Hindawi Publishing Corporation
The Scientific World Journal
Volume 2013, Article ID 653035, 28 pages
http://dx.doi.org/10.1155/2013/653035

*Research Article*

# The Convergence Coefficient across Political Systems

**Maria Gallego[1,2] and Norman Schofield[1]**

[1] *Center in Political Economy, Washington University, 1 Brookings Drive, Saint Louis, MO 63130, USA*
[2] *Department of Economics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, ON, Canada N2L 3C5*

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

Received 4 August 2013; Accepted 21 August 2013

Academic Editors: M. Kohl and J. Pacheco

Copyright © 2013 M. Gallego and N. Schofield. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Formal work on the electoral model often suggests that parties or candidates should locate themselves at *the electoral mean*. Recent research has found no evidence of such convergence. In order to explain nonconvergence, the stochastic electoral model is extended by including estimates of electoral valence. We introduce the notion of a convergence coefficient, $c$. It has been shown that high values of $c$ imply that there is a significant centrifugal tendency acting on parties. We used electoral surveys to construct a stochastic valence model of the the elections in various countries. We find that the convergence coefficient varies across elections in a country, across countries with similar regimes, and across political regimes. In some countries, the centripetal tendency leads parties to converge to the electoral mean. In others the centrifugal tendency dominates and some parties locate far from the electoral mean. In particular, for countries with proportional electoral systems, namely, Israel, Turkey, and Poland, the centrifugal tendency is very high. In the majoritarian polities of the United States and Great Britain, the centrifugal tendency is very low. In anocracies, the autocrat imposes limitations on how far from the origin the opposition parties can move.

## 1. Introduction

Work on modeling elections has often assumed that the policy space was restricted to one dimension or that there were at most two parties [1, 2]. The extensive formal literature on electoral competition has typically been based on the assumption that parties or candidates adopt positions in order to win and has inferred that parties will converge to the electoral median, under deterministic voting in one dimension, or to the electoral mean in stochastic models.

In this paper we offer a formal stochastic model of elections that emphasizes the importance of the idea of *valence* and use this notion to provide an explanation of why vote maximizing political leaders in some countries will not adopt convergent policy positions at the mean of the electoral distribution. In the standard spatial model, candidate *positions* matter to voters. However, as Stokes [3, 4] has emphasized, the nonpolicy evaluations, or *valences*, of candidates by the electorate are just as important. (See also Clarke et al. [5, 6], Scotto et al. [7], and Clarke et al. [8].)

The main objective of this paper is to examine whether parties locate close to or far from the electoral mean (the electoral mean is the mean of voters ideal policies dimension by dimension) of various countries. We use Schofield's [9] stochastic electoral model as a unifying framework that allows us to compare parties positions across different political systems. In this model, parties respond to their partisan constituencies after taking into account the anticipated electoral outcome and the positions of other parties. Voters decisions depend on parties' locations and on the party's valence, the voters' *overall common* evaluation of the ability of a party leader to provide good governance. Using this model we examine whether parties converge to electoral mean in several elections in various countries under different political systems and use convergence, or the lack thereof, to classify political systems.

To examine whether parties converge to the electoral mean in each country in a particular election, we test whether any party has an incentive to stay or move away from the electoral mean to increase its vote share. In formal voting theory, it is usual to define a "Nash equilibrium" as a vector of party positions with the property that no party may make a unilateral move so as to increase its vote share. We use a variant of this concept, that is a "local Nash equilibrium"

(LNE) where we consider only *marginal* moves from the position. One of the standard results in formal theory is the *mean voter theorem*, where the "Nash equilibrium" of a spatial voting game under vote maximization is one where all parties position themselves at the electoral mean. (For variants of the theorem see Enelow and Hinich [10–12]). We call such a vector the *electoral mean*.

To study each party's best response to the electoral situation they face, we use the results presented in Schofield [9]. Schofield identifies a *convergence coefficient*, denoted $c$, whose value determines whether vote maximizing parties converge or not to the electoral mean. This coefficient depends on various parameters of the model. In particular it depends on the *competence valences* of the party leaders. Using $j \in P = 1, \ldots, p$ to denote the parties, the valence of party $j$, $\lambda_j$, essentially measures the electoral perception of the "quality" of $j$, the voters' *overall common* evaluation of the ability of $j$'s leader to provide good governance. The valence terms, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)$, are assumed to be independent of the party's positions and can be estimated as the intercept term in the appropriate stochastic model of the voter utility function. As Sanders et al. [13] comment, valence theory is based on the assumption that "voters maximize their utilities by choosing the party that is best able to deliver policy success." These valence terms measure the *bias* in favor of one another of the party leaders [14].

The convergence coefficient, $c$, also depends on the weight that voters give to the policy differences they have with the various parties $\beta$. Lastly, $c$ depends on the variance/covariance matrix of the electoral distribution, $\sigma^2$. By its construction, $c \equiv c(\boldsymbol{\lambda}, \beta, \sigma^2)$ is dimensionless and thus independent of the units of measurement of the various parameters. We use the convergence coefficient to compare results across elections and countries and to classify political systems.

The convergence coefficient is a summary measure that provides an estimate of the centrifugal or centripetal forces acting on the parties. The *Valence Theorem*, presented in Section 2 (see Schofield [9] for the proof of this result) shows that if the policy space is two-dimensional and if $c(\boldsymbol{\lambda}, \beta, \sigma^2) < 1$, then the sufficient condition for convergence to the mean has been met and the "local Nash Equilibrium" (LNE) (the set of such local Nash Equilibria contains the set of Nash Equilibria) is one where all parties locate at the electoral mean. On the other hand, if $w$ is the dimension of the policy space and $c(\boldsymbol{\lambda}, \beta, \sigma^2) \geq w$, then the LNE, if it exists, will be one where at least one party will have an incentive to diverge from the electoral mean in order to maximize its vote share. Thus, the necessary condition for convergence to the mean is that $c(\boldsymbol{\lambda}, \beta, \sigma^2) < w$.

In essence, a high empirical convergence coefficient of an election is a convenient measure of the electoral incentive of a small, or low valence, party to move away from the electoral mean to its core constituency position. We can interpret a high value of the convergence coefficient as a measure of the *centrifugal* tendency exerted on parties pulling them away from the electoral mean. The convergence coefficient is therefore a convenient, simple, and intuitive way to examine whether parties will have an incentive to locate close to, or far from, the electoral mean. We will 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.

We used preelection polls to study elections in several countries operating under different political regimes. The factor analysis done on preelection surveys showed that for all elections the policy space was two-dimensional, except in Azerbaijan were it was one-dimensional. The position of voters along this two-dimensional space were then estimated and their voting intentions used to estimate party positions. We then ran a multinomial logit (MNL) model for the election using the estimated party and voter positions. The intercept of the MNL model gives the valences of each party/leader. Following Schofield's [9] formal model, we rank parties according to their valence. Using these MNL estimates we calculate the convergence coefficient of the election and examine whether the party with the lowest valence has an incentive to locate close to or far from the electoral origin.

When comparing the convergence coefficients across countries we observe that in countries with proportional representation the convergence coefficient is high and that in countries with plurality systems or in anocracies it is low. Thus, suggesting that we can use the valence theorem and its associated convergence result to classify electoral systems.

The convergence coefficients for the 2005 and 2010 elections in the UK were not significantly different from 1, meeting the necessary condition for convergence to the mean. For the 2000, 2004, and 2008 US presidential elections, the convergence coefficient is significantly below 1 in 2000 and 2004 thus meeting the sufficient and thus necessary condition for convergence and not significantly different from 1 in 2008, only meeting the necessary condition for convergence. We suggest that the centrifugal tendency in the majoritarian polities like the United States and the United Kingdom is very low.

In contrast, the convergence coefficient gives an indication that the centrifugal tendency in Israel, Poland, and Turkey is very high. In these proportional representation systems with highly fragmented polities the convergence coefficients are significantly greater than 2 failing to meet the necessary condition for convergence to the electoral mean.

In the anocracies of Georgia, Russia, and Azerbaijan, where the President/autocrat dominates and controls who can run in legislative elections, the convergence coefficient is not significantly different from the dimension of the policy space (2 for Georgia and Russia and 1 for Azerbaijan), failing the necessary condition for convergence. While the analysis Georgia and Azerbaijan show that not all parties converge to the mean, in Russia it is likely that they did. Thus, in Russia opposition parties found it difficult to diverge from the mean. Note that convergence in anocracies may not generate a stable equilibrium as any change in the valence of the autocrat and the opposition may cause parties to diverge from the mean and may even lead to popular uprising that bring about changes in the governing parties such as in Georgia in previous elections or in the Arab revolutions.

We can also classify polities using the *effective vote number* and the *effective seat number.* (Fragmentation can be identified with the *effective number*. That is, let $H_v$ (the Herfindahl index) be the sum of the squares of the relative vote shares and let $env = H_v^{-1}$ be the *effective number of party vote strength*. In the same way we can define *ens* as the effective number of party seat strength using shares of seats. See Laakso and Taagepera [15].) We examine how these two measures of fragmentation relate to the convergence coefficient for the polities we consider. The effective vote or seat numbers give an indication of the difficulty inherent in interparty negotiation over government. These two measures do not, however, address the fundamental aspect of democracy, namely, the 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 allow for a more comprehensive way of classifying polities and political systems precisely because it 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 Valence Theorem and the convergence result can help us understand why parties locate close to or far from the electoral mean and how, under some circumstances, these considerations lead to political fragmentation.

The next section presents Schofield's [9] stochastic formal model of elections and implications it has for convergence to the mean. Section 3.1 applies the model to the elections to two plurality polities: The United States and the United Kingdom. In Section 3.2 we apply the model to polities using proportional electoral systems, 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. In the appendix we estimate the confidence intervals for the convergence coefficient as well as determining whether the low valence party has an incentive to deviate from the electoral mean.

## 2. The Spatial Voting Model with Valence

Recent research on modelling elections has followed earlier work by Stokes [3, 4] and emphasized the notion of valence of political candidates. As Sanders et al. [13] 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 [16] 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.

Valence theory has led to a considerable theoretical literature on voting based on the assumption that valence plays an important role in the relationship between party positioning and the votes that parties receive. (Ansolabare and Snyder [17], Groseclose [18], Aragones and Palfrey [19, 20], Schofield [21], Schofield et al. [22], Miller and Schofield [23], Schofield and Miller [24], Peress [25]) Empirical work, based on multinomial logit (MNL) methods, has also shown the importance of electoral judgements in analyses of elections in the United States and the United Kingdom. (Clarke et al. [8, 26–28], Schofield [29], Schofield et al. [30, 31], Scotto et al. [7]) These empirical models of elections have a "probabilistic" component. That is, they all assume that "voter utility" is partly "Downsian" in the sense that it is based on the distance between party positions and voter preferred positions and partly due to valence. The estimates of a party's valence is assumed to be subject to a "stochastic error." In this paper we use the same methodology.

The pure "Downsian" spatial model of voting tends to predict that parties will converge to the center of the electoral distribution [10–12]. 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 indicates that voters have a bias towards party A and as a consequence more voters will choose A over B. The question for B is whether it can gain votes by moving away from the center. It should be obvious that 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 find the optimal party positions to the anticipated electoral outcome we use a Downsian vote model that has a valence component as presented in Schofield [9]. Let the set of parties be denoted by $P = 1, \ldots, p$. The positions of the $p$ parties (We will use candidate, party and agents interchangeably throughout the paper.) in $X \subseteq \mathbb{R}^w$ where $w$ is the dimension of the policy space it is given by the vector

$$\mathbf{z} = \left(z_1, \ldots, z_j, \ldots, z_p\right) \in X^p. \tag{1}$$

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

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

Here $u_{ij}^*(x_i, z_j)$ is the observable component of the utility voter $i$ derived from party $j$. The competence valence of candidate $j$ is $\lambda_j$, and the competence valence vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_p)$ is such that $\lambda_p \geq \lambda_{p-1} \geq \cdots \geq \lambda_2 \geq \lambda_1$, so that party 1 has the lowest valence. Note that $\lambda_j$ is the same for all voters and provides an estimate of the "quality" of party $j$ or its ability to govern. The term $\|x_i - z_j\|$ is simply the Euclidean distance between voter $i$'s position $x_i$ and candidate $j$'s position $z_j$. The coefficient $\beta$ is the weight given to this policy difference. The error vector $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_j, \ldots, \epsilon_p)$ has a Type I extreme value distribution, where the variance of $\epsilon_j$

is fixed at $\pi^2/6$. Note that $\beta$ has dimension $1/L^2$ , where $L$ is whatever unit of measurement 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

$$\begin{aligned}\rho_{ij}(\mathbf{z}) &= \Pr\left[u_{ij}\left(x_i, z_j\right) > u_{il}\left(x_i, z_l\right), \forall l \neq j\right] \\ &= \Pr\left[\epsilon_l - \epsilon_j < u_{ij}^*\left(x_i, z_j\right) - u_{il}^*\left(x_i, z_j\right), \forall l \neq j\right].\end{aligned} \tag{3}$$

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

Train [32] showed that when the error vector $\boldsymbol{\epsilon}$ has a Type I extreme value distribution, the probability $\rho_{ij}(z)$ has a Multinomial Logit (MNL) specification and can be estimated. Thus, for each voter $i$ and party $j$, the probability that voter $i$ chooses party $j$ at the vector $\mathbf{z}$ is given by

$$\rho_{ij}(\mathbf{z}) = \frac{\exp\left[u_{ij}^*\left(x_i, z_j\right)\right]}{\sum_{k=1}^{p} \exp u_{ik}^*\left(x_i, z_k\right)}. \tag{4}$$

Voters decisions are stochastic in this framework. (See, for example, the models of McKelvey and Patty [14]. Note that there is a problem with the independence of irrelevant alternatives assumption (IIA) which can be avoided using a probit model [33]. However, Quinn et al. [34] 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 probit, though they are less elegant.) 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 as follows:

$$V_j(\mathbf{z}) = \frac{1}{n} \sum_{i \in N} \rho_{ij}(\mathbf{z}). \tag{5}$$

We assume a 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 [35].

We assume that parties can estimate how their vote shares would change if they *marginally* move their policy position. The Local Nash Equilibrium (LNE) is that vector $\mathbf{z}$ of party positions such 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 SLNE to be a vector that *strictly* locally maximizes $V_j(\mathbf{z})$.

Using the estimated MNL coefficients we simulate these models and 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. (The electoral mean or origin is the mean of all voters' positions, $(1/n)\sum x_i$ normalized to zero, so that $(1/n)\sum x_i = 0$.) We call this vector *the electoral mean*.

Given the vector of policy position $\mathbf{z}$, and since the probability that voter $i$ votes for party $j$ is given by (4), the impact of a *marginal* change in $j$'s position on the probability that $i$ votes for $j$ is then

$$\left.\frac{d\rho_{ij}(\mathbf{z})}{dz_j}\right|_{\mathbf{z}_{-j}} = 2\beta\rho_{ij}\left(1 - \rho_{ij}\right)\left(x_i - z_j\right), \tag{6}$$

where $\mathbf{z}_{-j}$ indicates that we are holding the positions of all parties but $j$ is fixed. The effect that $j$'s change in position has on the probability that $i$ votes for $j$ depends on the weight given to the policy differences with parties, $\beta$; on how likely is $i$ to vote for $j$, $\rho_{ij}$, and for any other party, $(1 - \rho_{ij})$ and on how far apart $i$'s ideal policy is from $j$'s, $(x_i - z_j)$.

From (5), party $j$ adjusts its position to maximize its expected vote share, that is, $j$'s first order condition is

$$\left.\frac{dV_j(\mathbf{z})}{dz_j}\right|_{z_{-j}} = \frac{1}{n}\sum_{i \in N}\frac{d\rho_{ij}}{dz_j} = \frac{1}{n}\sum_{i \in N} 2\beta\rho_{ij}\left(1-\rho_{ij}\right)\left(x_i - z_j\right) = 0, \tag{7}$$

where the third term follows after substituting in (6). The FOC for party $j$ in (7) is satisfied when

$$\sum_{i \in N} \rho_{ij}\left(1 - \rho_{ij}\right)\left(x_i - z_j\right) = 0 \tag{8}$$

so that the *candidate* for party $j$'s vote maximizing policy (See Schofield [36] for the proof.) is

$$z_j^C = \sum_{i \in N} \alpha_{ij} x_i, \quad \text{where } \alpha_{ij} \equiv \frac{\rho_{ij}\left(1 - \rho_{ij}\right)}{\sum_{i \in N} \rho_{ij}\left(1 - \rho_{ij}\right)}, \tag{9}$$

where $\alpha_{ij}$ represents the weight that party $j$ gives to voter $i$ when choosing its candidate vote maximizing policy. This weight depends on how likely is $i$ to vote for $j$, $\rho_{ij}$, and for any other party, $(1 - \rho_{ij})$ relative to all voters. (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 $1/n$; that is, the weight $j$ gives each voter is just the inverse of the population size.) Note that $\alpha_{ij}$ may be nonmonotonic in $\rho_{ij}$. To see this exclude voter $i$ from the denominator of $\alpha_{ij}$. 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 for sure vote for $j$, $i$ receives a lower weight in $j$'s candidate position 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 on its position. This is the most common case. When $\sum_{a \in N-i} \rho_{aj}(1 - \rho_{aj}) > 2/3$, then $\alpha_{ij}$ increases in $\rho_{ij}$. 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 $\alpha_{ij}$ are endogenously determined in the model.

Note that since voter $i$'s utility depends on how far $i$ is from party $j$, the probability that $i$ votes for $j$ given in (4) and the expected vote share of the party given in (5) 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 (2) we see that

$$\left[u_{ik}^*(x_i, z) - u_{ij}^*(x_i, z)\right] = \left(\lambda_k - \lambda_j\right), \tag{10}$$

so the probability that $i$ votes for $j$ in (4) is given by

$$\begin{aligned}\rho_{ij}(\mathbf{z}) &= \frac{1}{\sum_{k=1}^{p} \exp\left[u_{ik}^*(x_i, z_k) - u_{ij}^*\left(x_i, z_j\right)\right]} \\ &= \left[\sum_{k=1}^{p} \exp\left(\lambda_k - \lambda_j\right)\right]^{-1}.\end{aligned} \tag{11}$$

Clearly, in this case, $\rho_{ij}(\mathbf{z}) = \rho_j(\mathbf{z})$ is independent of voter $i$'s ideal point. Thus, from (9), the weight given by $j$ to each voter is also independent of voter $i$'s position and given by

$$\alpha_j \equiv \frac{\rho_j\left(1 - \rho_j\right)}{\sum_{i \in N} \rho_j\left(1 - \rho_j\right)} = \frac{1}{n}, \tag{12}$$

so that $j$ gives each voter equal weight in its policy position. In this case, from (9), $j$'s candidate position is

$$z_j^C = \frac{1}{n} \sum_{i \in N} x_i, \tag{13}$$

that is, $j$'s candidate position is to locate 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.

Moreover, as (11) indicates when parties locate at the mean $\mathbf{z}_0$, only valence differences between parties matter in voters' choices. The probability that a generic voter votes for party 1 (the party with the lowest valence) is

$$\rho_1 \equiv \rho_1(\mathbf{z}_0) = \left[\sum_{k=1}^{p} \exp\left(\lambda_k - \lambda_1\right)\right]^{-1}. \tag{14}$$

Using this spatial model, Schofield [9] proved a *Valence Theorem* determining whether vote maximizing parties locate at the mean. The theorem showed that the spatial model is characterized by a *convergence coefficient* given by

$$c \equiv c\left(\boldsymbol{\lambda}, \beta, \sigma^2\right) = 2\beta\left[1 - 2\rho_1\right]\sigma^2. \tag{15}$$

The convergence coefficient depends on $\beta$, the weight given to policy differences; on $\rho_1$, the probability that a generic voter votes for the lowest valence party at the vector $\mathbf{z}_0$ and on $\sigma^2$, the *electoral variance* given by

$$\sigma^2 \equiv \operatorname{trace}(\nabla), \tag{16}$$

where $\nabla$ is the symmetric $w \times w$ *electoral covariance matrix*. ($\nabla$ is simply a description of the distribution of voter preferred points taken about the electoral mean.)

The convergence coefficient increases in $\beta$ and $\sigma^2$ (and on its product $\beta\sigma^2$) and decreases in $\rho_1$. As (14) indicates $\rho_1$ decreases if the valence differences between party 1 and the other parties increases, that is, when the difference between $\lambda_1$ and $\{\lambda_2, \ldots, \lambda_p\}$ increases.

The Valence Theorem allows us to characterize polities according to the value of their convergence coefficient. The theorem states that when the *sufficient* condition for convergence to the electoral mean is met, that is, 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 $c \geq w$, then there may exist a nonconvergent LNE. Note that in this case, there may indeed be no LNE. However, there will exist a mixed strategy Nash equilibrium (MNE). In either of these two cases we expect 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, $\rho_1$ in (14) 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 (with the lowest valence when all parties locate at the origin), $\rho_1$ in (14) 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 $c \geq w$, we expect parties to diverge from the electoral center. Indeed, we expect those 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, $\rho_1$ 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 convergence coefficient is a way of characterizing the Hessian (the $w$ by $w$ *second* derivatives of the vote share function) of party 1 with the lowest valence. The Hessian of

the vote share function of party 1 is given by the characteristic matrix

$$C_1 = 2\beta\left(1 - 2\rho_1\right)\nabla - I. \tag{17}$$

Here $I$ is a $w$ by $w$ identity matrix and the other terms are as before. The eigenvalues of $C_1$ determine whether the vote share function of party 1 will be at a maximum, minimum, or at a saddlepoint at the electoral mean. If $C_1$ shows that party 1 is at a minimum or at a saddlepoint at the mean then party 1 has an incentive to locate away from the mean to increase its vote share. When all parties are at the mean and $c < 1$, then all eigenvalues of the Hessian of the vote share function of the lowest valence party are negative indicating that the vote share function is at a maximum. The LNE must then be at the electoral mean.

For an arbitrary dimension, $w$, if $c(\boldsymbol{\lambda}, \beta, \sigma^2) \leq 1$ in (15), then $\operatorname{trace}(C_1) < 0$. In the two-dimensional case, if $c(\boldsymbol{\lambda}, \beta, \sigma^2) < 1$, then $\det(C_1)$ must be positive, implying that both eigenvalues of $C_1$ are negative. It then follows that all $\{C_j\}$ have negative eigenvalues, giving a SLNE and thus an LNE at the electoral mean. (This result follows from the application of the triangle inequality to the determinant. A parallel result can be obtained in more than two dimensions.)

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

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

## 3. MNL Models of the Elections of Various Countries

We use the framework of the spatial model presented in Section 2 as a unifying methodology that allows us to study convergence across elections, countries, and political regimes. The Valence Theorem leads to the convergence coefficient of the election, a summary statistic that determines whether parties converge to or diverge from the electoral mean. Using this formal multinomial (MNL) spatial model, we now estimate the convergence coefficient for the elections in various countries. For each MNL estimation we choose a baseline party and normalize its coefficients to zero, then estimate the coefficients of all other parties relative to those of the base party. Using these coefficients we estimate the convergence coefficient and the characteristic matrix of the low valence parties to determine whether these parties converge to or diverge from the electoral mean in each election for each country. (These elections were studied in depth elsewhere. In this paper, we present only the calculations leading to the convergence coefficient and estimate the confidence intervals for the convergence coefficients that were not provided in earlier work.)

We study convergence under three political regimes (plurality, proportional representation, and anocracy) and group countries according to the similarities of their political regimes. Under plurality rule, we examine elections in two Anglo-Saxon countries: the US and the UK; under proportional representation we study Israel, Turkey, and Poland; and under anocracy, Georgia, Russia, and Azerbaijan. Since we use the same unifying methodology for all countries we present the methodology for the first elections in detail then condense the analysis to its basic components for the remaining countries. For each country we give a general description of the analysis and direct the reader to the full analysis of each election in the detailed country paper. We summarize the results across countries in various tables.

*3.1. Convergence in Plurality Systems.* We begin our analysis by examining the United States and the United Kingdom. Elections in these countries are carried out under plurality rule. We show that the electoral system in these countries produces relatively low convergence coefficients. (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.)

*3.1.1. The 2000, 2004, and 2008 Elections in the United States.* We construct stochastic models of the 2000, 2004, and 2008 US presidential elections using survey data taken from the American National Election Surveys (ANES). The factor analysis done on ten survey questions taken from the ANES (See Schofield et al. [30, 31] for the list of survey questions and the factor loadings and the full analysis of the US elections.) led us to conclude that voters preferences can be represented along the economic ($E = x$-axis) and social ($S = y$-axis) dimensions 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. We interpreted greater values along this axis to mean more support for certain civil rights. Using the factor loadings we estimated each voter's position in these two dimensions. Figures 1, 2, and 3 give a smoothing of the estimated voter distribution of the 2000, 2004, and 2008 elections, respectively.

Voters' ideal points in the *2000 US election* are characterized by the following *electoral covariance matrix*:

$$\nabla_{2000}^{\mathrm{US}} = \begin{bmatrix} \sigma_E^2 = 0.58 & \sigma_{ES} = -0.20 \\ \sigma_{ES} = -0.20 & \sigma_S^2 = 0.59 \end{bmatrix}. \tag{18}$$

The trace of electoral covariance matrix is $\sigma_{\mathrm{US}\,2000}^2 \equiv \operatorname{trace}(\nabla_{\mathrm{US}}^{2000}) = \sigma_E^2 + \sigma_S^2 = 1.17$. Given the negative covariance between these two dimensions, $\sigma_{ES} = -0.20$, the correlation between these two factors is $-0.344$.

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

$$\lambda_{\mathrm{rep}}^{\mathrm{US2000}} = -0.43, \qquad \lambda_{\mathrm{dem}}^{\mathrm{US2000}} \equiv 0.0, \qquad \beta_{2000}^{\mathrm{US}} = 0.82. \tag{19}$$

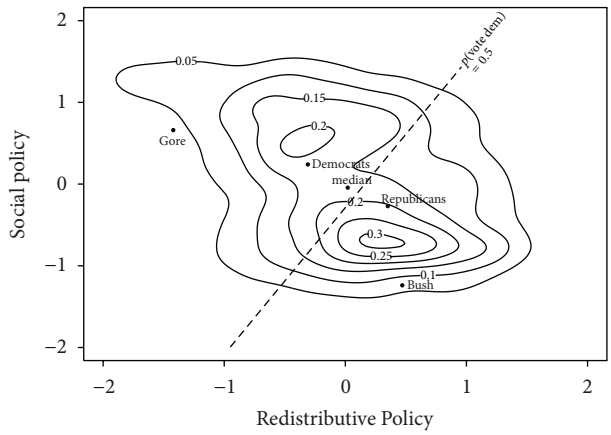



Figure 1: Distribution of voter ideal points and candidate positions in the 2000 US election.

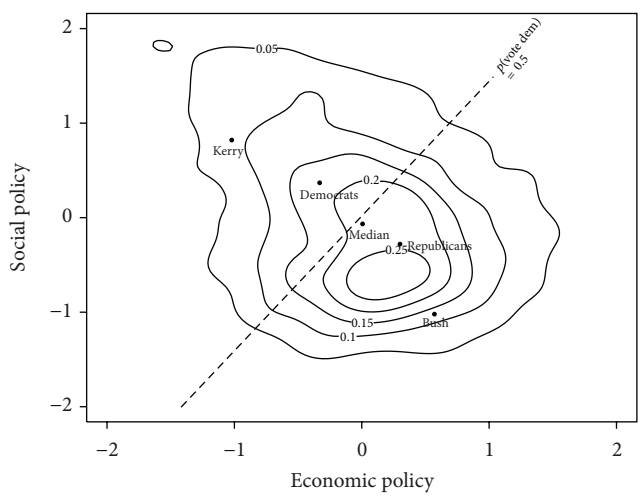



Figure 2: Distribution of voter ideal points and candidate positions in the 2004 US election.

Bush's competence valence, $\lambda_{\text{rep}}^{\text{US2000}} = -0.43$, measures the common perception that voters in the sample have on Bush's ability to govern and represents the nonpolicy component in the voter's utility function in (2). As seen in Table 1, for the 2000 election Bush has a statistically significant *lower* valence than Gore, the democratic (baseline) candidate. Bush's negative valence is an indication that voters regarded him as less able to govern than Gore, once policy differences are taken into account.

To find the convergence coefficient for this election, we assume that *all* parties locate at the electoral mean so that parties differ only in their valence terms (see Section 2). We can use (14) and the coefficients in (19) to estimate the probability that a typical US voter chooses to vote for the low valence Republican candidate, when both Bush and Gore locate at origin, $\mathbf{z}_0$; that is,

$$\begin{aligned}\rho_{\text{rep}}^{\text{US2000}} &= \left[\sum_{k=1}^{2} \exp\left(\lambda_k^{\text{US2000}} - \lambda_{\text{rep}}^{\text{US2000}}\right)\right]^{-1} \\ &= \left[1 + \exp(0.43)\right]^{-1} = 0.40.\end{aligned} \tag{20}$$

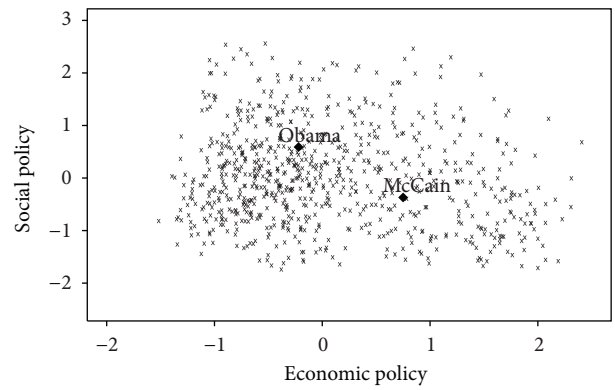



Figure 3: Distribution of voter ideal points and candidate positions in the 2008 US election.

We found the estimate for $\rho_{\text{rep}}^{\text{US2000}}$ using the MNL valence estimates. Note that since the central estimates of $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)$ given by the MNL regressions depend on the sample of voters surveyed then so does $\rho_1$. Thus, to make inferences from empirical models we need the 95% confidence bounds of $\rho_1$. In the introduction of the appendix we derive the methodology used to find the confidence bounds of $\rho_1$. The bounds of $\rho_1$ are calculated in Appendix A.1.

The results indicate that in the 2000 election, both candidates found it in their best interest to locate at the electoral mean. To see this, we compute the convergence coefficient using (15) and the electoral covariance matrix in (18) $\nabla_{\text{US}}^{2000}$ to determine whether the two parties converge to, or diverge from, the electoral mean.

Using (19) and (20) we have that $2\beta_{2000}^{\text{US}}(1 - 2\rho_{\text{rep}}^{\text{US2000}}) = 2 \times 0.82 \times 0.2 = 0.328$ and from (18) the trace is $\sigma_{\text{US2000}}^2 = 1.17$ so that using (15) the convergence coefficient for 2000 US election is

$$c_{\text{US}}^{2000} \equiv 2\beta_{2000}^{\text{US}}\left(1 - 2\rho_{\text{rep}}^{\text{US2000}}\right)\sigma_{\text{US2000}}^2 = 0.328 \times 1.17 = 0.384. \tag{21}$$

Appendix A.1 shows that $c_{\text{US}}^{2000}$ is significantly less than 1 implying that $c_{\text{US}}^{2000}$ meets the sufficient and thus necessary condition for convergence to the electoral mean given in Section 2.

To check whether Bush, the low valence candidate, has an incentive to stay at the electoral origin, $\mathbf{z}_0$, that is, whether Bush's vote share function is at a maximum at $\mathbf{z}_0$, we use the Hessian or characteristic matrix (of *second* order conditions) of Bush's vote share function using (17) at $\mathbf{z}_0$ as follows:

$$\begin{aligned}C_{\text{rep}}^{\text{US2000}} &= \left[2\beta_{2000}^{\text{US}}\left(1 - 2\rho_{\text{rep}}^{\text{US2000}}\right)\right]\nabla_{2000}^{\text{US}} - I \\ &= 0.328\begin{bmatrix} 0.58 & -0.20 \\ -0.20 & 0.59 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.81 & -0.06 \\ -0.06 & -0.81 \end{bmatrix}.\end{aligned} \tag{22}$$

Because the characteristic matrix for Bush $C_{\text{rep}}^{\text{US2000}}$ is estimated using the MNL coefficients of the 2000 US sample,

Table 1: MNL spatial model for countries with plurality systems.

| | United States[b] | | | | United Kingdom[c] | | |
|---|---|---|---|---|---|---|---|
| | Party | 2000 | 2004 | 2008 | Party | 2005 | 2010 |
| Var | | Est.[a] $\lvert t-\text{value}\rvert$ | Est.[a] $\lvert t-\text{value}\rvert$ | Est.[a] $\lvert t-\text{value}\rvert$ | | Est.[a] $\lvert t-\text{value}\rvert$ | Est.[a] $\lvert t-\text{value}\rvert$ |
| $\beta$ | | $0.82^{***}$ (14.9) | $0.95^{***}$ (14.21) | $0.85^{***}$ (14.16) | | $0.15^{***}$ (12.56) | $0.86^{***}$ (38.45) |
| Valence | $\lambda_{\text{rep}}$ | $-0.43^{***}$ (5.05) | $-0.43^{***}$ (5.05) | $-0.84^{***}$ (7.64) | $\lambda_{\text{Lab}}$ | $0.52^{***}$ (6.84) | $-0.04$ (1.31) |
| | | | | | $\lambda_{\text{Con}}$ | $0.27^{***}$ (3.22) | $0.17^{***}$ (4.50) |
| Base party | | Dem[b] | Dem[b] | Rep[b] | | Lib[c] | Lib[c] |
| $n$ | | 1,238 | 935 | 788 | | 1149 | 6218 |
| $LL$ | | −708 | −501 | −298 | | −1136 | −5490 |

[a]$^{*}$prob < 0.05; $^{**}$prob < 0.01; $^{***}$prob < 0.001.
[b]US: Rep: Republican; Dem: Democrats.
[c]UK: Lab: Labour; Con: Conservatives; Lib: Liberal Democrats.

Table 2: The convergence coefficient in plurality systems.

| | United States | | | United Kingdom | |
|---|---|---|---|---|---|
| | 2000 | 2004 | 2008 | 2005 | 2010 |
| | Weight of policy differences ($\beta$) | | | | |
| Est. $\beta$ (conf. Int.[a]) | 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$\nabla = \sigma^2$) | | | | |
| $\sigma^2$ | 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[b] | Dem[b] | Rep[b] | LibDem[c] | Labour[c] |
| Est. $\rho_1$ (conf. Int.[a]) | $\rho_{\text{Dem}} = 0.4$ (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 ($c \equiv c(\boldsymbol{\lambda}, \beta, \sigma^2) = 2\beta[1 - 2\rho_1]\sigma^2$) | | | | |
| Est. $c$ (conf. Int.[a]) | 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]Conf. Int.: confidence intervals.
[b]US: Dem: Democrats; Rep: Republican.
[c]UK: LibDem: Liberal Democrats.

$C_{\text{rep}}^{\text{US2000}}$ depends on the sample of voters surveyed. The confidence bounds on $C_{\text{rep}}^{\text{US2000}}$ in Appendix A.1 suggest that if Bush positions himself at the electoral origin, then with probability exceeding 95%, his vote share function would be at a maximum. We infer that, with probability exceeding 95%, the origin is an LNE for the spatial model for the 2000 US election. The valence differences between Bush and Gore are not large enough to cause either of them to move from the origin. The unique local Nash equilibrium was one where both candidates converge to the electoral origin in order to maximize their vote shares.

All the components needed to derive the convergence coefficient for 2000 US election and its confidence bounds are summarized in Table 2.

Bush faced Kerry as the democratic candidate in the *2004 US election*. The distribution of voters in 2004 gives the following electoral covariance matrix along the economic and social dimensions:

$$\nabla_{2004}^{\text{US}} = \begin{bmatrix} \sigma_E^2 = 0.58 & \sigma_{ES} = -0.177 \\ \sigma_{ES} = -0.177 & \sigma_S^2 = 0.59 \end{bmatrix}. \tag{23}$$

While the covariance between economic and social axes differs, the trace $\sigma_{\text{US2004}}^2 = \text{trace}\,(\nabla_{\text{US}}^{2004}) = \sigma_E^2 + \sigma_S^2 = 1.17$ is similar to that in the 2000 US election.

From Table 1, the MNL estimates of the spatial model for the 2004 US election are

$$\lambda_{\text{rep}}^{\text{US2004}} = -0.43, \qquad \lambda_{\text{dem}}^{\text{US2004}} \equiv 0.0, \qquad \beta_{2004}^{\text{US}} = 0.95. \tag{24}$$

Bush has a significantly lower valence ($\lambda_{\text{rep}}^{\text{US2004}} = -0.43$) than Kerry ($\lambda_{\text{dem}}^{\text{US2004}} \equiv 0.0$), the baseline candidate.

From (14) the probability that a US voter chooses Bush, the low valence candidate, when both Bush and Kerry are at the electoral origin, $\mathbf{z}_0$, is

$$\rho_{\text{rep}}^{\text{US2004}} = \left[\sum_{k=1}^{2} \exp\left(\lambda_k^{\text{US2004}} - \lambda_{\text{rep}}^{\text{US2004}}\right)\right]^{-1} = [1 + \exp(0.43)]^{-1} = 0.40. \tag{25}$$

The confidence bounds for $\rho_{\text{rep}}^{\text{US2004}}$ are given in Appendix A.1. Since Bush's valence, relative to that of his opponent, was similar in the two elections, it is not surprising that the probability of voting Republican is similar in the two elections, compare (20) and (25). From (15), $2\beta_{2004}^{\text{US}}(1-2\rho_{\text{rep}}^{\text{US2004}}) = 2 \times 0.95 \times 0.2 = 0.38$ and $\sigma_{\text{US2004}}^2 = 1.17$, so that the convergence coefficient of the 2004 election is

$$c_{\text{US}}^{2004} = 2\beta_{2004}^{\text{US}}\left[1 - 2\rho_{\text{rep}}^{\text{US2004}}\right]\sigma_{\text{US2004}}^2 = 0.38 \times 1.19 = 0.45. \tag{26}$$

Since $c_{\text{US}}^{2004} = 0.45$ is significantly less than 1 (see Appendix A.1), the sufficient condition for convergence given in Section 2 is met. Moreover, from (17) Bush's characteristic matrix is

$$\begin{aligned} C_{\text{rep}}^{\text{US2004}} &= \left[2\beta_{2004}^{\text{US}}\left(1 - 2\rho_{\text{rep}}^{\text{US2004}}\right)\right]\nabla_{2004}^{\text{US}} - I \\ &= 0.38\begin{bmatrix} 0.53 & -0.18 \\ -0.18 & 0.66 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.80 & -0.06 \\ -0.06 & -0.75 \end{bmatrix}. \end{aligned} \tag{27}$$

If Bush positions himself at the electoral origin, then with probability exceeding 95% (see Appendix A.1), his vote share function would be at a maximum. Bush, the low valence candidate, has then no incentive to move from the origin, $\mathbf{z}_0$. With probability exceeding 95%, the mean is an LNE for model of the 2004 US election.

Our analysis suggests that Obama's victory over McCain in the *2008 US election* was the result of an overall shift in the relative valences of the Democratic and Republican candidates as compared to those of the candidates in the 2000 and 2004 elections. The electoral covariance matrix for the sample in 2008 along the economic and social dimensions is

$$\nabla_{2008}^{\text{US}} = \begin{bmatrix} \sigma_E^2 = 0.80 & \sigma_{ES} = -0.127 \\ \sigma_{ES} = -0.127 & \sigma_S^2 = 0.83 \end{bmatrix}. \tag{28}$$

Relative to the two previous elections the "variance" of the electoral distribution $\sigma_{\text{US2008}}^2 = \text{trace}\,(\nabla_{2008}^{\text{US}}) = \sigma_E^2 + \sigma_S^2 = 1.63$ increased, while the covariance between these dimensions $\sigma_{ES} = -0.127$ decreased.

The MNL estimates of the spatial model given in Table 1 for the 2008 US election are

$$\lambda_{\text{rep}}^{\text{US2008}} = -0.84, \qquad \lambda_{\text{dem}}^{\text{US2008}} \equiv 0.0, \qquad \beta_{2008}^{\text{US}} = 0.85. \tag{29}$$

Obama, the baseline candidate, has a significantly higher valence than McCain.

From (14) the probability that a voter chooses McCain, when both candidates are at the origin, $\mathbf{z}_0$, is

$$\rho_{\text{rep}}^{\text{US2008}} = \left[\sum_{k=1}^{2} \exp\left(\lambda_k^{\text{US2008}} - \lambda_{\text{rep}}^{\text{US2008}}\right)\right]^{-1} = \left[1 + \exp(0.84)\right]^{-1} = 0.30. \tag{30}$$

From (15), $2\beta_{\text{US}}^{2008}(1 - 2\rho_{\text{dem}}^{\text{US2008}}) = 2 \times 0.85 \times 0.4 = 0.68$, and $\sigma_{\text{US2008}}^2 = 1.63$, so the convergence coefficient is

$$\begin{aligned} c_{\text{US}}^{2008} &= 2\beta_{2008}^{\text{US}}\left[1 - 2\rho_{\text{dem}}^{\text{US2008}}\right]\sigma_{\text{US2008}}^2 \\ &= 0.68 \times 1.63 = 1.11. \end{aligned} \tag{31}$$

Appendix A.1 shows that $c_{\text{US}}^{2008} = 1.11$ is significantly greater than 1 and significantly less than 2. The Valence Theorem then states that the necessary but not the sufficient condition for convergence has been met. To check whether the low valence candidate, McCain, has an incentive to move from the electoral mean, we examine McCain's characteristic matrix using (17) to get

$$\begin{aligned} C_{\text{rep}}^{\text{US2008}} &= \left[2\beta_{2008}^{\text{US}}\left(1 - 2\rho_{\text{rep}}^{\text{US2008}}\right)\right]\nabla_{2008}^{\text{US}} - I \\ &= 0.68\begin{bmatrix} 0.80 & -0.127 \\ -0.127 & 0.83 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.46 & -0.086 \\ -0.086 & -0.44 \end{bmatrix}. \end{aligned} \tag{32}$$

With probability exceeding 95% (see Appendix A.1) McCain's vote share function is at a maximum when he locates at the origin, and thus has no incentive to move. Thus, with probability exceeding 95%, the electoral origin is an LNE for the spatial model for the 2008 US election.

In conclusion, Table 2 illustrates that the convergence coefficient varies across elections in the same country even when there are only two parties. This is to be expected as from (15) the convergence coefficient depends on the "variance" of the electoral distribution, $\sigma^2 = \text{trace}(\nabla)$; on the weight voters give to differences with party's policies, $\beta$; and on the probability that a voter chooses the party with the lowest valence, $\rho_1$. The electoral distributions of the 2000 and 2004 are quite similar, as can be seen by comparing (18) and (23). Voters' preferences had however substantially changed by 2008, see (28). The electoral variance along both axes increased relative to 2000 and 2004. 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 in all three elections had no incentive to move from the origin.

*3.1.2. The 2005 and 2010 Elections in Great Britain.* We study the 2005 and 2010 elections in the UK using the British

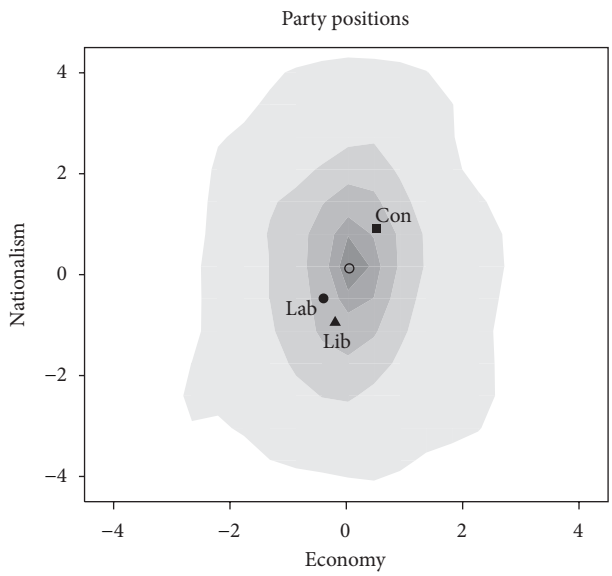



Figure 4: Electoral distribution and estimated party positions in Britain in 2005.

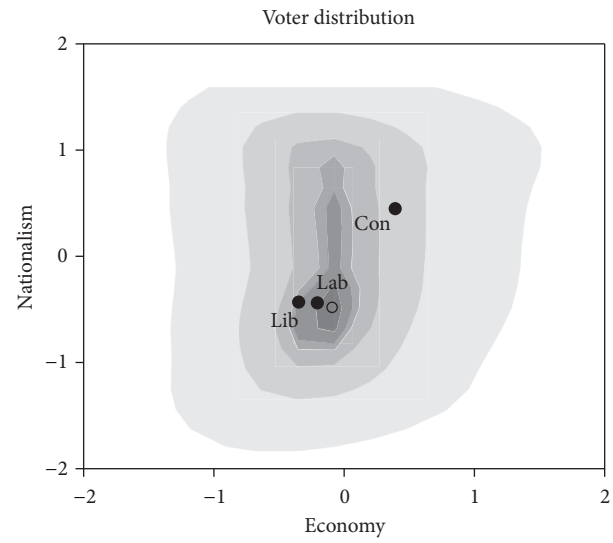



Figure 5: Voter and party positions in Britain in 2010.

Election Study (BES). (The full analysis of the 2005 and 2010 elections in Great Britain can be found in Schofield et al. [37].) The factor analysis conducted on the questions of the two surveys led us to conclude that the same two dimensions mattered in 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 in the *nationalism* dimension (Nat = $x$-axis). Items such as "tax/spend," "free market," "international monetary transfer," "international companies," and "worry about job loss overseas" have strong influence in the *economic* ($E = y$-axis) dimension with higher values indicating a promarket attitude. Figures 4 and 5 present the smoothed electoral distribution obtained from these analyses for the 2005 and 2010 elections.

The electoral covariance matrix for the *2005 UK election* is

$$\nabla_{2005}^{\mathrm{UK}} = \begin{bmatrix} \sigma_{\mathrm{Nat}}^2 = 1.646 & \sigma_{\mathrm{Nat}E} = 0.00 \\ \sigma_{E\mathrm{Nat}} = 0.067 & \sigma_E^2 = 3.961 \end{bmatrix}, \tag{33}$$

where $\sigma_{\mathrm{UK2005}}^2 \equiv \mathrm{trace}(\nabla_{2005}^{\mathrm{UK}}) = \sigma_{\mathrm{Nat}}^2 + \sigma_E^2 = 5.607$.

From Table 1, the MNL estimates of the spatial model for the 2005 UK are

$$\begin{aligned} \lambda_{\mathrm{Lab}}^{\mathrm{UK2005}} &= 0.52, \qquad \lambda_{\mathrm{Con}}^{\mathrm{UK2005}} = 0.27, \\ \lambda_{\mathrm{Lib}}^{\mathrm{UK2005}} &\equiv 0.0, \qquad \beta_{2005}^{\mathrm{UK}} = 0.15. \end{aligned} \tag{34}$$

Both the Labour (Lab) and the Conservative (Con) parties had a significantly higher valence than the Liberal Democrats (Lib), the baseline party.

From (14), the probability that a voter chooses the Liberal Democratic Party, the lowest valence party, when all parties locate at the origin, $\mathbf{z}_0$, is

$$\begin{aligned} \rho_{\mathrm{Lib}}^{\mathrm{UK2005}} &= \left[ \sum_{k=1}^{3} \exp\left(\lambda_k^{\mathrm{UK2005}} - \lambda_{\mathrm{Lib}}^{\mathrm{UK2005}}\right) \right]^{-1} \\ &= \left[1 + \exp(0.52) + \exp(0.27)\right]^{-1} = 0.25. \end{aligned} \tag{35}$$

Given that $2\beta_{2005}^{\mathrm{UK}}(1 - 2\rho_{\mathrm{Lib}}^{\mathrm{UK2005}}) = 2 \times 0.15 \times 0.5 = 0.15$ and since $\sigma_{\mathrm{UK2005}}^2 = 5.607$ in (33), from (15) the convergence coefficient, in Table 2, is

$$\begin{aligned} c_{\mathrm{UK}}^{2005} &= 2\beta_{2005}^{\mathrm{UK}} \left[1 - 2\rho_{\mathrm{Lib}}^{\mathrm{UK2005}}\right] \sigma_{\mathrm{UK2005}}^2 \\ &= 0.15 \times 5.607 = 0.84. \end{aligned} \tag{36}$$

Appendix A.1 shows that $c_{\mathrm{UK}}^{2005}$ is significantly less than 1 and thus meets the sufficient and necessary conditions for convergence given in Section 2. From (17) the characteristic matrix of the Liberal Democratic Party is

$$\begin{aligned} C_{\mathrm{Lib}}^{2005\mathrm{UK}} &= \left[2\beta_{2005}^{\mathrm{UK}}\left(1 - 2\rho_{\mathrm{Lib}}^{\mathrm{UK2005}}\right)\right] \nabla_{2005}^{\mathrm{UK}} - I \\ &= 0.15 \begin{bmatrix} 1.646 & 0.0 \\ 0.067 & 3.961 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.75 & 0.0 \\ 0.01 & -0.406 \end{bmatrix}. \end{aligned} \tag{37}$$

From the 95% confidence bounds in Appendix A.1, we conclude that if the LibDem locates at the origin, it is maximizing its vote share and has no incentive to vacate the center. Thus, with probability exceeding 95%, the origin is an LNE for the 2005 UK election.

The electoral covariance matrix for the *2010 UK election* is

$$\nabla_{2010}^{\mathrm{UK}} = \begin{bmatrix} \sigma_{\mathrm{Nat}}^2 = 0.601 & \sigma_{\mathrm{Nat}E} = 0.067 \\ \sigma_{E\mathrm{Nat}} = 0.067 & \sigma_E^2 = 0.861 \end{bmatrix}, \tag{38}$$

where $\sigma_{\mathrm{UK2010}}^2 \equiv \mathrm{trace}(\nabla_{2010}^{\mathrm{UK}}) = 1.462$, lower than in 2005.

From Table 1, the MNL estimates of the spatial model of the 2010 election are

$$\begin{aligned} &\lambda_{\mathrm{Lab}}^{\mathrm{UK2010}} = -0.04, \qquad \lambda_{\mathrm{Con}}^{\mathrm{UK2010}} = 0.17, \\ &\lambda_{\mathrm{Lib}}^{\mathrm{UK2010}} \equiv 0.0, \qquad \beta_{2010}^{\mathrm{UK}} = 0.86. \end{aligned} \tag{39}$$

Given the great popular discontent with Gordon Brown, the Labour leader, heading into the 2010 election, it is not surprising to find that both Conservatives and Liberal Democrats (the base party) had significantly higher valences than Labour.

From (14) the probability that a voter chooses Labour, when all parties locate at the origin, $\mathbf{z}_0$, is

$$\begin{aligned} \rho_{\mathrm{Lab}}^{\mathrm{UK2010}} &= \left[\sum_{k=1}^{3} \exp\left(\lambda_k^{\mathrm{UK2010}} - \lambda_{\mathrm{Lab}}^{\mathrm{UK2010}}\right)\right]^{-1} \\ &= \left[1 + \exp(0.21) + \exp(0.04)\right]^{-1} = 0.319. \end{aligned} \tag{40}$$

Since $2\beta_{2010}^{\mathrm{UK}}(1 - 2\rho_{\mathrm{Lab}}^{\mathrm{UK2010}}) = 2 \times 0.86 \times 0.362 = 0.622$ and $\sigma_{\mathrm{UK2010}}^2 = 1.462$ in (38), from (15) the convergence coefficient, in Table 2, is

$$\begin{aligned} c_{\mathrm{UK}}^{2010} &= 2\beta_{2010}^{\mathrm{UK}}\left[1 - 2\rho_{\mathrm{Lab}}^{2010}\right]\sigma_{\mathrm{UK2010}}^2 \\ &= 0.622 \times 1.462 = 0.91. \end{aligned} \tag{41}$$

The convergence coefficient $c_{\mathrm{UK}}^{2010} = 0.91$ is significantly less than 1 (see Appendix A.1), meeting the sufficient and thus necessary condition for convergence. From (17), Labour's characteristic matrix is

$$\begin{aligned} C_{\mathrm{Lab}}^{\mathrm{UK2010}} &= \left[2\beta_{2010}^{\mathrm{UK}}\left(1 - 2\rho_{\mathrm{Lab}}^{\mathrm{UK2010}}\right)\right]\nabla_{2010}^{\mathrm{UK}} - I \\ &= 0.622\begin{bmatrix} 0.601 & 0.067 \\ 0.067 & 0.861 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.63 & 0.042 \\ 0.042 & -0.46 \end{bmatrix}. \end{aligned} \tag{42}$$

If Labour, the low valence party, locates at the origin, then with probability exceeding 95%, its vote share function is at a maximum (see Appendix A.1) giving it no incentive to move from the mean. Thus, with probability exceeding 95%, the electoral origin is an LNE for the 2010 UK election.

The major shift in voters' preferences between the two elections led to very different electoral outcomes as evidenced by the electoral covariance matrices in (33) and (38). Voter dissatisfaction with the governing Labour leader led to a dramatic decrease in his competence valence and on the probability of voting Labour. 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 in 2010. The analysis for the UK elections shows 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 the one for the 2010 election.

The analysis of these two Anglo-Saxon countries illustrate that even under plurality rule the convergence coefficient varies from election to election and from country to country. The analysis for the 2010 UK election highlights that candidates' valences matter and that parties understand how their valence affects their electoral prospects and may adjust their positions to increase their votes. This section illustrates that under plurality, the convergence coefficient has low values that generally satisfy the necessary condition for convergence to the mean and is thus below the dimension of the policy space.

*3.2. Convergence in Proportional Systems?* We now estimate the convergence coefficients for three parliamentary countries using proportional representation: Israel, Turkey, and Poland. As is well known, these countries are characterized by multiparty elections in which generally no party wins a legislative majority leading then to coalitions governments. This section shows that these countries are characterized by very high convergence coefficients.

*3.2.1. The 1996 Election in Israel.* In the 1996, as in previous elections, Israel had approximately nineteen parties attaining seats in the Knesset. (These include parties on the left, on the center, on the right, as well as religious parties. On the left there is Labor, Merets, Democrat, Communists and Balad; those on 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.) There were small parties with 2 seats to moderately large parties such as Likud and Labor whose seat strengths lie in the range 19 to 44, out of a total of 120 Knesset seats. Since Likud and Labour compete for dominance of 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 and for these parties.

Schofield et al. [30] performed a factor analysis of the surveys conducted by Arian and Shamir [38] for 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; "religion" 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.

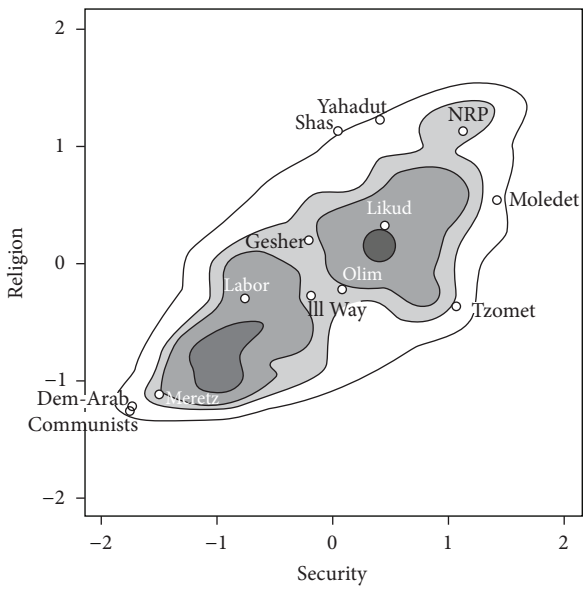



Figure 6: Party positions and voter distribution in Israel in the 1996 election.

Voter distribution along these two axes gives the following covariance matrix:

$$\nabla^{\mathrm{I}}_{!996} = \begin{bmatrix} \sigma^2_S = 1.00 & \sigma_{SR} = 0.591 \\ \sigma_{RS} = 0.591 & \sigma^2_R = 0.732 \end{bmatrix}, \tag{43}$$

giving a "variance" of $\sigma^2_{\mathrm{I}1996} \equiv \mathrm{trace}(\nabla^{\mathrm{I}}_{!996}) = 1.732$.

Only the seven largest parties are included in the MNL estimation. These include Likud, Labor, NRP, Moledat, Third Way (TW), and Shas with Meretz being the base party. From Table 2, the MNL coefficients for the 1996 election in Israel (I) are

$$\begin{aligned} \lambda^{\mathrm{I}1996}_{\mathrm{Lik}} &= 0.78, & \lambda^{\mathrm{I}1996}_{\mathrm{Lab}} &= 0.999, \\ \lambda^{\mathrm{I}1996}_{\mathrm{NRP}} &= -0.626, & \lambda^{\mathrm{I}1996}_{\mathrm{MO}} &= -1.259, \\ \lambda^{\mathrm{I}1996}_{\mathrm{TW}} &\equiv -2.291, & \lambda^{\mathrm{I}1996}_{\mathrm{Shas}} &= -2.023, \\ \lambda^{\mathrm{I}1996}_{\mathrm{Merezt}} &\equiv 0.0, & \beta^{\mathrm{I}}_{1996} &= 1.207. \end{aligned} \tag{44}$$

The $\beta$-coefficient and the valence estimates for all parties are significantly nonzero. The two largest parties, Likud and Labour, have significantly higher valences than the other smaller parties with Third Way (TW) having the smallest valence.

From (14), the probability that an Israeli votes for TW, when all parties locate at the mean is

$$\begin{aligned} \rho^{\mathrm{I}1996}_{\mathrm{TW}} &= \left[ \sum_{k=1}^{7} \exp\left[ \lambda^{\mathrm{I}1996}_j - \lambda^{\mathrm{I}1996}_{\mathrm{TW}} \right] \right]^{-1} \\ &= \left[ 1 + e^{3.071} + e^{3.29} + e^{1.665} \right. \\ &\quad \left. + e^{1.032} + e^{0.268} + e^{2.291} \right]^{-1} \simeq 0.014. \end{aligned} \tag{45}$$

Given that $2\beta^{\mathrm{I}}_{1996}(1 - 2\rho^{\mathrm{I}1996}_{\mathrm{TW}}) = 2 \times 1.207 \times 0.972 = 2.346$ and since $\sigma^2_{\mathrm{I}1996} = 1.732$ from (43), then using (15) we compute the convergence coefficient for Israel, in Table 4, as

$$\begin{aligned} c^{\mathrm{I}}_{1996} &= 2\beta^{\mathrm{I}}_{1996} \left(1 - 2\rho^{\mathrm{I}1996}_{\mathrm{TW}}\right) \sigma^2_{\mathrm{I}1996} \\ &= 2.346 \times 1.732 = 4.06. \end{aligned} \tag{46}$$

The 95% confidence intervals for $c^{\mathrm{I}}_{1996} = 4.06$ in Appendix A.2 confirm that the necessary condition is *not* satisfied as $c^{\mathrm{I}}_{1996} = 4.06$ is significantly higher than 2, the dimension of the policy space. Moreover, at the electoral mean the vote share function of Third Way is *not* at a maximum since its Hessian from (17)

$$\begin{aligned} C^{\mathrm{I}1996}_{\mathrm{TW}} &= 2\beta^{\mathrm{I}}_{1996} \left(1 - 2\rho^{\mathrm{I}1996}_{\mathrm{TW}}\right) \nabla^{\mathrm{I}}_{!996} - I \\ &= 2.346 \begin{bmatrix} 1.00 & 0.591 \\ 0.591 & 0.732 \end{bmatrix} - I \\ &= \begin{bmatrix} 1.346 & 1.386 \\ 1.386 & 0.717 \end{bmatrix} \end{aligned} \tag{47}$$

shows that if TW locates at the mean its vote share function is at a saddlepoint since $C^{\mathrm{I}1996}_{\mathrm{TW}}$ has one positive (2.453) and one negative (−0.39) eigenvalue. Appendix A.2 confirms that $C^{\mathrm{I}1996}_{\mathrm{TW}}$ has one negative and one positive eigenvalue at both its lower and upper bounds. Thus, with a high degree of certainty TW deviates from the mean to maximize its votes and the electoral mean is *not* a LNE for the 1996 Israeli election.

*3.2.2. The 1999 and 2002 Elections in Turkey.* We used factor analysis of electoral survey data of Veri Arastima for TUSES to study the 1999 and 2002 Turkish elections. (See Schofield et al. [39] for details of the estimation.) The analysis indicates that voters made decisions in a two-dimensional space during the two elections. Voters who support secularism or "Kemalism" are placed on the left of the Religious ($R = x$) axis and those supporting Turkish nationalism ($N = y$) to the north. 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 (FP) which was banned by the Constitutional Court in 2001 and the change of the name of the pro-Kurdish party from HADEP to DEHAP. (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 nonnationalistic.) 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 Vitue Party (VP), the Motherland Party (ANAP), the True Path Party (DYP), the Republican People's Party (CHP), and the People's Democratic Party (HADEP). A DSP minority government formed, supported by ANAP and DYP. This only lasted about 4 months and was replaced by a DSP-ANAP-MHP coalition, indicating the difficulty

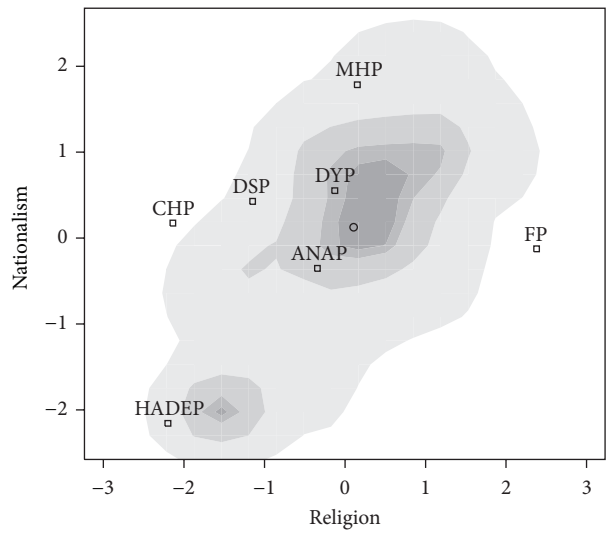



Figure 7: Party positions and voter distribution in the 1999 Turkish election.

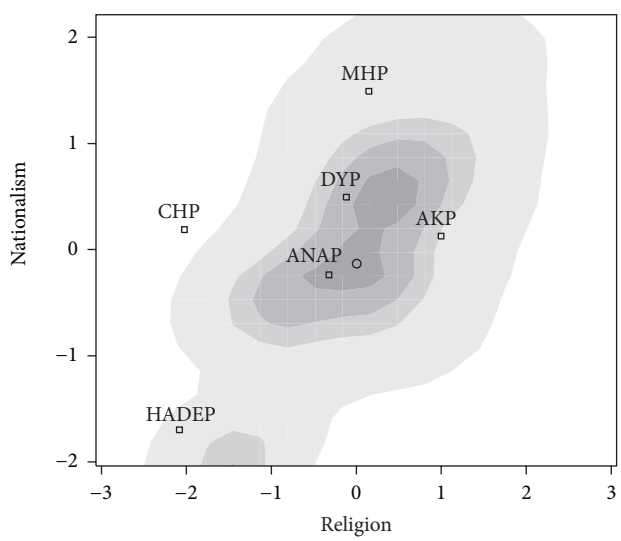



Figure 8: Party positions and voter distribution in Turkey in 2002.

of negotiating a coalition compromise across the disparate policy positions of the coalition members.

In *the 1999 election*, the electoral covariance matrix along the Religious ($R$) and Nationalism ($N$) axes is

$$\nabla^{\mathrm{T}}_{!999} = \begin{bmatrix} \sigma^2_R = 1.20 & \sigma_{RN} = 0.78 \\ \sigma_{NR} = 0.78 & \sigma^2_N = 1.14 \end{bmatrix}, \tag{48}$$

with $\sigma^2_{\mathrm{T}1999} \equiv \mathrm{trace}(\nabla^{\mathrm{T}}_{!999}) = 2.34$.

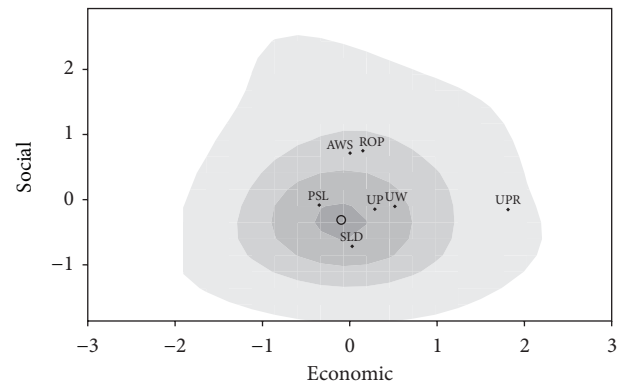



Figure 9: Voter distribution and party-positions in Poland in 1997.

Using DYP as the base party, from Table 3, the 1999 MNL coefficients are

$$\begin{aligned} \lambda^{\mathrm{T}1999}_{\mathrm{FP}} &= -0.16, & \lambda^{\mathrm{T}1999}_{\mathrm{MHP}} &= 0.66, \\ \lambda^{\mathrm{T}1999}_{\mathrm{DYP}} &\equiv 0.0, & \lambda^{\mathrm{T}1999}_{\mathrm{HADEP}} &= -0.071, \\ \lambda^{\mathrm{T}1999}_{\mathrm{ANAP}} &= 0.34, & \lambda^{\mathrm{T}1999}_{\mathrm{CHP}} &\equiv 0.73, \\ \lambda^{\mathrm{T}1999}_{\mathrm{DSP}} &= 0.72, & \beta^{\mathrm{T}}_{1999} &= 0.38. \end{aligned} \tag{49}$$

The $\beta$-coefficient and the valence estimates of DSP and MHP and CHP are significantly nonzero. The probability that a Turkish voter chooses FP with lowest valence in 1999, when all parties locate at the mean, $\rho^{\mathrm{T}1999}_{\mathrm{FP}}$ in (14), is

$$\begin{aligned} \rho^{\mathrm{T}1999}_{\mathrm{FP}} &= \left[ \sum_{k=1}^{7} \exp\left[\lambda^{\mathrm{T}1999}_j - \lambda^{\mathrm{T}1999}_{\mathrm{FP}}\right] \right]^{-1} \\ &= \left[1 + e^{0.82} + e^{0.16} + e^{0.09} \right. \\ &\quad \left. + e^{0.5} + e^{0.89} + e^{0.88}\right]^{-1} \simeq 0.08. \end{aligned} \tag{50}$$

Given that $2\beta^{\mathrm{T}}_{1999}(1 - 2\rho^{\mathrm{T}1999}_{\mathrm{FP}}) = 2 \times 0.38 \times 0.84 = 0.64$ and since $\sigma^2_{\mathrm{T}1999} = 2.34$ in (48), then using (15), Turkey's convergence coefficient in 1999, in Table 4, is

$$\begin{aligned} c^{\mathrm{T}}_{1999} &= 2\beta^{\mathrm{T}}_{1999}\left(1 - 2\rho^{\mathrm{T}1999}_{\mathrm{FP}}\right)\sigma^2_{\mathrm{T}1999} \\ &= 0.64 \times 2.34 = 1.49. \end{aligned} \tag{51}$$

The convergence coefficient is significantly higher that 1 and significantly lower than 2 (see Appendix A.2). From (17) FP's Hessian at the origin is

$$\begin{aligned} C^{\mathrm{T}1999}_{\mathrm{FP}} &= 2\beta^{\mathrm{T}}_{1999}\left(1 - 2\rho^{\mathrm{T}1999}_{\mathrm{FP}}\right)\nabla^{\mathrm{T}}_{!999} - I \\ &= 0.64 \begin{bmatrix} 1.20 & 0.78 \\ 0.78 & 1.14 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.24 & 0.448 \\ 0.448 & -0.27 \end{bmatrix}. \end{aligned} \tag{52}$$

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

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

[a]$^{*}$prob < 0.05; $^{**}$prob < 0.01; $^{***}$prob < 0.001.
[b]Israel: Lik: Likud; Lab: Labor; NRP: Mafdal; Mo: Moledet; TW: Third Way.
[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]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.

Table 4: The convergence coefficient in proportional systems.

| | Israel | Turkey | | Poland |
|---|---|---|---|---|
| | 1996 | 1999 | 2002 | 1997 |
| | Weight of policy differences ($\beta$) | | | |
| Central Est.[a] of $\beta$ (conf. Int.[b]) | 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$\nabla = \sigma^2$) | | | |
| $\sigma^2$ | 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[c] | FP[d] | ANAP[d] | ROP[e] |
| Central Est.[a] of $\rho_1$ (conf. Int.[b]) | $\rho^{\text{I}}_{\text{TW}} = 0.014$ (0.006, 0.034) | $\rho_{\text{FP}} = 0.08$ (0.046, 0.145) | $\rho^{\text{T}}_{\text{ANAP}} = 0.08$ (0.038, 0.133) | $\rho^{\text{P}}_{\text{ROP}} = 0.007$ (0.002, 0.022) |
| | Convergence coefficient ($c \equiv c(\boldsymbol{\lambda}, \beta, \sigma^2) = 2\beta[1 - 2\rho_1]\sigma^2$) | | | |
| Central Est.[a] of $c$ (conf. Int.[b]) | 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]Central Est.: central estimate.
[b]Conf. Int.: confidence intervals.
[c]Israel: TW: Third Way.
[d]Turkey: DYP: True Path Party.
[e]Poland: ROP: Movement for Reconstruction of Poland.

When at the electoral origin, FP's characteristic function shows that its vote share function is at a saddlepoint as the eigenvalues of $C_{\mathrm{FP}}^{\mathrm{T1999}}$ are $-0.74$ with minor eigenvector $(+1 - 1.116)$ and $+0.23$ with major eigenvector $(+1, +0.896)$. Moreover, as seen in Appendix A.2, the 95% confidence bounds show that at the lower bound of $C_{\mathrm{FP}}^{\mathrm{T1999}}$ FP has no incentive to move but it does at the upper bound. Since FP wants to move at the central estimate of $C_{\mathrm{FP}}^{\mathrm{T1999}}$ in (52) it is probable that in general FP wants to move away from the mean to increase its vote share. Moreover, since the convergence coefficient is significantly greater than 2, then with a high degree confidence, the electoral mean *cannot* be a LNE for Turkey in 1999.

The electoral covariance matrix of the *2002 Turkish election* is

$$\nabla_{2002}^{\mathrm{T}} = \begin{bmatrix} \sigma_R^2 = 1.18 & \sigma_{RN} = 0.74 \\ \sigma_{NR} = 0.74 & \sigma_N^2 = 1.15 \end{bmatrix}, \tag{53}$$

with $\sigma_{\mathrm{T2002}}^2 = \mathrm{trace}\,(\nabla_{2002}^{\mathrm{T}}) = 2.33$.

Note that the covariance matrix of 1999 in (48) and that of 2002 in (53) suggest few changes in the distribution of voters between these two election. Figures 8 and 9 suggest that there were few changes in party positions between these two elections. The basis of support for the AKP may be regarded as similar to that of the banned FP, suggesting that the leader of this party changed the party's position on the religion axis, adopting a much less radical position. One would think of this as generating political stability in Turkey. Yet, between 1999 and 2002, Turkey experienced two severe economic crises and in 2002, a 10% electoral cut-off rule was instituted. The crises and the cut-off rule changed the political landscape in Turkey. In the 2002 election, seven parties obtained less than 10% of the vote and won no seats. The AKP won 34% of the vote, and due to the cut-off rule, obtained a majority of the seats (363 out of 550).

Our analysis reflects this change in the political landscape. Using DYP as the base party, from Table 3, the 2002 MNL coefficients are

$$\begin{gathered} \lambda_{\mathrm{ANAP}}^{\mathrm{T2002}} = -0.31, \qquad \lambda_{\mathrm{MHP}}^{\mathrm{T2002}} = -0.12, \\ \lambda_{\mathrm{DYP}}^{\mathrm{T2002}} \equiv 0.0, \qquad \lambda_{\mathrm{HADEP}}^{\mathrm{T2002}} = 0.43, \\ \lambda_{\mathrm{AKP}}^{\mathrm{T2002}} = 0.78, \qquad \lambda_{\mathrm{CHP}}^{\mathrm{T2002}} \equiv 1.33, \qquad \beta_{2002}^{\mathrm{T}} = 1.52. \end{gathered} \tag{54}$$

The $\beta$-coefficient and the valences of AKP and CHP are significantly nonzero with ANAP having the lowest valence. The probability of voting ANAP, when parties locate at the mean, $\rho_{\mathrm{ANAP}}^{\mathrm{T20029}}$ in (14), is

$$\begin{aligned} \rho_{\mathrm{ANAP}}^{\mathrm{T2002}} &= \left[ \sum_{k=1}^{6} \exp\left[\lambda_j^{\mathrm{T2002}} - \lambda_{\mathrm{ANAP}}^{\mathrm{T2002}}\right] \right]^{-1} \\ &= \left[1 + e^{0.19} + e^{0.31} + e^{0.74} \right. \\ &\quad \left. + e^{1.09} + e^{1.164}\right]^{-1} \simeq 0.08. \end{aligned} \tag{55}$$

Given that $2\beta_{2002}^{\mathrm{T}}(1-2\rho_{\mathrm{ANAP}}^{\mathrm{T2002}}) = 2\times 1.52\times 0.84 = 2.55$ and since $\sigma_{\mathrm{T2002}}^2 = 2.33$ from (53), then using (15) we find that the 2002 convergence coefficient for Turkey, in Table 4, is

$$c_{2002}^{\mathrm{T}} = 2\beta_{2002}^{\mathrm{T}}\left(1 - 2\rho_{\mathrm{ANAP}}^{\mathrm{T20029}}\right)\sigma_{\mathrm{T2002}}^2 = 2.55 \times 2.33 = 5.94. \tag{56}$$

The political changes induced by the cut-off rule led to a higher convergence coefficient in 2002 relative to 1999 (increasing from a low of $c_{1999}^{\mathrm{T}} = 1.49$ in (51) to a high $c_{2002}^{\mathrm{T}} = 5.94$ in (56)). An indication that a more fractionalized polity emerged from this reform. The convergence coefficient of the 2002 election is significantly above 2, the dimension of the policy space (see Appendix A.2) giving ANAP an incentive to locate far from the mean. ANAP's characteristic matrix using (17) is

$$\begin{aligned} C_{\mathrm{ANAP}}^{\mathrm{T2002}} &= 2\beta_{2002}^{\mathrm{T}}\left(1 - 2\rho_{\mathrm{ANAP}}^{\mathrm{T2002}}\right)\nabla_{2002}^{\mathrm{T}} - I \\ &= 2.55\begin{bmatrix} 1.18 & 0.74 \\ 0.74 & 1.15 \end{bmatrix} - I \\ &= \begin{bmatrix} 2.01 & 1.88 \\ 1.88 & 1.93 \end{bmatrix}. \end{aligned} \tag{57}$$

When at the origin, $C_{\mathrm{ANAP}}^{\mathrm{T2002}}$ indicates that ANAP is minimizing its vote share since its eigenvalues are both positive (0.090 and 3.850). This together with the 95% confidence bounds in Appendix A.2 implies that there is a high probability that ANAP will vacate the center and that the mean is *not* an LNE for Turkey in 2002.

*3.2.3. The 1997 Polish Election.* In the election held in Poland in 1997 (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.) the following five parties won seats in the Sejm (lower house). The left-wing excommunist Democratic Left Alliance (SLD) and the agrarian Polish Peoples' Party (PSL), both of which have been the most frequent governing parties in the postcommunist period. The Freedom Union (UW) and the Solidarity Election Action (AWS) had grown out of the Solidarity movement. AWS combined various mostly right wing and Christian groups under one label, while UW was formed based on the liberal wing of Solidarity. The remaining party is the Movement for Reconstruction of Poland (ROP).

Applying factor analysis to questions from the Polish National Election Survey an economic and a social value dimensions were identified (see [40]). The *economic* dimension is influenced by issues such as privatization versus state ownership of enterprises, 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 decommunization versus equal rights for former nomenclature, and abortion rights regardless of situation versus no such rights regardless of situation are the most influential

issues in this *social values* dimension. The distribution of voters along these dimensions is seen in Figure 9. (See Schofield et al. [40] for details of the estimation.)

The covariance matrix for the *1997 Polish* (P) *election* is

$$\nabla^{\mathrm{P}}_{1997} = \begin{bmatrix} \sigma^2_E = 1.00 & \sigma_{ES} = 0.0 \\ \sigma_{SE} = 0.0 & \sigma^2_S = 1.00 \end{bmatrix}, \tag{58}$$

with variance $\sigma^2_{\mathrm{P}1997} = \mathrm{trace}(\nabla^{\mathrm{P}}_{1997}) = 2.00$.

From Table 3, the MNL coefficients for the 1997 election are

$$\begin{aligned} \lambda^{\mathrm{P}1997}_{\mathrm{UPR}} &= -2.3, & \lambda^{\mathrm{P}1997}_{\mathrm{UP}} &= -0.56, \\ \lambda^{\mathrm{P}1997}_{\mathrm{ROP}} &\equiv 0.0, & \lambda^{\mathrm{P}1997}_{\mathrm{PSL}} &= 0.07, \\ \lambda^{\mathrm{P}1997}_{\mathrm{UW}} &\equiv 0.73, & \lambda^{\mathrm{P}1997}_{\mathrm{SLD}} &= 1.40, \\ \lambda^{\mathrm{P}1997}_{\mathrm{AWS}} &= 1.92, & \beta^{\mathrm{P}}_{1997} &= 1.74. \end{aligned} \tag{59}$$

The $\beta$-coefficient and valence estimates for all parties except UP and PSL are significantly nonzero. The probability of voting UPR with lowest valence, in 1997, when parties locate at the mean, $\rho^{\mathrm{P}1997}_{\mathrm{TW}}$ in (14), is

$$\begin{aligned} \rho^{\mathrm{P}1997}_{\mathrm{UPR}} &= \left[ \sum_{k=1}^{6} \exp\left[\lambda^{\mathrm{P}1997}_j - \lambda^{\mathrm{P}1997}_{\mathrm{UPR}}\right] \right]^{-1} \\ &= \left[1 + e^{0.048} + e^{3.08} + e^{4.27} \right. \\ &\quad \left. + e^{3.77} + e^{2.42}\right]^{-1} \simeq 0.01. \end{aligned} \tag{60}$$

Given that $2\beta^{\mathrm{P}}_{1997}(1-2\rho^{\mathrm{P}1997}_{\mathrm{UPR}}) = 2\times 1.74\times 0.98 = 3.41$ and since $\sigma^2_{\mathrm{P}1997} = 2$ from (58), then using (15) the convergence coefficient for Poland, in Table 4, is

$$\begin{aligned} c^{\mathrm{P}}_{1997} &= 2\beta^{\mathrm{P}}_{1997}\left(1 - 2\rho^{\mathrm{P}1997}_{\mathrm{UPR}}\right)\sigma^2_{\mathrm{P}1997} \\ &= 3.41 \times 2 = 6.82. \end{aligned} \tag{61}$$

Appendix A.2 shows that $c^{\mathrm{P}}_{1997} = 6.82$ is significantly greater than 2 and thus fails the necessary condition for convergence to the mean. UPR's Hessian from (17) is

$$\begin{aligned} C^{\mathrm{P}1997}_{\mathrm{UPR}} &= 2\beta^{\mathrm{P}}_{1997}\left(1 - 2\rho^{\mathrm{P}1997}_{\mathrm{UPR}}\right)\nabla^{\mathrm{P}}_{1997} - I \\ &= 3.41 \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{bmatrix} - I \\ &= \begin{bmatrix} 2.41 & 0.0 \\ 0.0 & 2.41 \end{bmatrix}. \end{aligned} \tag{62}$$

The trace (= 3.82), the determinant (= 5.80), and the eigenvalues of $C^{\mathrm{I}}_{\mathrm{UPR}}$ (2.41, 1.41) are positive. The 95% confidence bound of $C^{\mathrm{I}}_{\mathrm{UPR}}$ in Appendix A.2 also shows positive eigenvalues at the lower and upper bounds of $C^{\mathrm{I}}_{\mathrm{UPR}}$. Thus, with a high degree of certainty UPR locates far from the origin to maximize its votes and the electoral mean is *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 2, the dimension of the policy space and are also much higher than that of the US 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 study elections in Georgia, Russia, and Azerbaijan. In these partial democracies or anocracies, (The term "partial democracy" has been applied to new democracies lacking the full array of democratic institutions present in western democracies (see [41].)) 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 give 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 to 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.* We use the postelection survey conducted by GORBI-GALLUP International from March 19 through April 3, 2008, to built a formal model of the 2008 election in Georgia (see [42]). The factor analysis done on the survey questions determined that there were two dimensions describing voters' attitudes towards democracy and the west. One dimension is strongly related with the respondents' attitude toward the US, the EU and NATO with larger values in the West ($W = y$-axis) dimension implying 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. The electoral distribution along these two dimensions is given in Figure 10. The points (S, G, P, N) in Figure 10 represent the estimated positions of the four candidates: Saakashvili (S), Gachechiladze (G), Patarkatsishvili (P), and Natelashvili (N). (See Schofield et al. [39] for details of the estimation.)

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

$$\nabla^{\mathrm{G}}_{2008} = \begin{bmatrix} \sigma^2_D = 0.82 & \sigma_{DW} = 0.03 \\ \sigma_{WD} = 0.03 & \sigma^2_W = 0.91 \end{bmatrix} \tag{63}$$

with $\sigma^2_{\mathrm{G}2008} \equiv \mathrm{trace}\,(\nabla^{\mathrm{G}}_{2008}) = 1.73$.

From Table 5, the MNL estimates of the 2008 election with Natelashvili as the base candidate are

$$\begin{aligned} \lambda^{\mathrm{G}2008}_{\mathrm{S}} &= 2.56, \quad \lambda^{\mathrm{G}2008}_{\mathrm{G}} = 1.50, \quad \lambda^{\mathrm{G}2008}_{\mathrm{P}} = 0.53, \\ \lambda^{\mathrm{G}2008}_{\mathrm{N}} &\equiv 0.0, \quad \beta^{\mathrm{G}}_{2008} = 0.78. \end{aligned} \tag{64}$$

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

All coefficients are significantly nonzero showing Natelashvili as having the lowest valence.

The probability that a Georgian votes for Natelashvili, when all candidates locate at the mean, is

$$\rho_{\mathrm{N}}^{\mathrm{G2008}} = \left[\sum_{k=1}^{4} \exp\left[\lambda_j^{\mathrm{G2008}} - \lambda_{\mathrm{N}}^{\mathrm{G2008}}\right]\right]^{-1} = \left[1 + e^{2.56} + e^{1.50} + e^{0.53}\right]^{-1} \simeq 0.05. \tag{65}$$

Given that $2\beta_{2008}^{\mathrm{G}}(1 - 2\rho_{\mathrm{N}}^{\mathrm{G2008}}) = 2 \times 0.78 \times 0.9 = 1.4$ and since $\sigma_{\mathrm{G2008}}^2 = 1.73$ from (63), then using (15) Georgia's the convergence coefficient, in Table 6, is

$$c_{2008}^{\mathrm{G}} = 2\beta^{\mathrm{G2008}}\left(1 - 2\rho_{\mathrm{N}}^{\mathrm{G2008}}\right)\sigma_{\mathrm{G2008}}^2 = 1.4 \times 1.73 = 2.42. \tag{66}$$

As shown in Appendix A.3, $c_{2008}^{\mathrm{G}}$ is *not* significantly different from 2 and thus fails the necessary condition for convergence to the mean. Natelashvili's Hessian or characteristic matrix, from (17), is

$$C_{\mathrm{N}}^{\mathrm{G2008}} = 2\beta_{2008}^{\mathrm{G}}\left(1 - 2\rho_{\mathrm{N}}^{\mathrm{G2008}}\right)\nabla_{2008}^{\mathrm{G}} - I = 1.4\begin{bmatrix} 0.82 & 0.03 \\ 0.03 & 0.91 \end{bmatrix} - I = \begin{bmatrix} 0.15 & 0.04 \\ 0.04 & 0.28 \end{bmatrix}. \tag{67}$$

Since the eigenvalues of $C_{\mathrm{N}}^{\mathrm{G2008}}$ are both positive (+0.139, +0.291), Natelashvili's vote share function is at a minimum when he is at the mean and has an incentive to move to increase his vote share. This together with the analysis of the 95% confidence intervals of $C_{\mathrm{N}}^{\mathrm{G2008}}$ in Appendix A.3 shows that with a high degree of certainty Natelashvili will locate far from the mean. This is not surprising since Georgians managed to induce three major changes in government through mass protests prior to this election. Thus, with a high degree of certainty Natelashvili locates far from the origin in this election and the electoral mean *cannot* be an LNE for the 2008 Georgian election.

*3.3.2. The 2007 Russian Election.* The analysis of the 2007 Russian election concentrates on four parties: the pro-Kremlin United Russia party (ER), Liberal Democratic Party (LDPR), Communist Party (CPRF), and Fair Russia (SR). Voters' ideological preferences were measured according to two questions taken from the survey conducted by VCIOM (Russian Public Opinion Research Center) in May 2007 (see [43]). The first dimension gives a measure of voters general (dis)satisfaction ($D = x$-axis). High values in this dimension correspond to negative feelings toward "justice," "labor" and, to a lesser extent, "order," "state," "stability," and "equality." Also, those with high values of the first axis tend to feel neutral toward order, elite, West, and non-Russians. The second dimension measures the voter's degree of economic liberalism ($E = y$-axis). High values correspond to positive feelings to "freedom," "business," "capitalism," "well-being," "success," and "progress," and to negative feelings toward "communism," "socialism," "USSR," and related concepts. The distribution of voter preferences along these two dimensions can be seen in Figure 11. (See Schofield and Zakharov [43] for details of the estimation.)

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

$$\nabla_{2007}^{\mathrm{R}} = \begin{bmatrix} \sigma_D^2 = 2.95 & \sigma_{DE} = 0.13 \\ \sigma_{ED} = 0.13 & \sigma_E^2 = 2.95 \end{bmatrix}, \tag{68}$$

with $\sigma_{\mathrm{R2007}}^2 \equiv \mathrm{trace}(\nabla_{2007}^{\mathrm{R}}) = 5.9$.

From Table 5, the MNL estimates of the spatial model for Russia are

$$\lambda_{\mathrm{SR}}^{\mathrm{R2007}} = -0.4, \qquad \lambda_{ER}^{\mathrm{R2007}} \equiv 0, \qquad \lambda_{\mathrm{LDPR}}^{\mathrm{R2007}} = 0.153, \qquad \lambda_{\mathrm{CPRF}}^{\mathrm{R2007}} = 1.971, \qquad \beta_{2007}^{\mathrm{R}} = 0.181. \tag{69}$$

Distance and all valences, except for that of the LDPR party, are significantly nonzero. When parties locate at the mean, the probability that a Russian votes for Fair Russia (SR) with lowest valence, from (14) is

$$\rho_{\mathrm{SR}}^{\mathrm{R2007}} = \left[\sum_{k=1}^{4} \exp\left[\lambda_j^{\mathrm{R2007}} - \lambda_{\mathrm{SR}}^{\mathrm{R2007}}\right]\right]^{-1} = \left[1 + e^{0.4} + e^{0.553} + e^{2.371}\right]^{-1} \simeq 0.07. \tag{70}$$

Given that $2\beta_{2007}^{\mathrm{R}}(1 - 2\rho_{\mathrm{SR}}^{\mathrm{R2007}}) = 2 \times 0.181 \times 0.86 = 0.31$ and since $\sigma_{\mathrm{R2007}}^2 = 5.9$ from (68), then using (15) Russia's convergence coefficient, in Table 6, is

$$c_{2007}^{\mathrm{R}} = 2\beta_{2007}^{\mathrm{R}}\left(1 - 2\rho_{\mathrm{SR}}^{\mathrm{R2007}}\right)\sigma_{\mathrm{R2007}}^2 = 0.31 \times 5.9 = 1.83. \tag{71}$$

Table 5: MNL spatial model in anocracies.

| | Georgia[c] | | Russia[b] | | Azerbaijan[d] | |
|---|---|---|---|---|---|---|
| | Party | 2008 | Party | 2007 | Party | 2010 |
| Var | | Est.[a] $\lvert t-\text{value}\rvert$ | | Est.[a] $\lvert t-\text{value}\rvert$ | | Est.[a] $\lvert t-\text{value}\rvert$ |
| $\beta$ | | $0.78^{***}$ (13.78) | | $0.181^{***}$ (12.08) | | $1.34^{***}$ (4.62) |
| Valance | $\lambda_S$ | $2.56^{***}$ (13.66) | $\lambda_{CPRF}$ | $1.971^{***}$ (17.79) | $\lambda_{YAP}$ | $1.30^{*}$ (2.14) |
| | $\lambda_G$ | $1.50^{***}$ (7.96) | $\lambda_{LDRP}$ | 0.153 (1.09) | | |
| | $\lambda_P$ | $0.53^{*}$ (2.51) | $\lambda_{SR}$ | $-0.404^{***}$ (2.50) | | |
| Base party | | N | | ER | | AXCP-MP |
| $n$ | | 676 | | 1004 | | 149 |
| $LL$ | | −533 | | −797 | | −11.5 |

[a]$^{*}$prob < 0.05; $^{**}$prob < 0.01; $^{***}$prob < 0.001.
[b]Georgia: S: Saakashvili, G: Gachechiladze, P: Patarkatsishvili, and N: Natelashvili.
[c]Rusia: ER: United Russia; CPRF: Communist Party; SR: Fair Russia; LDPR: Liberal Democratic Party.
[d]Azerbaijan: YAP: Yeni Azerbaijan Party AXCP-MP: Azerbaijan Popular Front Party (AXCP)-and Musavat (MP).

Table 6: The convergence coefficient in anocracies.

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

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

Since $c_{2007}^{R}$ is not significantly different from 2 (see Appendix A.3), the necessary condition for convergence is *not* met. The characteristic matrix or Hessian of Fair Russia (SR) from (17) is

$$\begin{aligned} C_{SR}^{R2007} &= 2\beta_{2007}^{R}\left(1 - 2\rho_{SR}^{R2007}\right)\nabla_{2007}^{R} - I \\ &= 0.31\begin{bmatrix} 2.95 & 0.13 \\ 0.13 & 2.95 \end{bmatrix} - I \\ &= \begin{bmatrix} -0.086 & 0.04 \\ 0.04 & -0.086 \end{bmatrix}. \end{aligned} \tag{72}$$

The eigenvalues are both negative (−0.126, −0.046), implying that at this central estimate Fair Russia is maximizing its vote share and thus has no incentive to vacate the origin. This conclusion holds at the lower 95% bound of $C_{SR}^{R2007}$ in Appendix A.3. However, at the upper bound of $C_{SR}^{R2007}$ Fair Russia is minimizing its vote share. It seems then that with the Russian President and his party exerting much influence over the election and Putin being so popular that Fair Russia is more likely to remain at the origin. (This result however highlights that unexpected political events could prompt Fair Russia to move from the origin.) It is then likely that the electoral mean is a LNE for the 2007 Russian election.

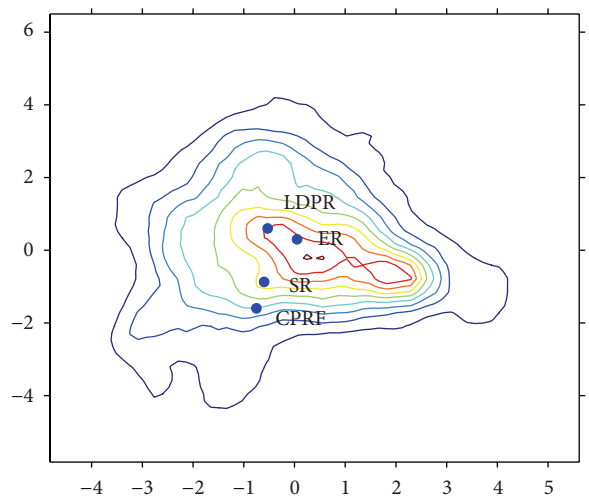



Figure 11: Party positions and voters distribution in the 2007 Russian election.

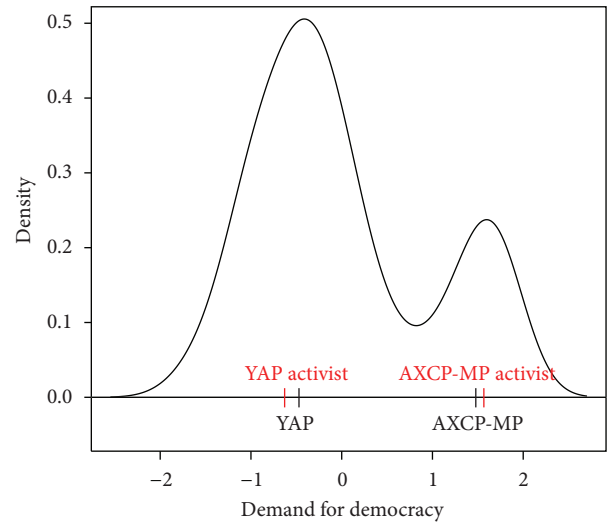



Figure 12: Voter distribution and activist positions in the 2010 Azerbaijani election.

*3.3.3. The 2010 Election in Azerbaijan.* In the 2010 election in Azerbaijan, 2,500 candidates filed application to run in the election, but only 690 were given permission by the electoral commission. The parties that competed in the election were the Yeni Azerbaijan Party (the party of the President, YAP), Civic Solidarity Party (VHP), Motherland Party (AVP), Azerbaijan Popular Front Party (AXCP), and Musavat (MP). Various small parties formed political blocks.

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 quasiopposition 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 3 seats and Ana Vaten kept the 2 seats they had in the previous legislature. For the first time, not a single candidate from the opposition Azerbaijan Popular Front (AXCP) or Musavat were elected.

We organized a small preelection survey of 2010 election in Azerbaijan allowing us to construct a model of the election (see [42]). 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 coalition.

The factor analysis showed that voters were only concerned with *one dimension*: the "demand for democracy" with higher values being associated with voters who had a negative evaluation of the current democratic situation in Azerbaijan, who did not think that free opinion is allowed, had a low degree of trust in key national political institutions, and expected that the 2010 parliamentary election would be undemocratic. Figure 12 shows the distribution of voters and the party positions at the mean of their supporters. (See [42] for details of the estimation.) In this one dimensional model the variance is

$$\sigma^2_{\text{A2010}} \equiv \text{trace}\left(\nabla^{2010}_{\text{G}}\right) = 0.93. \tag{73}$$

The *binomial* logit estimates for the 2010 election with AXCP-MP as the base party, in Table 5, are

$$\lambda^{\text{A2010}}_{\text{YAP}} = 1.30, \qquad \lambda^{\text{A2010}}_{\text{AXCP-MP}} \equiv 0.0, \qquad \beta^{\text{A}}_{2010} = 1.34. \tag{74}$$

All coefficients are significantly nonzero with AXCP-MP having the lowest valence. If these two parties locate at the mean, the probability that an Azerbaijani votes AXCP-MP from (14) is

$$\begin{aligned}\rho^{\text{A2010}}_{\text{AXCP-MP}} &= \left[\sum_{k=1}^{2} \exp\left[\lambda^{\text{A2010}}_{j} - \lambda^{\text{A2010}}_{\text{AXCP-MP}}\right]\right]^{-1} \\ &= \left[1 + e^{1.3}\right]^{-1} \simeq 0.21.\end{aligned} \tag{75}$$

Given that $2\beta^{\text{A}}_{2010}(1 - 2\rho^{\text{A2010}}_{\text{AXCP-MP}}) = 2 \times 1.34 \times 0.58 = 1.554$ and since $\sigma^2_{\text{A2010}} = 0.93$ from (73), then using (15) the convergence coefficient for Azerbaijan, in Table 6, is

$$\begin{aligned}c^{\text{A}}_{2010} &= 2\beta^{\text{A}}_{2010}\left(1 - 2\rho^{\text{A2010}}_{\text{AXCP-MP}}\right)\sigma^2_{\text{A2010}} \\ &= 1.554 \times 0.93 = 1.445.\end{aligned} \tag{76}$$

Given that $c^{\text{A}}_{2010}$ is not significantly different from 1, the dimension of the policy space (see Appendix A.3) and the necessary condition for convergence is *not* met. The one dimensional Hessian of AXCP-MP from (17) is

$$\begin{aligned}C^{\text{A2010}}_{\text{AXCP-MP}} &= 2\beta^{\text{A}}_{2010}\left(1 - 2\rho^{\text{A2010}}_{\text{AXCP-MP}}\right)\sigma^2_{\text{A2010}} - I \\ &= 1.554 \times 0.93 - 1 = 0.445.\end{aligned} \tag{77}$$

Clearly, $C^{\text{A2010}}_{\text{AXCP-MP}}$ has a single positive eigenvalue indicating the AXCP+MP is minimizing its vote share at the origin. The 95% bounds of $C^{\text{A2010}}_{\text{AXCP-MP}}$ in Appendix A.3 shows that this matrix has positive eigenvalues at the lower and upper bounds of the confidence interval. Thus, with a high degree of certainty AXCP+MP will deviate from the origin and 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 coefficient does not satisfy the necessary condition for convergence to the mean. That is, these convergence coefficients are not significantly different from the dimension of the policy space. 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 and so depends on the valence of all parties and on how dispersed voters are in the policy space. Thus any change in valence can substantially affect party positions.

## 4. Convergence across Political Systems

In the previous sections we used the unifying framework of Schofield's [9] stochastic electoral model outlined in Section 2 to study whether parties locate 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 US 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, $c^{\text{G}}_{2008} = 2.42$ in (66) which is not different from 2, suggesting that parties can diverge from the mean. (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.) The convergence coefficient of all three anocracies was not significantly different than the dimension of the policy space [2 for Georgia and Russia and 1 for Azerbaijan: $c^{\text{G}}_{2008} = 2.42$ given in (66), $c^{\text{Ru}}_{2007} = 1.83$ in (71), and $c^{\text{A}}_{2010} = 1.44$ in (76)]. These results suggest that convergence in anocracies is fragile and depends on the distribution of voters' preferences as well as on the valences of 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 signalling the lack of convergence of small valence parties to the electoral mean (from Table 7, Israel's $c^{\text{I}}_{1996} = 4.06$ in (46), Turkey's $c^{\text{T}}_{1999} = 1.49$ in (51) in 1999, and $c^{\text{T}}_{2002} = 5.94$ in (56) in 2002 and Poland's $c^{\text{P}}_{1997} = 6.82$ in (61)). Having no possibility of forming government, these small parties maximize their vote shares by locating closer to their core supporters. Elections lead to multiparty 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 multiparty (>3) legislatures by noting that the convergence coefficient $c \equiv c(\boldsymbol{\lambda}, \beta, \sigma^2) = 2\beta[1 - 2\rho_1]\sigma^2$ in (15) depends on fundamental characteristics of the electorate. These characteristics include the weight given by voters to the distance to the parties' positions, $\beta$; the electoral variance, $\sigma^2$ in (16); and the probability that a voter chooses the lowest valence party, $\rho_1$ in (14). Thus, in countries with many parties, the smallest low valence parties have little chance of receiving much support, a low $\rho_1$. If, in addition, voters care a lot about policy differences (a high $\beta$) and if the electorate is very dispersed (a high $\sigma^2$), then small parties will have an incentive to move towards their core supporters and away from the mean. That is, in highly fragmented polities where voters and correspondingly parties are very dispersed, we observe high convergence coefficients.

In essence, Schofield's [9] Valence theorem gives a simple summary statistic, the convergence coefficient, that measures the degree of fragmentation, or lack thereof, in each polity. Poland is an extreme case of this fragmentation and correspondingly has a very high convergence coefficient (see Table 7).

The are other measures of political fragmentation in the literature. The *effective number of party vote strength* (*env*) used by Laakso and Taagepera [15] serves to measure how many dominant parties there are in a polity a given election. To find the *env*, let the Herfindahl index of the election be given by

$$H_v = \sum_{j=1}^{p} v_j^2, \tag{78}$$

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 party size 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 $H_v$; that is,

$$env = H_v^{-1}. \tag{79}$$

Table 7: Convergence and fragmentation.

| | | | | | |
|---|---|---|---|---|---|
| | | | Plurality systems | | |
| Variable | | US | | Britain | |
| Political system | | Presidential | | Parliamentary | |
| Election year | 2000 | 2004 | 2008 | 2005 | 2010 |
| Conv. Coef.[a] (conf. Int.[b]) | 0.38 (0.2, 0.7) | 0.45 (0.2, 0.8) | 1.11 (0.7, 1.5) | 0.84 (0.5, 1.3) | 0.95 (0.9, 1.1) |
| Converge to mean | Yes | Yes | Yes | Yes | Yes |
| Number of parties[c] | 2 | 2 | 2 | 9 | 9 |
| | | President | | | |
| *env*[c] | 2.16 | 2.05 | 2.05 | | |
| | | House of Representatives | | House of Commons | |
| *env*[d] | 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.[a] (conf. Int.[b]) | 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 | | Likely | No | No |
| Number of parties[b] | 11 | | 9 | 10 | 7 |
| | Prime Ministers[e] | | | | |
| *env*[c] | 2.00 | | | | |
| | Knesset | | Parliament | | Sejm |
| *env*[c] | 5.84 | | 6.91 | 5.62 | 4.99 |
| *ens*[c] | 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.[a] (conf. Int.[b]) | 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) |
| Number of parties[c] | 8 | | 4 | | 7 |
| *env*[d] | 2.76 | | 1.88 | | 1.31 |
| | Parliamentary | | Duma (2007) | | National assembly (2010) |
| Number of parties[a] | 5 | | 7 | | 12 |
| *env*[d] | 2.56 | | 2.22 | | 4.74 |
| *ens*[d] | 1.55 | | 1.94 | | 2.27 |

[a]This is the central estimate of the convergence coefficient.
[b]Conf. Int.: confidence interval rounded to the nearest tenth.
[c]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.

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 US, the *env* and the *ens* at the Presidential and House levels are closely aligned. There is little variation between the *env* and *env* 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 US, the convergence coefficient provides more information than do *env* or *env*. 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 higher effective number of parties as measured by the *env* and *env*. All three measures, *c*, *env*, and *ens*, indicate that the United Kingdom became more fragmented in 2010. Thus, in the 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 has 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 that a way of interpreting the arguments of Duverger [44] and Riker [45] 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 nonnegligible 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 1990's.

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 that we study show that parties could either converge to or diverge from the mean under anocracy as the equilibrium is fragile. Changes in valences, for example, of the autocrat or in voters' preferences, can lead small valence opposition parties to diverge from the mean and to mount popular uprisings as happened in previous elections in Georgia or in 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, $\beta$; the perceived ability of parties or candidates to govern as captured by their valences $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)$; and the dispersion of voters' preferences in the policy space, $\sigma^2$. 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 Great Britain. 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 multiparty 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

In this paper, Schofield's [9] Valence Theorem together with multinomial logit models of elections are used as a unifying framework to compare the convergence properties of parties across elections, countries, and political systems. We found evidence to support the hypothesis that in countries with proportional representation parties located away from the electoral mean.

We relate the convergence coefficient to the effective number of parties according to both vote (*env*) and seat (*ens*)

shares and showed how the characteristics of the electorate and the political regime under which parties operate. Then, compare the convergence coefficient to the fractionalization measures provided by the *env* and *ens*. The advantage of the convergence coefficient is that it is a summary statistic that incorporates the preferences of voters, the valence of parties, and the dispersion of voters and parties in the policy space.

## Appendix

## A. Confidence Intervals

Schofield's [9] Valence Theorem, presented in Section 2, perfectly predicts whether parties converge to or diverge from the electoral origin. Convergence or divergence depends on the value of the convergence coefficient $c \equiv 2\beta[1-2\rho_1]\sigma^2$ in (15) and on the Characteristic matrix of party 1 with lowest valence, $C_1 = 2\beta(1-2\rho_1)\nabla - I$ in (17). Both $c$ and $C_1$ depend on $\beta$ and on $\rho_1 = [\sum_{k=1}^{p}\exp(\lambda_k - \lambda_1)]^{-1}$ in (14).

The central estimate of $\beta$ and of $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)$ given by the MNL regressions depend on the sample of voters surveyed as do $\rho_1$, $c$, and $C_1$. Thus, to make inferences from empirical models we need the 95% confidence bounds of these estimates. Using these bounds we assert with some degree of certainty whether parties converge to or diverge from the electoral mean or if there is a knife-edge unstable equilibrium.

To build these bounds, we could perform simulations of the election. For each simulation we could generate the value of $\beta$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)$, $\rho_1$, $c$ and $C_1$. Repeating the simulation many times would generate their distribution from which we could derive their 95% confidence bounds. Note that $c$ and $C_1$ increase in $\beta$ and decrease in $\rho_1$. So that given the electoral covariance matrix $\nabla$ and variance/trace $\sigma^2$ in (16) of an election, when in a simulation $\beta$ has a low value and $\rho_1$ a high one, the values of $c$ and $C_1$ are low with the opposite being true when $\beta$ is high and $\rho_1$ is low. Since we have not performed simulations for the elections in this study, we use these features of $c$ and $C_1$ to generate our confidence bounds.

Let $L$ identify the lower and $U$ the upper bounds of the 95% confidence intervals of any estimate. The MNL estimation for an election gives the confidence bounds of $\beta$ and $\lambda_1$, $(\beta^L, \beta^U)$ and $[\lambda_1^L, \lambda_1^U]$. To estimate the bounds on $\rho_1$ in (14), $[\rho_1^L, \rho_1^U]$, we use the bounds on $\lambda_1$ and Taylor's Theorem, which asserts that

$$\begin{aligned}\rho_1(\lambda_1 \pm h) &= \rho_1(\lambda_1) \pm h\frac{d\rho_1}{d\lambda_1} \\ &= \rho_1(\lambda_1) \pm h\rho_1(\lambda_1)[1-\rho_1(\lambda_1)] \\ &= \rho_1(\lambda_1)[1 \pm h(1-\rho_1(\lambda_1))] = [\rho_1^L, \rho_1^U].\end{aligned} \tag{A.1}$$

Using (15) and the bounds on $\beta$ and $\rho_1$ we build the confidence intervals for the convergence coefficient $c$ as follows. In (15) use $\beta^L$ and $\rho_1^U$ to get the lower bound of $c$, $c^L$, and use $\beta^U$ and $\rho_1^L$ for the upper bound of $c$, $c^U$. The 95% confidence interval of the convergence coefficient is then

$$[c^L, c^U] = [2\beta^L[1-2\rho_1^U]\sigma^2, 2\beta^U[1-2\rho_1^L]\sigma^2]. \tag{A.2}$$

Following a similar procedure we estimate the bounds for $C_1$ using (17) and the corresponding bounds of $\beta$ and $\rho_1$ to get the bounds for the Hessian of the lowest valence party

$$[C_1^L, C_1^U] = [2\beta^L[1-2\rho_1^U]\nabla - I, 2\beta^U[1-2\rho_1^L]\nabla - I]. \tag{A.3}$$

Clearly, the bounds for $c$ and $C_1$ must be similar to those generated by repeated simulations.

Using these procedures, we now derive the 95% confidence intervals for the central estimates of $\rho_1$, $c$, and $C_1$ for each of the elections studied (see summary in Tables 2, 4, and 6). We first derive the detail of the confidence bounds for the 2000 US election, then in less detail those of other elections. Table 7 gives the values needed to derive the confidence intervals for the convergence coefficient of the election.

### *A.1. Convergence in Plurality Systems*

#### *A.1.1. Confidence Bounds for the 2000, 2004, and 2008 US Elections*

*US 2000 Election*. From Table 1, the 95% confidence interval for $\beta_{2000}^{\text{US}} = 0.82$ are $[\beta_{2000}^{\text{US}L}, \beta_{2000}^{\text{US}U}] = [0.82 \pm 1.96 \times 0.06] = [0.71, 0.93]$. Using (A.1), the bounds for $\rho_{\text{rep}}^{\text{US2000}} = 0.4$ in (20) are $[\rho_{\text{rep}}^{\text{US2000}L}, \rho_{\text{rep}}^{\text{US2000}U}] = [0.35, 0.44]$. Using these bounds and (18), the bounds for the convergence coefficient for the 2000 US election in (21) from (A.2) are

$$\begin{aligned}&[c_{2000}^{\text{US}L}, c_{2000}^{\text{US}U}] \\ &\quad = [2(0.71)(1-2\times 0.44)(1.17), \\ &\qquad 2(0.93)(1-2\times 0.35)(1.17)] \\ &\quad = [0.20, 0.65].\end{aligned} \tag{A.4}$$

With 95% confidence, the convergence coefficient is below 1 meeting the sufficient and thus necessary condition for convergence to the mean. The bounds on Bush's characteristic matrix in (22) from (A.3) are

$$\begin{aligned}&[C_{\text{rep}}^{\text{US2000}L}, C_{\text{rep}}^{\text{US2000}U}] \\ &\quad = \left[2(0.71)(1-2\times 0.44)\begin{bmatrix}0.58 & -0.20\\ -0.20 & 0.59\end{bmatrix} - I,\right. \\ &\qquad \left. 2(0.93)(1-2\times 0.35)\begin{bmatrix}0.58 & -0.20\\ -0.20 & 0.59\end{bmatrix} - I\right] \\ &\quad = \left[\begin{bmatrix}-0.90 & -0.03\\ -0.03 & -0.90\end{bmatrix}, \begin{bmatrix}-0.68 & -0.11\\ -0.11 & -0.67\end{bmatrix}\right].\end{aligned} \tag{A.5}$$

Since the eigenvalues of the lower and upper bounds of $C_{\text{rep}}^{\text{US2000}}$ are negative $[C_{\text{rep}}^{\text{US2000}L} = (-0.87, -0.93)$, $C_{\text{Bush}}^{\text{US2000}U} = (-0.79, -0.57)]$, with 95% confidence Bush's vote share is at a maximum when all parties locate at the mean. Thus, with a high degree of certainty the origin is a LNE for the 2000 US election.

*US 2004 Election*. From Table 1, the 95% confidence bounds of $\beta_{2004}^{\text{US}} = 0.95$ is $[\beta_{2004}^{\text{US}L}, \beta_{2004}^{\text{US}U}] = [0.95 \pm 1.96 \times 0.07] =$

[0.82, 1.08]. Using (A.1), the bounds of $\rho_{\text{rep}}^{\text{US2004}} = 0.4$ in (25) are $[\rho_{\text{rep}}^{\text{US2004}L}, \rho_{\text{rep}}^{\text{US2004}U}] = [0.35, 0.44]$. The bounds for $c_{2004}^{\text{US}} = 0.38$ in (21) from (A.2) and for the characteristic matrix of Bush $C_{\text{rep}}^{2004}$ in (27) from (A.3) are

$$\begin{aligned}\left[c_{2004}^{\text{US}L}, c_{2004}^{\text{US}U}\right] &= [2\,(0.82)\,(1-2\times 0.44)\,(1.17)\,,\\ &\qquad 2\,(1.08)\,(1-2\times 0.35)\,(1.17)]\\ &= [0.23, 0.76]\,,\\ \left[C_{\text{rep}}^{\text{US2004}L}, C_{\text{rep}}^{\text{US2004}U}\right]&\\ &= \left[2\,(0.82)\,(1-2\times 0.44)\begin{bmatrix}0.58 & -0.18\\ -0.18 & 0.59\end{bmatrix} - I,\right.\\ &\qquad \left. 2\,(1.08)\,(1-2\times 0.35)\begin{bmatrix}0.58 & -0.18\\ -0.18 & 0.59\end{bmatrix} - I\right]\\ &= \left[\begin{bmatrix}-0.89 & -0.04\\ -0.04 & -0.88\end{bmatrix}, \begin{bmatrix}-0.62 & -0.12\\ -0.12 & -0.62\end{bmatrix}\right].\end{aligned} \tag{A.6}$$

The convergence coefficient is significantly below 1. Bush maximizes his vote share when located at the origin since the eigenvalues of the lower and upper bounds of $C_{\text{rep}}^{\text{US2004}}$ are negative $[C_{\text{rep}}^{\text{US2004}L} = (-0.87, -0.93)$, $C_{\text{rep}}^{\text{US2004}U} = (-0.79, -0.57)]$. Thus, with 95% confidence Bush does not want to move from the mean implying that with a great certainty the origin is a LNE for the 2004 US election.

*US 2008 Election*. From Table 1 the bounds of $\beta_{2008}^{\text{US}} = 0.85$ are $[\beta_{2008}^{\text{US}L}, \beta_{2008}^{\text{US}U}] = [0.85 \pm 1.96 \times 0.06] = [0.73, 0.97]$. Using (A.1), those of $\rho_{\text{rep}}^{\text{US2008}}$ in (30) are $[\rho_{\text{rep}}^{\text{US2008}L}, \rho_{\text{rep}}^{\text{US2080}U}] = [0.26, 0.35]$. So that the bounds for $c_{2008}^{\text{US}} = 1.1$ in (31) from (A.2) and for McCain's characteristic matrix $C_{\text{rep}}^{\text{US2008}}$ in (32) from (A.3) are

$$\begin{aligned}\left[c_{2008}^{\text{US}L}, c_{2008}^{\text{US}U}\right] &= [2\,(0.73)\,(1-2\times 0.35)\,(1.63)\,,\\ &\qquad 2\,(0.97)\,(1-2\times 0.26)\,(1.63)]\\ &= [0.71, 1.52]\,,\\ \left[C_{\text{rep}}^{\text{US2008}L}, C_{\text{rep}}^{\text{US2008}U}\right]&\\ &= \left[2\,(0.73)\,(1-2\times 0.35)\begin{bmatrix}0.80 & -0.13\\ -0.13 & 0.83\end{bmatrix} - I,\right.\\ &\qquad \left. 2\,(0.97)\,(1-2\times 0.26)\begin{bmatrix}0.80 & -0.13\\ -0.13 & 0.83\end{bmatrix} - I\right]\\ &= \left[\begin{bmatrix}-0.65 & -0.06\\ -0.06 & -0.64\end{bmatrix}, \begin{bmatrix}-0.26 & -0.12\\ -0.12 & -0.23\end{bmatrix}\right].\end{aligned} \tag{A.7}$$

The convergence coefficient is not statistically different from 1 and thus meets the necessary but not the sufficient condition for convergence. Since the eigenvalues of the lower and upper bounds of $C_{\text{rep}}^{\text{US2008}}$ are negative $[C_{\text{rep}}^{\text{US2008}L} = (-0.75, -0.59)$, $C_{\text{rep}}^{\text{US2008}U} = (-0.37, -0.12)]$, then with 95% confidence McCain stays at the origin. With a high degree of certainty the mean is an LNE for the 2008 US election.

*A.1.2. Confidence Bounds for the 2005 and 2010 UK Elections*

*UK 2005 Election*. From Table 1, the bounds of $\beta_{2005}^{\text{UK}} = 0.15$ are $[\beta_{2005}^{\text{UK}L}, \beta_{2005}^{\text{UK}U}] = [0.15 \pm 1.96 \times 0.01] = [0.13, 0.17]$. Using (A.1), those for $\rho_{\text{lib}}^{\text{UK2005}}$ in (35) are $[\rho_{\text{lib}}^{\text{UK2005}L}, \rho_{\text{lib}}^{\text{UK2005}U}] = [0.18, 0.32]$, so that those for $c_{2005}^{\text{UK}}$ in (36) from (A.2) and for the Liberal Democrats' characteristic matrix $C_{\text{lib}}^{\text{UK2005}}$ in (37) from (A.3) are

$$\begin{aligned}\left[c_{2005}^{\text{UK}L}, c_{2005}^{\text{UK}U}\right] &= [2\,(0.13)\,(1-2\times 0.32)\,(5.61)\,,\\ &\qquad 2\,(0.17)\,(1-2\times 0.18)\,(5.61)]\\ &= [0.51, 1.25]\,,\\ \left[C_{\text{lib}}^{\text{UK2005}L}, C_{\text{lib}}^{\text{UK2005}U}\right]&\\ &= \left[2\,(0.13)\,(1-2\times 0.32)\begin{bmatrix}1.65 & 0.00\\ 0.00 & 3.96\end{bmatrix} - I,\right.\\ &\qquad \left. 2\,(0.17)\,(1-2\times 0.18)\begin{bmatrix}1.65 & 0.00\\ 0.00 & 3.96\end{bmatrix} - I\right]\\ &= \left[\begin{bmatrix}-0.85 & 0.00\\ 0.00 & -0.64\end{bmatrix}, \begin{bmatrix}-0.63 & 0.00\\ 0.00 & -0.12\end{bmatrix}\right].\end{aligned} \tag{A.8}$$

With $c_{2005}^{\text{UK}}$ not significantly different from 1, the necessary but not the sufficient condition for convergence to the mean has been met. The eigenvalues of the bounds on $C_{\text{lib}}^{\text{UK2005}}$ are negative, $[C_{\text{lib}}^{\text{UK2005}L} = (-0.85, -0.64)$, $C_{\text{lib}}^{\text{UK2005}U} = (-0.37, -0.12)]$. With 95% confidence the LibDem locate at the origin and the mean is an LNE of the 2005 UK election.

*UK 2010 Election*. From Table 1, the bounds of $\beta_{2010}^{\text{UK}} = 0.86$ are $[\beta_{2010}^{\text{UK}L}, \beta_{2010}^{\text{UK}U}] = [0.86 \pm 1.96 \times 0.02] = [0.81, 0.90]$. Using (A.1), those for $\rho_{\text{lab}}^{\text{UK2010}}$ in (40) are $[\rho_{\text{lab}}^{\text{UK2010}L}, \rho_{\text{lab}}^{\text{UK2010}U}] = [0.29, 0.32]$. So that those for $c_{\text{UK}}^{2010}$ in (41) from (A.2) and for Labour's characteristic matrix $C_{\text{lab}}^{\text{UK2010}}$ in (42) from (A.3) are

$$\begin{aligned}\left[c_{\text{UK}}^{2010L}, c_{\text{UK}}^{2010U}\right] &= [2\,(0.81)\,(1-2\times 0.32)\,(1.46)\,,\\ &\qquad 2\,(0.90)\,(1-2\times 0.29)\,(1.46)]\\ &= [0.86, 1.10]\,,\\ \left[C_{\text{lib}}^{\text{UK2010}L}, C_{\text{lib}}^{\text{UK2010}U}\right]&\\ &= \left[2\,(0.81)\,(1-2\times 0.32)\begin{bmatrix}0.60 & 0.07\\ 0.07 & 0.86\end{bmatrix} - I,\right.\\ &\qquad \left. 2\,(0.90)\,(1-2\times 0.29)\begin{bmatrix}0.60 & 0.07\\ 0.07 & 0.86\end{bmatrix} - I\right]\\ &= \left[\begin{bmatrix}-0.65 & 0.04\\ 0.04 & -0.49\end{bmatrix}, \begin{bmatrix}-0.55 & 0.05\\ 0.05 & -0.35\end{bmatrix}\right].\end{aligned} \tag{A.9}$$

The convergence coefficient meets the necessary but not the sufficient condition for convergence to the mean as is not significantly different from 1. The eigenvalues of the bounds of $C_{\text{lib}}^{\text{UK2010}}$ are negative, $[C_{\text{lab}}^{\text{UK2010}L} = (-0.66, -0.48), C_{\text{lab}}^{\text{UK2015}U} = (-0.56, -0.34)]$. Thus, with 95% confidence Labour does not

want to move from the origin and the origin is an LNE of the model of the 2010 UK election.

*A.2. Convergence in Proportional Systems*

*A.2.1. Confidence Bounds for the 1996 Israeli Election.* From Table 3, the bounds of $\beta^{\mathrm{I}}_{1996} = 1.207$ are $[\beta^{IL}_{1996}, \beta^{IU}_{1996}] = [1.207 \pm 1.96 \times 0.065] = [1.076, 1.338]$. Using (A.1), those for $\rho^{\mathrm{I}1996}_{\mathrm{TW}}$ in (45) are $[\rho^{\mathrm{I}1996L}_{\mathrm{TW}}, \rho^{\mathrm{I}1996U}_{\mathrm{TW}}] = [0.006, 0.034]$, implying that those of $c^{\mathrm{I}}_{1996}$ in (46) from (A.2) and for the TW's characteristic matrix $C^{\mathrm{I}1996}_{\mathrm{TW}}$ in (47) from (A.3) are

$$\begin{aligned}
\left[c^{IL}_{1996}, c^{IU}_{1996}\right] &= [2\,(1.076)\,(1-2\times 0.034)\,(1.732)\,,\\
&\qquad 2\,(1.338)\,(1-2\times 0.006)\,(1.732)]\\
&= [3.474, 4.579]\,,\\
&\left[C^{\mathrm{I}1996L}_{\mathrm{TW}}, C^{\mathrm{I}1996U}_{\mathrm{TW}}\right]\\
&= \left[2\,(1.076)\,(1-2\times 0.034)\begin{bmatrix}1.00 & 0.591\\ 0.591 & 0.732\end{bmatrix} - I,\right.\\
&\qquad \left. 2\,(1.338)\,(1-2\times 0.006)\begin{bmatrix}1.00 & 0.591\\ 0.591 & 0.732\end{bmatrix} - I\right]\\
&= \left[\begin{bmatrix}1.006 & 1.185\\ 1.185 & 0.468\end{bmatrix}, \begin{bmatrix}1.644 & 1.563\\ 1.563 & 0.935\end{bmatrix}\right].
\end{aligned} \tag{A.10}$$

Since $c^{\mathrm{I}}_{1996}$ is significantly greater than 2, the necessary condition for convergence to the electoral mean is *not* met. The lower and upper bounds of $C^{\mathrm{I}1996}_{\mathrm{TW}}$ have one negative and one positive eigenvalue, $[C^{\mathrm{I}1996L}_{TW} = (-0.48, 1.95)$, $C^{\mathrm{I}1996U}_{\mathrm{TW}} = (-0.313, 2.892)]$, TW is at a saddle point at both bounds. Thus, with 95% confidence TW locates away from the origin and the origin *fails* to be a LNE for the 1996 Israeli election.

*A.2.2. Confidence Bounds for the 1999 and 2002 Turkish Elections*

*1999 Turkish Election.* From Table 3, the bounds of $\beta^{\mathrm{T}}_{1999} = 0.375$ are $[\beta^{TL}_{1999}, \beta^{TU}_{1999}] = [0.375 \pm 1.96 \times 0.088] = [0.203, 0.547]$. Using (A.1), those for $\rho^{\mathrm{T}1999}_{\mathrm{FP}}$ in (50) are $[\rho^{\mathrm{T}1999L}_{\mathrm{FP}}, \rho^{\mathrm{T}1999U}_{\mathrm{FP}}] = [0.046, 0.145]$, so that those of $c^{\mathrm{T}}_{1999}$ in (51) from (A.2) and for the FP's characteristic matrix $C^{\mathrm{T}1999}_{\mathrm{FP}}$ in (52) from (A.3) are

$$\begin{aligned}
\left[c^{TL}_{1999}, c^{TU}_{1999}\right] &= [2\,(0.203)\,(1-2\times 0.145)\,(2.34)\,,\\
&\qquad 2\,(0.547)\,(1-2\times 0.046)\,(2.34)]\\
&= [0.675, 2.234]\,,\\
&\left[C^{\mathrm{T}1999L}_{\mathrm{FP}}, C^{\mathrm{T}1999U}_{\mathrm{FP}}\right]\\
&= \left[2\,(0.203)\,(1-2\times 0.145)\begin{bmatrix}1.20 & 0.78\\ 0.78 & 1.14\end{bmatrix} - I,\right.\\
&\qquad \left. 2\,(0.547)\,(1-2\times 0.046)\begin{bmatrix}1.20 & 0.78\\ 0.78 & 1.14\end{bmatrix} - I\right]\\
&= \left[\begin{bmatrix}-0.654 & 0.225\\ 0.225 & -0.671\end{bmatrix}, \begin{bmatrix}0.192 & 0.775\\ 0.775 & 0.132\end{bmatrix}\right].
\end{aligned} \tag{A.11}$$

Since $c^{\mathrm{T}}_{1999}$ is significantly greater than 2, the necessary condition for convergence to the mean is *not* met. $C^{\mathrm{T}1999L}_{\mathrm{FP}}$ has two negative eigenvalues, $[C^{\mathrm{T}1999L}_{\mathrm{FP}} = (-0.888, -0.437)]$ indicating that at the lower bound FP has no incentive to move from the origin. However, $C^{\mathrm{T}1999U}_{\mathrm{FP}}$ has one negative and one positive eigenvalue $C^{\mathrm{T}1999U}_{\mathrm{FP}} = (-0.614, 0.938)$; thus, FP is at a saddlepoint at the upper bound and wants to move from the mean. At the central estimate of $C^{\mathrm{T}1999}_{\mathrm{FP}}$ given in (52) FP is also at a saddlepoint. It is more probable that FP wants to move and that the electoral mean is *not* a LNE of 1999 Turkish election.

*2002 Turkish Election.* From Table 3, the bounds of $\beta^{\mathrm{T}}_{2002} = 1.52$ are $[\beta^{TL}_{2002}, \beta^{TU}_{2002}] = [1.52 \pm 1.96 \times 0.12] = [1.285, 1.755]$. Using (A.1), those for $\rho^{\mathrm{T}2002}_{\mathrm{ANAP}}$ in (55) are $[\rho^{\mathrm{T}2002L}_{\mathrm{ANAP}}, \rho^{\mathrm{T}2002U}_{\mathrm{ANAP}}] = [0.038, 0.133]$, implying that those of $c^{\mathrm{T}}_{2002}$ in (56) from (A.2) and for the ANAP's characteristic matrix $C^{\mathrm{T}2002}_{\mathrm{ANAP}}$ in (57) from (A.3) are

$$\begin{aligned}
\left[c^{TL}_{2002}, c^{TU}_{2002}\right] &= [2\,(1.285)\,(1-2\times 0.133)\,(2.33)\,,\\
&\qquad 2\,(1.755)\,(1-2\times 0.038)\,(2.33)]\\
&= [4.338, 7.438]\,,\\
&\left[C^{\mathrm{T}2002L}_{\mathrm{ANAP}}, C^{\mathrm{T}2002U}_{\mathrm{ANAP}}\right]\\
&= \left[2\,(1.285)\,(1-2\times 0.133)\begin{bmatrix}1.18 & 0.74\\ 0.74 & 1.15\end{bmatrix} - I,\right.\\
&\qquad \left. 2\,(1.755)\,(1-2\times 0.038)\begin{bmatrix}1.18 & 0.74\\ 0.74 & 1.15\end{bmatrix} - I\right]\\
&= \left[\begin{bmatrix}-0.660 & 0.213\\ 0.213 & -0.669\end{bmatrix}, \begin{bmatrix}0.172 & 0.735\\ 0.735 & 0.142\end{bmatrix}\right].
\end{aligned} \tag{A.12}$$

Since $c^{\mathrm{T}}_{2002}$ is significantly greater than 2, the necessary condition for convergence to the mean has *not* been met. The eigenvalues of $C^{\mathrm{T}2002L}_{\mathrm{ANAP}}$ are all negative, $C^{\mathrm{T}2002L}_{\mathrm{ANAP}} = (-0.878, -0.451)$ so that at the lower bound ANAP remain at the mean. However, at $C^{\mathrm{T}2002U}_{\mathrm{ANAP}}$ there is one negative and one positive eigenvalue $C^{\mathrm{T}2002U}_{\mathrm{ANAP}} = (-0.578, 0.892)$, ANAP is at a saddlepoint and wants to move. At the central estimate of $C^{\mathrm{T}2002}_{\mathrm{ANAP}}$ in (57) the eigenvalues are both positive and ANAP is minimizing its vote share. There is a high likelihood that ANAP wants to move from the origin and that the electoral mean is *not* a LNE of 2002 Turkish election.

*A.2.3. Confidence Bounds for the 1997 Polish Election.* From Table 3, the bounds of $\beta^{\mathrm{P}}_{1997} = 1.739$ are $[\beta^{PL}_{1997}, \beta^{PU}_{1997}] = [1.739 \pm 1.96 \times 0.12] = [1.512, 1.966]$. Using (A.1), those for $\rho^{\mathrm{P}1997}_{\mathrm{UPR}}$ in (60) are $[\rho^{PL}_{1997}, \rho^{PU}_{1997}] = [0.002, 0.022]$, so that those of $c^{\mathrm{P}}_{1997}$ in (61) from (A.2) and for the UPR's characteristic matrix $C^{\mathrm{P}1997}_{\mathrm{UPR}}$ in (62) from (A.3) are

$$\begin{aligned}
\left[c^{PL}_{1997}, c^{PU}_{1997}\right] &= [2\,(1.512)\,(1-2\times 0.022)\,(2)\,,\\
&\qquad 2\,(1.966)\,(1-2\times 0.002)\,(2)]\\
&= [5.782, 7.833]\,,
\end{aligned}$$

$$\left[C_{1997}^{PL}, C_{1997}^{PU}\right] = \left[2\,(1.512)\,(1-2\times 0.022)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - I,\ 2\,(1.966)\,(1-2\times 0.002)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - I\right] = \left[\begin{bmatrix}1.891 & 0.000\\ 0.000 & 1.891\end{bmatrix}, \begin{bmatrix}2.916 & 0.000\\ 0.000 & 2.916\end{bmatrix}\right]. \tag{A.13}$$

With $c_{1997}^{P}$ significantly greater than 2, the necessary condition for convergence to the mean is *not* met. The eigenvalues of the bounds of $C_{1997}^{P}$ are positive, $[C_{UPR}^{P1997L} = (1.891, 1.891), C_{UPR}^{P1997L} = (2.916, 2.916)]$ as are those of the central estimate of $C_{1997}^{P}$ in (62). Thus, with a high probability UPR will not locate at the mean and the electoral mean is *not* a LNE of 1997 Polish election.

### A.3. Convergence in Anocracies

*A.3.1. Confidence Bounds for the 2008 Georgian Election.* From Table 5, the bounds of $\beta_{2008}^{G} = 0.78$ are $[\beta_{2008}^{GL}, \beta_{2008}^{GU}] = [0.78 \pm 1.96 \times 0.06] = [0.66, 0.89]$. Using (A.1), those for $\rho_{N}^{G2008} = 0.05$ in (65) are $[\rho_{N}^{G200L8}, \rho_{N}^{G2008U}] = [0.03, 0.07]$. So that those of $c_{2008}^{G}$ in (66) from (A.2) and for Natelashvili's characteristic matrix $C_{N}^{G2008}$ in (67) from (A.3) are

$$\left[c_{2008}^{GL}, c_{2008}^{GU}\right] = [2\,(0.66)\,(1-2\times 0.07)\,(1.73),\ 2\,(0.89)\,(1-2\times 0.03)\,(1.73)] = [1.99, 2.89],$$
$$\left[C_{N}^{G2008L}, C_{N}^{G2008U}\right] = \left[2\,(0.66)\,(1-2\times 0.07)\begin{bmatrix}0.82 & 0.03\\ 0.03 & 0.91\end{bmatrix} - I,\ 2\,(0.89)\,(1-2\times 0.03)\begin{bmatrix}0.82 & 0.03\\ 0.03 & 0.91\end{bmatrix} - I\right] = \left[\begin{bmatrix}-0.06 & 0.03\\ 0.03 & 0.05\end{bmatrix}, \begin{bmatrix}0.37 & 0.05\\ 0.05 & 0.52\end{bmatrix}\right]. \tag{A.14}$$

Since $c_{2008}^{G}$ is *not* statistically different from 2, the necessary condition for convergence is *not* met. The lower bound of $C_{N}^{G2008}$ has one negative and one positive eigenvalue $[C_{N}^{G2008L} = (-0.068, 0.058)]$, so that at the lower bound Natelashvili's vote share function is at a saddlepoint. The upper bound has two positive eigenvalues $[C_{N}^{G200U} = (0.355, 0.535)]$ so that at the upper bound Natelashvili is minimizing his vote share. At the central estimate of $C_{N}^{G2008}$ in (67) Natelashvili is also minimizing his vote share. Thus, with a high probability Natelashvili diverges from the mean and the mean *cannot* be a LNE of the 2008 Georgian election.

*A.3.2. Confidence Bounds for the 2007 Russian Election.* From Table 5, the bounds of $\beta_{2007}^{R} = 0.181$ are $[\beta_{2007}^{RL}, \beta_{2007}^{RU}] = [0.18 \pm 1.96 \times 0.01] = [0.15, 0.20]$. Using (A.1), those for $\rho_{SR}^{R2007} = 0.07$ in (70) are $[\rho_{SR}^{R2007L}, \rho_{SR}^{R2007U}] = [0.04, 0.12]$. So that those of $c_{2007}^{R}$ in (71) from (A.2) and for SR's characteristic matrix $C_{SR}^{R2007}$ in (72) from (A.3) are

$$\left[c_{2007}^{RL}, c_{2007}^{RU}\right] = [2\,(0.15)\,(1-2\times 0.12)\,(5.9),\ 2\,(0.15)\,(1-2\times 0.04)\,(5.9)] = [1.35, 2.28],$$
$$\left[C_{SR}^{R2007L}, C_{SR}^{R2007U}\right] = \left[2\,(0.15)\,(1-2\times 0.12)\begin{bmatrix}2.95 & 0.13\\ 0.13 & 2.95\end{bmatrix} - I,\ 2\,(0.2)\,(1-2\times 0.04)\begin{bmatrix}2.95 & 0.13\\ 0.13 & 2.95\end{bmatrix} - I\right] = \left[\begin{bmatrix}-0.33 & 0.03\\ 0.03 & -0.33\end{bmatrix}, \begin{bmatrix}0.14 & 0.05\\ 0.05 & 0.14\end{bmatrix}\right]. \tag{A.15}$$

With $c_{2007}^{R}$ not significantly different from 2, the necessary for convergence is *not* met. The lower bound of $C_{SR}^{R2007}$ has two negative eigenvalues, $[C_{SR}^{R2007L} = (-0.30, -0.36)]$ implying that at lower bound SR's vote share is at a maximum and SR stays at the origin. However, at the upper bound there are two positive eigenvalues, $[C_{SR}^{R2007U} = (0.09, 0.19)]$. Thus at the upper bound SR's vote share is at minimum and SR wants to move. At the central estimate of $C_{SR}^{R2007}$ in (72) SR also has two negative eigenvalues suggesting that SR wants to remain at the origin. So it seems more likely that SR will stay at the origin and that the mean is a LNE of the 2007 Russian election.

*A.3.3. Confidence Bounds for the 2010 Azerbaijani Election.* From Table 5 the bounds for $\beta_{2010}^{A} = 1.34$ are $[\beta_{2010}^{AL}, \beta_{2010}^{AU}] = [1.34 \pm 1.96 \times 0.29] = [0.77, 1.91]$. Using (A.1), those for $\rho_{AXCP\text{-}MP}^{A2010} = 0.21$ in (75) are $[\rho_{AXCP\text{-}MP}^{A2010L}, \rho_{AXCP\text{-}MP}^{A2010U}] = [0.08, 0.47]$. So that those of $c_{2010}^{A}$ in (76) from (A.2) and for AXCP-MP's characteristic matrix $C_{AXCP\text{-}MP}^{A2010}$ in (77) from (A.3) are

$$\left[c_{2010}^{AL}, c_{2010}^{AU}\right] = [2\,(0.77)\,(1-2\times 0.47)\,(0.93),\ 2\,(1.91)\,(1-2\times 0.08)\,(0.93)] = [0.085, 2.984],$$
$$\left[C_{AXCP\text{-}MP}^{A2010L}, C_{AXCP\text{-}MP}^{A2010U}\right] = [2\,(0.77)\,(1-2\times 0.47)\,(0.445) - 1,\ 2\,(1.91)\,(1-2\times 0.08)\,(0.445) - 1] = [0.037, 1.428]. \tag{A.16}$$

With $c_{2010}^{A}$ not significantly different from 1, the dimension of the policy space, the necessary and the sufficient (in this case

the same) conditions for convergence are *not* met. This one-dimensional characteristic matrix has positive eigenvalues at the lower and upper bounds as does the central estimate of $C^{A2010}_{AXCP\text{-}MP} = 0.445$ in (77). It is then very likely that AXCP-MP locates far from the origin and that the electoral mean is *not* an LNE for the 2010 election in Azerbaijan.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

Prepared for presentation at the Journées Louis-André Gérard-Varet, 24-28 June, Marseille, and for presentation at the joint LSE-WashU workshop on Comparative political economy, September, 2013. This paper is based on work supported by NSF grant 0715929 and a Weidenbaum Center grant. Earlier versions were completed while Gallego was a visitor at the Center and later while Schofield was the Glenn Campbell and Rita Ricardo-Campbell National Fellow at the Hoover Institution, Stanford.

## References

[1] A. Downs, *An Economic Theory of Democracy*, Harper and Row, New York, NY, USA, 1957.

[2] W. H. Riker and P. C. Ordeshook, *An Introduction to Positive Political Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1973.

[3] D. Stokes, "Spatial models and party competition," *The American Political Science Review*, vol. 57, pp. 368–377, 1963.

[4] D. Stokes, "Valence politics," in *Electoral Politics*, D. Kavanagh, Ed., pp. 141–164, Clarendon Press, Oxford, UK, 1992.

[5] H. Clarke, D. Sanders, M. Stewart, and P. Whiteley Oxford University Press, Oxford, UK, 2005.

[6] H. Clarke, D. Sanders, M. Stewart, and P. Whiteley, *Performance Politics and the British Voter*, Cambridge University Press, Cambridge, UK, 2009.

[7] T. J. Scotto, H. D. Clarke, A. Kornberg et al., "The dynamic political economy of support for Barack Obama during the 2008 presidential election campaign," *Electoral Studies*, vol. 29, no. 4, pp. 545–556, 2010.

[8] H. D. Clarke, T. J. Scotto, and A. Kornberg, "Valence politics and economic crisis: electoral choice in Canada 2008," *Electoral Studies*, vol. 30, no. 3, pp. 438–449, 2011.

[9] N. Schofield, "The mean voter theorem: necessary and sufficient conditions for convergent equilibrium," *Review of Economic Studies*, vol. 74, no. 3, pp. 965–980, 2007.

[10] J. M. Enelow and M. J. Hinich, "Nonspatial candidate characteristics and electoral competition," *Polish Journal of Ecology*, vol. 44, pp. 115–131, 1982.

[11] J. M. Enelow and M. J. Hinich, *The Spatial Theory of Voting*, Cambridge University Press, Cambridge, UK, 1984.

[12] J. M. Enelow and M. J. Hinich, "A general probabilistic spatial theory of elections," *Public Choice*, vol. 61, no. 2, pp. 101–113, 1989.

[13] D. Sanders, H. D. Clarke, M. C. Stewart, and P. Whiteley, "Downs, stokes and the dynamics of electoral choice," *British Journal of Political Science*, vol. 41, no. 2, pp. 287–314, 2011.

[14] R. D. McKelvey and J. W. Patty, "A theory of voting in large elections," *Games and Economic Behavior*, vol. 57, no. 1, pp. 155–180, 2006.

[15] M. Laakso and R. Taagepera, "Effective number of parties: a measure with applications to West Europe," *Competition and Political Science*, vol. 12, pp. 3–27, 1979.

[16] N. Schofield and I. Sened, *Multiparty Democracy: Elections and Legislative Politics*, Cambridge University Press, Cambridge, UK, 2006.

[17] S. Ansolabare and J. M. Snyder, "Valence politics and equilibrium in spatial election models," *Public Choice*, vol. 103, no. 3-4, pp. 327–336, 2000.

[18] T. Groseclose, "A model of candidate location when one candidate has a valence advantage," *American Journal of Political Science*, vol. 45, no. 4, pp. 862–886, 2001.

[19] E. Aragones and T. R. Palfrey, "Mixed equilibrium in a Downsian model with a favored candidate," *Journal of Economic Theory*, vol. 103, no. 1, pp. 131–161, 2002.

[20] E. Aragones and T. R. Palfrey, "Electoral competition between two candidates of different quality: the effects of candidate ideology and private information," *Social Choice and Strategic Decisions Studies in Choice and Welfare*, pp. 93–112, 2005.

[21] N. Schofield, "Valence competition in the spatial stochastic model," *Journal of Theoretical Politics*, vol. 15, no. 4, pp. 371–383, 2003.

[22] N. Schofield, G. Miller, and A. Martin, "Critical elections and political realignments in the USA: 1860–2000," *Political Studies*, vol. 51, no. 2, pp. 217–442, 2003.

[23] G. Miller and N. Schofield, "Activists and partisan realignment in the United States," *American Political Science Review*, vol. 97, no. 2, pp. 245–260, 2003.

[24] N. Schofield and G. Miller, "Elections and activist coalitions in the United States," *American Journal of Political Science*, vol. 51, no. 3, pp. 518–531, 2007.

[25] M. Peress, "The spatial model with non-policy factors: a theory of policy-motivated candidates," *Social Choice and Welfare*, vol. 34, no. 2, pp. 265–294, 2010.

[26] H. D. Clarke, A. Kornberg, J. MacLeod, and T. Scotto, "Too close to call: political choice in Canada, 2004," *Political Science and Politics*, vol. 38, no. 2, pp. 247–253, 2005.

[27] H. D. Clarke, A. Kornberg, T. Scotto, and J. Twyman, "Flawless campaign, fragile victory: voting in Canada's 2006 federal election," *Political Science and Politics*, vol. 39, no. 4, pp. 815–819, 2006.

[28] H. D. Clarke, A. Kornberg, and T. Scotto, *Making Political Choices*, Toronto University Press, Toronto, Canada, 2009.

[29] N. Schofield, "A valence model of political competition in Britain: 1992–1997," *Electoral Studies*, vol. 24, no. 3, pp. 347–370, 2005.

[30] N. Schofield, C. Claassen, U. Ozdemir, and A. Zakharov, "Estimating the effects of activists in two-party and multi-party systems: comparing the United States and Israel," *Social Choice and Welfare*, vol. 36, no. 3, pp. 483–518, 2011.

[31] N. Schofield, C. Claassen, M. Gallego, and U. Ozdemir, "Empirical and formal models of the US presidential elections in 2004 and 2008," in *The Political Economy of Institutions, Democracy and Voting*, N. Schofield and G. Caballero, Eds., pp. 217–258, Springer, Berlin, Germany, 2011.

[32] K. Train, *Discrete Choice Methods for Simulation*, Cambridge University Press, Cambridge, UK, 2003.

[33] J. K. Dow and J. W. Endersby, "Multinomial probit and multinomial logit: a comparison of choice models for voting research," *Electoral Studies*, vol. 23, no. 1, pp. 107–122, 2004.

[34] K. M. Quinn, A. D. Martin, and A. B. Whitford, "Voter choice in multi-party democracies: a test of competing theories and models," *American Journal of Political Science*, vol. 43, no. 4, pp. 1231–1247, 1999.

[35] J. E. Roemer, "A theory of income taxation where politicians focus upon core and swing voters," *Social Choice and Welfare*, vol. 36, no. 3, pp. 383–421, 2011.

[36] N. Schofield, "Equilibria in the spatial stochastic model of voting with party activists," *Review of Economic Design*, vol. 10, no. 3, pp. 183–203, 2006.

[37] N. Schofield, M. Gallego, and J. Jeon, "Leaders, voters and activists in the elections in Great Britain 2005 and 2010," *Electoral Studies*, vol. 30, no. 3, pp. 484–496, 2011.

[38] A. Arian and M. Shamir, *The Election in Israel: 1996*, SUNY Press, Albany, NY, USA, 1999.

[39] N. Schofield, M. Gallego, U. Ozdemir, and A. Zakharov, "Competition for popular support: a valence model of elections in Turkey," *Social Choice and Welfare*, vol. 36, no. 3, pp. 451–482, 2011.

[40] N. Schofield, J. S. Jeon, M. Muskhelishvili, U. Ozdemir, and M. Tavits, "Modeling elections in post-communist regimes: voter perceptions, political leaders and activists," in *The Political Economy of Institutions, Democracy and Voting*, N. Schofield and G. Caballero, Eds., pp. 259–301, Springer, Berlin, Germany, 2011.

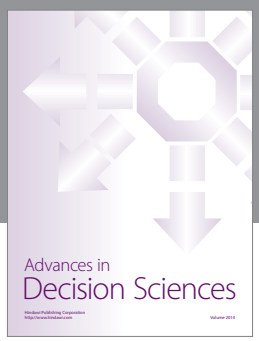
[41] D. L. Epstein, R. Bates, J. Goldstone, I. Kristensen, and S. O'Halloran, "Democratic transitions," *American Journal of Political Science*, vol. 50, no. 3, pp. 551–569, 2006.

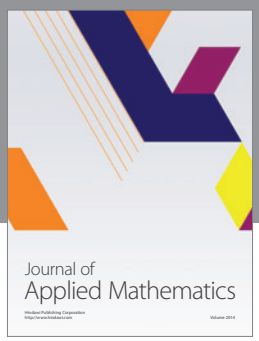
[42] N. Schofield, M. Gallego, J. Jeon, and M. Muskhelishvili, "Modelling elections in the Caucasus," *Journal of Elections, Public Opinion and Parties*, vol. 22, no. 2, pp. 187–214, 2012.

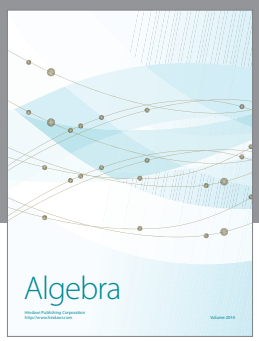
[43] N. Schofield and A. Zakharov, "A stochastic model of the 2007 Russian Duma election," *Public Choice*, vol. 142, no. 1-2, pp. 177–194, 2010.

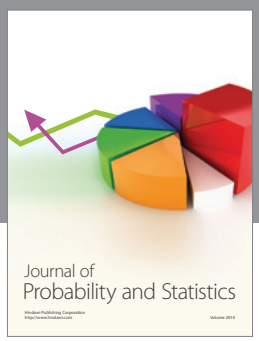
[44] M. Duverger, *Political Parties: Their Organization and Activity in the Modern State*, John Wiley & Sons, New York, NY, USA, 1954.

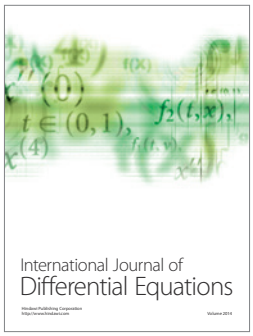
[45] W. H. Riker, *Democracy in the United States*, Macmillan, New York, NY, USA, 1953.

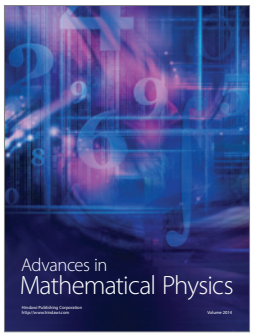
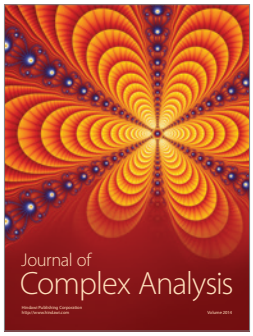
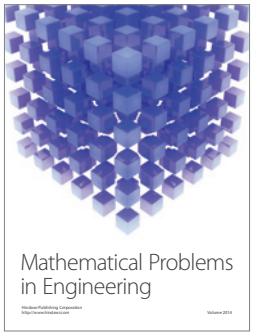
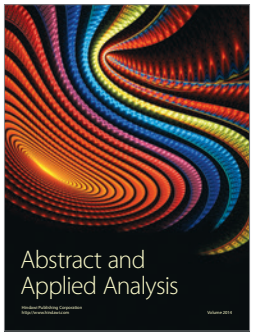
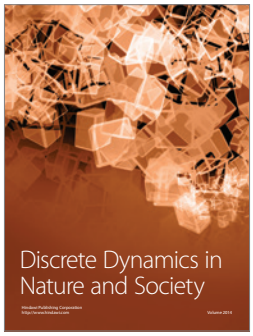
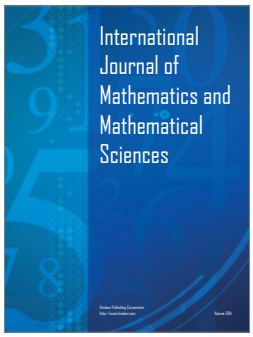
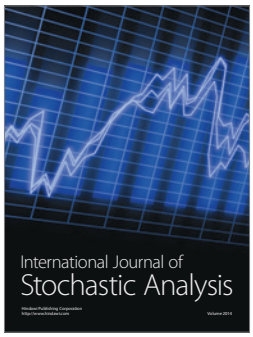










Hindawi

Submit your manuscripts at
http://www.hindawi.com