Hindawi Publishing Corporation
International Journal of Mathematics and Mathematical Sciences
Volume 2010, Article ID 562813, 30 pages
doi:10.1155/2010/562813

Research Article

Application of a Theorem in Stochastic Models of Elections

Norman Schofield,[1] Christopher Claassen,[1] Ugur Ozdemir,[1] and Alexei Zakharov[1,2]

[1] *Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA*
[2] *State University, Higher School of Economics, 117218 Novocheremushkinskaya St. 34-1, Apt. 68 Moscow, Russia*

Correspondence should be addressed to Norman Schofield, schofield.norman@gmail.com

Received 2 July 2009; Revised 2 December 2009; Accepted 11 February 2010

Academic Editor: Martin Bohner

Copyright © 2010 Norman Schofield et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Previous empirical research has developed stochastic electoral models for Israel, Turkey, and other polities. The work suggests that *convergence to an electoral center* (often predicted by electoral models) is a nongeneric phenomenon. In an attempt to explain nonconvergence, a formal model based on *intrinsic valence* is presented. This theory showed that there are necessary and sufficient conditions for convergence. The necessary condition is that a convergence coefficient c is bounded above by the dimension w of the policy space, while a sufficient condition is that the coefficient is bounded above by 1. This coefficient is defined in terms of the difference in exogenous valences, the "spatial coefficient", and the electoral variance. The theoretical model is then applied to empirical analyses of elections in the United States and Britain. These empirical models include sociodemographic valence and electoral perceptions of character trait. It is shown that the model implies convergence to positions close to the electoral origin. To explain party divergence, the model is then extended to incorporate activist valences. This extension gives a first-order *balance condition* that allows the party to calculate the optimal marginal condition to maximize vote share. We argue that the equilibrium positions of presidential candidates in US elections and by party leaders in British elections are principally due to the influence of activists, rather than the centripetal effect of the electorate.

1. Introduction

Electoral models based on the work of Hotelling [1] and Downs [2] suggest that parties will converge to an electoral center (at the electoral *median)* when the policy space has a single dimension. Although a pure strategy Nash equilibrium generically fails to exist in competition between two agents under majority rule in high enough dimension, there will exist mixed-strategy equilibria whose support is located near to the electoral center [3]. However, previous empirical research has developed stochastic electoral models for

Argentina, Israel, Russia, Turkey, and other polities [4–10], and has suggested that *divergence from the electoral center* is a generic property of electoral systems.

This paper presents the formal stochastic model based on *electoral valence* to explain nonconvergence of candidates in the 2008 elections in the United States, and in an earlier elction in Britain. The key idea is that the convergence result need not hold if there is an asymmetry in the electoral perception of the "quality" of party leaders [11, 12]. The average weight given to the perceived quality of the leader of the jth party is called the party's *intrinsic* (or *exogenous) valence.* In empirical models, a party's valence is assumed to be independent of the party's position, and adds to the statistical significance of the model. It is obtained from the intercept of the empirical model, and reflects a common perception of the quality of the candidate or party leader. In general, intrinsic valence reflects the overall degree to which the party or candidate is generally perceived to be able to govern effectively [13, 14].

We assume here that, in addition to intrinsic valence, there are three further kinds of valence. The first kind is a *sociodemographic valence*. Empirical models show that different subgroups in the electorate respond to leaders or candidates in different ways. These sociodemographic valences reflect the fact that particular party leaders have established specific political relationships with various political groups that are, at least in the short run, independent of the party's position.

The second type of valence is individual specific, and is defined by individual perception of the character traits of the candidates or party leaders.

The third kind of valence is called *activist (or endogenous) valence*. When party j adopts a policy position z_j, in the policy space, X, then the activist valence of the party is denoted as $\mu_j(z_j)$. Implicitly we adopt a model originally due to Aldrich [15]. In this model, activists provide crucial resources of time and money to their chosen party, and these resources are dependent on the party position. (For convenience, it is assumed that $\mu_j(z_j)$ is only dependent on z_j, and not on z_k, $k \neq j$, but this is not a crucial assumption.) The party then uses these resources to enhance its image before the electorate, thus affecting its overall valence. Although activist valence is affected by party position, it does not operate in the usual way by influencing voter choice through the distance between a voter's preferred policy position, say x_i, and the party position. Rather, as party j's activist support, $\mu_j(z_j)$, increases due to increased contributions to the party in contrast to the support $\mu_k(z_k)$ received by party k, then (in the model) all voters become more likely to support party j over party k.

However, activists are likely to be more extreme than the typical voter. By choosing a policy position to maximize activist support, the party will lose centrist voters. The party must therefore determine the "optimal marginal condition" to maximize vote share. The first result presented here gives this as a (first-order) *balance condition.* Moreover, because activist support is denominated in terms of time and money, it is reasonable to suppose that the activist function will exhibit decreasing returns. We point out that when these activist functions are sufficiently concave, then the model will exhibit a Nash equilibrium, where each party or political candidate adopts a position that maximizes its vote, in response to the positions adopted by the other agents.

This stochastic model is also applied to the case with intrinsic valence alone. For this model, it can be shown that the joint electoral origin satisfies the first-order condition for a Nash equilibrium. Because the vote share functions are differentiable, we make use of calculus techniques, and therefore use the notion of *local Nash equilibrium* (LNE). To determine whether the origin is an LNE, it is necessary to examine the Hessian of the vote share function of the political agent with lowest intrinsic valence. We thus obtain the necessary and sufficient conditions for the validity of the *mean voter theorem* that all agents should converge

to the electoral origin. The second result gives these conditions in terms of a "convergence coefficient" incorporating all the parameters of the intrinsic valence model. This coefficient, c, involves the differences in the intrinsic valences of the agents, and the "spatial coefficient" β. When the policy space, X, is assumed to be of dimension w, then the necessary condition for existence of a Nash equilibrium when all agents are located at the electoral origin is that the coefficient c is bounded above by w. When the necessary condition fails, then agents, in equilibrium, will adopt divergent positions.

In the next section we briefly sketch the nature of the local Nash equilibria in political games involving these different types of electoral valences. We focus on candidates in the 2008 US Presidential election, in order to illustrate our results. The formal models are presented in Section 3. There we formally introduce the notion of a *local Nash equilibrium*, and then show that the unique Nash equilibrium in the presidential campaign of 2008 should be where both candidates adopted positions very close to the electoral origin. Since the candidates did not adopt such convergent positions, we can estimate the effect of activists in this election. We then follow up with a brief analysis of the 1979 general election in Britain, and show again that the empirical model indicated convergence. Again, nonconvergence of the parties allows us to estimate the effect of activists. In the conclusion, we offer some general remarks about Madison's argument about the "probability of a fit choice."

2. Activist Support for the Parties

The main result of this paper can be applied to analysis of the equilibrium candidate positions $(z^*_{\text{dem}}, z^*_{\text{rep}})$ in a two-candidate game of vote maximization in a US election. It is shown here that the the first-order condition is given by a *balance equation*. This means that, for each party j = dem or rep, there is a *weighted electoral mean* for party j, given by the expression

$$z_j^{el} = \sum_{i=1}^{n} \varpi_{ij} x_i, \tag{2.1}$$

and which is determined by the set of voter preferred points $\{x_i\}$. Notice that the coefficients $\{\varpi_{ij}\}$ for candidate j will depend on the position of the other candidate, k. Define the *centripetal marginal electoral pull* for candidate j, at z_j, by

$$\frac{d\mathcal{E}_j^*}{dz_j}(z_j) = \left[z_j^{el} - z_j\right]. \tag{2.2}$$

The influence of activists on candidate j is given by *the marginal activist pull for party* j

$$\left[\left.\frac{d\mu_j}{dz_j}\right|_{z_j}\right]. \tag{2.3}$$

The first-order *balance equation* for equilibrium is that the position z_j^* for each j must satisfy the gradient equation

$$\frac{d\mathcal{E}_j^*}{dz_j}\left(z_j^*\right) + \frac{1}{2\beta}\left[\left.\frac{d\mu_j}{dz_j}\right|_{z_j}\right] = 0. \tag{2.4}$$

The locus of points satisfying this equation is called the *balance locus* for the party.

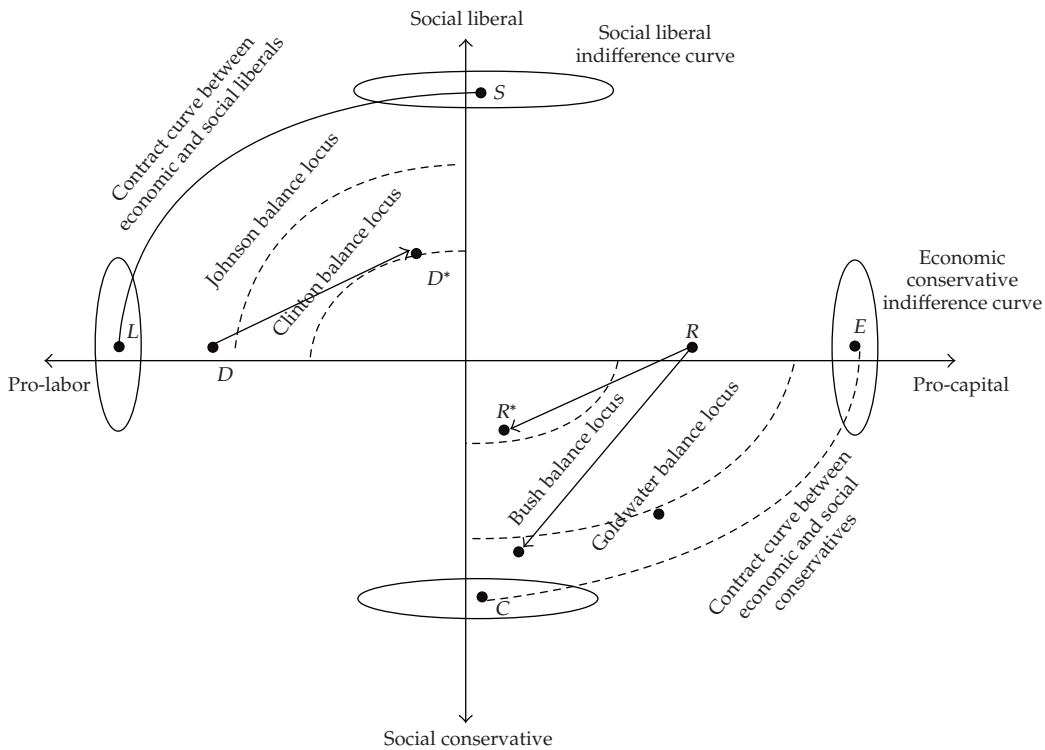

Figure 1: The balance loci in the United States.

To illustrate this model, consider Figure 1 which illustrates elections in the US. Our empirical analysis indicates that there are two dimensions, economic and social. Consider initial positions R and D, on either side of and approximately equidistant from the origin, as in the figure. Both social conservative activists, represented by C, and social liberal activists represented by S, would be indifferent between both parties. A Democratic candidate by moving to position D^* will benefit from activist support of the social liberals, but will lose some support from the economic liberal activists at L. The "contract curve" between the two activist groups, centered at L and S, represents the set of conflicting interests or "bargains" that can be made between these two groups over the policy to be followed by the candidate. In the figure, the indifference curves of the activist groups are shown to be eccentric, with economic activists much less concerned about social policy, and social activists less concerned about economic policy. Under this assumption, it can be shown that this contract curve is a *catenary* whose curvature is determined by the "eccentricities" of the utility functions of the activist groups. We therefore call this contract curve *the Democratic activist catenary*. It is obtained by shifting the appropriate activist catenary towards the weighted electoral mean of the party. The *marginal activist pull for party* j (at a position z_j) is a gradient vector, $d\mu_j/dz_j|_{z_j}$, which represents the marginal effect of the activist groups on the party's valence. The gradient term $(d\mathcal{E}_j^*/dz_j)(z_j)$ is the *marginal electoral pull of party* j (at z_j), and this pull is zero at $z_j^* = z_j^{el}$. Otherwise, it is a vector pointing towards z_j^{el}.

To illustrate, the pair of positions (D^*, R^*) in Figure 1 are equilibrium candidate positions that maximize each candidate's vote share.

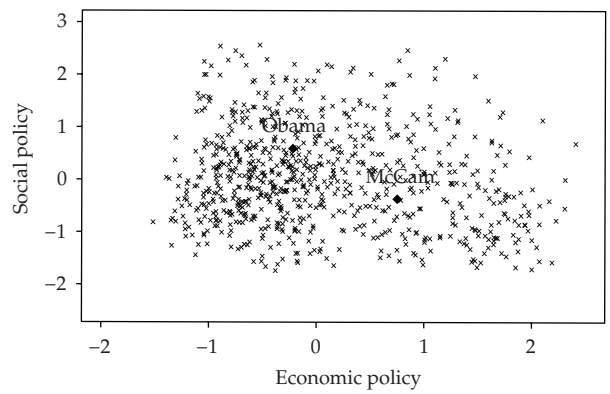

Figure 2: Distribution of voter ideal points and candidate positions.

The positioning of R^* in the lower right electoral quadrant in Figure 1 and of D^* in the upper left quadrant is meant to indicate the realignment that has occurred since the election victory of Kennedy over Nixon in 1960. By 1964 Lyndon Johnson had moved away from a typical New Deal Democratic position, L, to a position comparable to D^*. The long-term effect of this transformation was that by 2000; most of the southern states had become dominated by the Republican party. Empirical analysis of this election suggests that the intrinsic valence of Johnson was greater than that of Goldwater. (See [8, 16].) According to the activist model, this implies that Goldwater's dependence on activist support was greater than Johnson's. This is reflected in Figure 1, where the balance locus for Goldwater is shown to be further from the electoral origin than the balance locus for Johnson. From this we can infer the influence of activists on the two-candidates, thus providing an explanation why socially conservative activists responded so vigorously to the new Republican position adopted by Goldwater, and came to dominate the Republican primaries in support of his proposed policies. These characteristics of the balance solution appear to provide an explanation for Johnson's electoral landslide in 1964.

In this paper we shall apply the electoral model to account for the positions of Obama and McCain in the 2008 presidential election in the context of an electoral distribution, obtained from the American National Election Survey (ANES). Figure 2 shows the estimated voter distribution together with these estimated candidate positions.

We first present the formal stochastic model and then give the empirical analysis of this election.

3. The Formal Stochastic Model

Details of the spatial stochastic electoral models are published in [17, 18]. This model is an extension of the standard multiparty stochastic model [19] modified by inducing asymmetries in terms of valence.

We define a stochastic electoral model, which utilizes sociodemographic variables and voter perceptions of character traits. For this model we assume that voter i utility is given by the expression

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) \ - \beta\|x_i - z_j\|^2 + \varepsilon_j, \tag{3.1}$$

$$= u^*_{ij}(x_i, z_j) + \varepsilon_j. \tag{3.2}$$

Here $u^*_{ij}(x_i, z_j)$ is the observable component of utility, while $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ is the *intrinsic* valence vector, which we assume satisfies the ranking condition $\lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1$.

The political agents (who may be presidential candidates, in US elections, or party leaders, as in British elections) are denoted as $(1, \ldots, p)$. The points $\{x_i : i\epsilon N\}$ are the preferred policies, in a space X, of the voters and $\mathbf{z} = \{z_j : j\epsilon P\}$ are the positions, in X, of the agents. The term

$$\left\|x_i - z_j\right\|^2 = \Sigma_{t=1}^{w}\left(x_{it} - z_{jt}\right)^2 \tag{3.3}$$

is simply the Euclidean distance between x_i and z_j. The error vector $\boldsymbol{\varepsilon} = (\varepsilon_{1.}, \ldots, \varepsilon_j, \ldots, \varepsilon_p)$ is distributed by the type I extreme value distribution, as assumed in empirical conditional logit estimation. In empirical models, the valence vector $\boldsymbol{\lambda}$ is given by the intercept term for each agent in the model. The symbol $\boldsymbol{\theta}$ denotes a set of k-vectors $\{\theta_j : j\epsilon P\}$ representing the effect of the k different sociodemographic parameters (class, domicile, education, income, religious orientation, etc.) on voting for agent j while η_i is a k-vector denoting the ith individual's relevant "sociodemographic" characteristics. The compositions $\{(\theta_j \cdot \eta_i)\}$ are scalar products, called the *sociodemographic valences* for j.

The terms $\{(\boldsymbol{\alpha}_j \cdot \tau_i)\}$ are scalars giving voter i's perceptions and beliefs. These can include perceptions of the character traits of agent j, or beliefs about the state of the economy, and so forth. We let $\boldsymbol{\alpha} = (\alpha_p, \ldots, \alpha_1)$. A *trait score* can be obtained by factor analysis from a set of survey questions asking respondents about the traits of the agent, including "moral", "caring", "knowledgable", "strong", "honest", "intelligent", and so forth. The perception of traits can be augmented with voter perception of the state of the economy, in order to examine how anticipated changes in the economy affect each agent's electoral support.

The terms $\{\mu_j : j\epsilon P\}$ are the *activist valence functions* . The full model including activists is denoted as $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$.

Partial models are:

(i) *pure sociodemographic,* denoted as $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta})$, with only intrinsic valence and sociodemographic variables,

(ii) *pure spatial,* denoted as $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, with only intrinsic valence and β,

(iii) *joint spatial,* denoted as $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, with intrinsic valence, sociodemographic variables and β,

(iv) *joint spatial model with traits,* denoted as $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, without the activist components.

In all models, the probability that voter i chooses agent j, when agent positions are given by $\mathbf{z}$, is

$$\rho_{ij}(\mathbf{z}) = \Pr\left[\left[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)\right],\ \forall l \neq j\right]. \tag{3.4}$$

A strict local Nash equilibrium (LNE) for a model $\mathbb{M}$ is a vector, $\mathbf{z}$, such that each agent, j, chooses z_j to locally strictly maximize the expected vote share

$$V_j(\mathbf{z}) = \frac{1}{n}\Sigma_i \rho_{ij}(\mathbf{z}), \quad \text{subject to } \mathbf{z}_{-j} = \left(z_1, \cdots z_{j-1}, z_{j+1,} \cdots z_p\right). \tag{3.5}$$

In these models, political agents cannot know precisely how each voter will choose at the vector $\mathbf{z}$. The stochastic component as described by the vector $\boldsymbol{\varepsilon}$ is one way of modeling the degree of risk or uncertainty in the agents' calculations. Implicitly we assume that they can use polling information and the like to obtain an approximation to this stochastic model in a neighborhood of the initial candidate locations. For this reason we focus on LNE. (Halpern [20] gives some objections to the concept of Nash equilibrium, in terms of computability and the knowledge requirement of agents, and this provides some basis for our use of LNE. In the empirical work presented below, we find that LNE and PNE coincide.) Note, however, that as agents adjust position in response to information in search of equilibrium then the empirical model may become increasingly inaccurate.

A strict Nash equilibrium (PNE) for a model $\mathbb{M}$ is a vector $\mathbf{z}$ which globally strictly maximizes $V_j(\mathbf{z})$. Obviously if $\mathbf{z}$ is not an LNE then it cannot be a PNE.

It follows from [21] that, for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, the probability, $\rho_{ij}(\mathbf{z})$, that voter i, with ideal point, x_i, picks j at the vector, $\mathbf{z}$, of agent positions is given by

$$\rho_{ij}(\mathbf{z}) = \left[1 + \Sigma_{k \neq j}\left[\exp\left(f_{kj}\right)\right]\right]^{-1}, \tag{3.6}$$

where

$$f_{kj} = u^*_{ik}(x_i, z_k) - u^*_{ij}(x_i, z_j). \tag{3.7}$$

Thus

$$\begin{aligned} \frac{d\rho_{ij}(\mathbf{z})}{dz_j} &= -\left[1 + \Sigma_{k \neq j}\left[\exp\left(f_{kj}\right)\right]\right]^{-2} \frac{d}{dz_j}\left[\Sigma_{k \neq j}\left[\exp\left(f_{kj}\right)\right]\right] \\ &= \left\{2\beta\left(x_i - z_j\right) + \left.\frac{d\mu_j}{dz_j}\right|_{z_j}\right\}\left[\rho_{ij} - \rho_{ij}^2\right]. \end{aligned} \tag{3.8}$$

We use this gradient equation in the form of MATLAB algorithms, given in Appendices A and B to obtain the LNE. This equation shows that the first-order condition for $\mathbf{z}^*$ to be an LNE is given by

$$\begin{aligned} 0 = \frac{dV_j(\mathbf{z})}{dz_j} &= \frac{1}{n}\sum_{i \in N}\frac{d\rho_{ij}}{dz_j} \\ &= \frac{1}{n}\sum_{i \in N}\left[\rho_{ij} - \rho_{ij}^2\right]\left\{2\beta\left(x_i - z_j\right) + \left.\frac{d\mu_j}{dz_j}\right|_{z_j}\right\}. \end{aligned} \tag{3.9}$$

Hence

$$\begin{aligned} z_j^* &= \sum_i \frac{\left[\rho_{ij} - \rho_{ij}^2\right]x_i}{\sum_{k \in N}\left[\rho_{kj} - \rho_{kj}^2\right]} + \frac{1}{2\beta}\left.\frac{d\mu_j}{dz_j}\right|_{z_j} \\ &= \sum_{i=1}^{n}\varpi_{ij}x_i + \frac{1}{2\beta}\left.\frac{d\mu_j}{dz_j}\right|_{z_j}. \end{aligned} \tag{3.10}$$

This can be written as

$$0 = \left[z_j^{el} - z_j^*\right] + \frac{1}{2\beta}\frac{d\mu_j}{dz_j}\bigg|_{z_j}, \tag{3.11}$$

where

$$\begin{aligned} z_j^{el} &= \sum_{i=1}^{n} \varpi_{ij} x_i, \\ \varpi_{ij} &= \frac{\left[\rho_{ij} - \rho_{ij}^2\right]}{\sum_{k\in N}\left[\rho_{kj} - \rho_{kj}^2\right]}. \end{aligned} \tag{3.12}$$

Here z_j^{el} is the *weighted electoral mean* of agent j. Because this model is linear, it is possible to modify these weights to take account of the differential importance of voters in different constituencies. (For example, presidential candidates may attempt to maximize total electoral votes, so voters can be weighted by the relative electoral college seats of the state they reside in.) We can therefore write the first-order balance condition at an equilibrium, $\mathbf{z}^* = (z_1^*, \ldots, z_j^*, \ldots, z_p^*)$, as a set of *gradient balance conditions*

$$\frac{d\mathcal{E}_j^*}{dz_j}\left(z_j^*\right) + \frac{1}{2\beta}\frac{d\mu_j}{dz_j}\bigg|_{z_j^*} = 0. \tag{3.13}$$

The first term in this equation is the *centripetal marginal electoral pull* for agent j, defined at z_j by

$$\frac{d\mathcal{E}_j^*}{dz_j}(z_j) = \left[z_j^{el} - z_j\right]. \tag{3.14}$$

The second gradient term, $d\mu_j/dz_j|_{z_j}$, is the *centrifugal marginal activist pull for j, at z_j.*

To determine the LNE for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, it is of course necessary to consider the Hessians $dV_j^2(\mathbf{z})/dz_j^2$. These will involve the second-order terms $d^2\mu_j/dz_j^2$. In the next section, we suggest that there will be natural conditions under which these will be negative definite. Indeed if the eigenvalues are negative and of sufficiently large modulus, then we may expect the existence of PNE.

For the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, it is clear that when the agents adopt the same positions then $\rho_{kj} = \rho_j$ is independent of the voter suffix, k. Thus all $\varpi_{ij} = 1/n$ gives the first-order condition for an LNE. By a change of coordinates, it follows that $\mathbf{z}_0 = (0, \ldots, 0)$ is a candidate for an LNE. Note however that this argument does not follow for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, and generically $\mathbf{z}^{el} = (z_1^{el}, z_2^{el}, \ldots, z_p^{el}) \neq (0, \ldots, 0)$.

Since the valence functions are constant in the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, the marginal effects, $d\mu_j/dz_j$, will be zero. However, since the weights in the weighted electoral mean for each agent will vary from one individual to another, it is necessary to simulate the model to determine the LNE $\mathbf{z}^{el} = (z_1^{el}, z_2^{el}, \ldots, z_p^{el})$. Notice also that the marginal vote effect, $d\rho_{ij}/dz_j$, for a voter with $\rho_{ij}(\mathbf{z}) \simeq 1$ will be close to zero. Thus in searching for LNE, each agent will seek voters with $\rho_{ij}(\mathbf{z}) < 1$.

The necessary and sufficient second-order condition for LNE at $\mathbf{z}_0$ in the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, is determined as follows. When all agents are at the electoral origin, and agent 1 is, by definition, the lowest valence agent, then the probability that a generic voter picks agent 1 is given by:

$$\rho_1 = \left[1 + \Sigma_{k \neq 1}\left[\exp(\lambda_k - \lambda_1)\right]\right]^{-1}. \tag{3.15}$$

To compute the Hessian of agent 1, we proceed as follows:

$$\begin{aligned}
\frac{dV_1^2(\mathbf{z})}{dz_1^2} &= \frac{d}{dz_1}\frac{1}{n}\sum_{i\in N} 2\beta(z_1 - x_i)\left[\rho_{i1}^2 - \rho_{i1}\right] \\
&= \frac{2\beta}{n}\sum_{i\in N}\left[\left\{\left[\rho_{i1}^2 - \rho_{i1}\right]\frac{d}{dz_1}(z_1 - x_i)\right\} + \left\{(2\rho_{i1} - 1)\frac{d\rho_{i1}}{dz_1}.(z_1 - x_i)\right\}\right] \\
&= \frac{2\beta}{n}\sum_{i\in N}\left[\rho_{i1} - \rho_{i1}^2\right]\left\{2\beta(1 - 2\rho_{i1})(x_i - z_1)^T(x_i - z_1) - I\right\}.
\end{aligned} \tag{3.16}$$

Here I is the w by w identity matrix, and we use ^{T} to denote a column vector. When all agents are at the same position, then $\rho_{i1} = \rho_1$ is independent of i. Moreover,

$$\frac{1}{n}\sum_{i\in N}(x_i)^T(x_i) = \nabla_0 \tag{3.17}$$

is the w by w covariance matrix of the distribution of voter ideal points, taken about the electoral origin. Thus the Hessian of the vote share function of agent 1 at $\mathbf{z}_0 = (0, \ldots, 0)$ is given by

$$2\beta\left[\rho_1 - \rho_1^2\right]\{2\beta(1 - 2\rho_1)\nabla_0 - I\}. \tag{3.18}$$

Since $[\rho_1 - \rho_1^2] > 0,\ \beta > 0$, this Hessian can be identified with the w by w *characteristic matrix* for agent 1, given by

$$C_1 = 2\beta(1 - 2\rho_1)\nabla_0 - I. \tag{3.19}$$

Then the necessary and sufficient second-order condition for LNE at $\mathbf{z}_0$ is that C_1 has negative eigenvalues. (For convenience we focus on a strict local equilibrium associated with negative eigenvalues of the Hessian.)

It follows from this that a *necessary* condition for $\mathbf{z}_0 = (0, \ldots, 0)$ to be an LNE is that the trace of the matrix C_1 is strictly negative. (For a weak LNE we require the trace to be nonpositive.) In turn this means that a *convergence coefficient*, c, defined by

$$c = 2\beta(1 - 2\rho_1)\sigma^2 \tag{3.20}$$

satsfies the critical convergence condition, $c < w$. Here $\sigma^2 = \text{trace}(\nabla_0)$ is the sum of the variance terms on all axes.

A *sufficient* condition for convergence to $\mathbf{z}_0$ in the two-dimensional case is that $c < 1$.

When the necessary condition fails, then the lowest valence agent has a best response that diverges from the origin. In this case there is no guarantee of existence of a PNE.

We can also consider a model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta})$ where we use different coefficients $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_w)$ on the axes, so the spatial component has the form

$$\Sigma_{t=1}^{w} \beta_t (x_{it} - z_{jt})^2. \tag{3.21}$$

Then the characteristic matrix can be taken to be

$$\mathbf{C}_1 = 2(1 - 2\rho_1)\boldsymbol{\beta}\nabla_0\boldsymbol{\beta} - \boldsymbol{\beta}, \tag{3.22}$$

where $\boldsymbol{\beta}$ is the diagonal matrix of the β coefficients, while $\boldsymbol{\beta}\nabla_0\boldsymbol{\beta}$ is the covariance matrix where each axis is weighted by the β coefficients $\beta_1, \beta_2, \ldots, \beta_w$. The necessary condition is thus that $\text{trace}(\mathbf{C}_1) < 0$,

or

$$2(1 - 2\rho_1)\text{trace}(\boldsymbol{\beta}\nabla_0\boldsymbol{\beta}) < \beta_1 + \beta_2 \cdots + \beta_w. \tag{3.23}$$

Because the model is linear, we can obtain a similar result where there a multiple electoral groups, each weighting the axes differently.

In the empirical analyses, we can used Newton's method with gradient information to compute best responses, in order to determine LNE in the various models.

3.1. Application to the Case with Multiple Activist Groups

We adapt the model presented by Schofield and Cataife in [4], where there are multiple activist groups for each party.

(i) For each agent, j, let $\{A_j\}$ be a family of potential activists, where each $k \in A_j$ is endowed with a utility function, U_k, which is a function of the position z_j. The resources allocated to j by k are denoted as $R_{jk}(U_k(z_j))$. The total activist valence function for agent j is the linear combination

$$\mu_j(z_j) = \sum_{k \in A_j} \mu_{jk}(R_{jk}(U_k(z_j))), \tag{3.24}$$

where $\{\mu_{jk}\}$ are functions of the contributions $\{R_{jk}(U_k(z_j)\}$, and each μ_{jk} is a concave function of R_{jk}.

(ii) Assume that the gradients of the valence functions for j are given by

$$\frac{d\mu_{jk}}{dz_j} = a_k^* \frac{dR_{jk}}{dz_j} = a_k^* a_k^{**} \frac{dU_k}{dz_j}, \tag{3.25}$$

where the coefficients $\{a_k^*, a_k^{**}\} > 0$ are differentiable functions of z_j.

(iii) Under these assumptions, the first-order equation $d\mu_j/dz_j = 0$ becomes

$$\begin{aligned}\frac{d\mu_j}{dz_j} &= \sum_{k\in A_j} \frac{d}{dz_j}\left[\mu_{jk}\left(R_{jk}\left(U_k(z_j)\right)\right)\right] \\ &= \sum_{k\in A_j} \left(a_k^{**} a_k^*\right) \frac{dU_k}{dz_j} = 0.\end{aligned} \tag{3.26}$$

The *Contract Curve* generated by the family $\{A_j\}$ is the locus of points satisfying the gradient equation

$$\sum_{k\in A_j} a_k \frac{dU_k}{dz_j} = 0, \quad \text{where } \sum_{k\in A_j} a_k = 1 \quad \text{and all } a_k > 0. \tag{3.27}$$

The *Balance Locus* for agent j, defined by the family $\{A_j\}$, is the solution to the first-order gradient equation

$$\left[z_j^{el} - z_j^*\right] + \frac{1}{2\beta}\left[\sum_{k\in A_j} a_k \frac{dU_k}{dz_j}\right] = 0. \tag{3.28}$$

The simplest case, discussed in [4], is in two dimensions, where each agent has two supporting activist groups. In this case, the contract curve for each agent's supporters will, generically, be a one-dimensional arc. Miller and Schofield [22] also supposed that the activist utility functions were ellipsoidal, mirroring differing saliences on the two axes. As discussed earlier, in this case the contract curves would be *catenaries*, and the balance locus would be a one-dimensional arc. The balance solution for each agent naturally depends on the position(s) of opposed agent (s), and on the coefficients, as indicated above, of the various activists. The determination of the balance solution can be obtained by computing the vote share Hessian along the balance locus.

Since the activist valence function for agent j depends on the resources contributed by the various activist groups to this agent, we may expect the marginal effect of these resources to exhibit diminishing returns. Thus the activist valence functions can be expected to be concave in the activist resources, so that the Hessian of the overall activist valence, μ_j, can be expected to have negative eigenvalues. When the activist functions are sufficiently concave (in the sense that the Hessians have negative eigenvalues of sufficiently large modulus), then we may infer not only that the LNE will exist, but that they will be PNE.

If we associate the utilities $\{U_k\}$ with leaders of the activist groups for the agents, then the combination

$$\sum_{k\in A_j} a_k \frac{dU_k}{dz_j} \tag{3.29}$$

may be interpreted as the marginal utility of the candidate of party j, induced by the activist support.

To see this, suppose that each agent were to maximize the function

$$V_j(\mathbf{z}) = \delta\mu_j(z_j) + \frac{1}{n}\Sigma_i\rho_{ij}(\mathbf{z}), \tag{3.30}$$

where μ_j is no longer an activist function, but a policy-determined component of the agent's utility function, while δ is the weight given to the policy preference. See [23] for such a model of policy-motivated agents. Then the first-order condition is almost precisely as we obtained above, namely

$$\frac{d\mathcal{E}_j^*}{dz_j}\left(z_j^*\right) + \frac{\delta}{2\beta}\frac{d\mu_j}{dz_j}\bigg|_{z_j^*} = 0. \tag{3.31}$$

Here $(d\mu_j/dz_j)(z_j^*)$ is a gradient pointing towards the policy preferred position of the agent. Thus we can make the identity

$$\delta\frac{d\mu_j}{dz_j}\bigg|_{z_j} = \sum_{k\in A_j} a_k\frac{dU_k}{dz_j} \tag{3.32}$$

and infer that agent's marginal policy preference can be identified with a combination of the marginal preferences of the party activists. In principle such a model could be used to determine optimal resource-raising strategies in an environment as complex as a presidential election.

4. Methodology: A Spatial Model of the 2008 Election

The 2008 American National Election Study (ANES) introduced many new questions on political issues in addition to the existing set. Assignment of respondents into the "new" or "old" set was random, with 1,059 respondents assigned to the "new" condition and having completed the followup post-election interview. Due to both Hispanic and African-American voter oversampling and followup attrition, the postelection weights are used for all analyses. As with all survey data, there was missing data for most of the survey items used in this study (varying from 0 to 8.6% by item). We used multiple imputation to correct for missing data.

The post-election interviews asked repondents whom they voted for, if at all. Since we use a conditional logit model, which requires data for both respondents (which we have) and candidates (which we only have for the major party candidates), we deleted 7 observations where respondents claimed to have voted for a presidential candidate other than McCain or Obama. The final sample size was thus 788 respondents.

To create the two-dimensional policy space, 29 survey items were selected to broadly represent the economic and social policy dimensions of American political ideology (see Appendix C for question wording). Some issues were overrepresented amongst these item, with seven questions about abortion, four for gay rights and policies concerning aid for African-Americans, and two about immigration issues. To avoid the policy space measure becoming dominated by these issues, with abortion a particular concern, separate scales

Table 1: Factor loadings for abortion.

Question	Abortion Issues
(12) Health	0.79
(13) Fatal	0.56
(14) Incest	0.70
(15) Rape	0.68
(16) Birth defect	0.74
(17) Finance	0.66
(18) Gender	0.47
Eigenvalue	3.19
Cronbach's alpha	0.83

Table 2: Factor loadings for gay issues.

Question	Gay Issues
(19) Military	0.49
(20) Job discrimination	0.38
(21) Marriage	0.75
(22) Adoption	0.76
Eigenvalue	1.52
Cronbach's alpha	0.64

were estimated for each of these policy areas, either using confirmatory factor analysis or a simple average in the case of the two immigration items. (see Tables 1, 2, and 3.) Finally, a confirmatory factor analysis was run using these four scales in conjunction with the remaining 12 survey items. Only two factors achieved eigenvalues greater than one. Each factor corresponded closely to a priori conceptualizations of economic and social policy, with the possible exceptions of the equality and gun access items, which loaded more strongly on the economic rather than social dimensions. (see Table 4 for factor loadings.) These factor scores were used as measures of individual locations on the policy space.

The ANES also includes questions on seven qualities or traits associated with Obama and McCain. Confirmatory factor analysis run on the 14 items produced a two-factor solution which corresponded perfectly with the named candidate. The resulting factor scores were used as estimates of voter perceptions of the candidate's personal traits. (see Table 5.)

Respondents were coded as activists if they claimed to have donated money to a candidate or party and nonactivists if they donated money to no candidate. Table 6 gives the descriptive data for activists and nonactivists.

The survey also gave data on whether the respondent was African-American, female, working class, from the South, as well as the number of years of education and level of income. These data were used to construct the sociodemographic models of voting.

To calculate the presidential candidate positions, we took advantage of new survey questions which asked respondents to locate the positions of Obama and McCain on seven distinct issues.

These seven questions (government spending, universal health care, citizenship for immigrants, abortion when nonfatal, abortion when gender incorrect, aid to blacks, and liberal-conservative) were otherwise worded the same as the corresponding items from the policy issue questions.

Table 3: Factor loadings for Black issues.

Question	Black Issues
(25) Government aid	0.44
(26) Special favors	0.71
(27) Slavery discrimination	0.41
(28) Try harder	0.69
Eigenvalue	1.35
Cronbach's alpha	0.64

Table 4: Factor loadings for economic and social policy.

Question	Economic Policy	Social Policy
(1) Government services	0.53	0.12
(2) Universal health care	0.51	0.22
(4) Government bigger	0.50	0.14
(5) Government or market	0.56	
(9) Welfare spending	0.24	
(6) Less government	0.65	
(7) Equality	0.14	0.37
(8) Tax Companies	0.28	0.10
(12)–(18) Abortion scale		0.55
(11) Immigrant scale	0.12	0.25
(19)–(22) Gay scale		0.60
(23) Traditional values		0.53
(24) Gun access	0.36	
(25)–(28) Black scale	0.14	0.45
(29) Liberal v conservative	0.30	0.60
Eigenvalue	1.93	1.83

Table 5: Factor loadings for candidate traits scores 2008

Question	Obama Traits	McCain Traits
Obama Moral	.72	−.01
Obama Cares	.71	−.18
Obama Knowledgable	.61	−.07
Obama Strong	.69	−.13
Obama Honest	.68	−.09
Obama Intelligent	.61	.08
Obama Optimistic	.55	.00
McCain Moral	−.09	.67
McCain Cares	−.17	.63
McCain Knowledgable	−.02	.65
McCain Strong	−.10	.70
McCain Honest	−.03	.63
McCain Intelligent	.11	.68
McCain Optimistic	−.07	.57
Eigenvalue	3.07	3.00

Table 6: Descriptive data.

	Econ Mean	Policy s.e.	95% C.I.	Social Mean	Policy s.e.	95% C.I.	n
Activists							
Democrats	−0.20	0.09	[−0.38, −0.02]	1.14	0.11	[0.92, 1.37]	80
Republicans	1.41	0.13	[1.66, 1.16]	−0.82	0.09	[−0.99, −0.65]	40
Nonactivists							
Democrats	−0.17	0.03	[−0.24, −0.11]	0.36	0.04	[0.29, 0.44]	449
Republicans	0.72	0.06	[0.60, 0.84]	−0.56	0.05	[−0.65, −0.46]	219
Total							788

Table 7: Obama and McCain perceived positions.

Question	Obama	McCain
Estimated position on economic policy	−0.22	0.59
Estimated position on social policy	0.75	−0.37

We ran two linear regression models on the voter economic policy and social policy factor scores using only the seven policy items corresponding in wording to the seven candidate location items as predictors. The estimated coefficients from these two linear models enabled us to construct equations to map the data from the candidate location questions onto the complete voter policy space. These equations were able to predict the scores of the voter policy space fairly accurately. The coefficient of determination (R^2) for the economic and social policy equations were 0.63 and 0.75, respectively. To find McCain's ideal point, we simply took the average response for each of his seven candidate location questions, entered these into the economic and social policy prediction equations, and used the corresponding predicted values. We then repeated the process using Obama's candidate location questions. See Table 7 for the estimated positions of the two-candidates.

Figure 2 previously gave a plot of the voter distribution, while Figure 3 shows the perspective plot of the voter distribution. The plots of the activist positions are shown in Figure 4. Finally Figure 5 gives a smoothed contour plot of the probability density function of the voter distribution (The outer contour line is at the 0.05 level, while Democrat activists are denoted in red, and Republicans in blue.).

Figure 5 also shows the estimated threshold dividing likely Democrat candidate voters from Republican candidate voters. This partisan cleavage line was derived from a binomial logit model, designed to test the effects of each policy dimension on vote choice. We call this the *pure positional binomial logit model*.

According to the positional model, a voter i, with preferred position (x_i, y_i) is estimated to vote Republican with probability

$$\rho_{\text{rep}} = \frac{\exp(\lambda_r + bx_i + cy_i)}{1 + \exp(\lambda_r + bx_i + cy_i)}, \tag{4.1}$$

where $(\lambda_r, b, c) = (-0.74, 1.49, -1.80)$

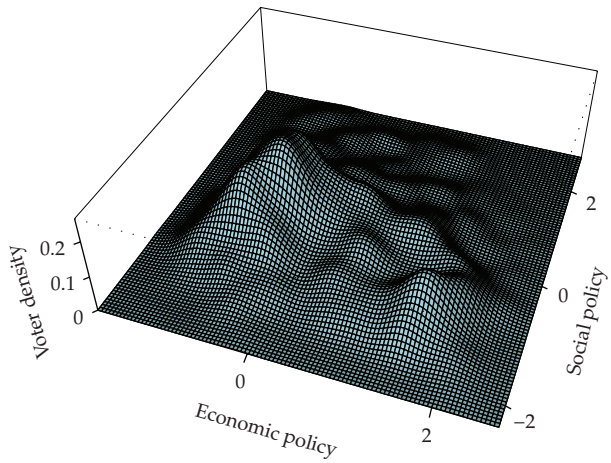

Figure 3: Perspective plot of the sample electorate in 2008.

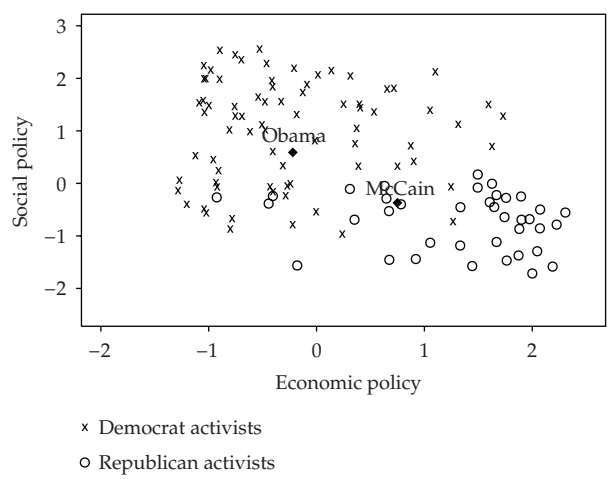

Figure 4: Distribution of activist ideal points and candidate positions.

That is, any voter with preferred point lying on the cleavage line has equal probability of picking one or other of the candidates. This cleavage line is given by the equation

$$y = 0.82x - 0.4. \tag{4.2}$$

This cleavage line misses the origin, and goes through the point $(0, -0.4)$, indicating the valence advantage of Obama. The coefficient λ_r is a measure of the negative relative valance of McCain with respect to Obama. This cleavage line is similar to those obtained by Schofield et al. [8].

The positional model does not allow us to estimate equilibria, so we now turn to the pure spatial model.

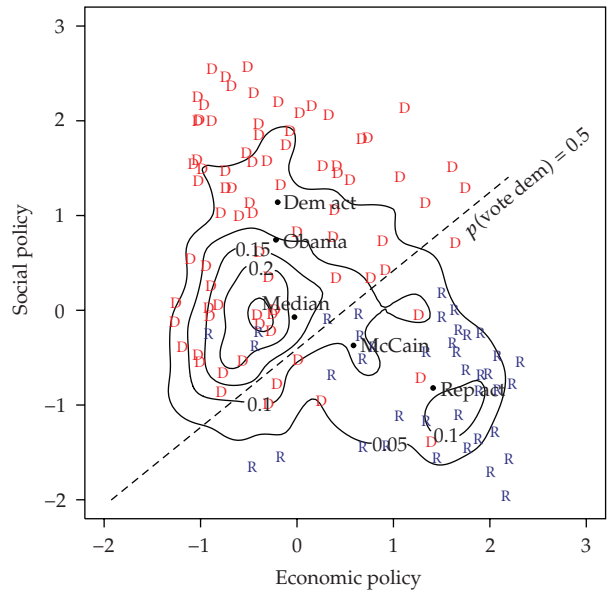

Figure 5: Smoothed voter distribution; candidate positions and activist positions in 2008. (Democrat activists are red and Republican activists are blue, together with cleavage line.)

4.1. Computation of Equilibria for the US 2008 Election

As above, we first assume that the utility of voter i for candidate j is given by the pure spatial model

$$u_{ij}(x_i, z_j) = \lambda_j - \beta \|x_i - z_j\|^2 + \varepsilon_j. \tag{4.3}$$

We assume that each candidate, j, chooses z_j to locally strictly maximize the expected vote share

$$V_j(\mathbf{z}) = \frac{1}{n} \Sigma_i \rho_{ij}(\mathbf{z}), \tag{4.4}$$

subject to the position(s) of the other candidates. We essentially assume therefore that candidates cannot know precisely how voters choose but they can estimate the relationship between their own position, that of the competing candidate, and the aggregate vote total. As we shall see, the induced candidate preference correspondences are convex valued, indicating existence of Nash equilibria. The local pure strategy Nash equilibria (LNE) can be computed as follows.

The electoral covariance matrix for the sample is given by

$$\nabla_0 = \begin{bmatrix} 0.80 & -0.127 \\ -0.127 & 0.83 \end{bmatrix}. \tag{4.5}$$

Table 8: β-spatial conditional logit models for USA (2008).

	(1) Spatial	(2) Sp. & Traits	(3) Sp. & Demog	(4) Full
McCain valence λ	−0.84***	−1.08***	−2.60**	−3.58***
	(0.11)	(0.13)	(0.93)	(1.05)
Distance β	0.85***	0.78***	0.86***	0.83***
	(0.06)	(0.07)	(0.07)	(0.08)
McCain traits		1.30***		1.36***
		(0.17)		(0.19)
Obama traits		−1.02***		−1.16***
		(0.15)		(0.18)
Age			−0.01	−0.01
			(0.01)	(0.01)
Female			0.29	0.44
			(0.23)	(0.26)
African-American			−4.16***	−3.79***
			(1.10)	(1.23)
Hispanic			−0.55	−0.23
			(0.41)	(0.45)
Education			0.15*	0.22***
			(0.06)	(0.06)
Income			0.03	0.01
			(0.02)	(0.02)
Working Class			−0.54*	−0.70**
			(0.24)	(0.27)
South			0.36	−0.02
			(0.24)	(0.27)
Observations	781	781	781	781
Loglikelihood	−298.63	−243.14	−250.25	−206.88
AIC	601.27	494.28	520.50	437.77
BIC	610.59	512.92	567.11	493.69

Standard errors in parentheses. *prob < .05; **prob < .01; ***prob < .001.
Vote for Obama is the baseline outcome.

The principal component of the electoral distribution is given by the vector $(1.0, -1.8)$ with variance 1.02, while the minor component is given by the orthogonal eigenvector $(1.8, 1.0)$ with variance 0.61.

All models in Table 8 are given with Obama as the base, so the results give the estimations of the probability of voting for McCain. The table also shows the loglikelihood, Akaike information criterion (AIC), and Bayesian information criterion (BIC) for the various models. Model (1) in Table 8 shows the coefficients for the β-spatial conditional logit model in 2008 to be

$$\left(\lambda_{\text{Obama}}, \lambda_{\text{McCain}}, \beta\right) = (0, -0.84, 0.85). \tag{4.6}$$

These parameters are estimated when the candidates are located at the estimated positions. We assume that the parameters of the model remain close to these values as we modify the candidates positions in order to determine the equilibria of the model.

According to the model $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, the probability that a voter chooses McCain, when the McCain and Obama positions are at the electoral origin, $\mathbf{z}_0 = ((0,0),(0,0))$, is

$$\rho_{\text{McCain}} = \left[1 + \exp(0.84)\right]^{-1} = [1 + (2.31)]^{-1} = 0.3. \tag{4.7}$$

Then $\beta(1 - 2\rho_{\text{McCain}}) = 0.85 \times 0.4 = 0.34$.

The characteristic matrix (essentially the Hessian of McCain's vote function at $\mathbf{z}_0$) is

$$\begin{aligned} C_{\text{McCain}} &= \left[2\beta\left(1 - 2\rho_{\text{McCain}}\right)\nabla_0\right] - I = [2 \times 0.34 \times \nabla_0] - I \\ &= (0.68)\nabla_0 - I \\ &= (0.68)\begin{bmatrix} 0.8 & -0.127 \\ -0.127 & 0.83 \end{bmatrix} - I = \begin{bmatrix} 0.54 & -0.086 \\ -0.086 & 0.56 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.46 & -0.086 \\ -0.086 & -0.44 \end{bmatrix}. \end{aligned} \tag{4.8}$$

The "convergence coefficient" is

$$c = 2\beta\left(1 - 2\rho_{\text{McCain}}\right)\text{trace}\nabla_0 = 2 \times (0.34) \times (1.63) = 1.1. \tag{4.9}$$

The *sufficient* condition for convergence to $\mathbf{z}_0$ is that $c < 1$. Thus our estimate for c exceeds this critical value for convergence. However, the necessary condition is satisfied, and the determinant of C_{McCain} is positive, while the trace is negative. Thus both of the eigenvalues of C_{McCain} are negative, and the origin is a *maximum* of McCain's vote share function. The best response functions of the candidates are well behaved, so the LNE is a PNE.

We also considered a spatial mode where the two axes had different coefficients, estimated to be $\beta_1 = 0.8$, $\beta_2 = 0.92$. Again, the determinant was found to be positive and trace negative, so the origin is also a maximum of McCains vote share function for this model. Simulation of these models confirmed that the joint origin was an LNE.

We now turn to the models with traits and sociodemographics. Table 8 also gives the various spatial models with these additional valences.

Comparison of the loglikelihoods for the pure spatial model and the model with traits shows that the perception of character traits is important for the statistical significance of the model. (We use the Bayes' factors, or difference in loglikelihoods as a measure of statistical difference between two models [24].) For example, the spatial model with traits has a very large Bayes' factor of 114 over the pure traits model, while the spatial model with traits and sociodemographics has a Bayes' factor of 150 over the traits model.

Like the pure spatial model, the induced preference correspondences in the joint model with sociodemographic valences are all convex valued, indicating existence of a PNE. Simulation of the spatial model with sociodemographic valences showed that the PNE was one where both candidates adopt the origin. Although the sociodemographic valences add significance to the model, they do not affect the equilibrium positions. On the other hand, simulation of the full model with traits showed that the PNE was one where the candidates adopted the positions $z_{\text{Obama}} = (+0.10, -0.07)$ and $z_{\text{McCain}} = (+0.13, -0.12)$.

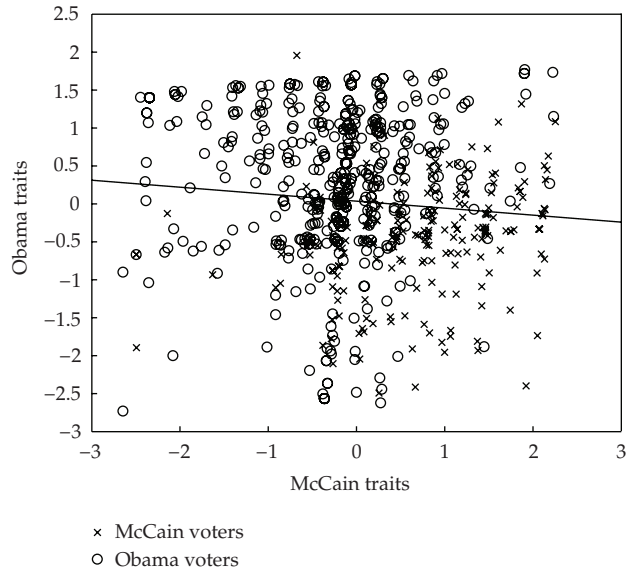

Figure 6: Correlation between Obama and McCain traits, showing vote choice.

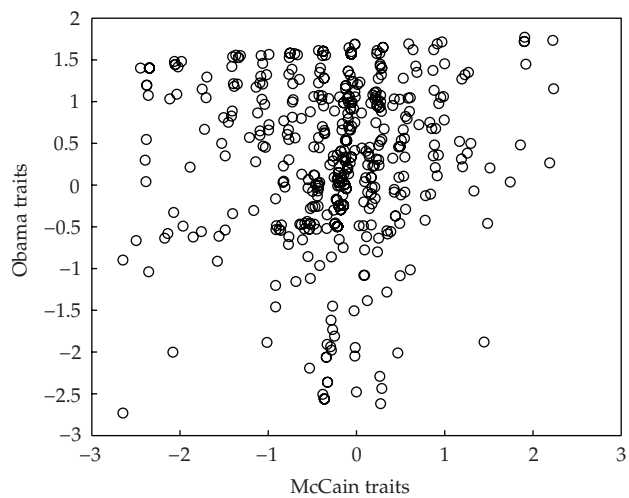

Figure 7: Scatterplot of Obama and McCain traits for Obama voters.

This equilibrium is only a slight perturbation from the joint origin. We can infer that though the traits add to the statistical significance of the stochastic model they do not significantly affect the equilibrium. Figures 6, 7, and 8 show the relationship of the perception of Obama and McCain traits. Figure 6 shows there is a slight negative correlation between these perceptions, while Figures 7 and 8 suggest that there are correlations between perceptions of candidate traits and vote choice. These weak correlations have only a slight effect on the strong convergence induced by the electoral pull.

We can therefore write $\mathbf{z}^{el} = (z^{el}_{\text{Obama}}, z^{el}_{\text{McCain}}) = ((+0.10, -0.07), (+0.13, -0.12))$, since the joint model with traits has no activist valence terms. The argument of Section 3 implies

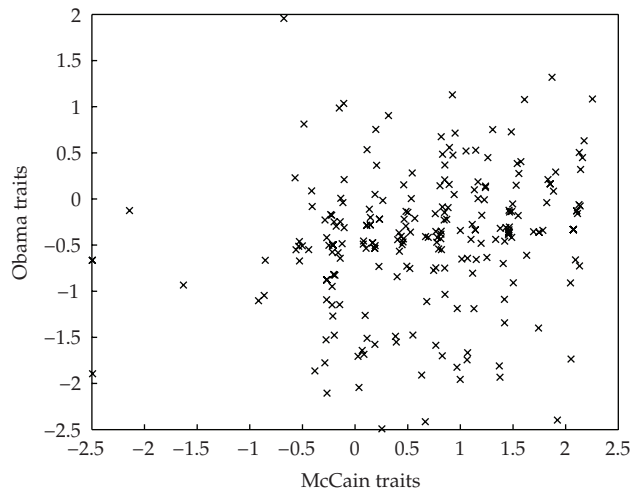

Figure 8: Scatterplot of Obama and McCain traits for McCain voters.

that $\mathbf{z}^{el}$ can be interpreted as the vector of "weighted electoral means" in a full model with activists. Assuming that the estimated candidate positions, $\mathbf{z}^*$, are in equilibrium with respect to the activist model, then by the balance condition, as given above, we obtain:

$$\begin{aligned}\mathbf{z}^* - \mathbf{z}^{el} &= \begin{bmatrix} & \text{McCain} & \text{Obama} \\ x & 0.59 & -0.22 \\ y & -0.37 & +0.75 \end{bmatrix} - \begin{bmatrix} & \text{McCain} & \text{Obama} \\ x & +0.13 & +0.10 \\ y & -0.12 & -0.07 \end{bmatrix} \\ &= \frac{1}{2\beta}\frac{d\boldsymbol{\mu}}{d\mathbf{z}}\Big|_{z^*} = \begin{bmatrix} & \text{McCain} & \text{Obama} \\ x & 0.46 & -0.32 \\ y & -0.25 & 0.82 \end{bmatrix}.\end{aligned} \tag{4.10}$$

Here

$$\frac{d\boldsymbol{\mu}}{d\mathbf{z}}\bigg|_{z^*} = \left(\frac{d\mu_{\text{mc}}}{dz_{\text{mc}}}\bigg|_{z^*_{\text{mc}}}, \frac{d\mu_{\text{ob}}}{dz_{\text{ob}}}\bigg|_{z^*_{\text{ob}}}\right) \tag{4.11}$$

is the pair of direction gradients, induced by activist preferences, acting on the two-candidates. The difference between $\mathbf{z}^*$ and $\mathbf{z}^{el}$ thus provides an estimate of the activist pull on the two-candidates. In this election, we estimate that activists pull the two-candidates into opposed quadrants of the policy space. The estimated distributions of activist positions for the two parties, in these two opposed quadrants (as given in Figure 4), are compatible with this inference. The means of these activist positions are:

$$\begin{bmatrix} & \text{Rep Act} & \text{Dem Act} \\ x & 1.41 & -0.2 \\ y & -0.82 & 1.14 \end{bmatrix}. \tag{4.12}$$

Miller and Schofield [16, 22] propose a model where activists have eccentric utility functions. If we assume that the Democrat activists tend to be more concerned with social policy and Republican activists with economic policy, then we have an explanation for the candidate shifts from the estimated equilibrium. Note in particular that the distribution of activist positions for the two parties looks very different from the voter positions. The latter is much more heavily concentrated near the electoral origin, while the former tends to be dispersed.

When the candidates are at their estimated positions, the estimated vote shares, according to the traits model, are $(V_{\text{Obama}}, V_{\text{McCain}}) = (0.68, 0.32)$. Since the actual vote shares are $(0.52, 0.48)$, it appears that the trait model may give a statistically plausible account for voter choice, but it does not provide, by itself, a good model of how candidates obtain votes. We suggest that the missing characteristic of this model of the election is due to the contributions of party activists.

Indeed, we suggest that the addition of activists to the model can account for the difference between convergent, equilibrium positions and divergent, estimated candidate positions, as obtained by Enelow and Hinich [25] and Poole and Rosenthal [26], respectively, in their various analyses of US elections.

The section on the formal model presented an extension where there are many activists for each candidate. This model suggests that the activist pulls on the two candidates will be particularly influenced by those activists who have more extreme policy preferences. This inference is corroborated by the above analysis, since it appears that the Democratic activists are more concerned with social policy, while the Republican activists are more concerned with economic policy.

Since the above equation is obtained from a first-order gradient condition, then as shown in Section 3, we could also interpret $(d\boldsymbol{\mu}/d\mathbf{z})(\mathbf{z})$ as the gradient obtained from a model where candidates have policy preferences derived from utility functions $(\mu_{\text{mc}}, \mu_{\text{ob}})$. Duggan and Fey [23] have explored such a model for the case of a deterministic vote model, and obtained symmetry conditions for equilibrium. However, in such a model of policy seeking candidates, a candidate must be willing to adopt a losing position because of strong preferences for particular policies. In the activist model presented here, candidates act as though they have policy preferences, but these are induced from activist preferences, and are compatible with vote maximizing strategies by the candidates.

5. The Election in Britain in 1979

Figure 9 shows the estimated positions of the three major parties in Britain in 1979, as obtained by Quinn et al. [27], with the electoral distribution obtained from the survey data from Eurobarometer [28] and the party positions obtained from the middle level Elites Study [29]. Tables 9(a) and 9(b) give the election results for five parties in Great Britain and five in Northern Ireland.

Using the pure spatial model as presented in Table 10 for just three parties in Great Britain, the coefficients are

$$\begin{gathered}(\lambda_{\text{LAB}}, \lambda_{\text{LIB}}, \lambda_{\text{CON}}, \beta)_{1979} = (-0.011, -1.574, 0.0, 0.272), \\ \rho_{\text{LIB}} = \frac{e^0}{e^0 + e^{1.563} + e^{1.574}} \simeq 0.094.\end{gathered} \tag{5.1}$$

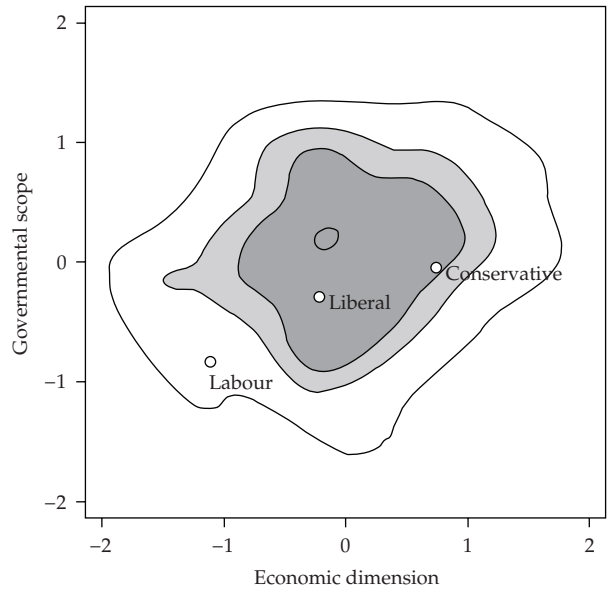

Figure 9: Smoothed electoral distribution in Britain and party positions in 1979.

Table 9

(a) Election in Great Britain (1979).

Party	Vote share %	Seats	Seat share %
Conservative Party	43.9	339	53.4
Labor Party	36.9	269	42.4
Liberal Party	13.8	1	1.7
Scottish National Party	1.6	2	0.6
Plaid Cymru	0.4	2	0.6
Total	96.6	623	98.7

(b) Election in Northern Ireland.

Party	Vote share %	Seats	Seat share %
Alliance	0.3	0	—
Ulster Unionists	0.8	5	0.7
United/Indep Unionists	0.1	2	0.3
Social Dem and Labor	0.4	1	0.15
Independent Republican	0.1	1	0.15
Total	1.7	9	1.3

Table 10: British pure spatial model for 1979, with respect to the conservative Party.

Party	Coefficient	Est	Lower 95% bound	Upper 95% bound
	SpatialCoeff β	0.272***	0.196	0.347
Labor	`valence`	−0.011	−0.243	0.221
Liberal Dems	`valence`	−1.574***	−1.887	−1.262
$n = 426$	Log-likelihood			
	= −377.8	***prob < .001		

Table 11: British Joint model in 1979 normalized with respect to the conservative party.

Party	Coefficient	Est	Lower 95% bound	Upper 95% bound
Labor	**Spatial Coeff** β	0.268***	0.190	0.346
	`income`	−0.102*	−0.203	−0.001
	`relig`	−0.233*	−0.443	−0.022
	`manlab`	+1.034***	0.514	1.554
	`stown`	0.105	−0.207	0.418
	`educ`	0.042	−0.057	0.142
	`valence`	0.574	−0.363	1.511
Liberal Dems	`income`	−0.111	−0.248	0.024
	`relig`	0.006	−0.286	0.299
	`manlab`	0.395	−0.379	1.170
	`stown`	−0.235	−0.666	0.195
	`educ`	0.121	−0.002	0.245
	`valence`	−0.883	−2.130	0.363
	Log-likelihood	***prob < .001		
	= −361.6	*prob < .0.05		

When all parties are located at the origin, the model suggests that the Liberals would gain just under 10% of the vote. In fact, in 1979 they gained 13.8%. The model suggests that the divergence of the two major parties from the origin allowed the Liberals to gain a further 4% of the vote. Since the electoral variance is 0.587 on the first (economic axis) and 0.444 on the second axis, with negligible covariance $(\sigma_1.\sigma_2) = -0.09$. we obtain

$$2\beta(1-2\rho_{\mathrm{LIB}}) = 2\times 0.272\times 0.81 = 0.44$$
$$C_{\mathrm{LIB}} = (0.44)\begin{bmatrix} 0.587 & -0.09 \\ -0.09 & 0.444 \end{bmatrix} - I = \begin{bmatrix} -0.73 & -0.04 \\ -0.04 & -0.80 \end{bmatrix}. \tag{5.2}$$

Both eigenvalues are clearly negative.

The "convergence coefficient" is

$$c = 2\beta(1-2\rho_{\mathrm{LIB}})\mathrm{trace}\nabla_0 = (0.44)\times(1.03) = 0.45. \tag{5.3}$$

The pure spatial model of the 1979 election in Britain implies that the electoral joint origin is a vote share maximizing equilibrium. We next consider a joint multinomial conditional logit model, $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\beta)$, with the sociodemographic variables used by Quinn et al. [27]. (These variables are denoted income, religion (relig),manual labor (manlab), size of town (stown) and education (educ), respectively, in Table 11.)

As Table 11 makes clear, only the group specific valence

$$(\theta_{\mathrm{Labor}}\cdot\eta_{\mathrm{manlab}}) = 1.034 \tag{5.4}$$

was statistically significant at the 1% level. The β-spatial coefficient was also significant at the 1% level. The loglikelihoods of the joint and pure spatial models were very similar, and the Bayes' factor of the joint models over the pure spatial model was +16.

According to the joint model, the weighted electoral mean for the Labor party should give greater weight to these voters who are manual laborers. Since these voters will tend to have preferred positions on the left of the economic dimension, we may infer that the Labor party activists will be positioned on the left of the economic dimension. However, the simulated LNE in the joint model was found to be the joint origin.

Thus the impacts of the sociodemographic valences on the simulated equilibrium are insignificant. Although these valences are useful in modeling the voting behavior of the electorate, they appear to have little significance on the policy positioning of the parties. If we assume that the party positions in Figure 9 are the LNE in the full activist model the, we obtain

$$\mathbf{z}^* = \begin{bmatrix} \text{Party} & \text{Labor} & \text{Liberal} & \text{Cons} \\ x\text{-axis} & -1.0 & -0.2 & 0.7 \\ y\text{-axis} & -0.9 & -0.4 & 0.0 \end{bmatrix}. \tag{5.5}$$

As in the analysis of the United States, we find that the overall effect of the activist groups on the two major parties is to pull these parties apart. This leaves the Liberal Democrats in the center. With low valence, they only gain about 14% of the vote.

6. Concluding Remarks

Valence, whether intrinsic or based on electoral perceptions of character traits, is intended to model that component of voting which is determined by the judgments of the citizens. In this respect, the formal stochastic valence model provides a framework for interpreting Madison's argument in *Federalist X* over the nature of the choice of Chief Magistrate in the Republic. Schofield [30] has suggested that Madison's argument may well have been influenced by Condorcet's work on the so-called "Jury Theorem" [31]. However, Madison's conclusion about the "probability of a fit choice" depended on assumption that electoral judgment would determine the political choice. The analysis presented here does indeed suggest that voters' judgments, as well as their policy preferences, strongly influence their political choice.

Condorcet's work has recently received renewed attention (McLennan [32]). This paper can be seen as a contribution to the development of a Madisonian conception of elections in representative democracies as methods of aggregation of both preferences and judgments. One inference from the work presented here does seem to belie Riker's arguments [33, 34] that there is no formal basis for populist democracy. Since voters' perceptions about candidate traits strongly influence their political decisions, the fundamental theoretical question is the manner by which these perceptions are formed. We argue that the low convergence coefficients in the majoritarian polities of the United States and Great Britain imply that the electorate is not polarized. Since candidates or party leaders do not adopt convergent positions, we can infer that democratic equilibria in these polities reflect the preferences of interest groups rather than the electorate at large.

On the other hand, empirical work on Israel [9] and Turkey [10] shows that the convergence coefficients in recent elections in these two polities are very large. (The estimates are 3.98 for Israel in 1996 and 5.4 for Turkey in 2002.) These estimates indicate that the polities in these countries are polarized. We can infer that parties in these polities diverge away from the electoral center, even in the absence of activism.

```
function [x,fval]=find_best_response
                  (n,m,k,z0,VV,beta,lambda,ff)
% REQUIRES MATLAB OPTIMIZATION TOOLBOX
% n: Number of voters.
% m: Number of parties
% z0: Initial positions
% k: Best response of which party?
% VV(2xn): Voter positions
% beta(1x1): Distance coefficient
% ff((m)x2): Where to fix the parties?
% lambda(nxm): Latent utility of the voters.

syms z1 z2;z=[z1;z2]; % Decision variables
svote=0;
di2=zeros(n,m);
for i=1:n
        den=0;
              di1=(VV(1,i)-z(1))^2+(VV(2,i)-z(2))^2;
                  % Decision variables
        for j=1:m
        di2(i,j)=(VV(1,i)-ff(j,1))^2+(VV(2,i)-ff(j,2))^2;
               if ~(j==k)
               den=den+exp(lambda(i,j)-beta*di2(i,j));
               end
        end
        V1=exp(lambda(i,k)-beta*di1)/
           (den+exp(lambda(i,k)-beta*di1)) ;
        svote=svote+V1;
end
V1=vpa(-svote/n,2);
fh=matlabFunction(V1,'vars',{z});
[x,fval] = fminsearch(fh, z0 ,optimset('TolX',1e-4));
           % Returns the best response
```

Algorithm 1

Appendices

A. Matlab Optimization Algorithm for Best Response

See Algorithm 1.

B. Matlab Optimization Algorithm for Local Nash Equilibrium

See Algorithm 2.

C. Question Wording for the 2008 American National Election Study Survey Items

(1) Do you think the government should provide more services than it does now, fewer services than it does now, or about the same number of services as it does now?

(2) Do you favor, oppose, or neither favor nor oppose the US government paying for all necessary medical care for all Americans?

```
function f=find_local_nash(n,m,VV,beta,lambda,ff,tol)
% n: Number of voters.
% m: Number of parties
% tol: Tolerance. The algorithm stops if the change is less than tol.
% VV(2xn): Voter positions
% beta(1x1): Distance coefficient
% ff((m)x2): Where to fix the rest of the parties?
% lambda(nxm): Latent utility of the voters.
               (Anything other than distance)

max_differ=10;
differ=zeros(m);

while abs(max_differ)>tol
              max_differ=0;
              for k=1:m
              z0=ff(k,:)';
              [x,fval]=find_best_response(n,m,k,z0,VV,beta,lambda,ff);
               differ(k)=max(abs(ff(k,:)-x'));
               if (differ(k)>tol)
                     max_differ=differ(k);
               end
               ff(k,:)=x';
               end

end
f=ff;\bigskip
```

Algorithm 2

(3) A proposal has been made that would allow people to put a portion of their Social Security payroll taxes into personal retirement accounts that would be invested in stocks and bonds. Do you favor this idea, oppose it, or neither favor nor oppose it?

I am going to ask you three questions, and ask you to choose which of two statements in these questions comes closer to your own opinion.

(4) One, the main reason that government has become bigger over the years is because it has gotten involved in things that people should do for themselves. Two, government has become bigger because the problems we face have become bigger.

(5) One, we need a strong government to handle today's complex economic problems. Two, the free market can handle these problems without government being involved.

(6) One, the less government, the better. Two, there are more things that government should be doing.

(7) This country would be better if we worried less about how equal people are. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?

(8) Do you think that big companies should pay a larger percent of their profits in taxes than small businesses do, that big companies should pay a smaller percent of their profits in taxes than small businesses do, or that big companies and small businesses should pay the same percent of their profits in taxes?

(9) Should federal spending on welfare programs be increased, decreased, or kept about the same?

(10) Do you favor, oppose, or neither favor nor oppose the US government making it possible for illegal immigrants to become US citizens?

(11) Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased a lot, increased a little, left the same as it is now, decreased a little, or decreased a lot?

(12) I would like to describe a series of circumstances in which a woman might want to have an abortion. For each one, please tell me whether you favor, oppose, or neither favor nor oppose it being legal for the woman to have an abortion in that circumstance. Staying pregnant would hurt the woman's health but is very unlikely to cause her to die.

(13) Staying pregnant could cause the woman to die.

(14) The pregnancy was caused by sex the woman chose to have with a blood relative.

(15) The pregnancy was caused by the woman being raped.

(16) The fetus will be born with a serious birth defect.

(17) Having the child would be extremely difficult for the woman financially.

(18) The child will not be the sex the woman wants it to be.

(19) Do you favor or oppose laws to protect homosexuals against job discrimination?

(20) Do you think homosexuals should be allowed to serve in the United States Armed Forces or don't you think so?

(21) Do you think gay or lesbian couples, in other words, homosexual couples, should be legally permitted to adopt children?

(22) Should same-sex couples be allowed to marry, or do you think they should not be allowed to marry?

(23) This country would have many fewer problems if there were more emphasis on traditional family ties. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?

(24) Do you think the federal government should make it more difficult for people to buy a gun than it is now, make it easier for people, or keep the rules the same?

(25) Some people feel that the government in Washington should make every effort to improve the social and economic position of blacks. Others feel that the government should not make any special effort to help blacks because they should help themselves. Where would you place yourself on this scale, or have not you thought much about this?

(26) Irish, Italians, Jewish, and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?

(27) Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?

(28) It is really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well off as whites. Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with this statement?

(29) We hear a lot of talk these days about liberals and conservatives. Where would you place yourself on a scale from liberal to conservative?

Acknowledgments

This paper is based on work supported by NSF Grant no. 0715929 and by a small grant from the Weidenbaum Center at Washington University. The first version was completed while Schofield was the Glenn Campbell and Rita Ricardo-Campbell National Fellow at the Hoover Institution, Stanford, 2009. Thanks are due to the very helpful comments of the anonymous referee.

References

[1] H. Hotelling, "Stability in Competition," *Economic Journal*, vol. 39, pp. 41–57, 1929.
[2] A. Downs, *An Economic Theory of Democracy*, Harper and Row, New York, NY, USA, 1957.
[3] J. S. Banks, J. Duggan, and M. Le Breton, "Bounds for mixed strategy equilibria and the spatial model of elections," *Journal of Economic Theory*, vol. 103, no. 1, pp. 88–105, 2002.
[4] N. Schofield and G. Cataife, "A model of political competition with activists applied to the elections of 1989 and 1995 in Argentina," *Mathematical Social Sciences*, vol. 53, no. 3, pp. 213–231, 2007.
[5] N. Schofield and I. Sened, "Multiparty competition in Israel, 1988–1996," *British Journal of Political Science*, vol. 35, no. 4, pp. 635–663, 2005.
[6] N. Schofield and I. Sened, *Multiparty Democracy: Elections and Legislative Politics*, Cambridge University Press, Cambridge, UK, 2006.
[7] N. Schofield and A. Zakharov, "A stochastic model of the 2007 Russian Duma election," *Public Choice*, vol. 142, pp. 177–194, 2009.
[8] N. Schofield, G. Miller, and A. Martin, "Critical elections and political realignments in the USA: 1860–2000," *Political Studies*, vol. 51, no. 2, pp. 217–240, 2003.
[9] N. Schofield, C. Claassen, U. Ozdemir, and A. V. Zakharov, "Estimating the effects of activists in two-party and muti-party systems: a comparison of the United States and Israel," *Social Choice and Welfare*. In press.
[10] N. Schofield, M. Gallego, U. Ozdemir, and A. V. Zakharov, "Competition for popular support: a valence model of elections in Turkey," *Social Choice and Welfare*. In press.
[11] D. Stokes, "Spatial models and party competition," *American Political Science Review*, vol. 57, pp. 368–377, 1963.
[12] D. Stokes, "Valence politics," in *Electoral Politics*, D. Kavanagh, Ed., Clarendon Press, Oxford, UK, 1992.
[13] E. Aragones and T. R. Palfrey, "Mixed equilibrium in a Downsian model with a favored candidate," *Journal of Economic Theory*, vol. 103, no. 1, pp. 131–161, 2002.
[14] E. Penn, "A model of far-sighted voting," *American Journal of Political Science*, vol. 53, pp. 36–54, 2009.
[15] J. H. Aldrich, "A Downsian spatial model with party activists," *American Political Science Review*, vol. 77, pp. 974–990, 1983.
[16] G. Miller and N. Schofield, "The transformation of the republican and democratic party coalitions in the U.S.," *Perspectives on Politics*, vol. 6, no. 3, pp. 433–450, 2008.
[17] N. Schofield, "Equilibria in the spatial stochastic model of voting with party activists," *Review of Economic Design*, vol. 10, no. 3, pp. 183–203, 2006.
[18] N. Schofield, "The mean voter theorem: necessary and sufficient conditions for convergent equilibrium," *Review of Economic Studies*, vol. 74, no. 3, pp. 965–980, 2007.
[19] R. D. McKelvey and J. W. Patty, "A theory of voting in large elections," *Games and Economic Behavior*, vol. 57, no. 1, pp. 155–180, 2006.
[20] J. Y. Halpern, "Beyond Nash equilibrium: a computer scientist looks at game theory," *Games and Economic Behavior*, vol. 45, pp. 114–131, 2003.
[21] K. E. Train, *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, UK, 2003.
[22] G. Miller and N. Schofield, "Activists and Partisan Realignment in the U.S.," *American Political Science Review*, vol. 97, pp. 245–260, 2003.
[23] J. Duggan and M. Fey, "Electoral competition with policy-motivated candidates," *Games and Economic Behavior*, vol. 51, no. 2, pp. 490–522, 2005.
[24] R. Kass and A. Raftery, "Bayes factors," *Journal of the American Statistical Assocociation*, vol. 90, pp. 773–795, 1995.

[25] J. M. Enelow and M. J. Hinich, "The location of american presidential candidates," *Mathematical and Computer Modelling*, vol. 12, no. 4-5, pp. 417–435, 1989.
[26] K. Poole and H. Rosenthal, "U.S. presidential elections 1968–1980: a spatial analysis," *American Journal of Political Science*, vol. 28, pp. 283–312, 1984.
[27] K. M. Quinn, A. D. Martin, and A. B. Whitford, "Voter choice in multi-party democracies: a test of competing theories and models," *American Journal of Political Science*, vol. 43, no. 4, pp. 1231–1247, 1999.
[28] J. Rabier and R. Inglehart, *Eurobarometer II April 1979. The Year of the Child in Europe*, Inter-University Consortium for Political and Social Research, Ann Arbor, Mich, USA, 1981.
[29] ISEIUM, "European Elections Study: European Political Parties' Middle Level Elites," Europa Institut, Mannheim, Germany, 1983.
[30] N. Schofield, "Evolution of the constitution," *British Journal of Political Science*, vol. 32, no. 1, pp. 1–20, 2002.
[31] N. Condorcet, *Essai sur l'application de l'analyse a la probabilite des decisions rendus a la pluralite des voix*, Imprimerie Royale, 1785.
[32] A. McLennan, "Consequences of the condorcet jury theorem for beneficial information aggregation by rational agents," *American Political Science Review*, vol. 92, no. 2, pp. 413–418, 1998.
[33] W. H. Riker, "Implications from the disequilibrium of majority rule for the study of institutions," *American Political Science Review*, vol. 74, pp. 432–446, 1980.
[34] W. H. Riker, *Liberalism against Populism*, W.H. Freeman, New York, NY, USA, 1982.