, and M. Le Breton (2002) Bounds for mixed strategy equilibria and the spatial model of elections. *Journal of Economic Theory*, 103: 88–105.

[14] Banks, J. S., J. Duggan, and M. Le Breton (2004) Social choice and electoral competition in the general spatial model. *Journal of Economic Theory,* forthcoming.

[15] Brown, R. F. (1970) *The Lefshetz Fixed Point Theorem* Glenview IL: Scott and Foreman.

[16] Clarke, H., M. Stewart and P. Whiteley (1998) New models for new labour: the political economy of labour support: January 1992 – April 1997. *American Political Science Review*, 92: 559–575.

[17] Coughlin, P. (1992) *Probabilistic Voting Theory* Cambridge: Cambridge University Press.

[18] Dierker, E. (1974) *Topological Methods in Walrasian Economics* Lecture Notes on Economics and Mathematical Sciences 92 Heidelberg: Springer.

[19] Downs, A. (1957) *An Economic Theory of Democracy* New York: Harper and Row.

[20] Duggan, J. (2000) Equilibrium equivalence under expected plurality and probability of winning maximization. Mimeo. University of Rochester.

[21] Fan, K. (1961) A generalization of Tychonoff's fixed point theorem. *Mathematische Annalen*, 42: 305–310.

[22] Groseclose, T. (2001) A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science*, 45: 862–886.

[23] Hinich, M. (1977) Equilibrium in spatial voting: the median voter result is an artifact. *Journal of Economic Theory*, 16: 208–219.

[24] Hirsch, M. W. (1976) *Differential Topology* Heidelberg: Springer.

[25] Konishi, H. (1996) Equilibrium in abstract political economies: with an application to a public good economy with voting. *Social Choice and Welfare*, 13: 43–50.
[26] Lin, T. M., J. Enelow and H. Dorussen (1999) Equilibrium in multicandidate probabilistic spatial voting. *Public Choice*, 98: 59–82.
[27] McKelvey, R. (1986) Covering, dominance, and institution-free properties of social choice. *American Journal of Political Science*, 30: 283–314.
[28] Michael, E. (1956) Continuous selection I. *Annals of Mathematics*, 63: 361–382.
[29] Miller, G. and N. Schofield (2003) Activists and partisan realignment in the United States. *American Political Science Review*, 97: 245–260.
[30] Milnor, J. (1963) *Morse Theory* Annals of Mathematics Studies, Series 51 Princeton: Princeton University Press.
[31] Nash, J. (1951) Non-cooperative games. *Annals of Mathematics*, 54: 289–295.
[32] Poole, K. and H. Rosenthal (1984) Presidential elections 1968-1980: a spatial analysis. *American Journal of Political Science*, 28: 283–312.
[33] Quinn, K., A. Martin and A. Whitford (1999) Voter choice in multiparty democracies. *American Journal of Political Science*, 43: 1231–1247.
[34] Saari, D. (1997) The generic existence of a core for q–rules. *Economic Theory*, 9: 219–260.
[35] Schofield, N. (1999) The C^1 topology on the space of smooth preference profile. *Social Choice and Welfare*, 16: 445–470.
[36] Schofield, N. (2003a) A valence model of political competition in Britain, 1992–1997. *Electoral Studies,* forthcoming.
[37] Schofield, N. (2003b) Valence competition in the spatial stochastic model. *Journal of Theoretical Politics*, 15: 371–383.
[38] Schofield, N. (2003c) *Mathematical Methods in Economics and Social Choice* Heidelberg: Springer.
[39] Schofield, N. (2004) Equilibrium in the spatial valence model of politics. *Journal of Theoretical Politics*, 16: 447–481.
[40] Schofield, N., A. Martin, K. Quinn and A. Whitford (1998) Multiparty electoral competition in the Netherlands and Germany: a model based on multinomial probit. *Public Choice*, 97: 257–293.
[41] Schofield, N., G. Miller and A. Martin (2003) Critical elections and political realignment in the US: 1860–2000. *Political Studies*, 51: 217–240.
[42] Schofield, N. and R. Parks (2000) Nash equilibrium in a spatial model of coalition bargaining. *Mathematical Social Sciences*, 39: 133–174.
[43] Schofield, N. and I. Sened (2002) Local Nash equilibrium in multiparty politics. *Annals of Operations Research*, 109: 193–210.
[44] Schofield, N. and I. Sened (2003) Multiparty competition in Israel. *British Journal of Political Science,* forthcoming.
[45] Schofield, N., I. Sened and D. Nixon (1998) Nash equilibrium in multiparty competition with 'stochastic' voters. *Annals of Operations Research*, 84: 3–27.

[46] Seyd, P. and P. Whitely (2002) *New Labour's Grassroots* Basingstoke UK: Macmillan.

[47] Smale, S. (1960) Morse inequalities for a dynamical system. *Bulletin of the American Mathematical Society*, 66: 43–49.

[48] Smale, S. (1974) Global analysis and economics I: Pareto optimum and a generalization of Morse theory. In M. Peixoto (ed) *Dynamical Systems* New York: Academic Press.

[49] Stokes, D. (1992) Valence politics. In D. Kavanagh (ed) *Electoral Politics* Oxford: Clarendon Press.