Empirical and Formal Models of the United States Presidential Elections in 2000 and 2004

Norman Schofield, Christopher Claassen, Maria Gallego, and Ugur Ozdemir

1 Introduction

The formal literature on two party electoral competition has typically been based on the assumption that parties or candidates adopt positions in order to win, and has inferred that parties will converge to the electoral median, under deterministic voting in one dimension (Downs 1957) or to the electoral mean in stochastic models.[1] An early empirical paper by Poole and Rosenthal (1984) provided some evidence that candidates in US presidential elections did not converge.

In the standard spatial model, only candidate *positions* matter to voters. However, as Stokes (1963, 1992) has emphasized, the non-policy evaluations, or *valences*, of candidates by the electorate are equally important. In empirical models, a party's valence is usually assumed to be independent of the party's position, and adds to the statistical significance of the model. In general, valence reflects the overall degree to which the party is perceived to have shown itself able to govern effectively in the past, or is likely to be able to govern well in the future (Penn 2009).

An extensive literature has developed over the last decade that considers deterministic or probabilistic voting models including valence or bias towards one or other of the candidates.[2]

[1]See the earlier work by Enelow and Hinich (1982, 1989), Erikson and Romero (1990) and more recent work by Duggan (2006), McKelvey and Patty (2006) and Patty et al. (2009).

[2]Ansolabehere and Snyder (2000), Groseclose (2001), Aragones and Palfrey (2002, 2005), Adams and Merrill (2002, 2005), Banks and Duggan (2005), Ashworth and Bueno de Mesquita (2009), Jessee (2009, 2010), Zakharov (2009), Serra (2010).

N. Schofield (✉), C. Claassen, and U. Ozdemir
Washington University in Saint Louis, St. Louis, MO, USA
e-mail: schofield.norman@gmail.com; classen@wustl.edu; uozdemir@artsci.wustl.edu

M. Gallego
Washington University in Saint Louis, St. Louis, MO, USA
and
Department of Economics, Wilfrid Laurier University, Waterloo, ON, Canada
e-mail: mgallego@wlu.ca

N. Schofield and G. Caballero (eds.), *Political Economy of Institutions, Democracy and Voting*, DOI 10.1007/978-3-642-19519-8_10,

This chapter offers a general model of elections based on the assumption that valence can be measured in a number of ways. The first kind is a fixed or *exogenous valence*, which for a party j is denoted λ_j. As in empirical work, we assume that λ_j is held constant at the time of an election, and so is independent of the party's position. Exogenous valence can be estimated as the intercept term in a stochastic model. Earlier work (Schofield and Sened 2006) has shown that, in models involving exogenous valence, if the valence differences are sufficiently large, then vote maximizing parties will not converge.

Here we construct stochastic models of the 2000 and 2004 US presidential elections involving exogenous valence and show that the valence of the two candidates were similar enough so that the unique Nash equilibrium was one where both candidates converge to the electoral origin in order to maximize vote share.[3]

Recent empirical work by Clarke et al. (2009a, b) has analyzed recent US presidential elections and British general elections. These works have shown that valence, as measured by the perceptions of the character traits of the candidates, or of party leaders, is a key element of these elections.[4]

In the empirical analysis we show that a voter's perception of each candidate's traits has a very significant impact on the probability that the voter chooses one candidate or the other. The simulation of the spatial model, based on both position and character traits, allows us to estimate what we call *Local Nash equilibria (LNE)*[5] to the vote maximizing game.[6] The LNE of this traits model is slightly perturbed from the electoral origin, so that the two candidates are located at the same position, slightly to the right on the economic axis, and at a neutral position on the social axis. These equilibrium positions differ from the estimated positions of the two candidates.

In order to account for the discrepancy between the estimated positions and the positions obtained by equilibrium analysis, we introduce a different kind of valence known as *activist valence*. When candidate j adopts a policy position z_j, in the policy space, X, then the activist valence of the party is denoted $\mu(z_j)$. Implicitly we adopt a model originally due to Aldrich (1983). In this model, activists provide crucial resources of time and money to their chosen candidate, and these resources are dependent on the candidate position.[7] The candidate can then use these

[3]The empirical analyses were based on the 2000 and 2004 American National Election Surveys (ANES).

[4]See also Clarke et al. (2005). Jesee (2009, 2010) has also examined partisan bias in the 2004 and 2008 Presidential elections.

[5]We focus on *local equilibria* because we consider that candidates will only be able to make small adjustments to their policy statements as the election nears.

[6]We focus on *vote maximizing* rather than maximizing the *probability of winning* because the former model is linear and would seem to more closely characterize the likely behavior of candidates adapting to electoral information obtained from polls and the like. As Patty (2002, 2007) has shown, these two classes of models differ in the equilibria.

[7]For convenience, it is assumed that $\mu(z_j)$ is only dependent on z_j, and not on z_k, $k \neq j$, but this is not a cucial assumption.

resources to enhance the candidate's image before the electorate, thus affecting the candidate's overall valence.

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 functions are sufficiently "concave" with respect to candidate positions,[8] then the activist vote maximizing model will exhibit a Nash equilibrium.

The difference we find between the estimated positions of the two candidates and those inferred to be equilibria from the full trait model gives us an estimate for the influence of activists.

Empirical analysis of the 2000 and 2004 US elections suggests that party activists tend to have more extreme policy positions than the typical voter. The problem for each candidate is that by choosing a position to maximize activist support, the candidate loses centrist voters and by choosing to be closer to centrist voters the candidate can loose activist support. The candidate must determine the trade-off between attracting resources from activists and appealing to the voters. This trade-off is captured by the "optimal marginal condition" that maximizes vote share. This is given as a (first order) *balance condition*.

Grossman and Helpman (1996), in their game theoretic model of activists, consider two distinct motives for interest groups:

> Contributors with an *electoral motive* intend to promote the electoral prospects of preferred candidates, [while] those with an *influence motive* aim to influence the politicians' policy pronouncements.

In our first activist model the term $\mu_j(z_j)$ influences every voter and thus contributes to the electoral motive for candidate j. In addition, the candidate must choose a position to balance electoral and activist support.

We argue that the influence of activists on the two candidates can be characterized in terms of activist gradients. For the two candidates, these gradients point into opposite quadrants of the policy space. We also obtained information from the American National Election Surveys on activists, namely those who contributed resources to one or other of the two parties. The mean positions of the two sets of party activists were shown to be compatible with our estimated party activist gradients.

Because each candidate is supported by multiple activists, we extend the activist model by considering a family of potential activists, $\{A_j\}$ for each candidate, j, where each $k \in A_j$ is endowed with a utility function, U_k, which depends on candidate j's position z_j, and the preferred position of the activist. The resources allocated to j by k are denoted $R_{jk}(U_k(z_j))$. Let $\mu_{jk}(R_{jk}(U_k(z_j)))$ denote the effect that activist k has on voters' utility. Note that the activist valence function for j is the same for all voters. With multiple activists, the *total activist valence function* for agent j is the linear combination

[8] We mean by this that the appropriate Hessians have negative eigenvalues of sufficient magnitude.

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

Bargains between the activists supporting candidate j then gives a *contract set* of activist support for candidate j, and this contract set can be used formally to determine the balance locus, or set of optimal positions for each candidate. This balance locus can then be used to analyze the pre-election contracts between each candidate and the family of activist support groups.

Consider now the situation where these contracts have been agreed, and each candidate is committed to a set of feasible contracts as outlined in Grossman and Helpman (1994, 1996, 2001). Suppose further that the activists have provided their resources. Then at the time of the election the effect of this support is incorporated into the empirical estimates of the various exogenous, sociodemographic and trait valences. Consequently, when we estimate these valences we also estimate the aggregate activist influence. The estimated positions of the candidates can then be regarded as incorporating policy preferences of the activists. Electoral models where candidates have policy positions, as proposed by Wittman (1977), Calvert (1985), Duggan and Fey (2005), and Duggan (2006) implicitly assume that candidates would be willing to accept defeat because of an adherence to particular policy positions.

We argue that it is more plausible that the estimated positions of the candidates are the result of maximizing candidate utility functions that balance the electoral consequences of position-taking with the necessity of obtaining activist resources to contest the election. This calculation requires an estimate of the degree to which these resources will influence the perceptions that the electorate has of the various valences associated with the model.

In the final version of the model we allow the activist valence function to be individual specific. The total resources available to candidate j are now denoted $\mathbf{R}_j(z_j)$, and these may be allocated to individuals, with resource m_{ij} targeted on voter, or "voter class", i by candidate j. Since m_{ij} will depend on z_j, we write this allocation as $m_{ij}(z_j)$, so the budget constraint is

$$\begin{aligned}\mathbf{R}_j(z_j) &= \sum_{k \in A_j} R_{jk}(U_k(z_j)) \\ &= \sum_{i \in N} m_{ij}(z_j).\end{aligned}$$

The optimization problem is now a more complex one, subject to this constraint. In actual fact candidates will generally not allocate resources to individuals *per se*, but to voter classes via media outlets in different regions, or "zip codes". Indeed, much of the action in political campaigns is concerned with the analysis of local data so as to determine how voters might be targeted in an optimal fashion. For example, the logit models in this chapter give sociodemographic analyses that would, in principle allow for the targeting of specific groups in the polity.

The general balance condition presented in the Technical Appendix specifies how these resources should be allocated throughout the polity.

A recent literature on elections has focussed on the effects of campaign expenditure on US election results (Coate 2004).[9] Herrera et al. (2008) suggest that electoral volatility forces candidates to spend more, while Ashworth and Bueno de Mesquita (2009) suppose that candidates buy valence so as to increase their election chances. Meirowitz (2008) notes that

> candidates and parties spending this money thought that it would influence the election outcome. Downsian models of competition cannot explain how candidates choose spending campaign levels or what factors influence these decisions.

Meirowitz proxies the choice of expenditure in terms of candidate choice of effort, but his model does not explicitly deal with the budget question.

Ansolabehere et al. (2003) provide an empirical analysis of Congressional and Presidential election campaign contributions up to 2000. They note that candidates, parties and organizations raised and spent about $3 billion in the 1999–2000 election cycle. However, the federal government at that time spent about $2 trillion, so the prize from influencing politics was of considerable value. They suggest that so little is spent in contributions relative to the possible gains because contributions are a consumption good, rather than an investment good. However, they do observe that the electoral motive is not insignificant: they suggest that the marginal impact of $100,000 spent in a House race is about 1% in vote share.

The essence of the model presented here is that it attempts to endogenize the question of the resource budget of candidates since the total resources used by candidates in seeking election victory come from the *implicit contracts* they can make with their supporting activists.[10] The activists must solve their own optimization problem by estimating the benefit they receive from the electoral and influence motives, in deciding what resources to make available to their chosen candidate.

Essentially there is an arms race between candidates over these resources due to a feedback mechanism between politics and economics. As the outcome of the election becomes more important, activists become increasingly aware that the resources they provide have become crucial to election victories, and they become more demanding of their chosen candidates. Because of the offer of resources, candidates are forced to move to more radical positions, and polarization in candidate opositions increases, even though there may be little change in the degree of polarization of the electorate.

In the conclusion we suggest that the results presented here lend some support to the activist model proposed by Miller and Schofield (2003, 2008) and elaborated in Schofield and Miller (2007). Changes in voter choice appear to result not only from changes in the distribution of electoral preferences, but from the shifts in electoral perceptions. In turn, these changes are the consequence of the shifting pattern of

---

[9] An earlier paper by Groseclose and Snyder (1996) looked at vote buying, but in the legislature.

[10] Snyder and Ting (2008) also consider the contracts between activists and candidates but assume that the policy space is one dimensional.

activist support for the candidates. Since the importance of electoral contributions has increased, this has enhanced the influence of activist groups.[11] The empirical and formal models presented here give a reason why electoral politics has become very polarized in the United States.[12] This polarization appears to have benefited the wealthy in society and may well account for the increase in inequality in income and wealth distribution that has occurred over the last decade. (Hacker and Pierson 2006, 2010; Pierson and Skocpol 2007).

As Miller and Schofield (2008) have argued, over the long run the coalition structure among activist groups for the parties will shift in response to exogenous shocks, leading to a shift in the activist coalitions. This may be the cause of the slow realignment that appears to have occurred over many decades in U.S. politics.[13]

In the next section we present the empirical methodology that was used, together with the computation method to find equilibria. Section 3 draws some conclusions from the analysis, including a number of inferences from the model relating to Madison's argument about the "probability of a fit choice" in the Republic. Section 4 is a Technical Appendix that gives the details of the spatial model that we deploy.

2 Methodology: Spatial Models of the 2000 and 2004 Elections

2.1 The 2000 Election

To construct a model of the 2000 election, we used survey data from the 2000 American National Election Study (ANES 2000). The survey is a nationally representative sample of the voting age population, with 1555 pre- and post-election respondents in 2000. The first step was to build up a map of the policy space and to assign each surveyed individual a two-dimensional ideal point on that space. Following Schofield et al. (2003), we constructed a two dimensional policy space based on economic and social issues. Exploratory factor analysis led to the ten survey items, reported in the Data Appendix from which two factors were extracted. The factor loadings per item are given in Table 1. Figure 1 gives a smoothing of the estimated voter distribution. Essentially, left on the economic (x) axis in this figure is pro-redistribution. The second social axis (y) is determined by attitudes to abortion and gays, so we interpret north on this axis to be in support of certain civil rights. Figure 2 gives a perspective plot of the electoral distribution.

Respondent's partisan choice was measured with the following question:

"Who do you think you will vote for in the election for President?"

[11]Indeed, Herrera et al. (2008) observe that spending by parties in federal campaigns went from 58 million dollars in 1976 to over 1 billion in 2004 in nominal terms.

[12]See the works by Fiorina et al. (2005), Fiorina and Abrams (2009) and McCarty et al. (2006) on polarization in the electorate and Layman et al. (2010) on polarization among activists. Schofield et al. (2011) gives similar results for the 2008 election.

[13]See also the earlier work by Sundquist (1973).

Table 1 Factor loadings from the American national election survey, 2000

Question	Social policy	Economic policy
1. Economic problems	0.02	0.32
2. Federal spending	0.09	0.36
3. Equality	0.21	0.50
4. African American	0.15	0.46
5. Immigrants	0.08	0.31
6. Liberal vs conservative	0.38	0.33
7. Guns	0.17	0.34
8. Abortion	0.51	0.02
9. Gays	0.65	0.18
10. Family	0.44	0.24
% Variance explained	11.2	11.0

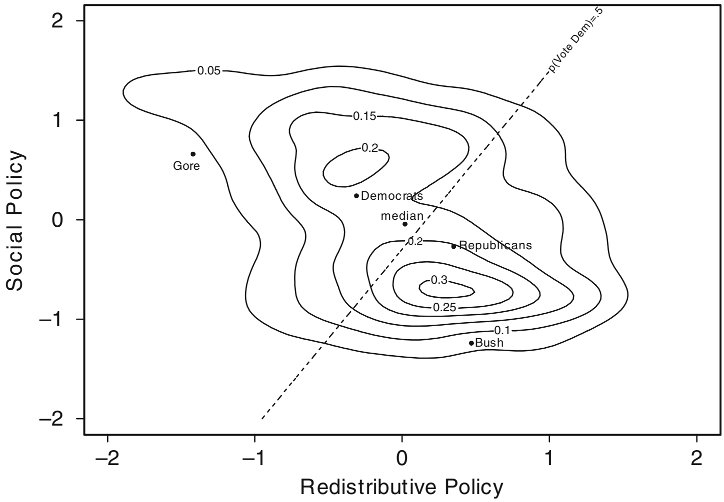


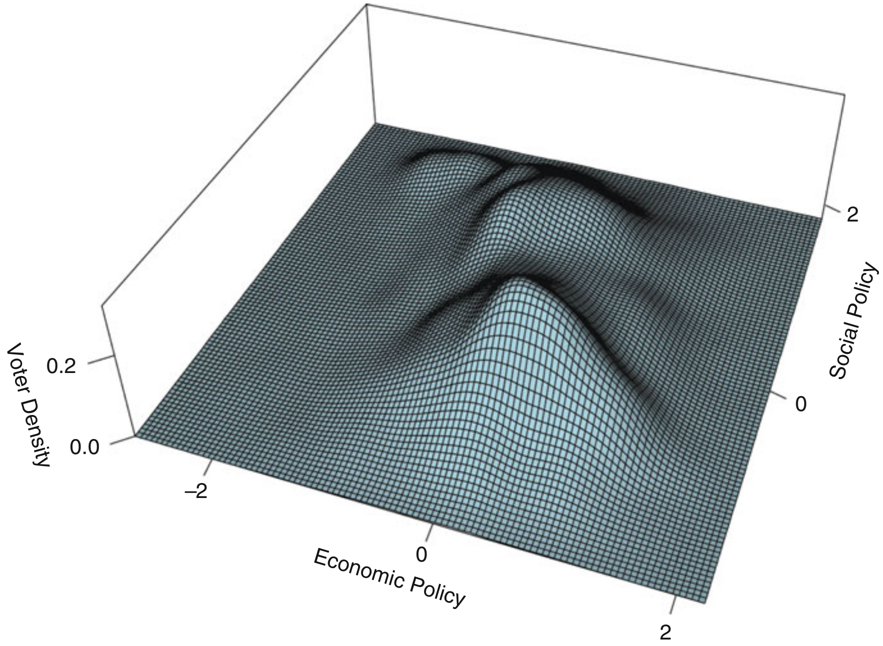
Fig. 1 Contour plot of the voter distribution in 2000 with the equiprobable cleavge line

Activism was measured with the following question:

"During an election year people are often asked to make a contribution to support campaigns. Did you give money to an individual candidate running for public office?" OR

"Did you give money to a political party during this election year?"

An activist is thus operationally defined here as someone who claimed to donate money to either a candidate or a party. Activists were then coded as Republican or Democrat depending on the party or partisan affiliation of the candidate to which they gave money. Of the sample, 4.5% ($n = 70$) were Republican activists, and 2.9% ($n = 47$), Democrat activists, in 2000. As Table 2 shows, the mean Democratic partisan position was $\left(x_{dem}^{part}, y_{dem}^{part}\right) = (-0.31, 0{,}24)$ with standard error (0.029, 0.03) while the activist mean was $\left(x_{dem}^{act}, y_{dem}^{act}\right) = (-0.54, 0.48)$ with standard error (0.10, 0.10). For the Republican partisans we find $\left(x_{rep}^{part}, y_{rep}^{part}\right) = (0.36, -0.27)$ with standard error (0.027, 0.03) while the activist mean was



This figure will be printed in b/w

Fig. 2 Perspective plot of the voter distribution in 2000

Table 2 Descriptive data 2000

	Econ mean	Policy std. err.	C.I	Social mean	Policy std. err.	C.I.	n
Activists							
Democrats	−0.54	0.10	[−0.74,−0.34]	0.48	0.10	[+0.28,+0.68]	47
Republicans	0.42	0.07	[+0.28,+0.56]	−0.30	0.08	[−0.46,−0.14]	70
Non-activists							
Democrats	−0.31	0.029	[−0.34,−0.28]	0.24	0.03	[+0.18,+0.30]	634
Republicans	0.36	0.027	[+0.3,+0.42]	−0.27	0.03	[−0.33.−0.21]	536

$\left(x_{rep}^{act}, y_{rep}^{act}\right) = (0.42, -0.30)$, with standard error (0.07, 0.08). The 95% confidence intervals on these estimates give some weak evidence that the activist mean positions are more extreme than the partisan mean positions. Figure 3 shows the mean activist and voter positions, and standard error bars with the voter bars the smaller.

The positions of the major presidential candidates, Bush and Gore, in 2000 were estimated in a similar fashion to that of the sampled individuals. These estimated responses of the candidates are given in Table 3. Scores were assigned to each candidate for each of the constituent survey items based on press reports of their

Fig. 3 Comparison of mean partisan and activist positions in 2000

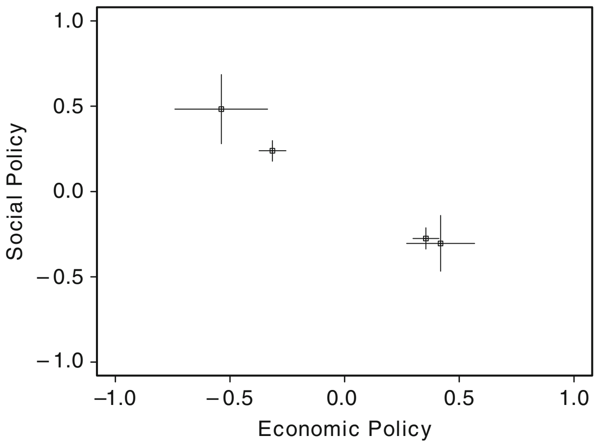


Table 3 Gore and Bush estimated responses, 2000

Question	Bush	Gore
1. Economic problems	3	1
2. Federal spending	3	2
3. Equality	3	1
4. African American	4	2
5. Immigrants	3	3
6. Liberal vs conservative	3	1
7. Guns	2	1
8. Abortion	3	2
9. Gays	3	1
10. Family	5	3
Estimated position: economic policy	+0.47	−1.42
Estimated position: social policy	−1.24	+0.66

campaign stances on various issues. The factor analysis was used to obtain estimated policy positions $z_{gore} = (x_{Gore}, y_{Gore}) = (-1.42, 0.66)$, and $z_{Bush} = (x_{Bush}, y_{Bush}) = (0.47, -1.24)$. Notice that the candidate positions are more extreme on both axes than the partisan and activist positions. In particular Gore's position x_{Gore} is well outside the confidence intervals of both partisans and activists on the economic axis, while Bush's position y_{Bush} lies outside the confidence intervals on the social axis. This suggests that both candidates were influenced by more radical activists.

The contour plot of Figure 1 includes an estimated cleavage line dividing likely Democrat candidate voters from Republican candidate voters. This partisan cleavage line was derived from a standard binomial logit model, designed to test the effects of each policy dimension on vote choice. We do not report the full results of the positional model here.

Because this logit model involves the preferred positions of voters, we refer to it as a *pure positional model*. Note however that the position of each candidates is implicit, rather than explicit in this model. Our estimates of the log-likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were all quite acceptable, and all coefficients were significant with probability < 0.01. A voter i, with preferred position (x_i, y_i) is estimated to vote Republican with probability

$$\rho_{rep} = \frac{[\exp(\lambda_r + bx_i + cy_i]}{1 + [\exp(\lambda_r + bx_i + cy_i]}. \tag{1}$$

The estimated coefficients in this model are $(\lambda_r, b, c) = (-0.24, 1.30, -0.70)$.

These empirical results suggest that economic policy (the *x-axis*) is a more salient dimension than social policy (the *y-axis*) in modeling vote choice for the Republican candidate.

According to this model, 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

$$y = 1.87x - 0.34. \tag{2}$$

Note that the cleavage line is very similarly positioned to the cleavage lines for the 1964 and 1980 elections, estimated by Schofield et al. (2003). Although this positional model has very significant coefficients, it takes the positions of Bush and Gore as exogenous, and so cannot be used to estimate the vote maximizing positions. We now present the formal stochastic model based on estimates of the candidate positions, and use this to explain the data just presented.

The pure spatial model, $\mathbb{M}(\lambda, \beta)$, is based on the voter utility assumption

$$u_{ij}(x_i, z_j) = \lambda_j - \beta\|x_i - z_j\|^2 + \varepsilon_j \tag{3}$$

$$= u^*_{ij}(x_i, z_j) + \varepsilon_j. \tag{4}$$

Here $u^*_{ij}(x_i, z_j)$ is the observable and $\{\varepsilon_j\}$ denote Type I extreme value errors, as discussed in Sect. 4.1. At a vector, $\mathbf{z}$, the probability that a voter i chooses candidate j is:

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

Table 4 presents the estimations of these spatial various models.[14]

We can compare these models using the differences in log likelihoods, as in Table 5.

[14]All models in Table 4 are given with Gore as the base, so the results give the estimations of the probability of voting for Bush.

Table 4 Spatial logit models for USA 2000 (Base = Gore)

Variable	(1) $\mathbb{M}(\boldsymbol{\lambda},\beta)$. Spatial	(2) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\alpha},\beta)$. Sp. and traits	(3) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\beta)$. Sp. and dem	(4) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\boldsymbol{\alpha},\beta)$. Full
Bush valence $\boldsymbol{\lambda}$	−0.43***	−0.69***	−0.39	0.48
	(5.05)	(5.64)	(0.95)	(0.72)
Spatial coeff. β	0.82***	0.35***	0.89***	0.38***
	(14.9)	(3.69)	(14.8)	(3.80)
Bush trait		3.559***		3.58***
		(13.84)		(13.60)
Gore trait		−3.22***		−3.15***
		(14.25)		(13.64)
Age			−0.14**	−0.22**
			(2.33)	(2.17)
Gender (F)			−0.139	−0.39
			(1.00)	(1.41)
African American			−1.57***	−1.45***
			(5.85)	(3.67)
Hispanic			−0.27	−0.23
			(0.77)	(0.49)
Class			−0.20	−0.12
			(1.30)	(0.47)
Education			0.18***	0.11
			(3.60)	(1.32)
Income			0.042**	−0.01
			(3.6)	(0.32)
Observations	1,238	1,238	1,238	1,238
Log likelihood (LL)	−707.8	−277.3	−661.3	−263.7
AIC	1,420	562.7	1,341	549.4
BIC	1,432	585.9	1,393	613.4

|*t – stat*| in parentheses
***prob < 0.001, **prob < 0.01, *prob < 0.05

Table 5 Differences in LL for US model comparisons in 2000

	$\mathbb{M}_2$	JPT	JST	ST	T
	JPT	na	4*	18***	25***
$\mathbb{M}_1$	JST	−4	na	14	21***
	ST	−18	−14	na	7***
	T	−25	−21	−7	na

JPT joint positional with traits, *JST* joint spatial with traits, *ST* pure spatial with traits, *T* pure traits
*** = highly significant

Note that the log likelihood of the pure spatial model given in Table 4(1) is −708, which we found to be very similar to the log likelihood of the pure positional model.

We use the equilibrium concept of local Nash equilibrium (LNE). This is simply a vector, **z**, such that each candidates vote share, $V_j(\mathbf{z}) = \frac{1}{n}\sum_i \rho_{ij}(\mathbf{z})$, is locally maximized.

To determine whether the joint origin, $\mathbf{z}_0 = (z_{Gore}, z_{Bush}) = (0, 0)$ is an equilibrium for the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, we need to examine the Hessians of the vote share functions.

The distribution of voter ideal points is characterized by *electoral covariance matrix*

$$\nabla_0 = \begin{bmatrix} 0.58 & -0.20 \\ -0.20 & 0.59 \end{bmatrix}.$$

The principal component of the electoral distribution is given by the vector (1.0, −3.05) with variance 0.785, while the minor component is given by the orthogonal eigenvector (1.0, 0.327) with variance 0.385. The correlation between these two factors is only −0.344.

Table 4(1) shows the intercept term λ_{Bush}, or exogenous valence for Bush in comparison to Gore, to be −0.43, while the β-coefficient is 0.82.

From the results in Sect. 4.1, it follows, according to the model $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, that the probability that a generic voter, i, chooses Bush, when both Bush and Gore are at the electoral origin, $\mathbf{z}_0$, is:

$$\begin{aligned}\rho_{Bush} &= \frac{\exp[u^*_{ibush}(x_i, z_{bush})]}{\exp[u^*_{igore}(x_i, z_{gore})] + \exp[u^*_{ibush}(x_i, z_{bush})]} \\ &= \frac{\exp[-0.43]}{\exp[0] + \exp[-0.43]} \\ &= [1 + \exp(0.43)]^{-1} \\ &= [1 + 1.54)]^{-1} = 0.40\end{aligned}$$

Section 4.1 shows that the Hessian of Bush's vote share function at $\mathbf{z}_0$ is given by the characteristic matrix

$$\begin{aligned}C_{Bush} &= [2\beta(1 - 2\rho_{Bush})]\,\nabla_0 - I \\ &= [2 \times 0.82 \times 0.2 \times \nabla_0] - I \\ &= (0.33)\,\nabla_0 - I \\ &= \begin{bmatrix} 0.19 & -0.07 \\ -0.07 & 0.195 \end{bmatrix} - I = \begin{bmatrix} -0.81 & -0.06 \\ -0.06 & -0.80 \end{bmatrix}\end{aligned}$$

The determinant is positive, and the trace negative, so both eigenvalues are negative, and the joint origin is a LNE of the pure spatial model. The *convergence coefficient*, c, is defined to be

$$c = 2\beta(1 - 2\rho_1)\,trace\,(\nabla_0) = 0.37.$$

The valence theorem in Sect. 4.1 shows that a sufficient condition for convergence to $\mathbf{z}_0$ in the pure spatial model is the condition $c < 1$. Using the coefficients of the pure spatial model, simulation of vote maximizing behavior confirmed that the joint origin was a LNE for the 2000 Presidential election. We also included the third party candidates, Nader and Buchanan, in the estimation, but the estimates of their valences were so low that they had no impact on the Local Nash Equilibrium at the joint origin. We also considered a model with different β coefficients on the two axes, but found again that the joint origin was a LNE.

Note however that $z_{Gore} = (-1.42, 0.66)$, while $z_{Bush} = (0.47, -1.24)$, so these estimated positions did not locally maximize the two candidates' vote shares, contingent on the validity of the pure spatial model with exogenous valence.

We now extend the model by adding the sociodemographic variables and electoral perception of traits, based on the utility assumption

$$u_{ij}(x_i, z_j) = \lambda_j + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta \left\| x_i - z_j \right\|^2 + \varepsilon_j \quad (5)$$

$$= u^*_{ij}(x_i, z_j) + \varepsilon_j \quad (6)$$

Here $\boldsymbol{\theta} = \{(\theta_j \cdot \eta_i)\}$ refers to sociodemographic characteristics while $\boldsymbol{\alpha} = \{(\alpha_j \cdot \tau_i)\}$ refer to the electoral perception of traits. We also obtained the results of a pure traits model, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha})$, and a pure sociodemographic model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta})$, respectively. These results showed that the pure sociodemographic model was statistically quite weak in comparison to the traits model.

Table 4 (models 2,3,4) gives the spatial models with traits, and sociodemographics, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \beta)$ and $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, respectively, as well as the model with both sociodemographics and traits, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$. The Table shows that a number of the sociodemographic coefficients were significantly different from zero in the models $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$ and $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, particularly those given by the categorical variable associated with African–American voters. Education is significant in $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, but not when traits are included. Moreover, the difference between the log-likelihoods of this joint model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, and the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, was a significant +46.

Table 4 shows that the models with traits are far superior to the models without traits. This can be seen from the comparison of the log-likelihoods of the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \beta)$ against $\mathbb{M}(\boldsymbol{\lambda}, \beta)$ and $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ against $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$. Note in particular that in the spatial model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \beta)$, with traits, both the spatial coefficient, β, and exogenous valence, $\boldsymbol{\lambda}$, are still significant. In the model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, the spatial coefficient remains significant, but the exogenous valence becomes insignificant. This suggests that the traits together with the sociodemographic variables provide a measure of candidate valence, but that the traits are weakly correlated with the sociodemographic variables.

For the model with traits, we found the difference between the maximum and minimum values of the Bush traits to be 4.5, while the difference for the Gore traits was 4.8. Since the coefficients on the Bush and Gore traits were 3.6 and 3.1 respectively, the magnitudes of these effects are highly significant in explaining voter behavior.[15]

[15]Table 4 shows that the pure traits model has an AIC of 574.8. This compares with an AIC of 664.3 for the trait model found by Clarke et al. (2009a).

We can justify the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ by comparing it with the joint positional model. This positional model with traits, denoted JPT, has a very similar AIC to the model presented by Clarke et al. (2009a), and correctly classifies approximately 90% of the voter choice. The difference in log-likelihoods between this model and the joint spatial model with traits, denoted JST in Table 9, is only +4. We can infer that both models give very statistically significant estimations of voter response. Notice also that both the difference between the log-likelihood of the pure spatial model with traits, denoted ST in Table 5, and of JST over the pure traits model are +7 and +21, respectively.[16] We can infer that though there may be some correlation between voter perception and voter preference, the significance of the model with traits is greatly enhanced by using spatial characteristics and the sociodemographics. As suggested by the related work by Clarke et al. (2009a), the spatial and traits models complement one another.

However, simulation of the model, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, showed that the unique local equilibrium was very close to the joint origin at

$$\mathbf{z}^{el} = \begin{bmatrix} & Bush & Gore \\ x & 0.027 & 0.027 \\ y & -0.02 & -0.02 \end{bmatrix}$$

At this LNE, we estimated the vote share for Bush to be 46% while the share for Gore was 54%. Although we find the traits model provides a statistically significant method of examining voter choice, it does not provide a satisfactory model of candidate positioning.

We now extend the model and assume voter utility is given by

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta||x_i - z_j||^2 + \varepsilon_j \tag{7}$$

$$= u^*_{ij}(x_i, z_j) + \varepsilon_j \tag{8}$$

where $\mu_j(z_j)$ is an activist term determined by candidate location. As Sect. 4.1 argues, the equilibrium, $\mathbf{z}^{el}$, of the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, provides an estimate of the weighted electoral means for the two candidates.

We now assume the estimated positions given in Fig. 1 comprise an LNE, $z^{\mathbb{V}*}$, of the model where each candidate has induced policy preferences from the supporting activists, as discussed in Sect. 4.3. Then the difference between the estimated positions and the weighted electoral means provides an estimate of the total activist pulls. Thus:

[16]The AIC results are similar. The AIC for the pure traits model is 575, which drops to 563 when β is added and drops further to 549 with the addition of the sociodemographics.

$$[\mathbf{z}^{\mathbb{V}*} - \mathbf{z}^{el}] = \begin{bmatrix} & Bush & Gore \\ x & 0.47 & -1.42 \\ y & -1.24 & +0.66 \end{bmatrix} - \begin{bmatrix} & Bush & Gore \\ x & 0.027 & 0.027 \\ y & -0.02 & -0.02 \end{bmatrix}$$
$$= \begin{bmatrix} & Bush & Gore \\ x & +0.44 & -1.45 \\ y & -1.22 & +0.68 \end{bmatrix}.$$

From the Activist Theorem 2 of Sect. 4.3, each term in this gradient equation is given by the expression

$$[z_j^{\mathbb{V}*} - z_j^{el}] = \frac{n^*\delta_j}{2(1-\delta_j)\beta}\frac{d\mu_j}{dz_j}(z_j^{\mathbb{V}*}) \equiv \frac{n^*\delta_j}{2(1-\delta_j)\beta}\sum_{k\in A_j} a_k \frac{dU_k}{dz_j} \text{ for } j \in P.$$

Here $j =$ Bush, Gore, and the $\frac{dU_k}{dz_j}$ terms are activist gradients pointing towards the preferred positions of the various activists for the two parties. The terms δ_j are weighting parameters, with $\delta_j = 0$ corresponding to a pure vote maximizing strategy.

Now the mean activist positions for the two parties are

2000	Rep Act	Dem Act
x	+0.42	−0.54
y	−0.30	+0.48

The gradient activist pulls and the mean activist positions are compatible if we assume that Republican activists strongly favor conservative social policies, while Democrat activists strongly favor liberal economic policies. As discussed above, Miller and Schofield (2003, 2008) have proposed a model of this kind, where each party has two classes of activists, economic and social. The set of bargains each set of party activists may make over the influence they exert on their parties is given by a one dimensional contract curve, as shown in Fig. 7 in Sect. 4.2, below. The set of possible optimal positions for each candidate is then given by a one dimensional balance locus. The actual optimal position will depend on the "eccentricity" of the utility functions of the activists, namely the trade off between activist marginal willingness to contribute and their demand for policy gains.

These estimates indicate that there is a tug of war between voters and activists over the location of the party candidates. The distribution of voters preferred points is concentrated in the electoral center, so the weighted electoral gradient of each candidate points towards the center, as is consistent with the standard spatial model. However, activists for the two parties are more concerned with social policy (for the Republicans), and economic policy (for the Democrats). There will be conflict between activists for each party as recent events have suggested. The overall effect draws the candidates into the opposing quadrants, as suggested by Fig. 1. We infer that activists with more radical policy preferences have a significant influence on the candidates. Indeed, comparison of the estimated vote share at $\mathbf{z}^{el}$ suggest that activist contributions helped the Bush campaign by increasing his vote share from the estimated 46% to about 50%.

Here we have shown that the electoral model requires additional terms of the form $\{\mu_j\}$. We have argued that these terms are due to activists.[17] Determining the precise form requires a solution of a complex activist-candidate bargaining game as proposed by Grossman and Helpman (1994, 1996, 2001) and Baron (1994). For our purpose we have assumed that the candidate-activist contracts have been concluded prior to the election, and the empirical estimates of the various valences are the consequence of the shifts in positions away from the electoral equilibria that we have estimated.

Note that in the version of the model given in Sect. 4.4, where candidates target voters, the precise theoretical equilibrium positions will depend not only on the activist positions, but on the willingness of voters to be persuaded.

2.2 *The Election of 2004*

We repeated the above analysis for the 2004 election contest between Bush and Kerry, using the same set of responses from the ANES 2004, as for 2000.

The positions of the major presidential candidates, Bush and Kerry, in 2004 were estimated in a similar fashion to that of the sampled individuals. Scores were assigned to each candidate for each of the constituent survey items based on press reports of their campaign stances on various issues. Factor loadings and descriptive statistics are given in Tables 6 and 7. These estimates were used to obtain estimated policy positions for each candidate, as in Table 8.

Figure 4 gives the contour plot of the electoral distribution in the policy space in 2004 while Fig. 5 gives a perspective plot. Again, left on the economic (x) axis is

Table 6 Factor loadings from the American national election survery, 2004

Question	Social policy	Economic policy
1. Economic problems	−0.02	0.39
2. Federal spending	0.04	0.39
3. Equality	0.30	0.43
4. African American	0.32	0.49
5. Immigrants	0.27	0.17
6. Liberal vs conservative	0.39	0.34
7. Guns	0.13	0.34
8. Abortion	0.57	−0.05
9. Gays	0.60	0.11
10. Family	0.62	0.21
% Variance explained	15.1	10.5

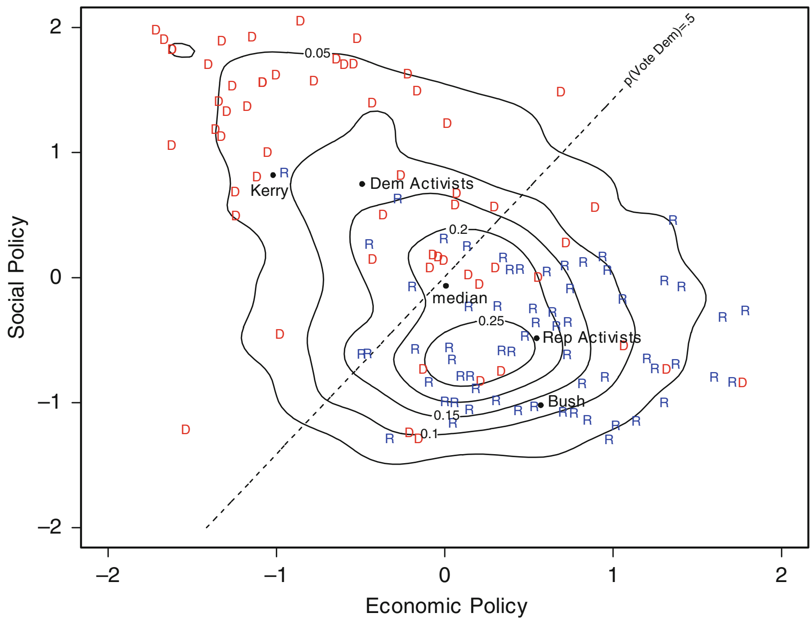
[17] As we show in Sect. 4.3, we can interpret these terms as policy preferences on the part of candidates, but induced from the policy preferences of activists.

Table 7 Descriptive data 2004

	Econ mean	Policy std.err	95% C.I	Social mean	Policy std.err	95% C.I	*n*
Activists							
Democrats	−0.49	0.07	[−0.63, −0.55]	0.75	0.14	[0.47, 1.03]	47
Republicans	0.55	0.06	[0.43, 0.67]	−0.48	0.06	[−0.6, −0.36]	63
Non-activists							
Democrats	−0.33	0.03	[−0.39, −0.27]	0.37	0.04	[0.29, 0.45]	413
Republicans	0.30	0.03	[0.24, 0.36]	−0.28	0.03	[−0.34, −0.22]	440

Table 8 Kerry and Bush estimated responses, 2004

Question	Bush	Kerry
1. Economic problems	3	1
2. Federal spending	3	2
3. Equality	3	1
4. African American	4	2
5. Immigrants	3	3
6. Liberal vs conservative	3	1
7. Guns	2	1
8. Abortion	3	2
9. Gays	3	1
10. Family	5	3
Estimated position: Social policy	−1.02	0.83
Estimated position: Economic policy	0.57	−1.30



This figure will be printed in b/w

Fig. 4 2004 estimated voter distribution, and activist positions (Democrat activists are denoted D and Republican activists are denoted R)

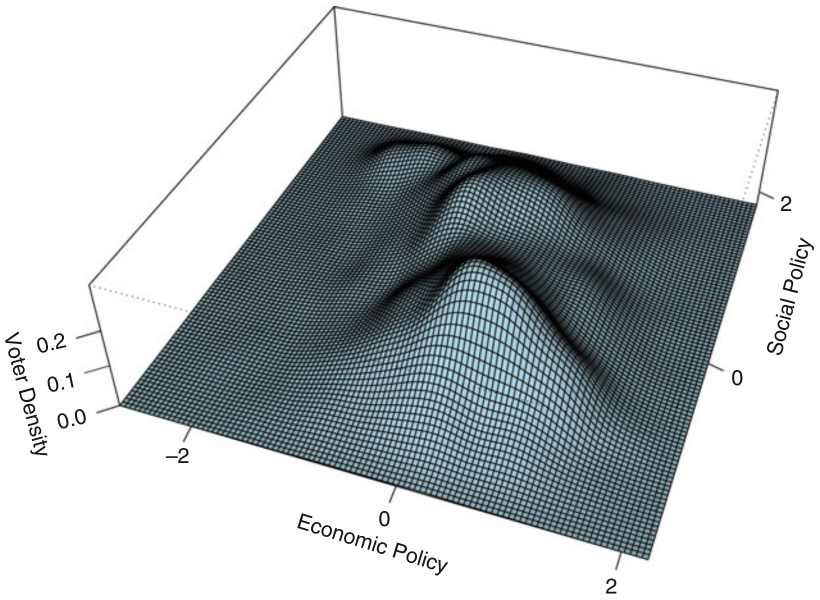


Fig. 5 Pespective plot of the sample electorate in 2004

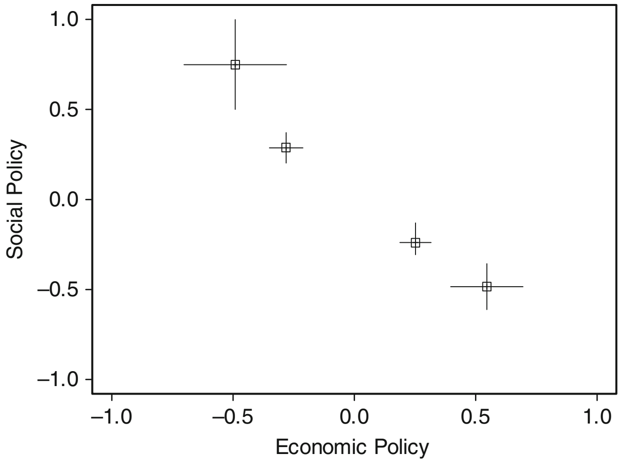


Fig. 6 Comparison of mean partisan and activist positions in 2004

pro-redistribution, while north on the social (y) axis is pro-individual civil rights. Figure 6 gives the estimated mean positions of Democrat and Republican partisans and activists for 2004, together with the the error bars of the estimates. The mean partisan positions are characterized by smaller error bars. Figure 4 is also annotated with the estimated positions of Republican and Democrat activists, and the two party presidential candidates.

Note that the positions of the presidential candidates are given by

$$\mathbf{z}^* = \begin{bmatrix} candidate & Bush & Kerry \\ x & 0.57 & -1.30 \\ y & -1.02 & +0.83 \end{bmatrix}$$

Figure 4 also shows the estimated threshold dividing likely Democrat candidate voters from Republican candidate voters. Again this partisan cleavage line was derived from a binomial logit model, designed to test the effects of each policy dimension on vote choice. The results were very similar to those obtained for 2000 and are not reported here.

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

$$\rho_{rep} = \frac{\exp(a + bx_i + cy_i)}{1 + \exp(a + bx_i + cy_i)} \tag{9}$$

where $(a, b, c) = (-0.20, 1.34, -0.93)$

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

$$y = 1.44x - 0.21. \tag{10}$$

which almost goes through the electoral origin. The effect of economic policy preferences is the stronger of the two dimensions in determining choice between the Democrat and Republican candidates.

The estimated position for Kerry is $z_{Kerry} = (-1.30, 0.83)$, while the Democrat activist and partisan mean positions are $z^{act}_{DEM} = (-0.49, 0.75)$ and $z^{part}_{DEM} = (-0.33, 0.37)$. Thus Kerry's position on the economic axis is a distance about +0.8 more extreme that the mean activist position on this axis.

Similarly, the estimated position for Bush is $z_{Bush} = (0.57, -1.02)$, while the Republican activist and partisan mean positions are $z^{act}_{REP} = (0.55, -0.48)$ and $z^{part}_{REP} = (0.30, -0.28)$. Thus Bush's position on the social axis is a distance about +0.5 more extreme that the mean activist position on this axis.

Note however, that there are Democratic and Republican activists located at far more extreme positions that the two candidates. It is consistent with the model of activists discussed in the Technical Appendix that more extreme activists will have a disproportionate effect on the candidate positions.

To use estimations of a spatial model, we consider the various logit models presented in Table 9.

The electoral covariance matrix obtained from the factor analysis is given by

$$\nabla_0 = \begin{bmatrix} 0.58 & -0.177 \\ -0.177 & 0.59 \end{bmatrix}.$$

Table 9 Spatial logit models for USA in 2004 (Base = Kerry)

Variable	(1) $\mathbb{M}(\boldsymbol{\lambda},\beta)$. Spatial	(2) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\alpha},\beta)$. Sp. and traits	(3) $\mathbb{M}$ $(\boldsymbol{\lambda},\boldsymbol{\theta},\beta)$. Sp. and dem	(4) $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\theta},\boldsymbol{\alpha},\beta)$ Full
Bush valence λ	−0.43***	−0.15	−1.72***	−0.670
	(5.05)	(1.00)	(3.50)	(0.70)
Spatial coeff. β	0.95***	0.47***	1.09***	0.475***
	(14.21)	(3.49)	(13.76)	(3.125)
Bush trait		4.18***		4.22***
		(11.49)		(11.40)
Kerry trait		−4.20***		−4.14***
		(11.58)		(11.13)
Age			−0.16**	0.03
			(2.61)	(0.25)
Gender (F)			0.08	−0.38
			(0.44)	(1.18)
African American			−1.62***	−1.13*
			(6.11)	(2.30)
Hispanic			−0.26	0.14
			(0.75)	(1.75)
Class			0.22	0.26
			(1.20)	(0.75)
Education			0.15***	0.136
			(2.37)	(1.12)
Income			0.056***	0.012
			(3.29)	(0.038)
Observations	935	935	935	935
Log likelihood (LL)	−501.7	−145.3	−448.2	−137.8
AIC	1,007	298.5	914.4	297.7
BIC	1,018	320.7	964.2	358.6

|*t* – *stat*| in parentheses

***prob < 0.001, **prob < 0.01, *prob < 0.05

The principal component of the electoral distribution is given by the vector (1.0, −1.04) with variance 0.765, while the minor component is given by the orthogonal eigenvector (1.0, 0.96) with variance 0.405. The total variance, $\sigma^2 = trace(\nabla_0)$ is 1.15, and the *electoral standard deviation* (esd) is $\sigma = 1.07$.

Table 9(1) shows the coefficients in 2004 for the pure spatial model to be

$$(\lambda_{Kerry}, \lambda_{Bush}, \beta) = (0, -0.43, 0.95).$$

According to the model $\mathbb{M}(\lambda, \beta)$, the probability that a voter chooses Bush, when both Bush and Kerry are at the electoral origin, $\mathbf{z}_0$, is

$$\rho_B = [1 + \exp(0.43)]^{-1} = [1 + 1.52)]^{-1} = 0.40.$$

Then from the valence theorem presented in Sect. 4.1, the Hessian for Bush, when both candidates are at the origin, is given by:

Table 10 Differences in LL for US model comparisons in 2004

	JPT	JST	ST	T
JPT	na	2	5	12
JST	−2	na	7	14
ST	−5	−7	na	7
T	−12	−14	−7	na

JPT joint positional with traits, *JST* joint spatial with traits, *ST* spatial with traits, *T* pure traits

$$\begin{aligned} C_{Bush} &= [2\beta(1-2\rho_{Bush})\nabla_0 = 2\times 0.95\times 0.2\times \nabla_0 - I] \\ &= (0.38)\nabla_0 - I \\ &= (0.38)\begin{bmatrix} 0.53 & -0.18 \\ -0.18 & 0.66 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.8 & -0.06 \\ -0.06 & -0.75 \end{bmatrix} \end{aligned}$$

It is obvious that both eigenvalues are negative. The convergence coefficient is $c = (0.38)\times(1.19) = 0.45$, and the joint origin is a LNE.

For the joint model given in Table 9(3), we cannot assert from first principles that $(z_{Bush}, z_{Kerry}) = (0, 0)$ satisfies the first order condition for a LNE. However, simulation of the model led us to infer that the joint origin was indeed a LNE of this model.

Although a number of the sociodemographic valences were significantly different from zero, we can infer that their magnitude is insufficient to perturb the equilibrium away from the joint origin.[18]

Note however that $z_{Kerry} = (-1.30, 0.83)$, while $z_{Bush} = (0.57, -1.02)$, so these estimated positions did not locally maximize the two candidates vote shares, contingent on the validity of the pure spatial model with exogenous valence.

We now extend the model by adding the electoral perception of traits.

It is also clear from Table 9 that the spatial models with traits (ST) and (JST) are far superior to all other spatial models without traits. Note in particular that in the joint spatial model with traits, the spatial coefficient is still significant, while the coefficient on the exogenous valence becomes insignificant. This is to be expected, as the traits are a substitute for the measure of candidate valence.

Table 10 gives the comparisons of the log-likelihoods between the joint positional (JPT) model with traits, and the spatial (ST) and joint spatial (JST) models with traits, as well as the pure traits model (T). There is only a minor difference of

[18]Notice that the difference between the loglikihoods of the joint spatial model and that of the pure spatial model is +53. It is possible that adding further demographic variables would change the LNE of the joint model. However, since the effect of sociodemographic variables is limited, it is unlikely that there would be any substantial effect on the LNE.

+2 in log-likelihood between JPT and JST, which we take as justification for the estimates of candidate positions in the stochastic model.[19] Note however that the AIC values for JST and ST are similar (297.7 versus 298.5). Finally, traits do not capture all the electoral characteristics. The AIC results are similar. The AIC for the pure traits model is 311, which drops to 298.5 when β is added and drops slightly more to 297.7 with the addition of the sociodemographics. The log-likelihood differences are 7 and 14 respectively.

Simulation of the joint spatial model with traits showed that the joint origin was not an equilibrium, but the LNE was very close to the joint origin:

$$\mathbf{z}^{el} = \begin{bmatrix} & Bush & Kerry \\ x & 0.03 & 0.03 \\ y & -0.021 & -0.021 \end{bmatrix}.$$

We now assume the estimated positions comprise an LNE of the full activist model, so

$$\mathbf{z}^{\mathbb{V}*} = \begin{bmatrix} candidate & Bush & Kerry \\ x & 0.57 & -1.30 \\ y & -1.02 & +0.83 \end{bmatrix}$$

Since $\mathbf{z}^{el} = (z^{el}_{Bush}, z^{el}_{Kerry})$ is an LNE from the joint model, with no activist valence terms, we infer that $\mathbf{z}^{el}$ is the vector of weighted electoral means. Thus by the balance condition, as given in Sect. 4.1:

$$\mathbf{z}^{\mathbb{V}*} - \mathbf{z}^{el} = \begin{bmatrix} & Bush & Kerry \\ x & 0.57 & -1.30 \\ y & -1.02 & +0.83 \end{bmatrix} - \begin{bmatrix} & Bush & Kerry \\ x & 0.03 & 0.03 \\ y & -0.021 & -0.021 \end{bmatrix}$$
$$= \begin{bmatrix} & Bush & Kerry \\ x & 0.54 & -1.33 \\ y & -1.0 & +0.85 \end{bmatrix}.$$

The difference between $\mathbf{z}^{\mathbb{V}*}$ 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 Fig. 4) are compatible with this inference. These estimates indicate that the more extreme economic activists exerted significant pulls on both candidates in 2004, drawing them into the opposite quadrants.

[19]Clarke et al. (2009a) obtained an AIC of 239 for a composite version of the model here called JPT. However, they used many more sociodemographic variables. The value of 297.7 for the AIC of the spatial model, JST, suggests that it is a valid model of electoral behavior.

$$\begin{bmatrix} 2004 & \text{Rep Act} & \text{Dem Act} \\ x & 0.55 & -0.49 \\ y & -0.48 & +0.75 \end{bmatrix}.$$

As in the analysis for 2000, if we assume that the Democrat activists tend to be more concerned with liberal economic policy and Republican activists tend to be more concerned with conservative social policy, then we have an explanation for the candidate shifts from the estimated equilibrium.

3 Concluding Remarks and Recent Events

Valence, whether exogenous or based on electoral perceptions of character traits, is intended to model that component of voting which is determined by the judgements 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 (Madison [1787],1999). Schofield (2002) has suggested that Madison's argument may well have been influenced by Condorcet's work on the so-called "Jury Theorem" (Condorcet 1785, McLennan 1998). 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' judgements, as well as their policy preferences, strongly influence their political choice.

This chapter 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 judgements. One inference from the work presented here does seem to belie Riker's arguments (1980, 1982) 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.

The empirical and formal models presented here do suggest that these perceptions are the result of the influence of activist groups. Changes in voter choice appear to result not only from changes in the electoral distribution, but from the shifts in electoral perceptions. In turn, these changes are the result of the competition between the candidates over activist support. As we noted in the introduction, the importance of electoral contributions has increased, and this has enhanced the influence of activist groups.

While the analysis presented here has focused on a presidential election, it can, in principle, be applied to congressional elections as well. In this case, instead of dealing with cooperation between activist groups for a single party, we could model competition between activist groups over candidate choice for a party. Recent events over the election for the New York 23rd congressional district show how contentious this competition can be. On November 1, 2009, conservative pressure forced the centrist Republican candidate, Dede Scozzafava, to drop out of the primary race and endorse the Democrat candidate, Bill Owens. On November 3, Owens won the election in what had been a Republican district since 1872.

In the mid term 2010 election cycle total campaign spending was about $4 billion, with Republican spending somewhat higher than that of the Democrats. The extremely high level of expenditure (especially for a midterm election) is particularly interesting because of the increasing degree of polarization mentioned in the Introduction. In this election the Democrats lost 63 seats in the House, leading to a Republican majority of 242–192. In the Senate the Democrats lost 6 seats but retained a majority of 53 (with 2 pro-Democrat Independents) to 47. President Obama was eventually able to strike a deal with the Republicans in December, 2010, to extend unemployment benefits and implement a 1-year payroll-tax cut for most workers. He was forced to accept the Republican demand for a continuation of Bush tax cuts even for the very wealthy.[20] The House speaker, Nancy Pelosi of California, accused Republicans of forcing Democrats "to pay a king's ransom in order to help the middle class". The bill passed the Senate on 15 December by 81–19, and at midnight on 16 December, 139 House Democrats voted with 138 House Republicans for the bill, against 112 Democrats and 36 Republicans. Obama immediately signed the bill.

After this initial compromise, the logjam seemed to have broken when Congress, on December 21, did approve a temporary spending bill up until March 2011.

However, on December 18, the "Dream Act" Bill to allow illegal immigrant students to become citizens failed on a Senate vote of 55–41. The Senate did vote 65–31 to repeal the "Don't Ask, Don't Tell" legislation, making it possible for gays to serve openly in the military. The House had previously approved this repeal by 250–175.

On December 20, the Senate voted 59–37 to reject an amendment to the new arms control treaty, "New Start," with Russia. The amendment would have killed the treaty because any change to the text would have required the United States and Russia to renegotiate the treaty. On December 22, the Senate voted 71–26 for the treaty. This treaty was seen as the most tangible foreign policy achievement of President Obama. Thirteen Republicans joined a unanimous Democratic caucus to vote in favor, exceeding the two-thirds majority required by the Constitution.

The Senate also voted for a $4.3 billion bill to cover medical costs for rescue workers after the 2001 terrorist attack. The House immediately voted for the bill 206–60, and it was sent to President Obama to sign into law. Congress also passed a defense authorization bill covering costs for Afghanistan and Iraq.[21]

As Obama said:

> I think it's fair to say that this has been the most productive post-election period we've had in decades, and it comes on the heels of the most productive 2 years that we've had in generations. If there's any lesson to draw from these past few weeks, it's that we are not

---

[20]This deal between the two opposed coalitions will have a serious effect on the overall U.S. debt.

[21]The bill did make it more difficult to transfer detainees from Guantánamo.

doomed to endless gridlock. We've shown in the wake of the November elections that we have the capacity not only to make progress, but to make progress together.

One of the first moves by the House in the new 112th Congress was to vote, on January 19, 2011, to repeal the Health Care Bill by a margin of 245–189. However, this repeal cannot pass the Democrat majority in the Senate.

A general inference from the model presented here is that the earlier debate about whether elections are chaotic, or are institutionally stabilized, needs to be recast. It may well be that social choice in the two party system, in the presence of activist conflict, need be neither chaotic nor equilibrating. Instead, as Miller and Schofield (2008) have argued, these activist conflicts appear to have induced a transformation of both parties in the United States. Over the long run, these continuing transformations induce a slow realignment of the political configuration.

4 Technical Appendix

4.1 The Stochastic Model of Elections

The electoral model that we deploy is an extension of the multiparty stochastic model of McKelvey and Patty (2006), modified by inducing asymmetries in terms of valence. The justification for developing the model in this way is the empirical evidence that valence is a natural way to model the judgements made by voters of party leaders or candidates for office.[22] There are a number of possible choices for the appropriate model for multiparty competition. The simplest one, which is used here, is that the utility function for the agent j is proportional to the anticipated vote share, V_j, of the party in the election.[23]

With this assumption, we can examine the conditions on the parameters of the stochastic model which are necessary for the existence of a pure strategy Nash equilibrium (PNE). Because the vote share functions are differentiable, we use calculus techniques to obtain conditions for positions to be locally optimal. Thus we examine what we call *local pure strategy Nash equilibria* (LNE). From the definitions of these equilibria it follows that a PNE must be a LNE, but not conversely.

The stochastic electoral model utilizes socio-demographic variables and voter perceptions of character traits. For this model we assume that voter i utility is given by the expression

[22] We can use the model either for party leaders or candidates for office, as in the United States. In the following we shall use the term *agents* to mean either one.

[23] For refining the model, and for empirical analysis, we could adapt the model so that parties choose positions to maximize their seat shares, relative to a given constituency structure. We adopt the simplifying vote share assumption in order to present the essential structure of the formal model.

$$u_{ij}(x_i, z_j) = \lambda_j + \mu_j(z_j) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta||x_i - z_j||^2 + \varepsilon_{ij} \quad (11)$$

$$= u^*_{ij}(x_i, z_j) + \varepsilon_j \quad (12)$$

Here $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ is the exogenous valence vector which we assume satisfies the ranking condition $\lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1$. The agents are labelled $(1, \ldots, p)$ and λ_j is the exogenous valence of agent or candidate j. The points $\{x_i : i \in N\}$ are the preferred policies of the voters, in the compact Euclidean space X, of finite dimension w. The vector $\mathbf{z} = \{z_j : j \in P\}$ gives the positions of the agents in the same space. The term

$$||x_i - z_j|| = \left[\sum\nolimits_{t=1}^{w} (x_{it} - z_{jt})^2\right]^{1/2}$$

is simply the Euclidean distance between x_i and z_j. We assume the error vector $\boldsymbol{\varepsilon}$ is distributed by the type I extreme value (Gumbel) distribution, as in empirical conditional logit estimation. This assumption means that all $\{\varepsilon_j\}$ are iid, so we write $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_j, .., \varepsilon_p)$. The variance of ε_j is fixed at $\frac{\pi^2}{6}$. As a result, the spatial parameter, β, has dimension L^{-2}, where L is whatever unit of measurement is used in X.

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 $\{\theta_j : j \in P\}$ 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 $\{(\alpha_j \cdot \tau_i)\}$ 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, etc. We let $\boldsymbol{\alpha} = (\alpha_p, ... \alpha_1)$. 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', etc. The perception of traits can be augmented with voter perception of the state of the economy, etc. in order to examine how anticipated changes in the economy affect each agent's electoral support.[24]

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

Partial models are:

1. Pure *sociodemographic*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \theta)$, with only exogenous valence and sociodemographic variables
2. *Pure spatial*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, with only exogenous valence and β

[24]See Clarke et al. (2009a) for the same empirical procedure.

3. *Joint spatial*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \beta)$, with exogenous valence, sociodemographic variables and β
4. *Spatial sociodemographic model with traits*, denoted $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, without the activist components

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

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

A (strict) local Nash equilibrium (LNE) for a model $\mathbb{M}$ is a vector, $\mathbf{z}$, such that each candidate, j, chooses z_j to locally (strictly)[25] maximize the expected vote share

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

In these models, political candidates 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 candidates' 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. Note however, that as candidates adjust positions 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 a LNE then it cannot be a PNE. Indeed there may not exist a PNE.

It follows from Train (2003) that, for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\mu}, \beta)$, the probability, $\rho_{ij}(\mathbf{z})$, that voter i, with ideal point, x_i, picks j at the vector, $\mathbf{z}$, of candidate positions is given by

$$\rho_{ij}(\mathbf{z}) = \left[1 + \sum\nolimits_{k \neq j}\left[\exp(f_{kj})\right]\right]^{-1}$$
$$\text{where } f_{kj} = u^*_{ik}(x_i, z_k) - u^*_{ij}(x_i, z_j).$$
$$\text{Thus } \frac{d\rho_{ij}(\mathbf{z})}{dz_j} = -\left[1 + \sum\nolimits_{k \neq j}\left[\exp(f_{kj})\right]\right]^{-2} \frac{d}{dz_j}\left[\sum\nolimits_{k \neq j}\left[\exp(f_{kj})\right]\right]$$
$$= \{2\beta(x_i - z_j) + \frac{d\mu_j}{dz_j}(z_j)\}[\rho_{ij} - \rho^2_{ij}].$$

We use this gradient equation to show that the first order condition for $\mathbf{z}^*$ to be a LNE is given by the following balance equations:

[25] We keep to strict equilibria to avoid non-generic problems when one eigenvalue is zero.

$$0=\frac{dV_j(z)}{dz_j}=\frac{1}{n}\sum_{i\in N}\frac{d\rho_{ij}}{dz_j}$$

$$=\frac{1}{n}\sum_{i\in N}[\rho_{ij}-\rho_{ij}^2]\left\{2\beta(x_i-z_j)+\frac{d\mu_j}{dz_j}(z_j)\right\}.$$

$$\text{Hence } z_j^*=\sum_i\frac{[\rho_{ij}-\rho_{ij}^2]x_i}{\sum_{k\in N}[\rho_{kj}-\rho_{kj}^2]}+\frac{1}{2\beta}\frac{d\mu_j}{dz_j}(z_j),$$

$$\text{or } z_j^*=\frac{1}{2\beta}\frac{d\mu_j}{dz_j}(z_j)+\sum_{i=1}^{n}\varpi_{ij}x_i.$$

$$\text{This can be written } 0=\left[z_j^{el}-z_j^*\right]+\frac{1}{2\beta}\frac{d\mu_j}{dz_j}(z_j)$$

$$\text{where } z_j^{el}=\sum_{i=1}^{n}\varpi_{ij}x_i.$$

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.[26]

A similar analysis holds for the full model $\mathbb{M}(\boldsymbol{\lambda},\boldsymbol{\mu},\boldsymbol{\theta},\boldsymbol{\alpha},\beta)$. We can therefore write the first order balance condition at an equilibrium, $\mathbf{z}^*=(z_1^*,..,z_j^*..z_p^*)$, as a set of *gradient balance conditions*

$$\frac{d\varepsilon_j^*}{dz_j}(z_j^*)+\frac{1}{2\beta}\frac{d\mu_j}{dz_j}(z_j^*)=0. \tag{13}$$

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

$$\frac{dE_j^*}{dz_j}(z_j)=\left[z_j^{el}-z_j\right]$$

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

Writing $\mathbf{z}^{el}=\left(z_1^{el},z_2^{el},..z_p^{el}\right)$, and $\frac{\mathbf{d}\boldsymbol{\mu}}{\mathbf{dz}}(\mathbf{z}^*)=\left(..,\frac{d\mu_j}{dz_j}(z_j^*),..\right)$, then in vector notation, the first order condition can be written

$$\left[\mathbf{z}^*-\mathbf{z}^{el}\right]=\frac{1}{2\beta}\frac{\mathbf{d}\boldsymbol{\mu}}{\mathbf{dz}}(\mathbf{z}^*)$$

[26] 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.

The vector $\mathbf{z}^*$ is said to be a *balance solution* if this vector equation is satisfied.

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 $\frac{dV_j^2(\mathbf{z})}{dz_j^2}$. These will involve the second order terms $\frac{d^2\mu_j}{dz_j^2}$. Schofield (2006) proved the following Activist Theorem. The theorem suggests that there will be natural conditions under which the second order terms $\frac{d^2\mu_j}{dz_j^2}$ will be negative definite. Indeed if the eigenvalues are negative and of sufficiently large modulus, then we may expect the existence of PNE.

Activist Theorem 1. *Consider the electoral model* $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$

(i) *The first order condition for* z^* *to be an LNE is that it is a balance solution.*
(ii) *If all activist valence functions are sufficiently concave,*[27] *then a balance solution will be a PNE.* □

For the pure spatial model, $\mathbb{M}(\boldsymbol{\lambda}, \beta)$, it is clear that when the candidate positions are identical, then $\rho_{kj} = \rho_j$, is independent of the voter suffix k. Thus all $\varpi_{ij} = \frac{1}{n}$ gives the first order condition for a LNE. By a change of coordinates, it follows that $\mathbf{z}_0 = (0, \ldots 0)$ is a candidate for a LNE.[28] Note however that this argument does not follow for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, and generically $\mathbf{z}^{el} = (z_1^{el}, z_2^{el}, ..z_p^{el}) \neq (0, \ldots 0)$.

Since the valence functions are constant in the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, the marginal effects, $\frac{d\mu_j}{dz_j}$, will be zero. However, since the weights in the weighted electoral mean for each candidate will vary from one individual to another, it is necessary to simulate the model to determine the LNE $\mathbf{z}^{el} = \left(z_1^{el}, z_2^{el}, ..z_p^{el}\right)$.[29] Notice also that the marginal vote effect, $\frac{d\rho_{ij}}{dz_j}$, for a voter with $\rho_{ij}(\mathbf{z}) \simeq 1$ will be close to zero. Thus in searching for LNE, each candidate will seek voters with $\rho_{ij}(z) < 1$.

For the pure spatial model, $\mathbb{M}(\lambda, \beta)$, we have shown that $\mathbf{z}_0$ satisfies the first order condition for LNE. The necessary and sufficient second order condition for LNE at $\mathbf{z_0}$ in the pure spatial model, $\mathbb{M}(\lambda, \beta)$, is determined as follows. When all candidates are at the electoral origin, and agent 1 is, by definition, the lowest valence candidate, then the probability that a generic voter picks candidate 1 is given:

$$\rho_1 = \left[1 + \sum\nolimits_{k \neq 1} [\exp(\lambda_k)]\right]^{-1} \tag{14}$$

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

---

[27]By this we mean that the eigenvalues of the activist functions are negative and of sufficient magnitude everywhere. That is to say, there exists $\alpha < 0$, such that all eigenvalues $< \alpha$ is sufficient to guarantee existence of a PNE.

[28]It is worth observing that if we use just distance rather than distance squared then the first order condition is essentially a counting rule, giving a median position of the candidates as equilibrium.

[29]This can be done using the gradient equation given below.

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

where I is the w by w identity matrix, and we use to denote scalar product. Again, when all candidates are at the origin then $\rho_{i1} = \rho_1$ is independent of i. Moreover,

$$\frac{1}{n}\sum_{i\in N}(x_i).(x_i) = \nabla_0 \tag{15}$$

is the w by w covariance matrix of the distribution of voter ideal points, taken about the electoral origin. Then setting $z_1 = 0$ in the above equation, we see that the Hessian of the vote share function of candidate 1 is given by

$$2\beta\left[\rho_1 - \rho_1^2\right][2\beta(1 - 2\rho_1)\nabla_0 - I] \tag{16}$$

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

$$C_1 = 2\beta(1 - 2\rho_1)\nabla_0 - I, \tag{17}$$

Following Schofield (2007), this shows that the necessary and sufficient second order condition for an LNE at $\mathbf{z}_0$ is that C_1 has negative eigenvalues.[30]

These second order conditions can be interpreted in terms of the trace and determinant of C_1. Schofield (2007) also shows that a necessary condition for $\mathbf{z}_0 = (0, \ldots 0)$ to be an LNE is that a convergence coefficient, c, defined by

$$c = 2\beta(1 - 2\rho_1)\sigma^2$$

satisfies the critical convergence condition, $c < w$. Here $\sigma^2 = trace(\nabla_0)$ is the sum of the variance terms on all axes. We state this as the Valence Theorem.

The Valence Theorem.

(i) *The joint origin* $\mathbf{z_0}$ *satisfies the first order condition to be a LNE for the model* $\mathbb{M}(\lambda, \beta)$.

[30] Strictly speaking, the condition is that the eigenvalues are non-positive. To avoid the degenerate case with a zero eigenvalue, we focus on a strict local equilibrium associated with negative eigenvalues of the Hessian.

(ii) *The necessary and sufficient second order condition for LNE at z_0 is that C_1 has negative eigenvalues.*[31]

(iii) *A necessary condition for z_0 to be a LNE for the model $\mathbb{M}(\lambda, \beta)$ is that $c(\boldsymbol{\lambda}, \beta) < w$.*

(iv) *A sufficient condition for convergence to z_0 in the two dimensional case is that $c < 1$.* □

Note that β has dimension L^{-2}, while σ^2 has dimension L^2 so c is dimensionless, and is therefore independent of the units of measurement.

In the two dimensional case, Schofield (2007) also shows that a *sufficient* condition for negative eigenvalues, and thus for convergence to $\mathbf{z}_0$, is that $c < 1$. Note that when $c \ll 1$, and the eigenvalues are both negative and of sufficient magnitude then we might expect that the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ will have an equilibrium $(z_1^{el}, z_2^{el}, ..z_p^{el}) \simeq \mathbf{0}$, that is close to the joint origin. We found this for the two elections discussed here.[32]

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

4.2 Extension to the Case with Multiple Activist Groups

We adapt the model presented in Schofield and Cataife (2007), where there are multiple activist groups for each party.

1. For each party candidate, 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 $R_{jk}(U_k(z_j))$. The total activist valence function for candidate j is the linear combination

$$\mu_j(z_j) = \sum_{k \in A_j} \mu_{jk}\left(R_{jk}\left(U_k(z_j)\right)\right). \tag{18}$$

   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}.

2. Assume the gradients of the valence functions for j are given by

$$\frac{d\mu_{jk}}{dz_j} = a_k^* \frac{dR_{jk}}{dz_j} = a_k^* a_k^{**} \frac{dU_k}{dz_j} \tag{19}$$

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

[31] In the usual way, the condition for an LNE is that the eigenvalues are negative semidefinite.

[32] Indeed, we found the same result for the 2008 election reported in Schofield et al. (2011).

3. Under these assumptions, the first order equation $\frac{d\mu_j}{dz_j} = 0$ becomes

$$\frac{d\mu_j}{dz_j} = \sum_{k \in A_j} \frac{d}{dz_j} \left[\mu_{jk} \left(R_{jk} \left(U_k (z_j) \right) \right) \right] \tag{20}$$

$$= \sum_{k \in A_j} \left(a_k^{**} a_k^* \right) \frac{dU_k}{dz_j} = 0. \tag{21}$$

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

$$\sum_{k \in A_j} a_k \frac{dU_k}{dz_j} = 0, \text{where } \sum_{k \in A_j} a_k = 1 \text{ and all } a_k > 0. \tag{22}$$

Here we normalize by setting $a_k = \frac{a_k^{**} a_k^*}{\sum_{m \in A_j} a_m^{**} a_m^*}$.

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

$$\left[z_j^{el} - z_j^* \right] + \frac{1}{2\beta} \left[\sum_{k \in A_j} a_k \frac{dU_k}{dz_j} \right] = 0. \tag{23}$$

The simplest case, discussed in Schofield and Cataife (2007) is in two dimensions, where each candidate has two activist groups. In this case, the contract curve for each candidate's supporters will, generically, be a one-dimensional arc. Miller and Schofield (2003) also supposed that the activist utility functions were ellipsoidal, mirroring differing saliences on the two axes. In this case the contract curves would be *catenaries*, and the balance locus would be a one dimensional arc. The balance solution for each candidate naturally depends on the position(s) of opposed candidate(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.

Figure 7 gives an illustration of the equilibrium balance solution for a Republican candidate, denoted $z_1^*(z_2)$ in the figure. Here z_2 denotes the position taken by the Democrat candidate.

Since the activist valence function for candidate j depends on the resources contributed by the various activist groups to this candidate, 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.

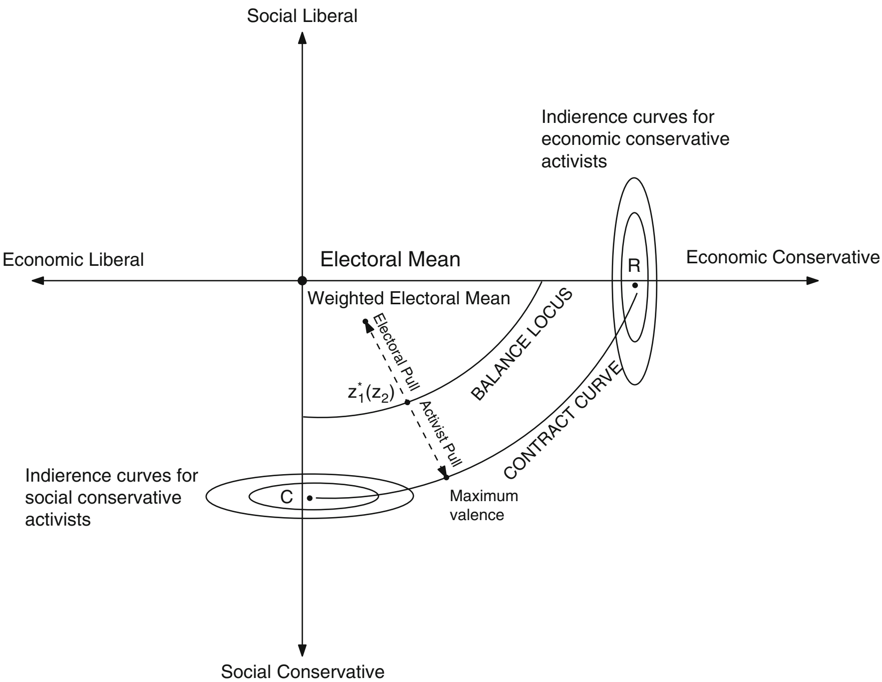


Fig. 7 Optimal Republican position

4.3 Policy Preferences of Candidates

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

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

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

To see this suppose that each candidate were to maximize the function $\mathbb{V}$, given by

$$\mathbb{V}_j(\mathbf{z}) = \delta_j \mu_j(z_j) + \frac{1-\delta_j}{n} \sum\nolimits_i \rho_{ij}(\mathbf{z})$$

where μ_j is no longer an activist function, but a policy determined component of the candidate's utility function, while $\delta_j \in [0,1]$ is the weight given to the policy preference. This model has been proposed by Wittman (1977), Calvert (1985), Duggan and Fey (2005) and Peress (2010).

If we let $z_j^{\mathbb{V}*}$ be the solution with these policy preferences, then the solutions for $\{z_j^{\mathbb{V}*}\}$ will depend on j, and so ρ_{ij} will depend on voters' identity, i.e., will depend on $\{x_i \in X\}_{i\in N}$. Thus ρ_{ij} cannot be written as ρ_j. The first order condition becomes

$$\frac{d\mathbb{V}_j(\mathbf{z})}{dz_j} = \delta_j \frac{d\mu_j}{dz_j}(z_j) + \frac{1-\delta_j}{n}\sum_{i=1}^{n} 2\beta(x_i - z_j)\left[\rho_{ij} - \rho_{ij}^2\right] = 0$$

$$\Rightarrow z_j^{\mathbb{V}*}(1-\delta_j)\sum_{k\in N}\left[\rho_{kj} - \rho_{kj}^2\right] = \frac{n\delta_j}{2\beta}\frac{d\mu_j}{dz_j} + (1-\delta_j)\sum_{i=1}^{n}\left[\rho_{ij} - \rho_{ij}^2\right]x_i.$$

so

$$z_j^{\mathbb{V}*} = \frac{n^*\delta_j}{2(1-\delta_j)\beta}\frac{d\mu_j}{dz_j} + \sum_{i\in N}[\varpi_{ij}]x_i = \frac{n^*\delta_j}{2(1-\delta_j)\beta}\frac{d\mu_j}{dz_j} + z_j^{el},$$

where

$$n^* = \frac{n}{\sum_{k\in N}\left[\rho_{kj} - \rho_{kj}^2\right]}.$$

The new "balance equation" becomes

$$\left[z_j^{el} - z_j^{\mathbb{V}*}\right] + \frac{n^*\delta_j}{2(1-\delta_j)\beta}\frac{d\mu_j}{dz_j}(z_j^*) = 0.$$

Here $\frac{d\mu_j}{dz_j}(z_j^*)$ is a gradient at z_j^* pointing towards the policy preferred position of the candidate.

Suppose now that all agents have contracted with their various activists groups, as given above by the multiple activist model, as above. Suppose further that the activists have provided resources which have been deployed to influence voters. If we now estimate the spatial sociodemographic model with traits, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, at the time of the election, then the effect of these resources will be incorporated in the parameters of the model. Simulation of this model will give $\mathbf{z}^{el}$. Suppose further that the agent is *committed* to the contract with the activists, so that the agent's equilibrium position, $z_j^{\mathbb{V}*}$, is that which is obtained from the model where the agent adopts a policy position induced from this contract. Comparing the above equation with the multiple activist model, we can make the identification

$$\left[z_j^{\mathbb{V}*} - z_j^{el}\right] = \frac{n^*\delta_j}{2(1-\delta_j)\beta}\frac{d\mu_j}{dz_j}(z_j^{\mathbb{V}*}) \equiv \frac{n^*\delta_j}{2(1-\delta_j)\beta}\sum_{k\in A_j} a_k \frac{dU_k}{dz_j} \text{ for } j \in P.$$

The weighted electoral mean z_j^{el} can be obtained for the model $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$ using simulation. This equation implies that the agent's marginal policy preference can be identified with a combination of the marginal preferences of the party activists. To solve this equation in detail requires first solving the contract game between activists and agents, as outlined in Grossman and Helpman (1994, 1996, 2001). For our purposes it is sufficient to use this reduced form, as we are interested in the difference $\left[z_j^{\mathbb{V}*} - z_j^{el}\right]$ between the estimated policy position, $z_j^{\mathbb{V}*}$ of agent j and its weighted electoral mean, z_j^{el}.

We now present these results as a Theorem.

Activist Theorem 2. *Suppose each candidate, j, is committed to a contract with a family of activists* $\{A_j\}$ with utility functions $\{U_k : k \in A_j\}$. *Let* z_j^{el} *be the estimated equilibrium position according to the spatial sociodemographic model*, $\mathbb{M}(\boldsymbol{\lambda}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \beta)$, at the time of the election. *Then the influence of the activists is given by the set of equations*

$$\left[z_j^{\mathbb{V}*} - z_j^{el}\right] = \frac{n^*\delta_j}{2(1-\delta_j)\beta}\frac{d\mu_j}{dz_j}(z_j^{\mathbb{V}*}) \equiv \frac{n^*\delta_j}{2(1-\delta_j)\beta}\sum_{k\in A_j} a_k \frac{dU_k}{dz_j} \text{ for } j \in P. \qquad \square$$

The advantage of this version of the result is that while the activist resources affect the voter probabilities, these are already included in the estimation of the model and the estimated weighted means $\{z_j^{el}\}$. Thus the effect of activist support is subsumed in the empirical estimates of the exogenous valences and the additional sociodemographics and trait valences. While this does not allow us to solve for the nature of the contracts, it does give an estimate of the nature of the contracts between agents and activists. A further advantage of the model is that it provides a rationale for agents to act as though they had policy preferences, while at the same time choosing policy options that are aimed at maximizing vote share.

4.4 *Extension of the Activist Model: Targeting Voters*

As before we let $\{A_j\}$ be the family of activist supporters for j and now write

$$\mathbf{R}_j(z_j) = \sum_{k\in A_j} R_{jk}\left(U_k(z_j)\right). \tag{24}$$

for the total resources obtained by agent j from the various activist groups in $\{A_j\}$. Here, we again assume the activist utilities are functions of z_j. These resources are denominated in terms of time (multiplied by the wage rate for labor) or money, so we can take the units to be currency, dollars say.

These resources are used to target the individual voters and the voter utility function is now

$$\begin{aligned} u_{ij}(x_i, z_j) &= \lambda_j + \mu_i(m_{ij}) + (\theta_j \cdot \eta_i) + (\alpha_j \cdot \tau_i) - \beta||x_i - z_j||^2 + \varepsilon_j \\ &= u_{ij}^*(x_i, z_j) + \varepsilon_j. \end{aligned}$$

Here $\mu_i(m_{ij})$ is the valence effect of the expenditure of resources, (m_{ij}) on the targeting of voter i, by agent j. We assume that the greater the resources m_{ij} spent on persuading voter i, the greater the implicit valence associated with candidate j, so $\frac{d\mu_i(m_{ij})}{dm_j} > 0$. We may also assume decreasing returns: $\frac{d^2\mu_i(m_{ij})}{dm_j^2} < 0$. Obviously we can partition the voters into different categories, in terms of their sociodemographic valences. Note that different agents may target the same voter or group of voters.

We assume that for each j the budget constraint is satisfied:

$$\mathbf{R}_j(z_j) = \sum_{k\in A_j} R_{jk}\left(U_k(z_j)\right) = \sum_{i\in N} m_{ij} \tag{25}$$

We further assume that j solves the optimization problem that we now construct. Since $\mathbf{R}_j(z_j)$ determines the budget constraint for j, we can write $m_{ij} \equiv m_{ij}(z_j)$, so

$$\mu_i(m_{ij}) \equiv \mu_i(m_{ij}(z_j)) \equiv \mu_{ij}(z_j).$$

We shall also assume that the solution to the optimization problem is smooth, in the sense that $\mu_{ij}(\text{-})$ is a differentiable function of z_j.

Then just as above, the first order condition gives a more general balance condition as follows:

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

So $z_j \sum_{i\in N}\left[\rho_{ij} - \rho_{ij}^2\right] = \sum_{i\in N}\left[\rho_{ij} - \rho_{ij}^2\right]\left\{x_i + \frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)\right\}$.

Hence $z_j^* = \frac{\sum_i\left[\left[\rho_{ij}-\rho_{ij}^2\right]\left[x_i+\frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)\right]\right]}{\sum_{k\in N}\left[\rho_{kj}-\rho_{kj}^2\right]}$ and $z_j^* = \sum_{i=1}^n \bar{\omega}_{ij}(x_i + \gamma_i)$ where $\gamma_i = \frac{1}{2\beta}\frac{d\mu_{ij}}{dz_j}(z_j)$ and $\bar{\omega}_{ij} = \frac{\left[\rho_{ij}-\rho_{ij}^2\right]}{\sum_{k\in N}\left[\rho_{kj}-\rho_{kj}^2\right]}$.

This can be written $\left[z_j^* - z_j^{el}\right] = \sum_{i=1}^n \bar{\omega}_{ij}\gamma_i$ where $z_j^{el} = \sum_{i=1}^n \bar{\omega}_{ij}x_i$.

When $\frac{d\mu_{ij}}{dz_j}(z_j) = \frac{d\mu_j}{dz_j}(z_j)$ this reduces to the previous result.

The difference now is that instead of there being a single *centrifugal marginal activist pull* $\frac{1}{2\beta}\frac{d\mu_j}{dz_j}(z_j)$ *there is an aggregrate activist pull*

$$\sum_{i=1}^n \bar{\omega}_{ij}\gamma_i = \frac{1}{2\beta}\sum_{i=1}^n \frac{\left[\rho_{ij} - \rho_{ij}^2\right]}{\sum_{k\in N}\left[\rho_{kj} - \rho_{kj}^2\right]}\frac{d\mu_{ij}}{dz_j}(z_j)$$

determined by the budget constraint.

Notice that the first order condition depends on the marginal terms, $\frac{d\mu_{ij}}{dz_j}(z_j)$, associated with policy positions, and these will depend on the marginal valence effects $\frac{d\mu_i(m_{ij})}{dm_j}$. Although these valence effects can be assumed to exhibit decreasing returns, these will vary across different classes of voters. The plausibility of existence of Nash equilibria turns on whether the induced second order terms $\frac{d^2\mu_{ij}}{dz_j^2}(z_j)$ have negative eigenvalues. The assumption of negative eigenvalues would give a version of the activist theorem.

Note also that if ρ_{ij} is close to 0 or 1, then $\bar{\varpi}_{ij}$ will be close to 0, so the optimal calculation will be complex, though in principle solvable. It is plausible the candidate should expend resources on pivotal voters for whom ρ_{ij} is close to 1/2.[33]

4.5 *Endogenizing Activist Support*

To sketch an outline of a general model to endogenize activist support, we first let the electoral mapping

$$\boldsymbol{\rho} : X^p \times \mathbb{B}^{n\times p} \to [0,1]^{n\times p}$$

specify the voter probabilities in terms of candidate positions in X^p and the voter distribution, in $\mathbb{B}^{n\times p}$, of resources $\{m_{ij}\}$ to all voters.[34] In principle $\boldsymbol{\rho}$ could be obtained by empirical analysis. We assume that $\boldsymbol{\rho}$ is common knowledge to agents and activists.

We then let

$$\mathbf{V} = V_1 \times .. \times V_p : X^p \times \mathbb{B}^{n\times p} \to [0,1]^p$$

be the *agent profile function*, mapping agent positions and voter distributions to vote shares, as given by the above models. Indeed, for a more general model we could consider multiparty systems where agents form beliefs about coalition behavior, as suggested in Schofield and Sened (2006). In this case the mapping would be

$$\mathbf{V} = V_1 \times .. \times V_p : X^p \times \mathbb{B}^{n\times p} \to \mathbb{R}^p.$$

We let the **k** activists have preferences over the positions taken by the p political agents and agent vote shares, so the *activist profile function* is a map

[33]Stokes (2005) make a somewhat similar inference, discussing clientist models of politics, where m_{ij} is simply a monetary bribe to i. Obviously the marginal benefit to a poor voter is greater than to a wealthy voter, under the usual assumption of decreasing marginal utility for money. Dal Bo (2007) also considers a model of bribery but does not consider income effect *per se*.

[34]It is reasonable to assume that the resource distributions lie in a compact ball, namely $\mathbb{B}^{n\times p}$.

$$\mathbf{U} : X^p \times [0,1]^p \to \mathbb{R}^k.$$

It is reasonable to suppose that both **V** and **U** are differentiable. We now regard the activists as principals who choose offers to make to the political agents. This offer can be regarded as a mapping

$$\mathbf{U}^* : X^p \to \mathbb{B}^p.$$

which specifies the provision of activist resources to each of the agents. Note that we assume that these principals are assumed to make inferences about how the agents will respond to the offer mapping, on the basis of common knowledge about the electoral mapping, $\boldsymbol{\rho}$.

The agents in turn choose a best response to $\mathbf{U}^*$. We seek an equilibrium to a game form which may be written

$$\begin{aligned} \mathbf{U}^* \otimes \mathbf{V} \quad &: \quad X^p \to X^p \times \mathbb{B}^p \to X^p \times \mathbb{B}^{n \times p} \to \mathbb{R}^{\mathrm{k}} \times [0,1]^p. \\ &: \quad (\mathbf{z}) \to (\mathbf{z}, \mathbf{U}^*(\mathbf{z})) \to (\mathbf{z}, \mathbf{m}) \to ((\mathbf{U}(\mathbf{z}, \mathbf{V}((\mathbf{z}, \mathbf{m})), \mathbf{V}((\mathbf{z}, \mathbf{m})) \end{aligned}$$

On the basis of the offer mapping, $\mathbf{U}^*$, the agents choose a position vector **z** and a distribution matrix, $m \in \mathbb{B}^{n \times p}$, such that (**z, m**) is a LNE for the agent profile function, **V**, subject to the constraint that **m** is compatible with the offer $\mathbf{U}^*(\mathbf{z})$.

This is an extremely complex dynamical game, and we do not attempt to explore the full ramifications of this model here.[35] Notice that the game form just presented attempts to endogenize activist choices based on an asumption of differentiability. It is quite possible that, in actual applications of the model, the activist offer mapping may be non differentiable, as activists may switch allegiance from one agent or party to another.[36]

Earlier results of Schofield (1978) and McKelvey (1979) had suggested chaos could be generic in electoral models. The model proposed here does not exhibit chaos. However, the application of a simpler version of this model (Schofield et al. 2003) to the historical development of the U.S. political economy suggests that the equilibria of the model are subject to both "circumferential" and "radial" transformations over time, as activists switch support, and candidates move nearer or further away from the origin.

[35]See Coram (2010) and Duggan and Kalandrakis (2011) for dynamical versions of bargaining models. Acemoglu and Robinson (2008) also develop a model based on Markov Perfect Equilibrium where the elite, the activists, have different preferences for the public good, in X and contribute to the de facto power, or political strength, of the political leader. However, they do not assume competing political leaders.

[36]The "matching" model proposed by Jackson and Watts (2010) embeds the Nash equilibrium within a coalition game, and would allow the principals to switch from one agent coalition to another.

Acknowledgments This chapter is based on work supported by NSF grant 0715929. This version was completed while Schofield was the Glenn Campbell and Rita Ricardo-Campbell National Fellow at the Hoover Institution, Stanford, 2010. A version was presented at the International Conference on the Political Economy of Democratic Institutions, Baiona, Spain, June 14–16, 2010. We thank Jon Eguia for very helpful comments.

Data Appendix

Question Wordings for the American National Election Surveys, for 2000 and 2004.

1. We need a strong government to handle today's complex economic problems [1]; or the free market can handle these problems without government being involved [3].
2. Should federal spending on welfare programs be increased [1], decreased [3], or kept about the same? [2].
3. This country would be better if we worried less about how equal people are. Do you agree strongly [5], agree somewhat [4], neither agree nor disagree [3], disagree somewhat [2], or disagree strongly [1] with this statement?
4. Many minorities overcame prejudice and worked their way up. African Americans should do the same without any special favors. Do you agree strongly [5], agree somewhat [4], neither agree nor disagree [3], disagree somewhat [2], or disagree strongly [1] with this statement?
5. 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 [1], increased a little [2], left the same as it is now [3], decreased a little [4], or decreased a lot [5]?
6. We hear a lot of talk these days about liberals and conservatives. Here is a three-point scale on which the political views that people might hold are arranged from extremely liberal to extremely conservative. liberal = [1], moderate = [2], all conservative = [3].
7. Do you think the federal government should make it more difficult [1] for people to buy a gun than it is now, make it easier [3] for people, or keep the rules the same [2].
8. Which one of the opinions on this page best agrees with your view [on abortion]? By law, abortion should never be permitted [3]; The law should permit abortion only in case of rape, incest, or when the woman's life is in danger [2]; The law should permit abortion for reasons other than rape, incest, or danger to the woman's life, but only after the need for the abortion has been clearly established [2]; By law, a woman should always be able to obtain an abortion as a matter of personal choice [1].
9. Do you think gay or lesbian couples, in other words, homosexual couples, should be legally permitted to adopt children? Yes [1], No [3].

10. This country would have many fewer problems if there were more emphasis on traditional family ties. Do you agree strongly [5], agree somewhat [4], neither agree nor disagree [3], disagree somewhat [2], or disagree strongly [1] with this statement?

References

Acemoglu D, Robinson J (2008) Persistence of power, elites, and institutions. Am Econ Rev 98:267–293

Adams J, Merrill S III (2002) Centrifugal incentives in multi-candidate elections. J Theor Polit 14:275–300

Adams J, Merrill S III (2005) Policy seeking parties in a parliamentary democracy with proportional representation: a valence-uncertainty model. Br J Polit Sci 39:539–558

Aldrich J (1983) A Downsian spatial model with party activists. Am Polit Sci Rev 77:974–990

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

Ansolabehere S, de Figueiredo JM, Snyder J (2003) Why is there so little money in U.S. politics? J Econ Perspect 17:105–130

Aragones E, Palfrey T (2002) Mixed equilibrium in a Downsian model with a favored candidate. J Econ Theory 103:131–161

Aragones E, Palfrey T (2005) Spatial competition between two candidates of different quality: the effects of candidate ideology and private information. In: Austen-Smith D, Duggan J (eds) Social choice and strategic decisions: eassays in honor of Jeffrey Banks. Springer, Berlin

Ashworth S, Bueno de Mesquita E (2009) Elections with platform and valence competition. Games Econ Behav 67:191–216

Banks J, Duggan J (2005) Probabilistic voting in the spatial model of elections: the theory of office motivated candidates. In: Austen-Smith D, Duggan J (eds) Social choice and strategic decisions: essays in honor of Jeffrey Banks. Springer, Berlin

Baron DP (1994) Electoral competition with informed and uninformed voters. Am Polit Sci Rev 88:33–47

Calvert RL (1985) Robustness of the multidimensional voting model: candidates, motivations, uncertainty and convergence. Am J Polit Sci 29:69–85

Clarke HD, Sanders D, Stewart M, Whiteley P (2005) Political choice in Britain. Oxford University Press, Oxford

Clarke HD, Kornberg A, Scotto T (2009a) Making political choices. Toronto University Press, Toronto

Clarke HD, Sanders D, Stewart M, Whiteley P (2009b) Performance politics and the British voter. Cambridge University Press, Cambridge

Coate S (2004) Political competition with campaign contributions and informative advertising. J Eur Econ Assoc 2:772–804

Condorcet N (1785) Essai sur l'application de l'analyse a la probabilite des decisions rendus a la pluralité des voix. Imprimerie Royale, Paris

Coram A (2010) Resource spending over time in an electoral competition. Elect Stud 29:497–508

Dal Bo E (2007) Bribing voters. Am J Polit Sci 51:789–803

Downs A (1957) An economic theory of democracy. Harper and Row, New York

Duggan J (2006) Candidate objectives and electoral equilibrium. In: Weingast BR, Wittman DA (eds) The Oxford handbook of political economy. Oxford University Press, Oxford

Duggan J, Fey M (2005) Electoral competition with policy-motivated candidates. Games Econ Behav 51:490–522

Duggan J, Kalandrakis T (2011) A Newton collocation method for solving dynamic bargaining games. Soc Choice Welfare 36:611–651
Enelow JM, Hinich M (1982) Nonspatial candidate characteristics and electoral competition. J Polit 44:115–131
Enelow JM, Hinich M (1989) The location of American presidential candidates. Math Comput Model 12:417–435
Erikson RS, Romero DW (1990) Candidate equilibrium and the behavioral model of the vote. Am Polit Sci Rev 84:1103–1126
Fiorina M, Abrams SJ (2009) Disconnect: the breakdown of representation in American politics. University of Oklahoma Press, Norman
Fiorina M, Abrams SJ, Pope JC (2005) Culture war?: the myth of a polarized America. Pearson Longman, New York
Groseclose T (2001) A model of candidate location when one candidate has a valence advantage. Am J Polit Sci 45:862–886
Groseclose T, Snyder J (1996) Buying supermajorities. Am Polit Sci Rev 90:303–315
Grossman GM, Helpman E (1994) Protection for Sale. Am Econ Rev 84:833–850
Grossman GM, Helpman E (1996) Electoral Competition and Special Interest Politics. Rev Econ Stud 63:265–286
Grossman GM, Helpman E (2001) Special interest groups. MIT Press, Cambridge
Hacker JS, Pierson P (2006) Off center: the republican revolution and the erosion of American democracy. Yale University Press, New Haven
Hacker JS, Pierson P (2010) Winner-take-all politics: how washington made the rich richer – and turned its back on the middle class. Simon and Schuster, New York
Herrera H, Levine D, Martinelli C (2008) Policy platforms, campaign spending and voter participation. J Public Econ 92:501–513
Jackson M, Watts A (2010) Social games: matching and the play of finitely repeated games. Games Econ Behav 70:170–191
Jessee SA (2009) Spatial voting in the 2004 presidential election. Am Polit Sci Rev 103:59–81
Jessee SA (2010) Partisan bias, political information and spatial voting in the 2008 presidential election. J Polit 72:327–340
Layman GC et al (2010) Activists and conflict extension in American party politics. Am Polit Sci Rev 104:324–346
Madison J ([1787],1999) The Federalist No. 10. In: Rakove J (ed) James Madison: writings. The Library of America, New York
McCarty N, Poole K, Rosenthal H (2006) Polarized America. MIT Press, Cambridge
McKelvey RD (1979) General conditions for global intransitivities in formal voting models. Econometrica 47:1085–1112
McKelvey RD, Patty J (2006) A theory of voting in large elections. Games Econ Behav 57:155–180
McLennan A (1998) Consequences of the Condorcet jury theorem for beneficial information aggregration by rational agents. Am Political Sci Rev 92:413–418
Meirowitz A (2008) Electoral contests, incumbency advantages and campaign finance. J Polit 70:681–699
Miller G, Schofield N (2003) Activists and partisan realignment in the U.S. Am Polit Sci Rev 97:245–260
Miller G, Schofield N (2008) The transformation of the republican and democratic coalitions in the U.S. Perspect Polit 6:433–450
Patty J (2002) Equivalence of objectives in two-candidate elections. Public Choice 112:151–166
Patty J (2007) Generic difference of expected vote share maximization and probability of victory maximization in simple plurality elections with probabilistic voters. Soc Choice Welfare 29:149–173
Patty J, Snyder JM, Ting M (2009) Two's company, three's an equilibrium: strategic voting and multicandidate elections. Quart J Polit Sci 4:251–278

Penn E (2009) A model of far-sighted voting. Am J Polit Sci 53:36–54
Peress M (2010) The spatial model with non-policy factors: a theory of policy motivated candidates. Soc Choice Welfare 34:265–294
Pierson P, Skocpol T (eds) (2007) The transformation of American politics: activist government and the rise of conservatism. Princeton University Press, Princeton
Poole K, Rosenthal H (1984) U.S. presidential elections 1968–1980: a spatial analysis. Am J Polit Sci 28:283–312
Riker WH (1980) Implications from the disequilibrium of majority rule for the study of institutions. Am Polit Sci Rev 74:432–446
Riker WH (1982) Liberalism against populism. Freeman, San Francisco
Schofield N (1978) Instability of simple dynamic games. Rev Econ Stud 45:575–594
Schofield N (2002) Evolution of the constitution. Br J Polit Sci 32:1–20
Schofield N (2006) Equilibria in the spatial stochastic model with party activists. Rev Econ Des 10:183–203
Schofield N (2007) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980
Schofield N, Cataife G (2007) A model of political competition with activists with an application to the elections of 1989 and 1995 in Argentina. Math Soc Sci 53:213–231
Schofield N, Miller G (2007) Elections and activist coalitions in the U.S. Am J Polit Sci 51:518–531
Schofield N, Sened I (2006) Multiparty democracy: elections and legislative politics. Cambridge University Press, Cambridge
Schofield N, Miller G, Martin A (2003) Critical elections and political realignments in the U.S.: 1860–2000. Pol Stud 51:217–240
Schofield N, Claassen C, Ozdemir U, Zakharov AV (2011) Estimating the Effects of activists in two-party and multi-party systems: comparing the United States and Israel. Soc Choice Welfare 36:483–518
Serra G (2010) Polarization of what? A model of elections with endogenous valence. J Polit 72:426–437
Snyder J, Ting M (2008) Interest groups and the electoral control of politicians. J Public Econ 92:482–500
Stokes D (1963) Spatial models and party competition. Am Polit Sci Rev 57:368–377
Stokes D (1992) Valence politics. In: Kavanagh D (ed) Electoral politics. Clarendon, Oxford
Stokes SC (2005) Perverse accountability: a formal model of machine politics with evidence from Argentina. Am Polit Sci Rev 99:315–325
Sundquist JL (1973) Dynamics of the party system: alignment and realignment of political parties in the United States. The Brookings Institution, Washington, DC
Train K (2003) Discrete choice methods for simulation. Cambridge University Press, Cambridge
Wittman D (1977) Candidates with policy preferences: a dynamic model. J Eco Theory 14:180–189
Zakharov AV (2009) A model of candidate location with endogenous valence. Public Choice 138:347–366