JOURNAL OF THEORETICAL POLITICS

Article

Do parties converge to the electoral mean in all political systems?

Journal of Theoretical Politics
2015, Vol. 28(2) 287–329

DOI: 10.1177/0951629815586878
jtp.sagepub.com
SAGE

Maria Gallego
Department of Economics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, Canada, N2L 3C5

Norman Schofield
Center in Political Economy, Washington University, USA

Abstract

Many formal models suggest that parties or candidates should locate at *the electoral mean*. Yet, there is no consistent evidence of such convergence across political systems. Schofield's (2007) Valence Theorem proves that when valence differences across parties are large, there is non-convergence to the mean. Convergence to the mean depends on the value of the convergence coefficient, *c*. When *c* is high there is significant centrifugal tendency acting on the parties and when *c* is low there is a significant centripetal tendency acting on the parties.

In this paper we apply the stochastic valence model of elections in various countries under different political regimes and use the convergence coefficient of these elections to classify political systems. Our results show that the convergence coefficient varies across elections in a country, across countries using the same political system and across political regimes. For countries using proportional representation, namely Israel, Turkey and Poland, the centrifugal tendency is very high and parties move away from the mean. In the majoritarian polities of the United States and the UK, parties are located at the mean, as the centrifugal tendency is very low. In anocracies, the autocrat imposes limitations on how far from the origin the opposition parties can move but the equilibrium is fragile.

Keywords

Convergence coefficient; local Nash equilibrium; stochastic vote model; the heart; valence

Corresponding author:
Norman Schofield, Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA.
Email: schofield.norman@gmail.com

1. Introduction

The political economy literature highlights that institutions matter. Understanding how institutions shape agents decisions has shown that agents make different decisions under different political institutions. That is, agents respond to the incentives created by the institution under which they operate.

In this article we use the stochastic electoral model in Schofield (2007) to study elections under different political systems. The idea is to use Schofield's model as a unifying framework to study elections in different countries operating under different political systems. We apply the formal model to study parties' positions in anticipation of the electoral outcome under different political regimes. Schofield's model examines whether parties locate close to or far from the electoral mean.[1] In this model, parties respond to voters' preferences after taking into account the anticipated electoral outcome and the positions of other parties. Voters' decisions depend on parties' positions and valence. Here a party's valence is the voters'*overall common* evaluation of the ability of a party leader to provide good governance.

The model in Schofield (2006 and 2007) is used to examine whether parties converge to the electoral mean in several elections in various countries under different political systems and uses convergence, or the lack thereof, to classify political systems. Schofield proved the existence of the 'convergence coefficient' whose value gives a summary measure of the centrifugal or centripetal forces acting on the parties. This coefficient depends on the *competence valences* of the party leaders, the weight voters give to policy differences with parties and the variance of the electoral distribution. The effect these parameters have on the convergence coefficient depend on the political regime in which the election takes place. The convergence coefficient is dimensionless and can be used to compare parties' positions across elections and political systems.

The convergence coefficient determines whether party 1, with the lowest valence, has an incentive to stay or move away from the electoral mean to increase its vote share when all parties locate at the electoral mean. If party 1 has no incentive to move from the mean, the *Valence Theorem* in Schofield (2007), presented in Section 2,[2] proves that no other party will want to move either. In this case, the electoral mean is a 'local Nash equilibrium' (LNE)[3] as no party wants to make a *unilateral deviation* to increase its vote share. The theorem shows that when valence differences are small, the *centripetal* political forces induce parties to locate at the mean. The theorem also proves that when valence asymmetries between parties are large, the election is one where small, or low valence, parties have an incentive to move away from the electoral mean to increase their vote share. In this case, the high value of the convergence coefficient is an indication of the *centrifugal* tendency exerted on parties, pulling them away from the electoral mean. The value of the convergence coefficient is therefore a simple and intuitive way to summarize whether parties are located close to, or far from, the electoral mean. In this paper we show that there is a strong connection between the values of the convergence coefficient and the nature of the political system under which parties operate.

The Valence Theorem shows that convergence, or lack thereof, to the electoral mean is determined by the convergence coefficient. Specifically, the theorem shows that if in a two-dimensional space $c < 1$, then the sufficient condition for convergence to the mean is met and the LNE is one where all parties locate at the electoral mean. On the other hand, if $c \geqslant w$, where w is the dimension of the policy space, the LNE, if it exists, will be one where at least one party has an incentive to diverge from the mean to increase its vote share. Thus, the necessary condition for convergence to the mean is that $c < w$. The convergence coefficient is then a summary measure of the centrifugal or centripetal forces acting on the parties.

Using the convergence results from Schofield (2007), we study each party's best response to the electoral situation they face in various elections in different countries. The convergence coefficient determines whether parties converge to the electoral mean in a particular election. The cross-election, cross-country values of the convergence coefficients illustrate that while the convergence coefficient in countries operating under proportional representation is high, that of countries with plurality systems or in anocracies[4] is low. This suggests that we can use Schofield's Valence Theorem and its associated convergence result to classify electoral systems.

The results indicate that in plurality systems, like the United States and the UK, the centrifugal tendency is low (significantly less that one) thus meeting the necessary conditions for convergence to the mean. However, for Israel, Poland and Turkey, the centrifugal tendency is very high. In these proportional representation systems with highly fragmented polities the convergence coefficients are significantly greater than two, failing to meet the necessary condition for convergence to the mean.

In the anocracies of Georgia, Russia and Azerbaijan, the convergence coefficient fails the necessary condition for convergence. The analyses of Georgia and Azerbaijan show that not all parties converge to the mean. In Russia, however, parties found it difficult to diverge from the mean. Convergence in anocracies may not generate a stable equilibrium as changes in the valence of the autocrat and/or opposition parties may cause parties to diverge from the mean and may even lead to popular uprisings that bring about changes in the ruling parties as happened in Georgia in previous elections or in the Arab revolutions.

Political systems are also classified using the *effective vote number* and the *effective seat number*.[5] We explore how these two fragmentation measures relate to our convergence measure. The effective vote or seat numbers give an indication of the difficulty inherent in inter-party negotiation over government. These two measures do not, however, address the fundamental aspect of democracy, namely, electoral preferences for policy. Since convergence involves both preferences, in terms of the electoral covariance matrix and the effect of the electoral system, we argue that the Valence Theorem and the associated convergence coefficient give a more comprehensive classification of polities and political systems, as they are is derived from the fundamental characteristics of the electorate. That is, while, we can use the effective vote and seat number to identify which polities are fragmented, the convergence result help us to understand why parties are located close to or far from the electoral mean and how, under some circumstances, these considerations lead to political fragmentation.

The next section presents the stochastic formal model of elections in Schofield (2007) and the implications it has for convergence to the mean. Section 3.1 applies the model to the elections in two plurality polities: the United States and the UK. In Section 3.2 we apply the model to polities using proportional representation, namely Israel, Turkey and Poland. Section 3.3 considers the convergence coefficients for three 'anocracies': Azerbaijan, Georgia and Russia. Comparisons between different fragmentation measures and the convergence coefficient are examined in Section 4. Section 5 concludes the paper.

2. The spatial voting model with valence

We model elections following Stokes (1963, 1992) who emphasizes the valence of political candidates. As Sanders et al. (2011) comment, valence theory extends the spatial or Downsian model of elections by considering not just the policy positions of parties but also

> the parties' rival attractions in terms of their perceived ability to handle the most serious problems that face the country… [Thus] voters maximize their utilities by choosing the party that they think is best able to deliver policy success.

Schofield and Sened (2006) have also argued that

> Valence relates to voters' judgments about positively or negatively evaluated conditions which they associate with particular parties or candidates. These judgements could refer to party leaders' competence, integrity, moral stance or 'charisma' over issues such as the ability to deal with the economy and politics.

A considerable number of electoral models assume valence plays an important role in the relationship between party positioning and the votes that parties receive (Ansolabare and Snyder, 2000; Aragones and Palfrey, 2002, 2005; Groseclose, 2001; Peress, 2010; Schofield, 2003, 2004). The empirical multinomial logit (MNL) models have also shown the importance of electoral judgements in analysis of elections in the United States, Canada and the UK (Clarke et al., 2005a,b, 2006, 2009a,b, 2011; Schofield, 2005; Schofield et al., 2011a,b,c,e; Scotto et al., 2010). In these empirical electoral models, the probability that a voter chooses a party depends on the utility the voter derives from each party. The voter's utility is partly 'Downsian' (depends on the distance between the party's position and voter's preferred or ideal position) but also depends on the party's valence (with each party's valence subject to a 'stochastic error'). In this paper, we follow this methodology.

The *pure* 'Downsian' spatial model of voting tends to predict that parties converge to the center of the electoral distribution (Enelow and Hinich, 1982, 1984, 1989); however, when valence is included, the prediction is very different. To see this suppose there are two parties, A and B, and both choose the same position at the electoral center, but A has much higher valence than B. This higher valence captures voters' bias towards A indicating that more voters will choose A over B.

The question for B is whether it can gain votes by moving away from the center. The optimal position of both A and B will depend on the various estimated parameters of the model. To answer this question we now present the details of the spatial model.

2.1. The theoretical model

To study the optimal party positions to the anticipated electoral outcome we use the Downsian vote model with valence in Schofield (2007). Let the set of parties be denoted by $P = 1, \ldots, p$. The positions of the p parties[6] in $X \subseteq \mathbb{R}^w$, where w is the dimension of the policy space, is given by the vector

$$\mathbf{z} = (z_1, \ldots, z_j, \ldots, z_p) \in X^p$$

Denote voter i's ideal policy by $x_i \in X$ and i's utility by $u_i(x_i, z) = (u_{i1}(x_i, z_1), \ldots, u_{ip}(x_i, z_p))$ where

$$u_{ij}(x_i, z_j) = \lambda_j - \beta\left\|x_i - z_j\right\|^2 + \epsilon_j = u^*_{ij}(x_i, z_j) + \epsilon_j \tag{1}$$

Here $u^*_{ij}(x_i, z_j)$ is the observable component of the utility voter i derives from party j. The competence valence of candidate j, λ_j, is the same across voters and provides an estimate of the voters' belief of j's 'quality' ability to govern. The competence valence vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ is such that $\lambda_p \geqslant \lambda_{p-1} \geqslant \cdots \geqslant \lambda_2 \geqslant \lambda_1$, so that party one has the lowest valence. The term $||x_i - z_j||$ is simply the Euclidean distance between the position of voter i, x_i, and the position of candidate j, z_j. The coefficient β is the weight given to this policy difference. The error vector $\epsilon = (\epsilon_1, \ldots, \epsilon_j, \ldots, \epsilon_p)$ has a Type I extreme value distribution where the variance of ϵ_j is fixed at $\pi^2/6$. Note that β has dimension $1/L^2$, where L is whatever unit of measurement is used in X.

Since voter behavior is modeled by a probability vector, the probability that voter i chooses party j when parties position themselves at $\mathbf{z}$ is

$$\begin{aligned}\rho_{ij}(\mathbf{z}) &= \Pr[u_{ij}(x_i, z_j) > u_{il}(x_i, z_l), \text{ for all } l \neq j] \\ &= \Pr[\epsilon_l - \epsilon_j < u^*_{ij}(x_i, z_j) - u^*_{il}(x_i, z_j), \text{ for all } l \neq j]\end{aligned}$$

Here Pr stands for the probability operator generated by the distribution assumption on ϵ. Thus, the probability that i votes for j is given by the probability that $u_{ij}(x_i, z_j) > u_{il}(x_i, z_l)$, for all $l \neq j \in P$, i.e. that i gets a higher utility from j than from any other party.

Train (2003) showed that when the error vector ϵ has a Type I extreme value distribution, the probability $\rho_{ij}(z)$ has a MNL specification and can be estimated. Thus, for each voter i and party j, the probability that i chooses j at $\mathbf{z}$ is given by

$$\rho_{ij}(\mathbf{z}) = \frac{\exp[u^*_{ij}(x_i, z_j)]}{\sum_{k=1}^{p} \exp u^*_{ik}(x_i, z_k)} \tag{2}$$

Voters decisions are stochastic in this framework.[7] Even though parties cannot perfectly anticipate how voters will vote, they can estimate the *expected* vote share of party j as the average of these probabilities

$$V_j(\mathbf{z}) = \frac{1}{n}\sum_{i\in N} \rho_{ij}(\mathbf{z}) \tag{3}$$

Each party's objective is to find the position that maximizes its expected vote share, as desired by 'Downsian' opportunists. On the other hand, the party may desire to adopt a position that is preferred by the base of the party supporters, namely the 'guardians' of the party, as suggested by Roemer (2001 and 2011).

We assume that parties can estimate how their vote shares would change if they *marginally* move their policy position. The LNE is the vector $\mathbf{z}$ of party positions such that that no party may shift position by a small amount to increase its vote share. More formally, a LNE is a vector $\mathbf{z} = (z_1, \ldots, z_j, \ldots, z_p)$ such that each vote share $V_j(\mathbf{z})$ is weakly locally maximized at the position z_j. To avoid problems with zero eigenvalues we also define a (*strict* local Nash Equilibrium) SLNE to be a vector that *strictly* locally maximizes $V_j(\mathbf{z})$.

Using the estimated MNL coefficients we simulate these models. We then relate any vector of party positions, $\mathbf{z}$, to a vector of vote share functions $V(\mathbf{z}) = (V_1(\mathbf{z}), \ldots V_p(\mathbf{z}))$, predicted by the particular model with p parties. Moreover, we can examine whether in equilibrium parties position themselves at the *electoral mean*.[8]

Given the vector of policy position $\mathbf{z}$, and since the probability that voter i votes for party j is given by equation (2), the impact of a *marginal* change in j's position on the probability that i votes for j is then

$$\left.\frac{d\rho_{ij}(\mathbf{z})}{dz_j}\right|_{\mathbf{z}_{-j}} = 2\beta\rho_{ij}(1-\rho_{ij})(x_i - z_j) \tag{4}$$

where $\mathbf{z}_{-j}$ indicates that the positions of all parties but j are held constant. The effect that the change in position of j has on the probability that i votes for j depends on the weight given to the policy differences of parties, β, on how likely it is that i votes for j, ρ_{ij}, or for any other party, $(1-\rho_{ij})$, and on how far apart the ideal policy of i is from the ideal policy of j, $(x_i - z_j)$.

From equation (3), party j adjusts its position to maximize its expected vote share, i.e. the first-order condition (FOC) of j is

$$\left.\frac{dV_j(\mathbf{z})}{dz_j}\right|_{\mathbf{z}_{-j}} = \frac{1}{n}\sum_{i\in N}\frac{d\rho_{ij}}{dz_j} = \frac{1}{n}\sum_{i\in N} 2\beta\rho_{ij}(1-\rho_{ij})(x_i - z_j) = 0 \tag{5}$$

where the third term follows after substituting in equation (4). The FOC for party j in equation (5) is satisfied when

$$\sum_{i\in N} \rho_{ij}(1-\rho_{ij})(x_i - z_j) = 0$$

so that the *candidate* for the vote maximizing policy of party j is

$$z_j^C = \sum\nolimits_{i \in N} \alpha_{ij} x_i \quad \text{where} \quad \alpha_{ij} \equiv \frac{\rho_{ij}(1-\rho_{ij})}{\sum_{i \in N} \rho_{ij}(1-\rho_{ij})} \tag{6}$$

where α_{ij} represents the weight that party j gives to voter i when choosing its candidate vote maximizing policy. We call these values critical as we do not know whether the vote share function is at a maximum, minimum or a saddle point. This weight depends on how likely i is to vote for j, ρ_{ij}, and for any other party, $(1-\rho_{ij})$, relative to all voters.[9] Note that α_{ij} may be non-monotonic in ρ_{ij}. To see this, exclude voter i from the denominator of α_{ij}. When $\sum_{a \in N-i} \rho_{aj}(1-\rho_{aj}) < 2/3$ then $\alpha_{ij}(\rho_{ij}=0) < \alpha_{ij}(\rho_{ij}=1) < \alpha_{ij}(\rho_{ij}=1/2)$. Thus, if i will definitely vote for j, i receives a lower weight in the candidate position of j than a voter who will only vote for j with probability 1/2 (an 'undecided' voter). Party j caters then to 'undecided' voters by giving them a higher weight in j's policy weight and thus a higher weight in its position. When $\sum_{a \in N-i} \rho_{aj}(1-\rho_{aj}) > 2/3$, then α_{ij} increases in ρ_{ij}. If j expects a large enough vote share (excluding voter i), it gives a core supporter (a voter who votes for sure for j) a higher weight in its policy position than it gives other voters as there is no risk of doing so. The weights α_{ij} are endogenously determined in the model.

Note that since the utility of voter i depends on how far i is from party j, the probability that i votes for j given in (2) and the expected vote share of the party given in (3) are influenced by the voters and parties positions in the policy space. That is, in the empirical models estimated below, the positions of voters and parties in the policy space, together with the valence estimates, influence voters electoral choices.

Recall that we are interested in finding whether parties converge to or diverge from the electoral mean. Suppose that *all* parties locate at the same position, $z_k = z$ for all $k \in P$. Thus, from (1) we see that

$$\left[u_{ik}^*(x_i, z) - u_{ij}^*(x_i, z)\right] = (\lambda_k - \lambda_j)$$

so the probability that i votes for j in (2) is given by

$$\rho_{ij}(\mathbf{z}) = \frac{1}{\sum_{k=1}^{p} \exp[u_{ik}^*(x_i, z_k) - u_{ij}^*(x_i, z_j)]} = \left[\sum\nolimits_{k=1}^{p} \exp(\lambda_k - \lambda_j)\right]^{-1} \tag{7}$$

Clearly, in this case, $\rho_{ij}(\mathbf{z}) = \rho_j(\mathbf{z})$ is independent of voter i's ideal point. Thus, from equation (6), the weight given by j to each voter is also independent of the position of voter i and given by

$$\alpha_j \equiv \frac{\rho_j(1-\rho_j)}{\sum_{i \in N} \rho_j(1-\rho_j)} = \frac{1}{n}$$

Therefore, j gives each voter equal weight in its policy position. In this case, from equation (6), the candidate position of j is

$$z_j^C = \frac{1}{n} \sum\nolimits_{i \in N} x_i$$

i.e. the candidate position of j is located at the electoral mean which we have placed at the electoral origin. Let $\mathbf{z}_0 = (0,\ldots,0)$ be the vector of party positions when all parties are at the electoral mean.

As indicated by equation (7), when parties locate at the mean, only the valence differences between parties matter in the voters' choices. The probability that a *generic* voter votes for party 1 (the party with the lowest valence) is

$$\rho_1 \equiv \rho_1(\mathbf{z}_0) = \left[\sum\nolimits_{k=1}^{p} \exp(\lambda_k - \lambda_1)\right]^{-1} \tag{8}$$

Using this model, Schofield (2007) proved a *Valence Theorem* determining whether the vote maximizing parties located at the mean depend on the value of the *convergence coefficient*

$$c \equiv c(\lambda, \beta, \sigma^2) \equiv 2\beta[1 - 2\rho_1]\sigma^2 \tag{9}$$

which depends on β, the weight given to policy differences, on ρ_1, the probability that a generic voter votes for the lowest valence party at the vector $\mathbf{z}_0$ and on σ^2, the *electoral variance* given by

$$\sigma^2 \equiv \mathrm{trace}(\nabla) \tag{10}$$

where ∇ is the symmetric $w \times w$ *electoral covariance matrix*.[10]

The convergence coefficient increases in β and σ^2 (and on its product $\beta\sigma^2$) and decreases in ρ_1. As equation (8) indicates, ρ_1 decreases if the valence differences between party 1 and the other parties increases, i.e. when the difference between λ_1 and $\{\lambda_2,\ldots,\lambda_p\}$ increases.

The Valence Theorem allows us to characterize polities according to the value of their convergence coefficients. The theorem states that when the *sufficient* condition for convergence to the electoral mean in a two dimensional policy space is met, i.e. when $c < 1$, the LNE is one where all parties adopt the same position at the mean of the electoral distribution. A *necessary* condition for convergence to the electoral mean is that $c < w$, where w is the dimension of the policy space. If $c \geqslant w$, then there *may* exist a non-convergent LNE. Note that in this case there may indeed be no LNE. However, there will exist a mixed strategy Nash equilibrium. In either of these two cases we expect that at least one party will *diverge* from the electoral mean.

Note that c is dimensionless, because $\beta\sigma^2$ has no dimension. In a sense $\beta\sigma^2$ is a measure of the polarization of the preferences of the electorate. Moreover, ρ_1 in equation (8) is a function of the distribution of beliefs about the competence of party leaders, which is a function of the difference $(\lambda_k - \lambda_1)$.

When some parties have a low valence, so the probability that a generic voter votes for party 1 (the lowest valence when all parties locate at the origin), ρ_1 in equation (8) will tend to be small because the valence differences between party 1 and the other parties is *sufficiently large*. Thus, vote maximizing parties will *not all* converge to the electoral mean. In this case c will be close to $2\beta\sigma^2$. If $2\beta\sigma^2$ is large because, for example, the electoral variance is large, then c will be large, suggesting $c > w$. In this case, the low valence party has an incentive to move away from the

origin to increase its vote share. This implies the existence of a *centrifugal* force pulling some parties away from the origin.

Thus, for $\beta\sigma^2$ sufficiently large so that $c \geqslant w$, we expect parties to diverge from the electoral center. Indeed, we expect the parties that exhibit the lowest valence to move further away from the electoral center, implying that the centrifugal force on parties will be significant. Thus, in fragmented polities with a polarized electorate, the nature of the equilibrium tends to maintain this *centrifugal* characteristic.

On the contrary, in a polity where there are no very small or low valence parties, then ρ_1 will tend to 1/2 and so c will be small. In a polity with small $\beta\sigma^2$ and with low valence differences, so that $c < 1$, we expect all parties to converge to the center. In this case, we expect this *centripetal* tendency to be maintained.

The *Valence Theorem* asserts that if $c(\lambda, \beta, \sigma^2) > w$ then the party with the lowest valence has an incentive to move away from the electoral mean to increase its vote share. In this case other low valence parties may also find it advantageous to vacate the center. The value of the convergence coefficient allows us to identify which parties have an incentive to move away from the electoral mean. The convergence coefficient gives a simple and intuitive way to identify whether a low valence party should vacate the electoral mean.

In the next section, we present the convergence coefficient of various elections in different countries.

3. MNL models of the elections of various countries

The framework of the formal MNL spatial model presented in Section 2 is used as a unifying methodology to determine whether parties converge to or diverge from the electoral mean in different elections in the same country, and a comparison is then made of the convergence across countries and across political regimes. In the empirical MNL estimations of different elections carried out below, the coefficients of the chosen baseline party are normalized at zero with those of all other parties measured relative to those of the baseline party. Following the formal model, the MNL coefficients are then used to estimate the convergence coefficient and the Valence Theorem used to determine if parties converge to the electoral mean in each election.[11]

Convergence is studied for countries operating under three political regimes (plurality rule, proportional representation and anocracy). The focus of this paper is not on individual elections as the detailed analysis of each election was carried out elsewhere, but rather to look at the 'bigger picture' by comparing convergence across a series of elections in the same country for countries operating under the same political regime and then to examine differences in convergence across political regimes.

3.1. Convergence in plurality systems

We first examine convergence under plurality rule for elections in the United States and the UK and show that in these countries plurality is associated with relatively low convergence coefficients. [12]

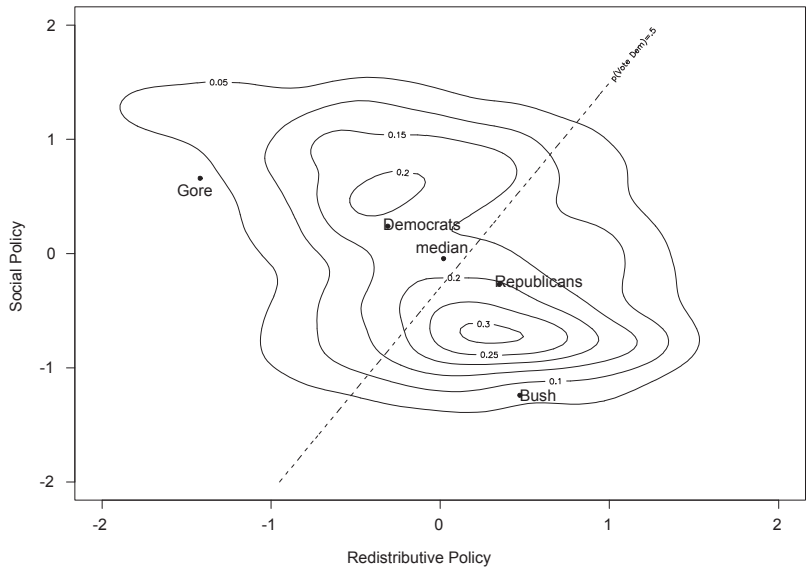

Figure 1. Distribution of voter ideal points and candidate positions in the 2000 US election.

3.1.1. The 2000, 2004 and 2008 elections in the United States. George W Bush won the 2000 election against Albert A Gore and won his re-election bid against John Kerry in 2004. In 2008 voters elected Barack Obama, the first black President, who ran against John McCain. The 2008 election re-drew the US electoral map. The election did not generate the usual division with the Democrats winning in the northeast and the far west and Republicans winning the Midwest and South. Rather several states that previously consistently voted Republican voted Democrat instead. Indiana and Virginia, which had not voted Democrat since 1964, cast majorities for Obama as did North Carolina and Colorado, who had not voted Democrat since 1976 and 1992, respectively.

Schofield et al. (2011b,c,d) construct stochastic models of the 2000, 2004 and 2008 US presidential elections using survey data taken from the American National Election Surveys (ANES). Using the factor analysis done on ten ANES questions[13] we deduced that voters preferences can be represented along an economic ($E = x$-axis) and a social ($S = y$-axis) dimension for all three elections. Voters located on the left of the economic axis are pro-redistribution. The social axis is determined by attitudes to abortion and gays. Greater values along this axis are interpreted to mean more support for certain civil rights. Voters' positions along the two dimensions are estimated using the factor loadings generated by the factor analysis. Figures 1–3 give a smoothing of the estimated voter distribution of the 2000, 2004 and 2008 elections and illustrate the 2008 electoral shift that occurred relative to the previous two elections.

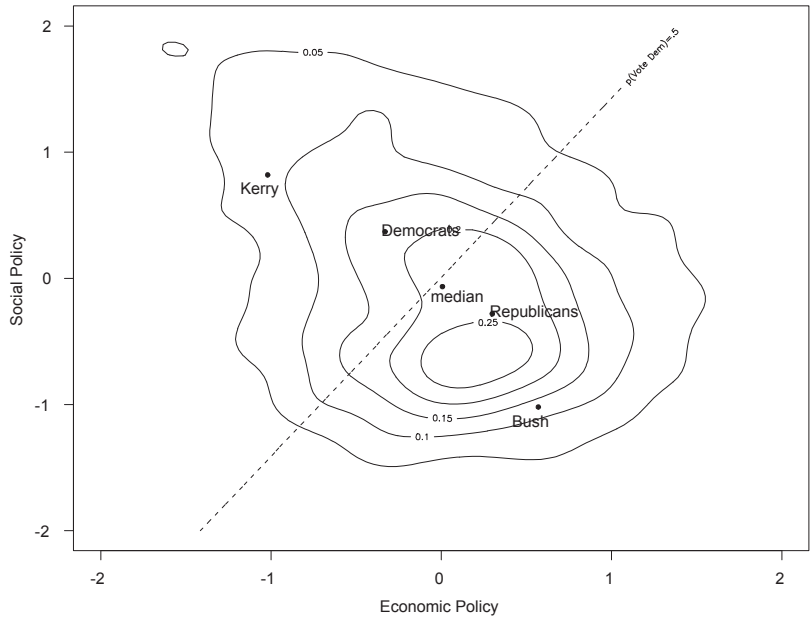

Figure 2. Distribution of voter ideal points and candidate positions in the 2004 US election.

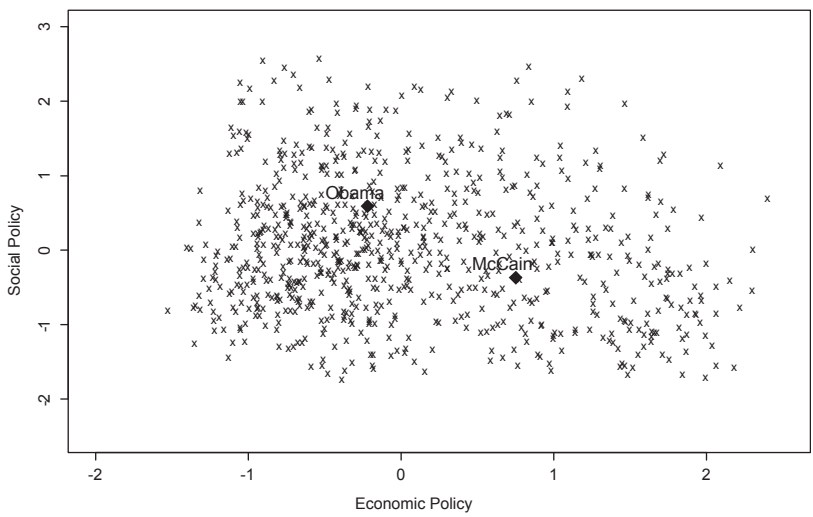

Figure 3. Distribution of voter ideal points and candidate positions in the 2008 US election.

The sizable electoral shift in 2008 is also observed in the *electoral covariance matrices* associated with Figures 1–3. In particular, while the 2000 electoral covariance matrix

$$\nabla_{2000}^{\mathrm{US}} = \begin{bmatrix} \sigma_E^2 = 0.58 & \sigma_{ES} = -0.20 \\ \sigma_{ES} = -0.20 & \sigma_S^2 = 0.59 \end{bmatrix} \tag{11}$$

is quite similar to that in 2004

$$\nabla_{2004}^{\mathrm{US}} = \begin{bmatrix} \sigma_E^2 = 0.58 & \sigma_{ES} = -0.177 \\ \sigma_{ES} = -0.177 & \sigma_S^2 = 0.59 \end{bmatrix} \tag{12}$$

they stand in sharp contrast with that in 2008 election

$$\nabla_{2008}^{\mathrm{US}} = \begin{bmatrix} \sigma_E^2 = 0.80 & \sigma_{ES} = -0.127 \\ \sigma_{ES} = -0.127 & \sigma_S^2 = 0.83 \end{bmatrix} \tag{13}$$

The *trace* or '*variance*' of these electoral covariance matrix show that while the variances of 2000 and 2004 are indistinguishable from each other, they are substantially below that of 2008

$$\sigma_{\mathrm{US2000}}^2 \equiv \mathrm{trace}(\nabla_{\mathrm{US}}^{2000}) = \sigma_E^2 + \sigma_S^2 = 1.17, \qquad \sigma_{\mathrm{US2004}}^2 = \mathrm{trace}(\nabla_{\mathrm{US}}^{2004}) = 1.17,$$
$$\sigma_{\mathrm{US2008}}^2 = \mathrm{trace}(\nabla_{2008}^{\mathrm{US}}) = 1.63$$

In particular, the 2008 electoral 'variance' $\sigma_{\mathrm{US2008}}^2 = \mathrm{trace}(\nabla_{2008}^{US}) = 1.63$ increased relative to 2000 and 2004, and the covariance between dimensions in 2008, $\sigma_{ES} = -0.127$, decreased relative to 2000 and 2004.

Using the spatial model presented in Section 2, we estimate the MNL model of the 2000 election. The coefficients for the US 2000 shown in Table 1 are

$$\lambda_{\mathrm{rep}}^{\mathrm{US2000}} = -0.43, \quad \lambda_{\mathrm{dem}}^{\mathrm{US2000}} \equiv 0.0, \quad \beta_{2000}^{\mathrm{US}} = 0.82 \tag{14}$$

Bush's 2000 competence valence, $\lambda_{\mathrm{rep}}^{\mathrm{US2000}} = -0.43$, measures the common perception that voters in the sample have on Bush's ability to govern and represents the non-policy component in the voter's utility function in equation (1). The results given in Table 1 suggest that in the 2000 election Bush had a statistically significant *lower* valence than Gore, the Democratic (baseline) candidate, an indication that voters regarded Bush as less able to govern than Gore, once policy differences are taken into account.

The similarity of the 2000 and 2004 elections led to similar MNL estimates for the 2004 election. Compare

$$\lambda_{\mathrm{rep}}^{\mathrm{US2004}} = -0.43, \quad \lambda_{\mathrm{dem}}^{\mathrm{US2004}} \equiv 0.0, \quad \beta_{2004}^{\mathrm{US}} = 0.95 \tag{15}$$

with those given in (14) (Table 1). Thus, in 2004, voters continued to believe that Bush was significantly less able to govern (had a significantly lower valence,

Table 1. MNL Spatial model for countries with plurality systems.

	United States[a]				United Kingdom[b]		
	Party	2000	2004	2008	Party	2005	2010
Var		Estimate[c] |t–value|	Estimate |t–value|	Estimate |t–value|		Estimate |t–value|	Estimate |t–value|
β		0.82*** (14.9)	0.95*** (14.21)	0.85*** (14.16)		0.15*** (12.56)	0.86*** (38.45)
Valence	λ_{rep}	–0.43*** (5.05)	–0.43*** (5.05)	–0.84*** (7.64)	λ_{Lab}	0.52*** (6.84)	–0.04 (1.31)
					λ_{Con}	0.27*** (3.22)	0.17*** (4.50)
Base party		Dem[a]	Dem[a]	Rep[a]		Lib[b]	Lib[b]
n		1238	935	788		1149	6218
Log-Likelihood *LL*		–708	–501	–298		–1136	–5490

[a]US: Rep = Republican, Dem = Democrats.
[b]UK: Lab = Labour, Con = Conservatives, Lib = Liberal Democrats.
[c]***Probability < 0.001.

$\lambda_{\text{rep}}^{\text{US2004}} = -0.43$) than his opponent, Kerry ($\lambda_{\text{dem}}^{\text{US2004}} \equiv 0.0$, the baseline candidate), once policy differences are taken into account.

To examine whether candidates located themselves at the electoral mean we use the formal model of Section 2 and the MNL coefficients given in (14) to estimate the convergence coefficient for the 2000 US election.

To do so we need to estimate the probability that a voter votes for Bush, the low valence candidate. Using (8) and the coefficients in (14) the probability that in 2000 a voter chooses the low valence Republican (rep) candidate, when both Bush and Gore locate at the origin, $\mathbf{z}_0$, is

$$\rho_{\text{rep}}^{\text{US2000}} = \left[\sum\nolimits_{k=1}^{2} \exp(\lambda_k^{\text{US2000}} - \lambda_{\text{rep}}^{\text{US2000}})\right]^{-1} = [1 + \exp(0.43)]^{-1} = 0.40 \qquad (16)$$

Since the central estimates of $\lambda = (\lambda_1, \ldots, \lambda_p)$ given by the MNL regressions depend on the sample of voters surveyed then so does $\rho_{\text{rep}}^{\text{US2000}}$. Thus, to make inferences from empirical models Table 2 gives the 95% confidence bounds of $\rho_{\text{rep}}^{\text{US2000}}$.

Given the similarity of Bush's valences in the 2000 and 2004 elections, its not surprising to find that the probability that a US voter chooses Bush, the low valence candidate, in 2004 when both Bush and Kerry locate at the electoral origin, $\mathbf{z}_0$, given in Table 2

$$\rho_{\text{rep}}^{\text{US2004}} = 0.40 \qquad (17)$$

is similar to that in 2000 in equation (16).

Table 2. The convergence coefficient in plurality systems.

	United States[a]			United Kingdom[b]	
	2000	2004	2008	2005	2010
Weight of policy differences (β)					
Estimate β (Conf. Int.)[c]	0.82 (0.71, 0.93)	0.95 (0.82, 1.08)	0.85 (0.73, 0.97)	0.15 (0.13, 0.17)	0.86 (0.81, 0.90)
Electoral variance (trace$\nabla = \sigma^2$)					
σ^2	1.17	1.17	1.63	5.607	1.462
Probability of voting for lowest valence party (party 1, $\rho_1 = \left[\sum_{k=1}^{p} \exp(\lambda_k - \lambda_1)\right]^{-1}$)					
	Dem	Dem	Rep	LibDem	Labour
Estimate ρ_1 (Conf. Int.)	$\rho_{\text{Dem}} = 0.4$ (0.35, 0.44)	$\rho_{\text{Dem}} = 0.4$ (0.35, 0.44)	$\rho_{\text{rep}} = 0.3$ (0.26, 0.35)	$\rho_{\text{Lib}} = 0.25$ (0.18, 0.32)	$\rho_{\text{Lab}} = 0.32$ (0.29, 0.32)
Convergence coefficient ($c \equiv c(\lambda, \beta, \sigma^2) = 2\beta[1 - 2\rho_1]\sigma^2$)					
Estimate c (Conf. Int.)	0.38 (0.2, 0.65)	0.45 (0.23, 0.76)	1.1 (0.71, 1.52)	0.84 (0.51, 1.25)	0.98 (0.86, 1.10)

[a]US: Dem = Democrats; Rep = Republican.
[b]UK: LibDem = Liberal Democrats
[c]Conf. Int. = confidence intervals.

Using the central values of $\rho_{\text{rep}}^{\text{US2000}}$, β_{2000}^{US} and σ_{US2000}^2, we estimate the value of the *convergence coefficient* for the 2000 election, c_{US}^{2000}, using (9) as

$$c_{\text{US}}^{2000} \equiv 2\beta_{2000}^{\text{US}}(1 - 2\rho_{\text{rep}}^{\text{US2000}})\sigma_{\text{US2000}}^2 = 0.328 \times 1.17 = 0.384 \tag{18}$$

Its confidence interval indicates that c_{US}^{2000} is significantly less than one (Table 2) implying that if candidates locate at the electoral mean in the 2000 they meet the sufficient and thus the necessary condition for convergence to the electoral mean given in Section 2.

Given that the convergence coefficient of the 2004 election

$$c_{\text{US}}^{2004} = 0.45 \tag{19}$$

is significantly less than one (Table 2), the sufficient and necessary conditions for convergence given in Section 2 are also met in the 2004 election.

Thus, with a high degree of certainty candidates found it in their best interest to locate at the electoral mean in the 2000 and 2004 elections. Our formal model predicts that the valence differences between Bush and Gore in 2000, and between Bush and Kerry in 2004 were not large enough to cause either candidate to move

away from the origin in either election. The unique LNE in the 2000 and 2004 elections is one where candidates converge to the electoral origin in each election and remain there as they are maximizing their vote shares.

The 2008 US election stands in sharp contrast to the 2000 and 2004 elections. Our analysis suggests that Obama's victory over McCain was not only the result of shifts in voters' preferences but also the result of shifts in the relative valences of the Democratic and Republican candidates as compared to the valences of the candidates in the previous two elections.

The MNL estimates of the 2008 election, given in Table 1

$$\lambda_{\text{rep}}^{\text{US2008}} = -0.84, \quad \lambda_{\text{dem}}^{\text{US2008}} \equiv 0.0, \quad \beta_{2008}^{\text{US}} = 0.85 \tag{20}$$

suggest that McCain had a significantly *lower* valence than Obama, the baseline candidate.

Using equation (8), the probability that a voter chooses the Republican candidate, McCain, when both candidates are located at the origin, $\mathbf{z}_0$

$$\rho_{\text{rep}}^{\text{US2008}} = 0.30 \tag{21}$$

is lower than that of Bush both in 2000 in equation (16) and in 2004 in equation (17). The central estimate of the probability of voting for the Democratic high valence candidate, $\rho_{\text{dem}}^{\text{US2008}} = 1 - \rho_{\text{rep}}^{\text{US2008}} = 1 - 0.3 = 0.7$, was higher in 2008 than in the previous two elections (Table 2).

The higher probability of voting for Obama coupled with a higher electoral 'variance' gives a convergence coefficient in 2008

$$c_{\text{US}}^{2008} = 1.11 \tag{22}$$

which is higher than in 2000 (equation (18)) and 2004 (equation (19)) . Since c_{US}^{2008} is not significantly greater than one and thus is significantly less than two (Table 2), the necessary, but *not* the sufficient, condition for convergence has been met suggesting that no candidate had an incentive to move from the mean. Thus, with a probability exceeding 95%, the electoral origin is a LNE for the 2008 US election.

In conclusion, the convergence coefficient varies across elections in the same country even when there are only two parties. This is to be expected as, from equation (9), the convergence coefficient depends on the 'variance' of the electoral distribution, $\sigma^2 = \text{trace}\ (\nabla)$, on the weight voters give to the differences of party's policies, β and on the probability that a voter chooses the lowest valence party, ρ_1.

The analysis of these three consecutive US elections illustrates that even though the electoral distributions of the 2000 and 2004 are quite similar, the shift in voters' preferences in 2008 brought about a change in the presidency. The electoral variance in 2008 increased along both axes relative to 2000 and 2004. Moreover, while the 2000 and 2004 convergence coefficients are indistinguishable from each other, the 2008 coefficient is significantly different from that in 2000 and 2004. In spite of these differences, candidates found it in their interest to locate at the electoral mean in all three elections.

3.1.2. The 2005 and 2010 elections in the UK. The 2008 financial crisis, which occurred between the 2005 and 2010 UK elections, greatly affected the UK as the country was battling severe unemployment, foreclosures and a national budget deficit. Despite making strides in 2009 to stimulate the economy, Prime Minister Gordon Brown, who took over from Tony Blair as leader of the Labour Party in June 2007, saw the UK's budget deficit increase. This coupled with media reports on the expense account abuse by Members of Parliament, including some cabinet members which caused some to resign, led to a decrease in Brown's public ratings. In 2010, the Labour Party lost its majority in the Commons leading Brown to resign as Labour Leader and to soon after step down as prime minister.

We study these two elections to examine how the large differences in popularity between the two Labour Prime Ministers translates into valence differences across Blair and Brown. The questions from the British Election Study (BES),[14] used in the factor analysis, led us to conclude that the same two dimensions mattered for voter choices in the two elections. The first factor deals with issues on 'EU membership', 'Immigrants', 'Asylum seekers' and 'Terrorism'. A voter who feels strongly about nationalism has a high value on the *nationalism* dimension ($Nat = x$-axis). Items such as 'tax/spend', 'free market', 'international monetary transfers', 'international companies' and 'worry about job loss overseas' have a strong influence on the *economic* ($E = y$-axis) dimension with higher values indicating a pro-market attitude. Figures 4 and 5 present the smoothed electoral distribution for the 2005 and 2010 elections.

The notable electoral shift that occurred between these two elections, attributed to Brown's rising unpopularity heading into the 2010 election relative to Prime Minister Blair's popularity in 2005, is also apparent in the electoral covariance matrices of these two elections. The 2005 covariance matrix

$$\nabla^{\mathrm{UK}}_{2005} = \begin{bmatrix} \sigma^2_{Nat} = 1.646 & \sigma_{NatE} = 0.00 \\ \sigma_{ENat} = 0.067 & \sigma^2_{E} = 3.961 \end{bmatrix} \tag{23}$$

is not only quite different from that of the 2010 election

$$\nabla^{\mathrm{UK}}_{2010} = \begin{bmatrix} \sigma^2_{Nat} = 0.601 & \sigma_{NatE} = 0.067 \\ \sigma_{ENat} = 0.067 & \sigma^2_{E} = 0.861 \end{bmatrix} \tag{24}$$

but the 2005 electoral 'variance', $\sigma^2_{\mathrm{UK2005}} \equiv \mathrm{trace}(\nabla^{\mathrm{UK}}_{2005}) = \sigma^2_{\mathrm{Nat}} + \sigma^2_{E} = 5.607$, is much larger than that in 2010, $\sigma^2_{\mathrm{UK2010}} \equiv \mathrm{trace}(\nabla^{\mathrm{UK}}_{2010}) = 1.462$.

The 2005 valence estimates for the Labour (Lab) and the Conservative (Con) parties (Table 2) were significantly higher ($\lambda^{\mathrm{UK2005}}_{\mathrm{Lab}} = 0.52, \lambda^{\mathrm{UK2005}}_{\mathrm{Con}} = 0.27$) than that of the Liberal Democrats (Lib, $\lambda^{\mathrm{UK2010}}_{\mathrm{Lib}} \equiv 0.0$), the baseline party. Things changed dramatically by 2010. Given Brown's deep unpopularity, it is no surprise that the Conservatives and Liberal Democrats (the base party) had significantly higher valences ($\lambda^{\mathrm{UK2010}}_{\mathrm{Con}} = 0.17, \lambda^{\mathrm{UK2010}}_{\mathrm{Lib}} \equiv 0.0$) than Labour ($\lambda^{\mathrm{UK2010}}_{\mathrm{Lab}} = -0.04$) in 2010. The sizable fall in Labour's valence between these two elections reflects Brown's 2010 unpopularity relative to Blair's in 2005.

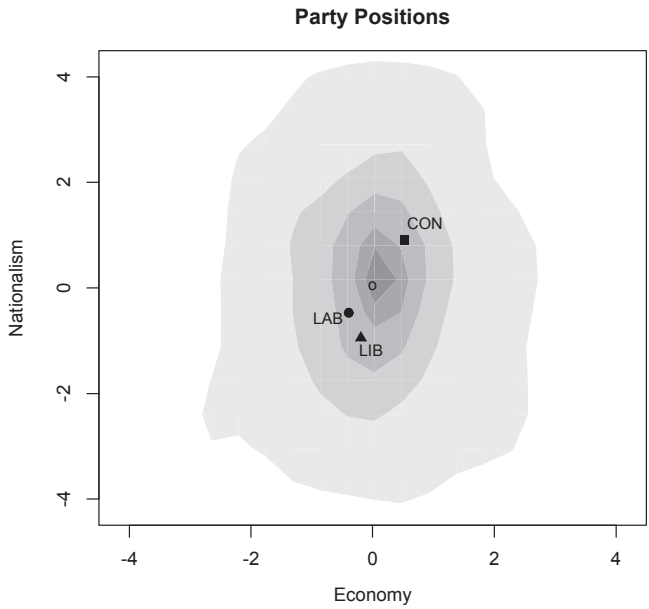

Figure 4. Electoral distribution and estimated party positions in the UK in 2005.

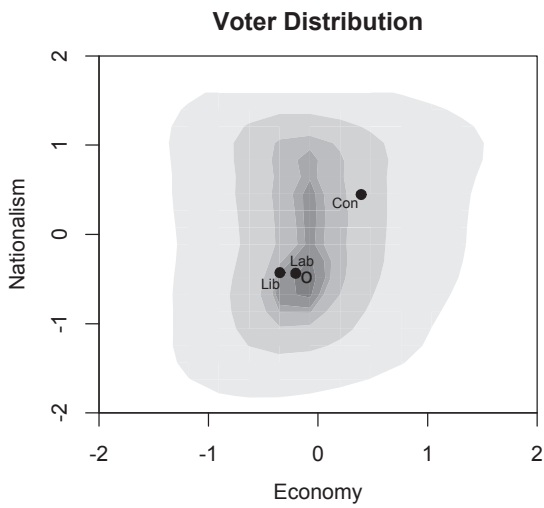

Figure 5. Voter and party positions in the UK in 2010.

The valence shifts between elections affect the probability of voting for all parties. Whereas in 2005 the probability of voting Liberal Democrat, with lowest valence when all parties locate at the mean, from equation (8), is

$$\rho_{\text{Lib}}^{\text{UK2005}} = \left[\sum\nolimits_{k=1}^{3} \exp(\lambda_k^{\text{UK2005}} - \lambda_{\text{Lib}}^{\text{UK2005}})\right]^{-1} = [1 + \exp(0.52) + \exp(0.27)]^{-1} = 0.25 \tag{25}$$

in 2010 as the Liberal Democrats valence rises relative to 2005, this probability rises to

$$\rho_{\text{Lib}}^{\text{UK2010}} = \left[\sum\nolimits_{k=1}^{3} \exp(\lambda_k^{\text{UK2010}} - \lambda_{\text{Lib}}^{\text{UK2010}})\right]^{-1} = [1 + \exp(-0.04) + \exp(0.17)]^{-1} = 0.318 \tag{26}$$

Moreover, while the probability of voting Labour in 2005, when all parties locate at the origin, $\mathbf{z}_0$, is

$$\rho_{\text{Lab}}^{\text{UK2005}} = \left[\sum\nolimits_{k=1}^{3} \exp(\lambda_k^{\text{UK2005}} - \lambda_{\text{Lab}}^{\text{UK2005}})\right]^{-1} = [1 + \exp(0.27 - 0.52) + \exp(0 - 0.52)]^{-1} = 0.42 \tag{27}$$

in 2010, as Labour's valence falls, making it the low valence party, this probability decreases to

$$\rho_{\text{Lab}}^{\text{UK2010}} = \left[\sum\nolimits_{k=1}^{3} \exp(\lambda_k^{\text{UK2010}} - \lambda_{\text{Lab}}^{\text{UK2010}})\right]^{-1} = [1 + \exp(0.21) + \exp(0.04)]^{-1} = 0.319 \tag{28}$$

In spite of the large differences between the two elections, their convergence coefficients

$$c_{\text{UK}}^{2005} = 0.84 \quad \text{and} \quad c_{\text{UK}}^{2010} = 0.91 \tag{29}$$

are both significantly less than one (see Table 2), indicating that the sufficient and thus necessary condition for convergence stated in Section 2 were met in both elections. Thus, all three parties found it in their interests to locate at the electoral mean in both elections. With probability exceeding 95%, the electoral origin is then a LNE in both British elections.

The major shift in voters' preferences between the two elections led to Labour wining a majority in 2005 and to a hung Parliament in 2010 when Brown lost the election. Our results suggest that voter dissatisfaction with Brown led to a dramatic decrease in his competence valence and on the probability of voting Labour in 2010. Even though the electoral variance fell in 2010 relative to 2005, the increase in the convergence coefficient meant that this lower variance was more than compensated by the lower probability of voting Labour, the low valence party, in 2010 (compare equation (27) with (28)).

The analysis of the UK elections gives an indication that the convergence coefficient reflects not only changes in the electoral distribution but also changes in

voters' valence preferences as the convergence coefficient of the 2005 election is substantially lower than that of the 2010 election.

The elections in these two Anglo-Saxon countries illustrate that even under the *plurality rule* the convergence coefficient varies from election to election and from country to country. The 2010 UK election highlights the fact that candidates' valences matter as they affect parties' electoral prospects. This section also suggests that under plurality rule the values of the convergence coefficient tend to be low, generally below the dimension of the policy space, thus satisfying the necessary condition for convergence to the mean.

3.2. Convergence in proportional systems?

We examine convergence to the electoral mean for three parliamentary countries using proportional representation electoral systems: Israel, Turkey and Poland. As is well known, these countries are characterized by multi-party elections in which generally no party wins a legislative majority leading to coalition governments. This section shows that these countries are characterized by very high convergence coefficients.

3.2.1. The 1996 election in Israel. The 1996 Israeli election made the leader of the Likud Party, Benjamin Netanyahu, Prime Minister of Israel. This election marked the first time a Prime Minister was directly elected by voters in Israel.

As in previous elections, there were approximately 19 parties attaining seats in the Knesset in 1996.[15] There were small parties with two seats to moderately large parties such as Likud and Labor whose seat strengths lay in the range 19 to 44, out of a total of 120 Knesset seats. Since Likud and Labour compete for dominance of the coalition government, these large parties must maximize their seat strength. Moreover, Israel uses a highly proportional electoral system with close correspondence between seat and vote shares. Thus, one can consider vote shares as the maximand for these parties.

Schofield et al. (2011c) performed a factor analysis of the surveys conducted by Arian and Shamir (1999) to study the 1996 Israeli election. The two dimensions identified by the factor analysis were security ($S = x$-axis) and religion ($R = y$-axis). 'Security' refers to attitudes toward peace initiatives, and 'religion' refers to the significance of religious considerations in government policy. A voter on the left of the security axis is interpreted as supporting negotiations with the PLO, while higher values on the religious axis indicates support for the importance of the Jewish faith in Israel. The distribution of voters is shown in Figure 6.

Voter distribution along these two axes gives the following covariance matrix

$$\nabla^{\mathrm{I}}_{1996} = \begin{bmatrix} \sigma^2_S = 1.00 & \sigma_{SR} = 0.591 \\ \sigma_{RS} = 0.591 & \sigma^2_R = 0.732 \end{bmatrix} \tag{30}$$

giving a 'variance' of $\sigma^2_{\mathrm{I}1996} \equiv \mathrm{trace}(\nabla^{\mathrm{I}}_{1996}) = 1.732$. To estimate the valences of the parties, only the seven largest parties are included in the MNL estimation, Likud,

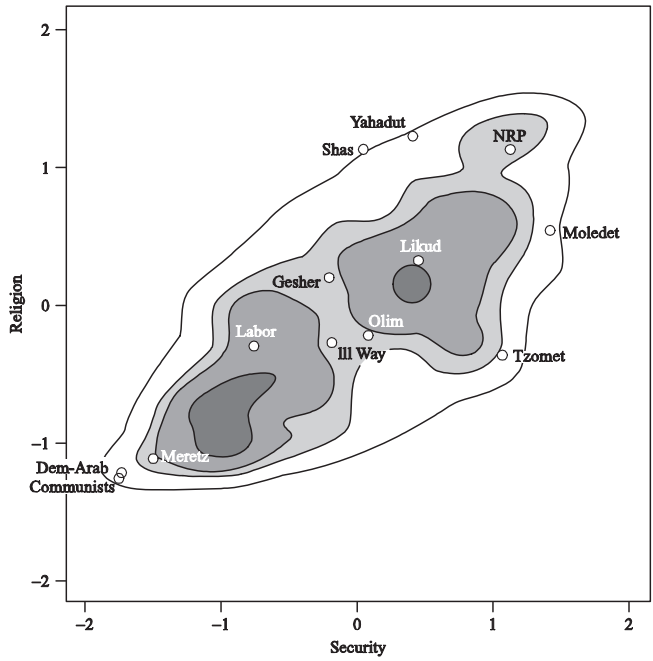

Figure 6. Party positions and voter distribution in Israel in the 1996 election.

Labor, NRP, Moledat, Third Way (TW) and Shas, with Meretz being the base party. From Table 3, the MNL coefficients for the 1996 election in Israel (I) are

$$\begin{aligned} &\lambda_{\text{Lik}}^{\text{I1996}} = 0.78, \quad \lambda_{\text{Lab}}^{\text{I1996}} = 0.999, \quad \lambda_{\text{NRP}}^{\text{I1996}} = -0.626, \quad \lambda_{\text{MO}}^{\text{I1996}} = -1.259, \\ &\lambda_{\text{TW}}^{\text{I1996}} \equiv -2.291, \quad \lambda_{\text{Shas}}^{\text{I1996}} = -2.023, \quad \lambda_{\text{Merezt}}^{\text{I1996}} \equiv 0.0, \quad \beta_{1996}^{\text{I}} = 1.207 \end{aligned} \tag{31}$$

The two largest parties, Likud and Labor, have significantly higher valences than the other smaller parties with TW having the smallest valence (Table 3).

From Table 4 and equation (8), our model predicts that the probability that an Israeli votes for TW, when all parties are located at the mean

$$\rho_{\text{TW}}^{\text{I1996}} = 0.014 \tag{32}$$

is very low implying that TW would locate away from the mean to increase its vote share. The analysis performed in Schofield et al. (2011c) also shows that other small parties would also locate far from the mean. This explains why Israel's 1996 convergence coefficient

$$c_{1996}^{\text{I}} = 4.06 \tag{33}$$

is high as its 95% confidence interval (Table 4) confirms that c_{1996}^{I} is significantly higher than two, the dimension of the policy space.

Table 3. MNL spatial model for countries with proportional systems.

	Israel[a]		Turkey[b]			Poland[c]	
Var	Party	1996	Party	1999	2002	Party	1997
Distance		Estimate [d] |t–value|		Estimate |t–value|	Estimate |t–value|		Estimate |t–value|
β		1.207*** (18.43)		0.375***(4.26)	1.52***(12.66)		1.739***(15.04)
Valence	λ_{Lik}	0.777*** (4.12)	λ_{DSP}	0.724***(4.73)		λ_{SLD}	1.419*** (7.47)
	λ_{Lab}	0.999*** (6.06)	λ_{MHP}	0.666***(4.53)	–0.12 (0.66)	λ_{PSL}	0.073 (0.33)
	λ_{NRP}	–0.626*** (2.53)	λ_{FP}	–0.159 (0.90)		λ_{AWS}	1.921***(11.05)
	λ_{MO}	–1.259*** (4.38)	λ_{ANAP}	0.336***(2.19)	–0.31 (1.63)	λ_{UW}	0.731*** (3.67)
	λ_{TW}	–2.291*** (8.30)	λ_{CHP}	0.734***(4.12)	1.33*** (7.40)	λ_{UP}	–0.56*** (2.13)
	λ_{Shas}	–2.023*** (6.45)	λ_{HADEP}	–0.071 (0.30)	0.43* (2.0)	λ_{UPR}	–2.348***(4.69)
			λ_{AKP}		0.78*** (5.2)		
Base party		Meretz		DYP	DYP		ROP
n		922		635	483		660
LL		–777		–1183	–737		–855

[a]Israel: Lik = Likud; Lab = Labor; NRP = Mafdal; Mo = Moledet; TW = Third Way.
[b]Turkey: DSP = Democratic Left Party; MHP = Nationalist Action Party; FP = Virtue Party; ANAP = Motherland Party; CHP = Republican People's Party; HADEP = People's Democracy Party; DYP = True Path Party.
[c]Poland: SLD = Democratic Left Alliance; PSL = Polish People's Party; UW = Freedom Union; AWS = Solidarity Election Action; UP = Labor Party; UPR = Union of Political Realism; ROP = Movement for Reconstruction of Poland; SO = Self Defense; PiS = Law and Justice; PO = Civic Platform; LPR = League of Polish Families; DEM = Democratic Party; SDP = Social Democracy of Poland.
[d]*prob: < 0.05; *** Probability < 0.001.

Table 4. The convergence coefficient in proportional systems.

	Israel[a]	Turkey[b]		Poland[c]
	1996	1999	2002	1997
Weight of policy differences (β)				
Central est. of β (Conf. Int.)[d]	1.207 (1.076, 1.338)	0.375 (0.203, 0.547)	1.520 (1.285, 1.755)	1.739 (1.512, 1.966)
Electoral variance (trace $\nabla = \sigma^2$)				
σ^2	1.732	2.34	2.33	2.00
Probability of voting for lowest valence party (party 1, $\rho_1 = \left[\sum_{k=1}^{p} \exp(\lambda_k - \lambda_1)\right]^{-1}$)				
	TW	FP	ANAP	ROP
Central est. of ρ_1 (Conf. Int.)	$\rho^{I}_{TW} = 0.014$ (0.006, 0.034)	$\rho_{FP} = 0.08$ (0.046, 0.145)	$\rho^{T}_{ANAP} = 0.08$ (0.038, 0.133)	$\rho^{P}_{ROP} = 0.007$ (0.002, 0.022)
Convergence coefficient ($c \equiv c(\lambda, \beta, \sigma^2) = 2\beta[1 - 2\rho_1]\sigma^2$)				
Central est. of c (Conf. Int.)	4.06 (3.474, 4.579)	1.49 (0.675, 2.234)	5.75 (4.388, 7.438)	5.99 (5.782, 7.833)

[a]Israel: TW = Third Way.
[b]Turkey: DYP = True Path Party.
[c]Poland: ROP = Movement for Reconstruction of Poland.
[d]Central Est. = Central estimate; Conf. Int. = Confidence intervals.

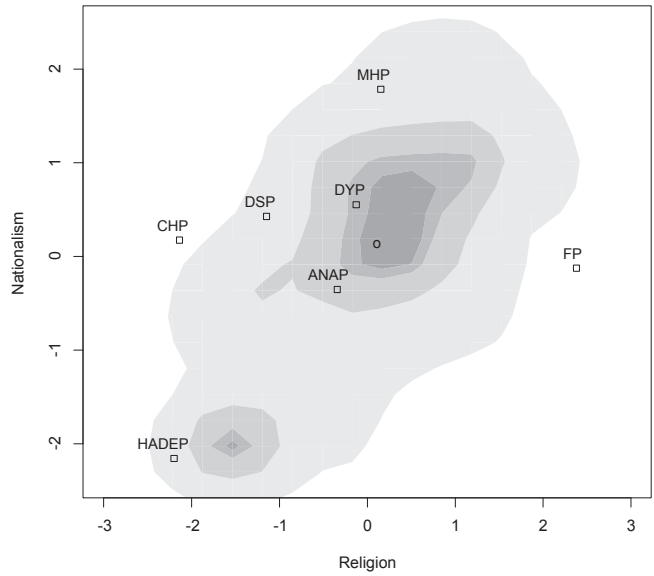

Figure 7. Party positions and voter distribution in the 1999 Turkish election.

Thus, the necessary condition for convergence to the mean is *not* satisfied. Therefore with a high degree of certainty the electoral mean is *not* a LNE for the 1996 Israeli election.

3.2.2. The 1999 and 2002 Elections in Turkey. Between the 1999 and 2002 elections, Turkey experienced two severe economic crises and a 10% electoral cut-off rule was instituted. The political changes induced by the two crises and the cut-off rule led us to examine their joint effect on the 2002 election relative to that in 1999.

We used factor analysis of electoral survey data of Veri Arastima (obtained from http.veriarastima.com.tr) to study the 1999 and 2002 Turkish elections.[16] The analysis indicates that voters made decisions in the same two-dimensional space in both elections. Voters who support secularism or 'Kemalism' are placed on the left of the Religious ($R = x$) axis and those supporting Turkish nationalism ($N = y$) with axis. Figures 7 and 8 give the distribution of voters along these two dimensions surveyed in these two elections.

Minor differences between these two figures include the disappearance of the Virtue Party which was banned by the Constitutional Court in 2001, and the change of the name of the pro-Kurdish party from People's Democratic Party (HADEP) to Demokratik Halk Partisi (DEHAP). [17] The most important change is the emergence of the new Justice and Development Party (AKP) in 2002, essentially substituting for the outlawed Virtue Party.

The parties included in the analysis of the 1999 election are the Democratic Left Party (DSP), the National Action party (MHP), the Virtue Party, the Motherland Party (ANAP), the True Path Party (DYP), the Republican People's Party (CHP) and HADEP. A DSP minority government formed, supported by the ANAP and

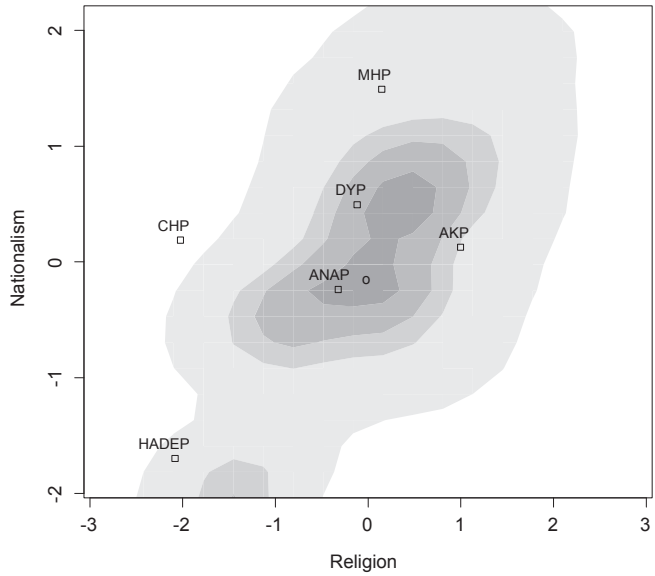

Figure 8. Party positions and voter distribution in Turkey in 2002.

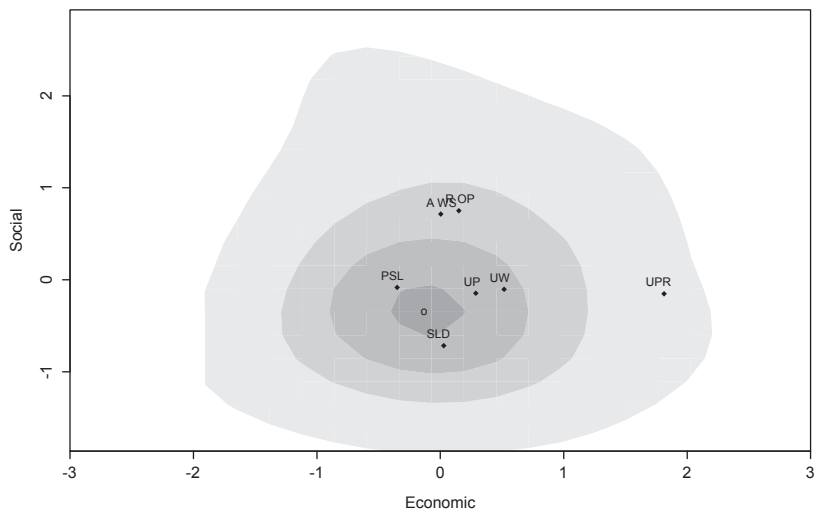

Figure 9. Voter distribution and party positions in Poland in 1997.

DYP. This only lasted about four months, and was replaced by a DSP–ANAP–MHP coalition, indicating the difficulty of negotiating a coalition compromise across the disparate policy positions of the coalition members.

Figures 8 and 9 and the analysis of the electoral covariance matrices of the two elections suggest that there were few changes in party positions between these two

elections. The 1999 electoral covariance matrix along the Religious (R) and Nationalism (N) axes given by

$$\nabla^{\mathrm{T}}_{1999} = \begin{bmatrix} \sigma^2_R = 1.20 & \sigma_{RN} = 0.78 \\ \sigma_{NR} = 0.78 & \sigma^2_N = 1.14 \end{bmatrix} \tag{34}$$

with $\sigma^2_{\mathrm{T}1999} \equiv \mathrm{trace}(\nabla^{\mathrm{T}}_{1999}) = 2.34$, is not very different from that in 2002 given by

$$\nabla^{\mathrm{T}}_{2002} = \begin{bmatrix} \sigma^2_R = 1.18 & \sigma_{RN} = 0.74 \\ \sigma_{NR} = 0.74 & \sigma^2_N = 1.15 \end{bmatrix} \tag{35}$$

with $\sigma^2_{\mathrm{T}2002} = \mathrm{trace}(\nabla^{\mathrm{T}}_{2002}) = 2.33$.

While the MNL coefficients of the 1999 election (Table 3) indicate that DSP, MHP, ANAP and CHP have significantly higher valences than FP, HADEP and DYP (the base party), the MNL estimates show that party valences had substantially changed by 2002. In particular, it is now CHP, HADEP and AKP that have significantly higher valences than DSP (the base party), MHP, FP and ANAP (with lowest valence).

The probability that a Turkish voter chooses the lowest valence when all parties locate at the mean party, FP in 1999, $\rho^{\mathrm{T}1999}_{\mathrm{FP}}$ in equation (8), is

$$\rho^{\mathrm{T}1999}_{\mathrm{FP}} = 0.08 \tag{36}$$

(note that, by 2002, FP has disappeared). Moreover, in spite of the valence differences across elections, $\rho^{\mathrm{T}1999}_{\mathrm{FP}}$ is indistinguishable from that of ANAP in 2002

$$\rho^{\mathrm{T}2002}_{\mathrm{ANAP}} = 0.08 \tag{37}$$

which is also indistinguishable from the probability of voting for ANAP in 1999

$$\rho^{\mathrm{T}1999}_{\mathrm{ANAP}} = \left[\sum\nolimits_{k=1}^{7} \exp[\lambda^{\mathrm{T}1999}_j - \lambda^{\mathrm{T}1999}_{\mathrm{ANAP}}]\right]^{-1} = \left[1 + e^{0.19} + e^{1.64} + e^{0.74} + e^{1.09}\right]^{-1} \simeq 0.08$$

According to our model, ANAP did not meet the 10% cut-off rule in 2002.

The economic crisis and the implementation of the 10% cut-off rule in 2002 led to radical political changes in Turkey. The support base of the AKP in 2002 is similar to that of the banned FP, suggesting that the leader of this party changed the party's position along the religion axis, adopting a much less radical position. One would think of this as generating political stability in Turkey. Yet, in the 2002 election, the 10% cut-off rule lead seven parties to win no seats in the Meclis (the Turkish Parliament) as they failed to meet the cut-off requirement. This 10% cut-off rule allowed the AKP, with only 34% of the vote, to obtain a 66% majority of the seats (363 out of 550).

The differences between these two elections can also be seen when examining parties' convergence to the mean. Turkey's 1999 convergence coefficient, using (9)

$$c^{\mathrm{T}}_{1999} = 1.49 \tag{38}$$

is not significantly different from two (Table 4), the dimension of the policy space, implying that the necessary condition for convergence given in Section 2 has *not* been met. The detailed analysis carried out in Schofield et al. (2011f) suggests that small parties found it in their interest to locate away from the electoral mean. Thus, the electoral mean was *not* a LNE for Turkey in 1999.

Turkey's convergence coefficient in 2002

$$c^{\mathrm{T}}_{2002} = 5.94 \tag{39}$$

is significantly *above* two, the dimension of the policy space (Table 4) giving all parties an incentive to locate away from the mean. The electoral mean is then also *not* a LNE for Turkey in 2002.

The dramatic increase in the convergence coefficient relative to 1999 gives a clear indication that a more fractionalized polity emerged from the electoral reform implementing the 10% cut-of rule.

3.2.3. The 1997 Polish election. The 1997 Polish election[18] marked the first time meaningful party programs appeared in the campaign. It also reconstituted the Polish political landscape. Parties' campaign platforms aimed at reforming the state, building strong local governments, implementing open market directives, to improve economic conditions, and legislation that would lead to stable economic and financial laws.

Five parties won seats in the Sejm (lower house) in 1997. The Solidarity Electoral Action (AWS) and several minor parties, based on the Solidarity opposition movement, won a decisive victory, with 33.8% of the votes and 201 of the 460 Sejm seats. Second place went to the Democratic Left Alliance (SLD), a center-left coalition whose roots were remanent of the previous communist regime, with 27.1% of the votes and 164 seats. The Freedom Union (UW) party came in third with 13.4% of the votes and 60 seats. The Polish Peasant Party (PSL), the 1993–97 coalition partner of the SLD, obtained 7.3% and 27 seats. The Movement for the Rebirth of Poland (ROP), another party with Solidarity roots, received 5.6% and six seats. The remaining two seats went to Silesian Germans, exempt from the 5% threshold as a national minority. Among those who did not clear the threshold was the leftist Labor Union (UP) with 4.7% of the vote.

The factor analysis carried out on questions from the Polish National Election Survey identified an economic and a social value dimensions (see Schofield et al., 2011g). The *economic* dimension is influenced by issues such as 'privatization versus state ownership of enterprizes', 'fighting unemployment versus keeping inflation and government expenditure under control', 'proportional versus flat income tax', 'support versus opposition to state subsidies to agriculture' and 'state versus individual social responsibility'. The 'separation of church and state versus the influence of church over politics', 'complete de-communization versus equal rights for former nomenclature', and 'abortion rights versus no such rights, regardless of situation' are the most influential issues along the *social values* dimension. Figure 9 gives the distribution of voters along these dimensions.

The covariance matrix for the 1997 Polish (P) election

$$\nabla^{\mathrm{P}}_{1997} = \begin{bmatrix} \sigma^2_E = 1.00 & \sigma_{ES} = 0.0 \\ \sigma_{SE} = 0.0 & \sigma^2_S = 1.00 \end{bmatrix} \tag{40}$$

with variance $\sigma^2_{\mathrm{P}1997} = \mathrm{trace}(\nabla^{\mathrm{P}}_{1997}) = 2.00$, is symmetric about the mean as suggested by Figure 9.

The MNL coefficients for the 1997 election given in Table 3 suggest that the prominence of the SLD and the AWS along with UW gave them a significantly higher valences than either ROP (the base party) and PSL, which in turn has significantly higher valences than UP and UPR.

Using the valence estimates in Table 4 and equation (8), the probability of voting UPR with lowest valence, in 1997, when parties locate at the mean, $\rho^{\mathrm{P}1997}_{\mathrm{TW}}$ given in Table 4

$$\rho^{\mathrm{P}1997}_{\mathrm{UPR}} = 0.01, \tag{41}$$

is low as expected in a multi-party political system.

In addition, the convergence coefficient for Poland in 1997, in Table 4

$$c^{\mathrm{P}}_{1997} = 6.82, \tag{42}$$

is significantly greater than two and thus fails the necessary condition for convergence to the electoral mean. The lack of convergence is a symptom of a system where smaller parties, who understand the fragmented nature of the Polish system, prefer to locate closer to their core constituency and away from the electoral mean in order to increase their vote shares. With a high degree of certainty, the electoral mean is then *not* a LNE for 1997 Polish election.

Summarizing, in this section we examined three countries that use proportional representation. Their convergence coefficients are significantly higher than two, the dimension of the policy space, being also much higher than the value of the convergence coefficients of the United States and the UK. A high convergence coefficient signals then a high degree of political fractionalization in these multi-party parliamentary democracies.

3.3. Convergence in anocracies

We now examine elections in Georgia, Russia and Azerbaijan. In these partial democracies, or anocracies,[19] the president/autocrat holds regular presidential and legislative elections while exerting undue influence on the elections. Anocracies lack important democratic institutions such as freedom of the press. Autocrats hold regular elections in an attempt to give their regime legitimacy. The autocrat 'buys' legitimacy by rewarding their supporters and opposition members with well paid legislative positions and gives legislators the ability to influence policies. Opposition parties participate in elections to become known political entities. This allows them to regularly communicate with voters. Their objective is oust the autocrat either in a future election or through popular uprisings. We assume that

opposition parties maximize their vote share even when understanding that there is little chance of ousting the autocrat in the election.

3.3.1. The 2008 Georgian election. President Mikheil Saakashvili's decision to step down on November 25, 2007, in response to a series of wide-spread anti-government demonstrations, brought the presidential election forward to January 5, 2008. A plebiscite on whether there should be spring or fall legislative elections and on whether Georgia should join NATO was also attached to the election. Saakashvili won with 53.7% of the votes, a sharp contrast with his 2004 96% vote share, another confirmation of the increasing public dissatisfaction with Saakashvili during this period. In spite of the opposition's claims of electoral fraud, international observers declared the 2008 election as the first genuinely competitive presidential election in Georgia's history.

The sweeping demands for increased democratization led the government to implement important changes to the electoral code in late November and mid-December. The reforms included the addition of opposition party representatives to the electoral commissions. The Organization for Security and Cooperation in Europe, which monitored the election, commented that the new electoral rules were implemented in a disorganized and inconsistent manner across Georgia. However, in spite of these irregularities, exit polls concurred with the electoral outcome.

We used the post-election survey conducted by GORBI Gallup International from 19 March through to 3 April 2008, to build a formal model of the 2008 election in Georgia (see Schofield et al., 2012). The factor analysis carried out on the survey questions determined that there were two dimensions describing voters' attitudes towards democracy and the west. One dimension is strongly related to the respondents' attitude toward the US, the EU and NATO with larger values along the 'West' ($W = y$-axis) dimension signaling a stronger anti-Western attitude. Along the Democracy ($D = x$-axis) dimension larger values are associated with negative judgements on the current state of democratic institutions in Georgia, coupled with a demand for more democracy. Figure 10 gives Georgia's electoral distribution along these two dimensions along with the estimated positions of Saakashvili (S), Gachechiladze (G), Patarkatsishvili (P) and Natelashvili (N).

The 2008 electoral covariance matrix in the Democracy (D) and West (W) axes,

$$\nabla^{\mathrm{G}}_{2008} = \begin{bmatrix} \sigma^2_D = 0.82 & \sigma_{DW} = 0.03 \\ \sigma_{WD} = 0.03 & \sigma^2_W = 0.91 \end{bmatrix} \tag{43}$$

with $\sigma^2_{\mathrm{G2008}} \equiv \mathrm{trace}(\nabla^{\mathrm{G}}_{2008}) = 1.73$, accords with the bi-polar electoral distribution seen in Figure 10.

Given the public's declining trust in Natelashvili's ability to govern, we were not surprised that the MNL estimates indicate that all other candidates had significantly larger valences than Natelashvili (Table 5). Moreover, our analysis also suggests that if all candidates were to locate at the electoral mean, the probability of Georgians voting for Natelashvili in 2008 (Table 6) is small

$$\rho^{\mathrm{G2008}}_{\mathrm{N}} = 0.05. \tag{44}$$

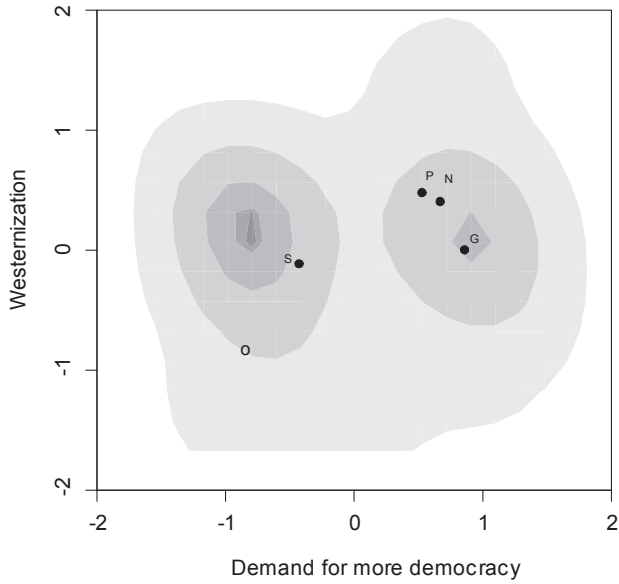

Figure 10. Voter distribution and candidate positions in the 2008 Georgian election.

Table 5. MNL spatial model in anocracies.

	Georgia[a]		Russia[b]		Azerbaijan[c]	
	Party	2008	Party	2007	Party	2010
Var		Estimate[d] |t–value|		Estimate |t–value|		Estimate |t–value|
β		0.78*** (13.78)		0.181*** (12.08)		1.34*** (4.62)
Valance	λ_S	2.56*** (13.66)	λ_{CPRF}	1.971*** (17.79)	λ_{YAP}	1.30* (2.14)
	λ_G	1.50*** (7.96)	λ_{LDRP}	0.153 (1.09)		
	λ_P	0.53* (2.51)	λ_{SR}	–0.404*** (2.50)		
Base party		N		ER		AXCP–MP
n		676		1004		149
LL		–533		–797		–11.5

[a]Georgia: S = Saakashvili, G = Gachechiladze, P = Patarkatsishvili and N = Natelashvili.
[b]Rusia: ER = United Rusia, CPRF = Communist Party, SR = Fair Rusia and LDPR = Liberal Democratic Party.
[c]Azerbaijan: YAP = Yeni Azerbaijan Party and AXCP–MP = Azerbaijan Popular Front Party and Musavat.
[d]* Probability$<$ 0.05, *** probability$<$ 0.001.

Since Natelashvili won the election, our analysis suggests that Natelashvili, the low valence candidate, positioned himself away from the mean to increased his votes. Georgia's convergence coefficient using equation (9)

$$c^{G}_{2008} = 2.42 \tag{45}$$

Table 6. The convergence coefficient in anocracies.

	Georgia[a] 2008	Russia[b] 2007	Azerbaijan[c] 2010
Weight of policy differences (β)			
Estimate β (Conf. Int.[d])	0.78 (0.66, 0.89)	0.181 (0.15, 0.20)	1.34 (0.77, 1.91)
Electoral variance (trace$\nabla = \sigma^2$)			
σ^2	1.73	5.90	0.93
Probability of voting for lowest valence party (party one, $\rho_1 = \left[\sum_{k=1}^{p} \exp(\lambda_k - \lambda_1)\right]^{-1}$)			
	N	SR	AXCP-MP
Estimate ρ_1 (Conf. Int.)	$\rho_N^G = 0.05$ (0.03, 0.07)	$\rho_{SR}^R = 0.07$ (0.04, 0.12)	$\rho_{AXCP-MP} = 0.21$ (0.08, 0.47)
Convergence coefficient ($c \equiv c(\lambda, \beta, \sigma^2) = 2\beta[1 - 2\rho_1]\sigma^2$)			
Estimate c (Conf. Int.)	2.42 (1.99, 2.89)	1.83 (1.35, 2.28)	1.44 (0.085, 2.984)

[a]Georgia: N = Natelashvili.
[b]Russia: SR = Fair Russia.
[c]Azerbaijan: AXCP–MP = Azerbaijan Popular Front Party (AXCP) and Musavat (MP). The estimates for Azerbaijan are less precise because the sample is small.
[d]Conf. Int. = confidence intervals.

is *not* significantly different from two (Table 6), implying that the necessary condition for convergence to the mean is not met, further re-enforcing the conclusion that all presidential candidates located far from the electoral mean and an indication that Georgia's electoral mean *cannot* be a LNE for the 2008 election.

3.3.2. The 2007 Russian election. When heading into the 2007 Duma election, President Vladimir Putin popularity ratings were outstandingly high, ranging between 70% and 85%. In 2007, Russians associated Russia's economic resurgence and more powerful presence on the world stage to his leadership. Putin's rhetoric directly enhanced Russians' national pride.

Moreover, it was widely expected that Putin's United Russia party would win a sweeping majority in the Duma's parliamentary elections. To guarantee this Putin increased the Duma cut-off rule from 5% to 7% prior to the election and in addition, as reported by the Economist the day after the election, put

> The entire government machine, the court system, the prosecution service, the police, the state media and even the central commission itself… to produce a landslide victory.

International observers produced a critical report of the election outlining the media bias in favor of Putin and his party as well as the prevalent harassment of

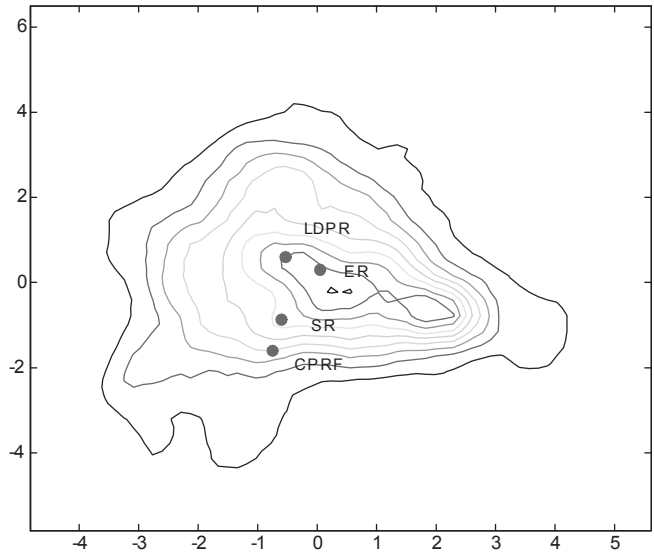

Figure 11. Party positions and voters distribution in the 2007 Russian election.

opposition parties. The Organization for Security and Cooperation in Europe said that the election had not met their standards.

Putin's United Russia party won a massive 64.1% of the vote with the Communists coming in a distant second at 11.6%. The ultranationalist Liberal Democratic party took only 8.2% of the votes with Just Russia barely passing the new 7% threshold.

The analysis carried out in Schofield and Zakharov (2010) concentrates on four parties: the pro-Kremlin United Russia party (ER), the Liberal Democratic Party (LDPR), the Communist Party (CPRF) and the Fair Russia (SR). Voters' ideological preferences were measured using the survey conducted by the Russian Public Opinion Research Center in May 2007. High values on voters' general (dis)satisfaction ($D = x$-axis) correspond to negative feelings toward 'justice', 'labor' and, to a lesser extent, 'order', 'state', 'stability' and 'equality'. Also, those with high values of the D-axis tend to feel neutral towards 'order', 'elite', 'West' and 'non-Russians'. Voter's degree of economic liberalism (E), along the y-axis, is coded such that high values correspond to positive feelings on 'freedom', 'business', 'capitalism', 'well-being', 'success' and 'progress' and negative ones reflect negative attitudes towards 'communism', 'socialism', 'USSR' and other related concepts.

The 2007 electoral covariance matrix along the (dis)satisfaction (D) and economic liberalism (E) axes

$$\nabla^{\mathrm{R}}_{2007} = \begin{bmatrix} \sigma^2_D = 2.95 & \sigma_{DE} = 0.13 \\ \sigma_{ED} = 0.13 & \sigma^2_E = 2.95 \end{bmatrix} \tag{46}$$

with electoral variance $\sigma^2_{\mathrm{R}2007} \equiv \mathrm{trace}(\nabla^{\mathrm{R}}_{2007}) = 5.9$ measures voters' preferences along these two dimensions shown in Figure 11.

The MNL estimates (Table 5) suggest that only the CPRF has a higher valence than Putin's ER (the base party) with Fair Russia (SR) having a significantly lower valence than Putin's party.

The probability that a Russian votes for SR with lowest valence, from equation (8)

$$\rho^{\mathrm{R}2007}_{\mathrm{SR}} = \left[\sum\nolimits_{k=1}^{4} \exp[\lambda^{\mathrm{R}2007}_{j} - \lambda^{\mathrm{R}2007}_{\mathrm{SR}}]\right]^{-1} = \left[1 + e^{0.4} + e^{0.553} + e^{2.371}\right]^{-1} \simeq 0.07 \tag{47}$$

is at the new cut-off rule. Moreover, from (9), Russia's convergence coefficient

$$c^{\mathrm{R}}_{2007} = 1.83 \tag{48}$$

is *not* significantly different from two (Table 6). Even though the necessary condition for convergence is *not* met, the analysis done in Schofield and Zakharov (2010) shows that Fair Russia, with the lowest valence, maximizes its vote share by remaining close to the electoral mean. This result, however, highlights that unexpected political events could prompt Fair Russia to move from the origin.

Since the United Russia party remained near the electoral mean capitalizing on Putin's popularity, our analysis suggests that Fair Russia may also prefer to locate close to the mean to maintain its vote share and thus close to Putin's party, not only because of Putin's popularity but also because of Putin's undue influence on the election. It is then likely that the electoral mean is a LNE for the 2007 Russian election.

3.3.3. The 2010 election in Azerbaijan. The Azerbaijani government amended the electoral code a few months before the 2010 parliamentary elections. The reform reduced the campaign period to 23 days, ended the allocation of limited state funding to candidates and gave parties unequal access to resources. The reform was a strong signal of President Ilham Aliyev's increased authoritarian control over the country's institutions and political life.

The observers from the Organization for Security and Co-operation in Europe (OSCE), the OSCE Parliamentary Assembly, the Parliamentary Assembly of the Council of Europe and the European Parliament assessed the 2010 election. They commented that 'Certain conditions necessary for a meaningful and competitive election were lacking in these elections'. In particular, by making it difficult for opposition candidates to register, the government consciously prevented opposition parties from mounting effective campaigns. In addition, the wide-ranging restrictions imposed on the media and on freedom of assembly created a restrictive political environment during the campaign. The biased media coverage, the unbalanced access and misuse of government resources coupled with interference by local authorities in favor of ruling party candidates gave a grossly uneven advantage to the President's Yeni Azerbaijan Party. Not only was the campaign atmosphere tense but the non-transparent financial transactions and infringements by

government officials led to numerous complains on voting irregularities, in particular on ballot stuffing, multiple voting and interference in the voting by the executive authorities.

In the 2010 Azerbaijani election, 2500 candidates filed applications to run in the election, but only 690 were given permission by the electoral commission. The parties competing in the election included the Yeni Azerbaijan Party (YAP), the party of the President, the Civic Solidarity Party (VHP), the Motherland Party (AVP), the Azerbaijan Popular Front Party (AXCP) and Musavat (MP). Even though small parties formed political blocks, opposition forces remained divided and weak as they ran against each other.

President Ilham Aliyev's ruling Yeni Azerbaijan Party took a majority of 72 out of 125 seats. Nominally independent candidates, who were aligned with the government, received 38 seats, and 10 small opposition or quasi-opposition parties took 10 seats. The Democratic Reforms Party, Great Creation, the Movement for National Rebirth, Umid, Civic Welfare, Adalet (Justice) and the Popular Front of United Azerbaijan, most of which were represented in the previous parliament, won one seat a piece. Civic Solidarity retained its three seats, and Ana Vaten kept the two seats they had in the previous legislature. For the first time, not a single candidate from the opposition AXCP or MP was elected.

We organized a small pre-election survey of 2010 election in Azerbaijan allowing us to construct a model of the election (see Schofield et al., 2012). For VHP and AVP, the estimation of their party positions was very sensitive to inclusion or exclusion of one respondent. Thus, we used only the small subset of 149 voters who completed the factor analysis questions and intended to vote for YAP or the AXCP–MP block.

The factor analysis showed that the election was fought along *one dimension*: the 'demand for democracy'. Along this dimension higher values are associated with voters who have a negative evaluation of the current democratic situation in Azerbaijan, who did not think that free opinion was allowed, had a low degree of trust in key national political institutions and expected that the 2010 parliamentary election would be undemocratic. Figure 12 gives the distribution of voters and party positions at the mean of their supporters. In this one-dimensional model the variance is

$$\sigma^2_{\mathrm{A2010}} \equiv \mathrm{trace}(\nabla^{\mathrm{A}}_{2010}) = 0.93. \tag{49}$$

The *binomial* logit estimates for the 2010 election with AXCP–MP as the base party, in Table 5, indicate that the opposition AXCP–MP block has a significantly lower valence than the President's Yeni Azerbaijan Party. We conjecture that is mainly due to all the restrictions placed on voters and candidates during the campaign.

Were these two parties to locate at the mean, the probability of voting for the AXCP–MP block from equation (8) is

$$\rho^{\mathrm{A2010}}_{\mathrm{AXCP-MP}} = 0.21. \tag{50}$$

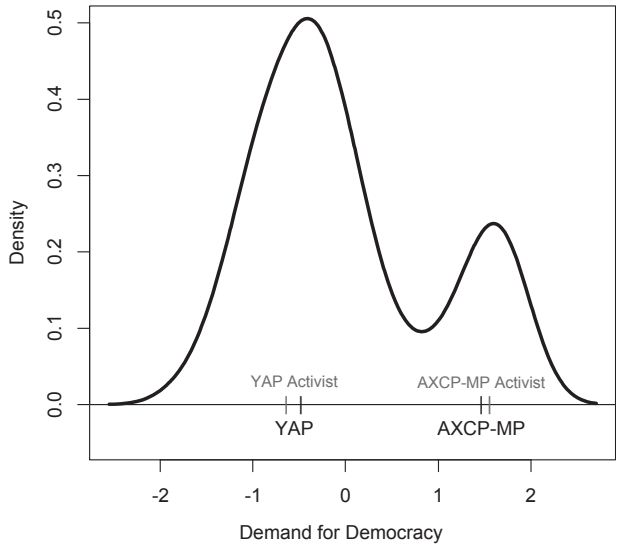

Figure 12. Voter distribution and activist positions in the 2010 Azerbaijani election.

This is not surprising given all the pre-election irregularities that may have biased the polling sample.

Moreover, Azerbaijan's convergence coefficient, using equation (9), given by

$$c^{\mathrm{A}}_{2010} = 1.445 \tag{51}$$

is not significantly different from one, the dimension of the policy space (Table 6), implying that the necessary condition for convergence is *not* met. Thus, with a high degree of certainty AXCP–MP will locate away from the electoral mean as there is no risk in doing so given their belief on their low ability of defeating the President's parliamentary candidates. The electoral mean is *not* a LNE for the 2010 election in Azerbaijan.

This section illustrates that for the three anocracies that we consider the convergence coefficients are not significantly different from the dimension of the policy space and do not satisfy the necessary condition for convergence to the mean. As a consequence, parties are at a knife-edge equilibrium. Under some conditions, parties converge to the mean, under others they diverge. Which equilibrium materializes depends on how popular or unpopular the president/autocrat and his party are, on the valence of all parties and on how dispersed voters are in the policy space. This illustrates that changes in a party's valence can substantially affect party positions and perhaps the outcome of the election.

4. Convergence across political systems

In the previous sections we used the unifying framework of the stochastic electoral model in Schofield (2007), outlined in Section 2, to study whether parties locate

Table 7. Convergence and fragmentation.

	Plurality systems				
Variable	US			UK	
Political system	Presidential			Parliamentary	
Election year	2000	2004	2008	2005	2010
Conv. Coef.[a]	0.38	0.45	1.11	0.84	0.95
(Conf. Int.[b])	(0.2,0.7)	(0.2,0.8)	(0.7,1.5)	(0.5,1.3)	(0.9,1.1)
Converge to mean	Yes	Yes	Yes	Yes	Yes
No. of parties	2	2	2	9	9
President					
env[c]	2.16	2.05	2.05		
	House of Representatives			House of Commons	
env	2.25	2.18	2.18	3.61	3.74
ens[d]	2.02	2.00	2.00	2.47	2.58
Proportional representation					
	Israel	Turkey		Poland	
Political system	Fragmented	Fragmented	Cut off	Fragmented	
Election year	1996	1999	2002	1997	
Conv. Coef. (Conf. Int.)	3.98 (3.5,4.6)	1.49 (0.7,2.2)	5.94 (4.4,7.4)	6.82 (5.8,7.8)	
Converge to mean No	No	Likely	No	No	
No. of parties	11	9	10	7	
	Prime ministers[e]				
env	2.00				
	Knesset	Parliament		Sejm	
Env	5.84	6.91	5.62	4.99	
Ens	5.89	6.35	2.29	6.77	
Anocracies plurality					
	Georgia	Russia		Azerbaijan	
Political system	Presidential	Presidential		Presidential	
Election year	2008	2007		2010	
Conv. Coef. (Conf. Int.)	2.42 (2.0,2.9)	1.83 (1.4,2.3)		1.44 (0.1,3.0)	
Converge to mean	No	Likely		No	
	President	President (2008)		President (2008)	
No. of parties	8	4		7	
env	2.76	1.88		1.31	
	Parliamentary	Duma (2007)		National Assembly (2010)	
No. of parties	5	7		12	
env	2.56	2.22		4.74	
ens	1.55	1.94		2.27	

[a]Conv. Coef. = the central estimate of the convergence coefficient.
[b]Conf. Int. = confidence interval rounded to the nearest tenth.
[c]Based on the number of parties who won votes in the election.
[d]Based on the number of parties who obtained seats in the election.
[e]This was the first time the Prime Minister was elected on a ballot separate from the Knesset.

near or far from the electoral mean for countries with plurality and proportional representation systems and in anocracies. Using this framework we estimated the convergence coefficient for various elections in different countries. We will now use this dimensionless coefficient to compare convergence to the electoral mean across elections, countries and political systems. We can then illustrate the use of the convergence coefficient to classify political systems. Table 7 presents a summary of the convergence coefficients across elections, countries and political systems that we now discuss.

As Table 7 indicates, the two countries using plurality systems (the United States and the UK) studied in Section 3.1 meet the conditions for convergence to the mean. Thus, suggesting that plurality rule imposes a strong centripetal tendency that keeps parties close to the mean. Our analysis suggests that in countries with plurality systems the convergence coefficient will be low, at or below the dimension of the policy space.

Of the anocratic countries that we studied in Section 3.3, Georgia seems to have the highest convergence coefficient, $c^{G}_{2008} = 2.42$ in equation (45) which is not significantly different from two, suggesting that parties can diverge from the mean.[20] The convergence coefficient of all three anocracies was not significantly different from the dimension of the policy space (two for Georgia and Russia and one for Azerbaijan: $c^{G}_{2008} = 2.42$ given in equation (45), $c^{R}_{2007} = 1.83$ in equation (48) and $c^{A}_{2010} = 1.44$ in equation (51)). These results suggest that convergence in anocracies is fragile and depends on the distribution of voters' preferences, as well as on the valences and thus on the popularity of both the autocrat and the opposition parties.

The countries with proportional systems studied in Section 3.2 have convergence coefficients that are significantly above their two-dimensional policy space signaling the lack of convergence of small valence parties to the electoral mean (from Table 7, Israel's $c^{I}_{1996} = 4.06$, Turkey's $c^{T}_{1999} = 1.49$ in 1999 and $c^{T}_{2002} = 5.94$ in 2002 and Poland's $c^{P}_{1997} = 6.82$). Having no possibility of forming a government, these small parties maximize their vote shares by locating closer to their core supporters. Elections lead to multi-party legislatures producing a highly fragmented party system where coalition governments are the norm. Note that changes to the electoral process in Turkey between 1999 and 2002 forced parties to move from locating close to the mean in 1999 to diverging towards their partisan constituencies so as to increase their vote shares in 2002. These results suggest that in countries with proportional systems, with highly fragmented political parties, divergence from the mean is the norm.

We can explain the lack of convergence to the mean in proportional systems with multi-party (>3) legislatures by noting that the convergence coefficient $c \equiv c(\lambda, \beta, \sigma^2) = 2\beta[1-2\rho_1]\sigma^2$ in (9) depends on fundamental characteristics of the electorate and on the party's valences, that is, on voters' beliefs on the party's ability to govern. These characteristics include the weight given by voters to their distance from parties' positions, β, the electoral variance, σ^2 in equation (10), and the probability that a voter chooses the lowest valence party when all parties locate at the mean, ρ_1 in equation (8) which in turn depends only on the valence of the other parties relative to that of the lowest valence party.

Thus in countries with many parties, the smallest low valence parties have little chance of receiving much support if they locate at the mean, i.e. a low ρ_1. If, in addition, voters care a lot about policy differences (a high β) and if the electorate is very dispersed (a high σ^2), then small parties will have an incentive to move towards their core supporters and away from the mean in order to increase the probability that the core supporters vote for them. That is, in highly fragmented polities where voters and correspondingly parties are very dispersed, we observe high convergence coefficients.

In essence, the Valence Theorem in Schofield (2007) gives a simple summary statistic, the convergence coefficient, that measures the degree of fragmentation, or lack thereof, in each polity by taking into account the characteristic of the electorate and voters' beliefs on each party's ability to govern, the party's valence. Poland, an extreme case of this fragmentation, has a very high convergence coefficient (see Table 7).

Other measures of political fragmentation in the literature include the *effective number of party vote strength* (*env*) used by Laakso and Taagepera (1979) to measure how many dominant parties there are in a polity in a given election. To find the *env*, let the Herfindahl index of the election be given by

$$H_v = \sum_{j=1}^{p} v_j^2$$

where v_j is the vote share of party j for $j = 1,\ldots,p$. This Herfindahl index H_v gives a measure of the size of parties in an election, and measures how competitive the election was. Laakso and Taagepera's *effective number of party vote strength* is then the inverse of H_v, i.e.

$$env = H_v^{-1}$$

In the same way we can define the *effective number of party seat strength* (*ens*) using seat shares instead of vote shares giving us a measure of the strength of parties in a legislature.

We calculate the *env* and *ens* for each election we consider (see Table 7) using all the parties that obtained votes in each election and exclude parties that ran in the election but that got no votes. We now compare the level of fragmentation given by the *env* and *ens* with that given by the convergence coefficient for each country and each election under the three political systems that we studied.

We first examine countries with plurality rule. In Table 7 we see that for the United States, the *env* and the *ens* at the Presidential and House levels are closely aligned. There is little variation between the *env* and *ens* indices in the three elections. According to these indices there is essentially no change in political fragmentation across these three elections. The convergence coefficient, however, rises in 2008 relative to 2000 and 2004 indicating that in 2008 the dispersion among voters was higher than in the previous two elections. For the United States, the convergence coefficient provides more information than do *env* or *ens*. For the UK, the convergence coefficient shows that the electorate was more dispersed in 2010 than in 2005 (see Tables 2 and 7). This dispersion led to the first minority government

since 1974 which resulted in a higher effective number of parties as measured by the *env* and *ens*. All three measures, *c*, *env* and *ens*, indicate that the UK became more fragmented in 2010. Thus, in countries using plurality, the convergence coefficient tends to provide more information than the *env* and *ens* numbers do as the convergence coefficient takes into account the degree of dispersion among the electorate and the valence of parties.

Polities with high convergence coefficients (Israel, Turkey in 2002 and Poland in Table 7), had a large number of parties competing in these elections. The greater the number of parties obtaining votes, and thus effectively competing in the election, led to large *env* values. These elections produced highly fragmented legislatures leading to very high *ens* values. Having a large number of effective parties competing in the election and greater effective number of parties in the legislature does not necessarily translate into a higher convergence coefficient. The convergence coefficient is lower for Israel with a larger number of effective parties (higher *env* and *ens*) than for Poland with fewer parties. Changes in the Turkish electoral system between 1999 and 2002, in which a minimum cut-off rule was instituted, led to a high *env* but a low *ens*. Small parties were, however, able to gain enough votes leading to a high convergence coefficient, an indication that these parties would disperse themselves in the policy space. The *env* and *ens* values of the 2002 Turkish election show high party fragmentation but no legislative fragmentation. This shows that these three measures of fragmentation provide different information about a particular election.

The convergence coefficient suggests a way of interpreting the arguments of Duverger (1954) and Riker (1953) on the effects of proportional electoral methods on electoral outcomes: the strong *centrifugal* tendency pulling all parties away from the electoral mean towards their core constituency. This tendency will be particularly strong for small, or low valence, parties. In particular, even small parties in such a polity can assign a non-negligible probability to becoming a member of a coalition government, and it is this phenomenon that maintains the fragmentation of the party system. For example, in Poland, no party can obtain a majority, and parties and coalitions regularly form and dissolve. In general, the convergence coefficients in Poland were of the order of 6.0 in the elections in the 1990s.

For countries using proportional representation, while the *env* and *ens* give a measure of electoral and legislative dispersion, the convergence coefficient provides a measure that summarizes dispersion across voters and parties in the policy space.

In the anocratic countries studied, the convergence coefficient seems in line with the *env* in presidential elections but going in the opposite direction in parliamentary elections (see Table 7). In these countries, the convergence coefficient does not meet the necessary condition for convergence to the mean. These countries show that parties could either converge to or diverge from the mean under anocracy as the equilibrium is fragile. Changes in valences, e.g. of the autocrat or in voters' preferences, can lead small valence opposition parties to diverge from the mean and to mount popular uprisings as has happened in previous elections in Georgia or in the recent Arab uprisings.

The convergence coefficient reflects information that the *env* and *ens* cannot capture as it reflects the preferences of the electorate through the policy weight, β, the perceived ability of parties or candidates to govern as captured by their valences $\lambda = (\lambda_1,\ldots,\lambda_p)$, and the dispersion of voters' preferences in the policy space, σ^2, all of which are not taken into account in the *env* and *ens*. Moreover, *env* and *ens* have nothing to say about the dispersion in parties' positions relative to the mean.

The analysis carried out in this section suggests that there is an inverse relationship between the degree of fractionalization in a polity and the convergence coefficient. By our interpretation of the nature of the convergence coefficient, the convergence effect in presidential elections in the United States is stronger than in parliamentary elections in the UK. That is, our results suggest that democratic presidential systems have fewer parties and a low convergence coefficient. Parliamentary democracies operating under plurality rule tend to have more parties than presidential democracies and a somewhat higher convergence coefficient. Parliamentary democracies operating under proportional representation tend to have multi-party legislatures and high convergence coefficients. Anocratic countries tend to have multiple parties competing in the election but low convergence coefficients as opposition parties remain close to the electoral mean when presidents/autocrats have high valences and diverge when they do not.

5. Conclusion

The analysis carried out in this paper illustrates the application of Valence Theorem in Schofield (2007) together with MNL models of elections. The model in Schofield (2007) serves as a unifying framework that allows us to do a cross-election, cross-country and cross-political system comparison of party's tendencies to converge to the electoral mean in these elections. The evidence presented supports the hypothesis that the fractionalized political system of countries with proportional representation led parties to locate far from the electoral mean.

The convergence coefficient is also compared to the effective number of parties as measured according to the vote (*env*) and seat (*ens*) shares. Our analysis highlighted the importance that not only the characteristics of the political system but also the electorate have on the party's location decision. We show that the convergence coefficient is a simple, intuitive summary statistic that incorporates the preferences of voters, the valence of parties and the dispersion of voters and parties in the policy space. As such the convergence coefficient takes into account more information than do the *ens* and *env* measures developed by Laakso and Taagepera (1979).

Funding

This work was supported by the NSF (grant numbers 0715929 and 1459286) and by the Weidenbaum Center.

Notes

1. The electoral mean is the mean of the voters ideal policies dimension by dimension.
2. See Schofield (2007) for the proof of this result.
3. The set of such local Nash equilibria contains the set of Nash equilibria.
4. Anocracies are those countries in which the president/autocrat governs alongside a legislature and exerts undue influence in the elections and in the legislature.
5. Fragmentation can be identified with the *effective number*. Let H_v (the Herfindahl index) be the sum of the squares of the relative vote shares and $env = H_v^{-1}$ be the *effective number of the party vote strength*. Similarly define *ens* as the effective number of the party seat strength using seat shares. See Laakso and Taagepera (1979).
6. We will use candidate, party and agents interchangeably throughout this paper.
7. See, for example, the models of McKelvey and Patty (2006). Note that there is a problem with the independence of irrelevant alternatives assumption which can be avoided using a probit model (Dow and Endersby, 2004). However Quinn et al. (1999) have shown that probit and logit models tend to give very similar results. Indeed the results given here for the logit model carry through for the probit model, although they are less elegant.
8. The electoral mean or origin is the mean of all voters' positions, $(1/n)\Sigma x_i$ normalized to zero, so that $(1/n)\Sigma x_i = 0$.
9. For example if all voters are equally likely to vote for j, say with probability v, then the weight party j gives to voter i in its vote maximizing policy is $1/n$, i.e. the weight j gives each voter is just the inverse of the size of the electorate.
10. ∇ is simply a description of the distribution of voter preferred points taken about the electoral mean.
11. Since these elections were studied elsewhere, here we present the convergence coefficients and their confidence intervals.
12. Relative to the convergence coefficient of other countries included in this study. In Section 4 we discuss how the values of the convergence coefficient are related to the political systems under which the countries operate.
13. See Schofield et al. (2011b,c) for the list of survey questions, the factor loadings and the full analysis of these US elections.
14. The full analysis of the 2005 and 2010 elections in the UK can be found in Schofield et al. (2011a).
15. These include parties on the left, on the center, on the right, as well as religious parties. On the left there are Labor, Meretz, Democrat, Communists and Balad; those in the center include Olim, Third Way, Center, Shinui; those on the right Likud, Gesher, Tsomet and Yisrael. The religious parties are Shas, Yahadut, NRP, Moledet and Techiya.
16. See Schofield et al. (2011f) for details of the estimation.
17. For simplicity, the pro-Kurdish party is denoted HADEP in the various Figures and Tables. Notice that the HADEP position in Figures 8 and 9 is interpreted as secular and non-nationalistic.
18. In this election Poland used an open-list proportional representation electoral system with a threshold of 5% nationwide vote for parties and 8% for electoral coalitions. Votes are translated into seats using the D'Hondt method.
19. The term 'partial democracy' has been applied to new democracies lacking the full array of democratic institutions present in western democracies (see Epstein et al. 2006).

20. Note that prior to 2008 Georgians had already brought about three major political changes through mass popular revolt. This rebellious 'tradition' may give opposition candidates the ability to position themselves away from the mean.

References

Aragones E and Palfrey TR (2002) Mixed equilibrium in a Downsian model with a favored candidate. *Journal of Economic Theory* 103(1): 131–161.

Aragones E and Palfrey TR (2005) Electoral competition between two candidates of different quality: the effects of candidate ideology and private information. In: Austen-Smith D and Duggan J (eds) *Social Choice and Strategic Essays in Honor of Jeffrey Banks*. Berlin: Springer, pp.93–112.

Arian A and Shamir M (1999). *The Election in Israel*: 1996. Albany: SUNY Press.

Clarke HD (1991) *Absent Mandate: Interpreting Change in Canadian Elections*. New York: Gage.

Clarke HD, Kornberg A, MacLeod J, et al. (2005a) Too close to call: political choice in Canada, 2004. *PS: Political Science and Politics* 38: 247–253.

Clarke HD, Kornberg A, Scotto TJ, et al. (2006a) Flawless campaign, fragile victory: voting in Canada's 2006 federal election. *PS: Political Science and Politics,* 39(4): 815–819.

Clarke HD, Kornberg A and Scotto T (2009a) *Making Political Choices*. Toronto: Toronto University Press.

Clarke HD, Sanders D, Stewart M, et al. (2005b) *Political Choice in Britain*. Oxford: Oxford University Press.

Clarke HD, Sanders D, Stewart M, et al. (2006b) Taking the bloom off New Labour's rose: party choice and voter turnout in Britain, 2005. *Journal of Elections, Public Opinion and Parties,* 16(1): 3–36.

Clarke HD, Sanders D, Stewart M, et al. (2009b) *Performance Politics and the British Voter*. Cambridge: Cambridge University Press.

Clarke HD, Scotto T and Kornberg A (2011) Valence politics and economic crisis: electoral choice in Canada 2008. *Electoral studies* 30(3): 438–449.

Dow JK and Endersby J (2004) Multinomial probit and multinomial Multinomial probit and multinomial logit: a comparison of choice models for voting research. *Electoral studies* 23: 107–122.

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

Duverger M (1954) *Political Parties: Their Organization and Activity in the Modern State*. New York: Wiley.

Enelow JM and Hinich MJ (1982) Nonspatial candidate characteristics and electoral competition. *The Journal of Politics* 44: 115–131

Enelow JM and Hinich MJ (1984) *The Spatial Theory of Voting*. Cambridge: Cambridge University Press.

Enelow JM and Hinich MJ (1989) A general probabilistic spatial theory of elections. *Public Choice* 61: 101–114.

Epstein D, Bates R, Goldstone J, et al. (2006) Democratic transitions. *American Journal of Political Science* 50: 551–568.

Gallego M, Schofield N, McAlister K, et al. (2014) The variable choice set logit model applied to the 2004 Canadian election. *Public Choice* 158 (3-4): 427–463.

Groseclose T (2001) A model of candidate location when one candidate has a valence advantage. *American Journal of Political Science* 45(4): 862–886.

Laakso M and Taagepera R (1979) Effective number of parties: a measure with applications to West Europe. *Comparative Political Studies* 12: 3–27.

McKelvey RD and Patty JW (2006) A theory of voting in large elections. *Games and Economic Behavior* 57: 155–180.

Peress M (2010) The spatial model with non-policy factors: a theory of policy-motivated candidates. *Social Choice and Welfare* 34(2): 265–294.

Quinn K, Martin A and Whitford A (1999) Voter choice in multiparty democracies. *American Journal of Political Science* 43: 1231–1247.

Riker WH (1953) *Democracy in the United States*. New York: New York.

Roemer JE (2001) *Political Competition: Theory and Applications*. Cambridge, MA: Harvard University Press.

Roemer JE (2011) A theory of income taxation where politicians focus upon core and swing voters. *Social Choice and Welfare* 36(3–4), 383–422.

Sanders D, Clarke HD, Stewart MC, et al. (2011) Downs, stokes and the dynamics of electoral choice. *British Journal of Political Science* 41: 287–314.

Schofield N (2003) Valence competition in the spatial stochastic model. *Journal of Theoretical Politics* 15(4): 371–383.

Schofield N (2004) Equilibrium in the spatial valence model of elections. *Journal of Theoretical Politics* 16(4):447–481.

Schofield N (2005) A valence model of political competition in Britain: 1992–1997. *Electoral Studies* 24: 347–370.

Schofield N (2006) Equilibria in the Spatial Stochastic Model of Voting with Party Activists. *Review of Economic Design* 10: 183–203.

Schofield N (2007) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. *The Review of Economic Studies* 74: 965–980.

Schofield N and Sened I (2006) *Multiparty Democracy: Elections and Legislative Politics*. Cambridge: Cambridge University Press.

Schofield N and Zakharov A (2010) A stochastic model of the 2007 Russian Duma election. *Public Choice* 142: 177–194.

Schofield N, Gallego M and Jeon JS (2011a) Leaders, voters and activists in the elections in Great Britain 2005 and 2010. *Electoral Studies* 30: 484–496.

Schofield N, Claassen C, Gallego M, et al. (2011b) Empirical and formal models of the US Presidential Elections in 2004 and 2008. In: Schofield N and Caballero G (eds) *The Political Economy of Institutions, Democracy and Voting*. Berlin: Springer, pp.217–258.

Schofield N, Claassen C, Ozdemir U, et al. (2011c). Estimating the effects of activists in two-party and multi-party systems: comparing the United States and Israel. *Social Choice and Welfare* 36: 483–518.

Schofield N, Gallego M, Ozdemir U, et al. (2011f) Competition for popular support: a valence model of elections in Turkey. *Social Choice and Welfare* 36: 451–482.

Schofield N, Jeon JS, Muskhelishvili M, et al. (2011g) Modeling elections in post-communist regimes: voter perceptions, political leaders and activists. In: Schofield N and Caballero G (eds) *The Political Economy of Institutions, Democracy and Voting*. Berlin: Springer, pp.259–301.

Schofield N, Gallego M, Jeon JS, et al. (2012). Modeling elections in the Caucasus. *Journal of Elections. Public Opinion and Parties* 22(2): 187–214.

Scotto TJ, Clarke HD, Kornberg A, et al. (2010) The dynamic political economy of support for Barack Obama during the 2008 presidential election campaign. *Electoral Studies* 29: 545–556.

Stokes D (1963) Spatial models and party competition. *American Political Science Review* 57: 368–377.

Stokes D (1992). Valence politics. In: Kavanagh D (ed.) *Electoral Politics*. Oxford: Clarendon Press, pp.141–164.

Train K (2003). *Discrete Choice Methods for Simulation*. Cambridge: Cambridge University Press.