University *of* Leicester

DEPARTMENT OF ECONOMICS

A Model of Party Formation and Competition

Dan Ladley, University of Leicester, UK

James Rockey, University of Leicester, UK

Working Paper No. 10/17
May 2010
Updated September 2010

A Model of Party Formation and Competition

Daniel Ladley

Department of Economics
University of Leicester*

James Rockey

Department of Economics
University of Leicester

September 2010

Abstract

This paper investigates the behaviour of a Citizen-Candidate Model in a simple framework with many large constituencies, many policy dimensions, and endogenous coalition formation. A model is simulated in which districts elect representatives who themselves interact to form parties. Competition between parties of different sizes and with different platforms is an emergent property of the model which leads to stable equilibria. The results demonstrate how the number of policy dimensions and representatives elected per electoral district influence the number, size, and relative locations of parties and consequently the possible equilibria. These results are obtained using a new algorithm for identifying and comparing equilibria found by simulation. Comparison with election data shows strong correspondence between the model's results and observed outcomes, including variation consistent with a form of Duverger's law.

*Email: Ladley-dl110@le.ac.uk, Rockey-jcr12@le.ac.uk

1 Introduction

A key feature of most modern representative democracies is that political competition is dominated by political parties. These parties vary a great deal - some are small and ideologically cohesive, others large collections of politicians with quite different views. Moreover, the size-distribution of parties varies meaningfully across democracies. The United States for most of its history has been dominated by two large parties whereas many European democracies have many parties of different sizes. This variation is important, in part because the set of possible governing coalitions and hence policy outcomes is contingent on the size-distribution of parties. As such, one might ask why do we observe this variation, and how does this variation depend on particular national characteristics?

Both of these questions have been the focus of much scholarly attention, but the approach of this paper is different to most previous work. There are three key characteristics of the model. Firstly, political parties are simply voluntary coalitions of elected politicians formed for mutual (electoral) benefit.[1] Secondly, a key feature of politics, in practice, is that not only are politicians themselves heterogenous but so are the electoral districts they represent, something which we also model explicitly. Finally, both the elected and their electors may have multidimensional preferences, that is any two of them may agree on some issues but disagree profoundly, on another, unrelated, issue. These features as well as the endogenous formation of parties means analysis of this model would be challenging analytically. The approach taken instead is to solve a dynamic computational analogue of the citizen-candidate model, of Osborne and Slivinski (1996) (henceforth OS) and Besley and Coate (1997) (henceforth BC), incorporating these three features and in which politicians may or may not form or join political parties.

To understand the variation in party size and the structure of party systems, two related questions are asked. Firstly, how does the form and the number of equilibria of the model vary with electoral system? The results are both intuitively sensible and empirically realistic. The second question is how does the number of policy dimensions affect outcomes, and evidence is provided, consistent with the results of Levy (2004), that moving from one to

[1] In reality, political parties may perform other important roles. These are not studied here, partly because of pronounced national differences in the other functions of parties.

multiple policy dimensions gives rise to more scope for agreement between politicians with different preferences. That is, that preferences in two or more dimensions are associated with fewer, but larger parties.

The second part of the paper studies how party structure depends on the electoral system. There are many reasons for variation in party structure, but a key source of variation is the proportionality of the electoral system.[2] A specific version of this claim is often referred to as Duverger's Law, which may be stated as 'elections using a plurality rule give rise to two party systems'. The results indeed find support for Duverger's law, and using data on post-1945 elections evidence is provided for a great deal of similarity between the size-distribution of parties predicted by the model, and observed outcomes.

2 Previous Literature

The version of the citizen-candidate model we study incorporates aspects of both those of OS and BC. The two papers have different intentions, a key focus of OS is the effects of different electoral systems. BC consider the more general question of the existence of equilibria with endogenous candidacy. An important difference in the models is that OS assume a continuum of sincere voters whilst BC model a finite number of strategic voters. Like the latter we model a (large) finite number of citizens, however voting is sincere as in OS.[3] The general form of BC's model doesn't assume a particular number of policy dimensions, or Euclidean preferences. There are, however, other fairly restrictive assumptions, and focuses largely on the existence of at least one equilibrium rather than the precise number or form. Although, the papers have different emphases and apply their models to different questions it is straightforward to identify the key similarities and differences between them. These two papers share a framework in which politicians emerge endogenously from the population of voters, they also both assume that politicians are unable to credibly commit to a policy different to the one

[2]Other sources of variation include further differences in the form of the electoral system, or specific electorates' preferences, for example, several countries, such as Canada, Spain and the UK, have well established minority parties associated with minority linguistic, national, or cultural groups.

[3]This assumption is important since as OS model the population as a continuum, sincere voting is amongst the strategic equilibria. This result hinges on the zero measure and hence the influence of individual voters, and is therefore not true for the finite population we model, even as the population becomes large.

they prefer. These assumptions of endogenous candidates and a lack of credible commitment have been characteristic of much of the subsequent literature, as has the distinction between sincere and strategic voters. It is useful to briefly consider more recent developments on this latter distinction, before moving to the issues of credible commitment, and political parties.

The motivation of Morelli (2004) would seem similar to that of this paper. Morelli's specific objective is to provide a framework where the *'Duvergian predictions can be studied even when the electorate is divided into multiple districts and candidates and parties are separate entities.'* He finds support for the Duvergian hypothesis, that plurality electoral systems lead to two party systems, and his setup incorporates what he claims are the necessary features of *'strategic voters, strategic parties, and strategic candidates, within and across districts'.* As will be discussed below, in his model political parties provide a means of coordinating voters within and between districts as well as a method by which coalitions of heterogeneous candidates can commit to a shared policy-platform. However, Morelli (2004)'s emphasis on strategic agents whilst conceptually different to our approach may matter little in practice. Dutta, Jackson and Breton (2001) show that outcome of all democratic voting procedures depend on the candidacy decision of those who don't (cannot) win the election in question. One of the most important contributions of Morelli (2004) is that often, but not always, with endogenous candidacy equilibrium rational (strategic) voting behaviour is sincere.[4] But, crucially, Morelli shows that "*the equilibrium policy outcome is not affected by whether voters are expected to be sincere or strategic. Thus, the sincere vs. strategic voting issue is irrelevant for welfare analysis.*"

Political parties have many roles in a democracy, and a variety of these have been modelled. These are surveyed by Merlo (2006) and Dhillon (2005) include parties as representing specific constituencies or groups (e.g. Snyder and Ting (2002) or Roemer (1999)) or voter coordination devices Morelli (2004). In Osborne and Tourky (2008) parties are modelled as a cost-sharing technology. A key result of their model is that costly voting implies that a single hegemonic party is never an equilibrium. As they note, this contrasts with the results of Morelli (2004), where under plurality rule only one-party stands in equilibrium. They consider

[4]Specifically, he shows that under the plurality rule that *'equilibrium [strategic] voting behaviour is always sincere'.* But that in a proportional representation system if there is no party with more than half of the votes there will always be some voters who in equilibrium vote strategically

the extension of their model to the case of n policy dimensions and suggest that there will be at most 2^n parties. Intuitively, this result is very much only an upper bound on the number of (effective) parties we should expect as the number of policy-dimensions increases and leaves the question of how the number of parties varies with the dimensionality of the policy space largely unanswered. In the model of Levy (2004) parties are devices they allow politicians to credibly campaign on a platform known not to correspond to their most-favoured as party membership provides a complete contracting mechanism.[5][6] In the model below, this role of political parties emerges endogenously - candidates seeking re-election often stand with platforms (which would be implemented if their party were to win the election), different from their preferred policy if this changes the implemented policy sufficiently in their favour.

There is a small computational literature which analyzes political parties and their behaviour. The first paper of which we are aware to apply a computational approach to voting is that of Tullock and Campbell (1970) who analyzed computationally the problem of cyclical majorities in small committees with multi-dimensional preferences. They found that the impact of additional preference dimensions beyond two was small. Although our setting is different, the results of our model suggest similarly that the key difference is between having one or more than one dimension. A key early contribution was that of Kollman, Miller and Page (1992) who in contrast to much of the previous rational choice literature, studied the behaviour of boundedly rational parties. They argued that the, sometimes incomplete, platform convergence predicted by analytic models was robust to non-fully rational parties. This type of question, involving understanding the behaviour of a large number of boundedly rational agents clearly lends itself to simulation-based approaches. Many of the results of this paper are obtained by simulating our model many times, and analyzing the distribution of results. Our approach is

[5]When this role of political parties is important in equilibrium varies depending on the context. In the model of Levy (2004) the commitment device provided by political parties is unimportant with one policy dimension but allows for stable equilibria to exist in the case of multiple policy dimensions which wouldn't otherwise be possible.In Morelli (2004) the commitment technology allows parties comprised of candidates with different preferred policies to stand, but in his setting it is rarely important.

[6]There is a venerable literature that studies the case when politicians are only concerned by office, and as such can credibly commit to any policy position, that began with Downs (1957). An alternative literature building on the work of Wittman (1977) considered politicians motivated by ideology or policy who as such can only credibly commit to their preferred policy. Alesina (1988) analyzed the case when politicians care about both winning and the policy they then implement. Emphasizing that the tension between the two gave rise to a dynamic-inconsistency problem because in a one-period game voters cannot control politicians once they have been elected, i.e. there is an incomplete-contract problem.

therefore similar to that of Kollman, Miller and Page (1997) who study a Tiebout type model.

Recent work has studied the dynamics of party behaviour and in particular the interaction between different types of party. Laver (2005) investigates the dynamical properties of a democratic system. In particular, he shows that the interaction of parties distinguished by different behaviours. Some parties are ideological and as such don't move around the policy space, others move towards the median voter's preferences, the position of the largest party, or move randomly repeating moves that were successful. This gives rise to interesting and realistic party dynamics, without necessarily any stable equilibria. This framework is then applied to Irish politics where

> Starting from expert survey estimates of party and voter positions, and party decision rules, the model generates time series of individual party sizes, variations in these, and the cross-sectional variation of sizes between parties that look similar to published opinion poll series[...].

Laver and Schilperoord (2007) extend this analysis by endogenizing the birth and death of political parties. In this model, parties that fail to obtain a certain vote share 'die'. Parties are 'born' when citizens are sufficiently dissatisfied with the absence of a party sharing their views that they found one. Parties are often born at the extremes of the policy space and move to the centre where they often then die. The computational approach of Laver (2005), Laver and Schilperoord (2007) is very different to the one pursued here. Their focus is on the dynamic properties of competition between parties with pre-specified behavioural rules, over time. Whilst, the model presented here is also dynamic, the focus is its steady state and in particular the equilibrium distribution of parties as the type of electoral system and number of policy dimensions varies. Similarly, whilst both are compared to empirical data, Laver (2005) use time-series data, whereas we focus on the cross-section.

The paper proceeds as follows. The next section presents the model, and discusses the results from simulating a single constituency. Section 4 proceeds to consider the equilibria reached when many constituency are considered simultaneously. Section 5 compares the statistical properties of the simulation results to those of a dataset describing results of elections for a variety of countries since 1945. The Final Section concludes.

3 Model

We consider a discrete time model of repeated elections in which a population of J individuals are split between D districts. These individuals are assumed to have policy preferences defined on the N-dimensional unit hypercube, $[0,1]^N \in \mathbb{R}^N$. Individual $j \in J$ has an ideal point within this space denoted $A_j = [a_{j1}, .., a_{jk}, .., a_{jN}]$ where a_{jk} is their preferred point in dimension k.

Individual j's utility depends on the distance between their preferences and the policy implemented as a result of the election. We define:

$$-|(W - A_j)| = -\sum_{k=1}^{N} |(w_k - a_{jk})| \tag{1}$$

where $W = [w_1, ..., w_N]$ is the implemented policy and the distance between two points a and b is denoted as $|a - b|$.

Distances between the implemented policy W and the individual's ideal point A_j are defined as the 'Manhattan' distance.[7] This choice represents a desire for the total divergence to represent the summation of differences in each dimension. Other norms, would seem to require further assumptions about how individuals weight differences across dimensions.[8] However, the results presented in this paper are not dependant on this assumption and are robust to the use of the Euclidean Norm.

The model presented here reflects the key feature of the work by OS and BC in that there is no distinction between politicians and voters: any citizen can choose to stand for office in any election. After each election individual j receives utility U^j contingent on whether they stood and the outcome of the election:

[7] Properly, the Manhattan distance is an L^1 norm.

[8] For example, in the case of the Euclidean norm its not clear that a divergence of 0.25 in each of two dimensions should be equal to a divergence of 0.35 on one dimension and 0 on the other.

$$U^j(W) = \begin{cases} -|W - A_j| - \kappa + \gamma & \text{if she is elected} \\ -|W - A_j| - \kappa & \text{if she is not elected} \\ -|W - A_j| & \text{if she does not stand} \end{cases} \tag{2}$$

Where κ is the cost of standing and γ is a rent derived by the citizen from holding office.[9]

Individuals determine whether they will stand for election based on past utilities received from their actions. Specifically, we assume that the probability of an individual standing for election at a particular point in time, t, is given by the ratio of their past utilities from having stood for office and not having stood:

$$P_{jt}^{stand} = \frac{U_{jt}^{Run}}{U_{jt}^{NoRun} + U_{jt}^{Run}} \tag{3}$$

Where U_{jt}^{Run} and U_{jt}^{NoRun} are measures of individual j's utility from running or not running at time t as defined below. After each election each individual, given an implemented policy W, calculates $U_{j,t+1}^{Run}$ and $U_{j,t+1}^{NoRun}$ as follows:

(4)

	Stood For Office	Didn't Stand
$U_{j,t+1}^{Run}$	$\beta U_{j,t}^{Run} + (1 - U^j(W)/N)$	$\beta U_{j,t}^{Run}$
$U_{j,t+1}^{NoRun}$	$\beta U_{j,t}^{NoRun}$	$\beta U_{j,t}^{NoRun}(1 - U^j(W)/N)$

Where $U_{j,0}^{Run} = U_{j,0}^{NoRun} = 1 \Leftrightarrow P_{j,0}^{stand} = 0.5$.[10] In our model individuals are not able to observe the counterfactual of what would have happened if they had reversed their standing

[9] In this model individuals place equal weight on distances in each dimension. It is worth noting, however, that the results of the model change little if individuals are assumed to vary the weights they assign to different policy dimensions. This complication comes at the price of removing any intuitive spatial interpretation of the policy space as, for example, the ideological distance from Person 1 to Person 2 will be perceived differently by each voter. Accordingly, we do not focus on this aspect.

[10] The values of $U_{j,0}^{NoRun}$ and $U_{j,0}^{NoRun}$ have a limited effect on model behaviour, larger values can dramatically increase the time until convergence and setting either value to be less than or equal to zero can cause obvious convergence problem.

decision therefore in line with the work of Rustichini (1999) we adopt a linear learning rule.[11]

3.1 Elections

In each time step a single election is conducted. In the first stage of each election all individuals within each constituency simultaneously declare whether they will stand for office. This results in the set C_t^d of standing candidates in each district d. Every individual j within the population then simultaneously votes for a candidate within their district, d^*, given by:

$$(5) \qquad \arg\min_k |(A_k - A_j)| \text{ where } k \in C^{d^*} \text{and } j \in d^*.$$

i.e. individuals vote for the candidate who's ideal point is closest to their own. This implicitly assumes sincere voting, each individual votes for the candidate who if elected and who's policy were implemented would maximize their utility.

Votes for each candidate are counted and the m individuals from each district with the highest number of votes are elected to office. If $|C^d| < m$ all members of C^d are elected, however, all non-standing members of d receive $-N - \kappa - 1$ utility for this election. This large negative utility is of greater magnitude than the lowest utility an individual may receive if they stand for election and so ensures that once the model has converged there will be at least m individuals standing for election in each district.[12]

3.2 Coalition formation

Once elections have taken place in each of the D districts the set of elected representatives together determine the policy to be implemented. We make no assumption about the existence, or otherwise, of coalitions, elected candidates either start their own coalition or join an existing one.[13] In the spirit of Levy (2004) and Morelli (2004), if they seek reelection,

[11]Experimentation with the β parameter showed that low values did not guarantee convergence therefore a high value was employed for all experiments discussed within this paper. Beyond this relationship the exact value had relatively little impact on the results.

[12]This payoff is analogous to the negative infinity payoff received when insufficient candidates stand for election in the OS model.

[13]Note, whilst every representative is assigned to a coalition, given that a coalition can have a membership of 1, this is equivalent to allowing individuals not to join a coalition.

all members of a coalition stand on a common electoral platform. Elected representatives are constrained to stand on their party's platform, but derive benefit from being in a larger party. The aim is to obtain as parsimonious as possible a representation of the benefits of party membership, whatever their origins, and its costs. It is argued that in context of our computational approach this abstraction captures the key thrust of the Osborne and Tourky (2008) model of parties as a cost-sharing technology.[14] After individuals have joined coalitions and the coalition dynamics described below have occurred, the preferred policy of the largest coalition is implemented. Representatives prefer larger parties, since in general larger parties are more likely to influence the choice of policy. These assumptions are considered to be a minimal way of representing a coordination technology for representatives.[15] A key simplifying assumption is that we don't consider post-election coalition formation. This has been the subject of much study, and Dhillon (2005) provides an excellent review. We define the preferred policy of a coalition to be the mean of the ideal points of its members. The process of coalition formation proceeds as follows. Initially each newly elected representatives start a new coalition of which they are, at this point, the only member. All returning representatives remain in their previous coalition, whether or not all previous members have been re-elected. Once, all representatives belong to a coalition (possibly with a total membership of 1)[16], candidates assess whether their current coalition best represents their interests. We assume that representatives employ a heuristic of the following functional form:

$$V_r^j = \frac{\#r}{|(A_j - \mu_r)\iota| + \eta} \tag{6}$$

Where, ι is a $N \times 1$ vector of ones, η is small[17], r is a coalition, $\#r$ the number of members in that coalition, and μ_r is that coalition's current policy. This heuristic is used to determine individuals satisfaction with membership of a particular coalition. Representatives

[14] In fact, modelling explicitly a cost-sharing technology *à la* Osborne and Tourky (2008), in addition to the existing preference for larger parties, does not meaningfully alter the results presented below.

[15] Note, that this technology as defined does not preclude the existence of a loose or non-existent party structure.

[16] Coalitions with no-members are assumed to no-longer exist.

[17] Specifically, η is parameterized as 0.02.

face a trade-off: membership of a larger coalition increases the likelihood that an individual's preferences will have some influence on the implemented policy. However, casual observation suggests that individuals dislike belonging to the same party as those very ideologically distant from themselves[18]. Individuals trade off the increased chance of being elected with potentially sacrificing the proximity to their preferred platform.

The composition of coalitions changes through a process of splitting and merging.[19] These processes identify whether there are subsets of coalitions that would be better off as separate coalitions or whether there exists pairs of parties which would be better off if they merged. As such it is a coalition-stability concept.

In order to conduct the splitting analysis principle groupings are found within each party using the k-means algorithm as first proposed by Lloyd (1982) and as interpreted by Hartigan and Wong (1979). This algorithm is widely used to identify clusters in multi-dimensional data. In essence it searches for the allocation of observations to clusters and the means of those clusters that minimizes the total sum of the squared distances between cluster midpoints and the points in each cluster, across all clusters. Here, we employ it to partition each coalitions into two groups who each consider whether it is in their interest to leave the coalition. In particular, the j candidates are partitioned into z sets (here $z = 2$). This collection of sets $G = \{G_1, .., G_z\}$ is chosen so to minimize the total within group variance, across all groups. That is:

$$\underset{G}{\operatorname{argmin}} \sum_{i=1}^{z} \sum_{A_j \in G_i} |(A_j - \mu_i)|^2 \tag{7}$$

The algorithm to do this proceeds in the following steps:

1. Initially two 'centers' P_1 and P_2 are chosen at random within the policy space.
2. Each member of the coalition identifies which centre they are closest to producing two groups G_1 and G_2
3. Set P_1 equal to the mean of the ideal points of G_1 and similarly for P_2 and G_2
4. Repeat from 2 until the centres no longer change.

[18]See Baylie and Nason (2008).

[19]We considered an additional process whereby individuals could unilaterally change coalition if under the above metric it was beneficial to do so. It was found that this did not effect the distribution of results.

This algorithm is not deterministic, it is dependent on the initially chosen centres and may find different clusters each time it runs. This is advantageous for this model as it allows a more thorough testing of the stability of each coalition as different groups consider seceding. Once the groups have been identified each group must determine whether to secede. Their decision is again based on satisfaction of individuals in the cluster with their continued membership. The average utility of the members of each group is calculated as a combined party and as separate coalitions. If the average utility of either group is higher after seceding then the coalition splits. The decision to secede is a unilateral one, a group does not need permission to leave a coalition.

Similarly each coalition considers if it would be better off merging with another party chosen at random. If the average utility of the members of both parties are greater as a combined unit than as separate grouping the two coalitions merge into one. In this case it being necessary that both groupings increase their utility for the merger to occur.

In both cases the average utility of members of the groups are employed in decision making. Consequently there may be one or more members of each group which disagree with the decision. In the long term, however, this dissatisfaction does not persist, the potential for coalitions to further split and merge, or the citizen potentially no-longer standing, ensures that eventually each individual is happy with their final position.

The above process occurs after each election, each coalition (in random order) first tests whether it would be beneficial if it splits and then tests whether it would be beneficial if it merged with one other randomly determined coalition. Once, the membership of the coalitions has been established it is assumed that the preferences of the coalition with the most representatives are implemented. This is an abstraction, for instance it is not necessarily the case, as in observed democracies, that the largest coalition contains a majority of representatives. However, the focus here is on the electoral process and not on the process of government policy formation. All individuals in all districts therefore receive payoffs based on the implemented policy of the largest party.

3.3 Equilibrium

The above process is repeated until the model converges to a stable outcome, an equilibrium. In this model we determine equilibrium as the state in which the composition of all coalitions is fixed. This implies that in all districts the same individuals are elected, resulting in the continuation of the same coalitions from the previous election and that no individual or group finds it beneficial to move from their current coalition. Practically the simulation is halted after $20,000$ elections in which the coalitions positions and memberships remain unchanged.

It is important to note that the outcome of the model is sensitive to the initial distribution of preferences, which might be expected. It is also, potentially path-dependent, and potentially sensitive to the standing decisions of candidates during the convergence process. These are determined by the particular seed of the random number generator. As such, which of the likely multiple equilibria the model converges to is partially stochastic. However, these differences tend to be small, and notably the focus is on the overall statistical properties of the equilibria obtained for many runs for each combination of parameters, with different seeds. Not the particular outcome of a given model.

To improve both the motivation and explanation of our approach, it is useful to first consider the relationships between 'computational equilibrium' with agents who are initially naive but then learn, and their strategic counterparts as modeled by, *inter alia*, Besley et al. The model presented here is run for a large number of time steps such that there is convergence. Individuals standing for election and party coalition composition at the end of the experiment have reminded fixed for a long period. As such the model can be viewed to have converged to an equilibrium. As the number of agents in a given constituency becomes large and the number of periods for which the agents behaviour is simulated also becomes large, then the expected results become an increasingly good approximation to an analytical equilibrium such as those obtained by OS.

Flowchart 1 shows the decisions made by citizens as the above model proceeds, whilst to ease readability Flowchart 2 expands on the details of the party dynamics box of Flowchart 1. As described above the model commences with the setting of the preference distributions and continues until convergence. Each election starts with citizens declaring their candidacy

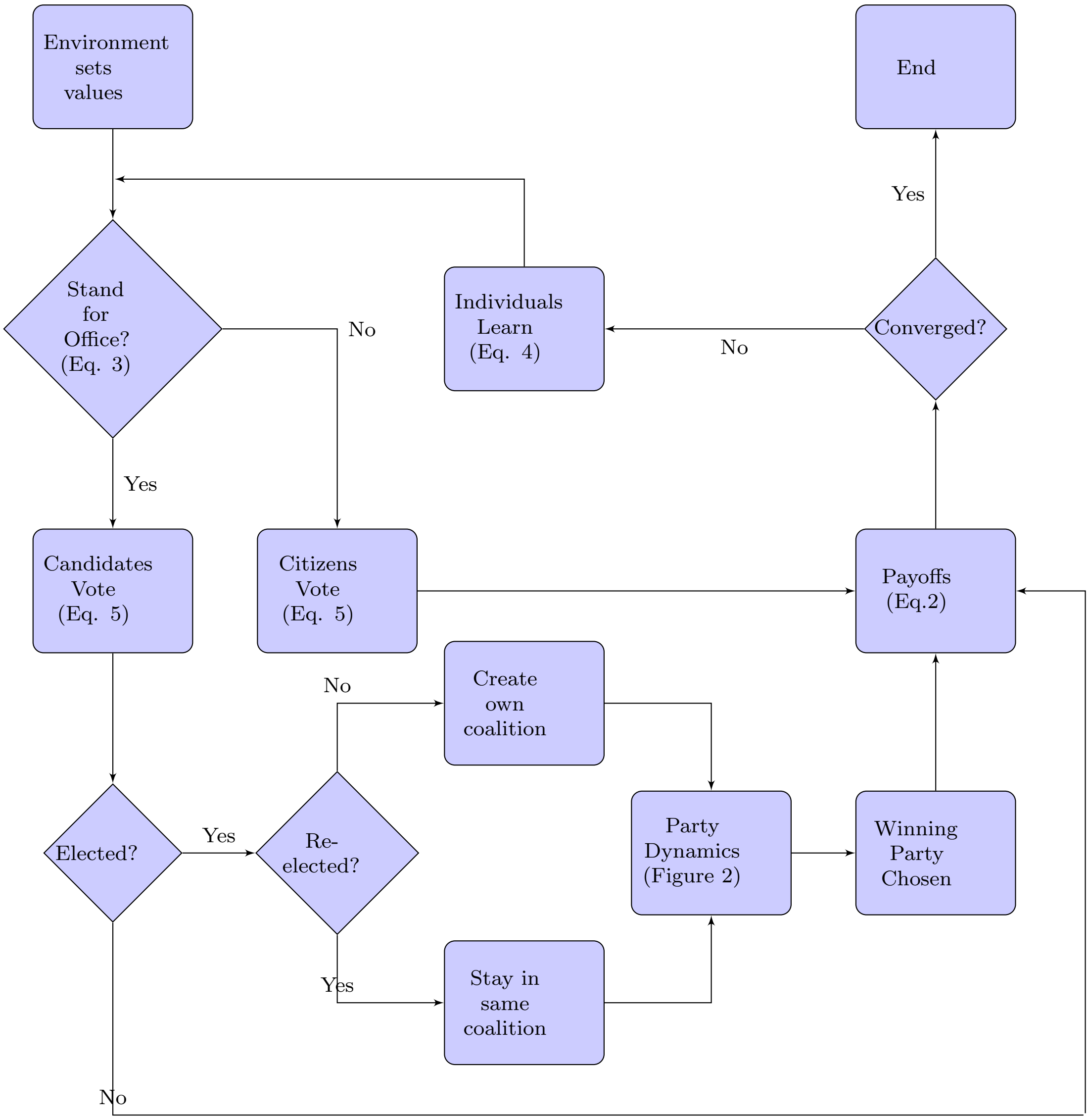


Figure 1: Flowchart depicting the order of a citizen's choices within the model.

and finishes when individuals calculate their payoffs. If the model has not converged citizens learn according to the rules described above and another election is called. It is worth emphasizing that the two elements below, 'citizens vote' and 'candidates vote' occur together and the results are amalgamated in order to determine those in office. Flowchart 2 shows the inter-coalition procedures, unlike Flowchart 1 this is not done from an individual perspective, rather it considers a top-down view of the model.

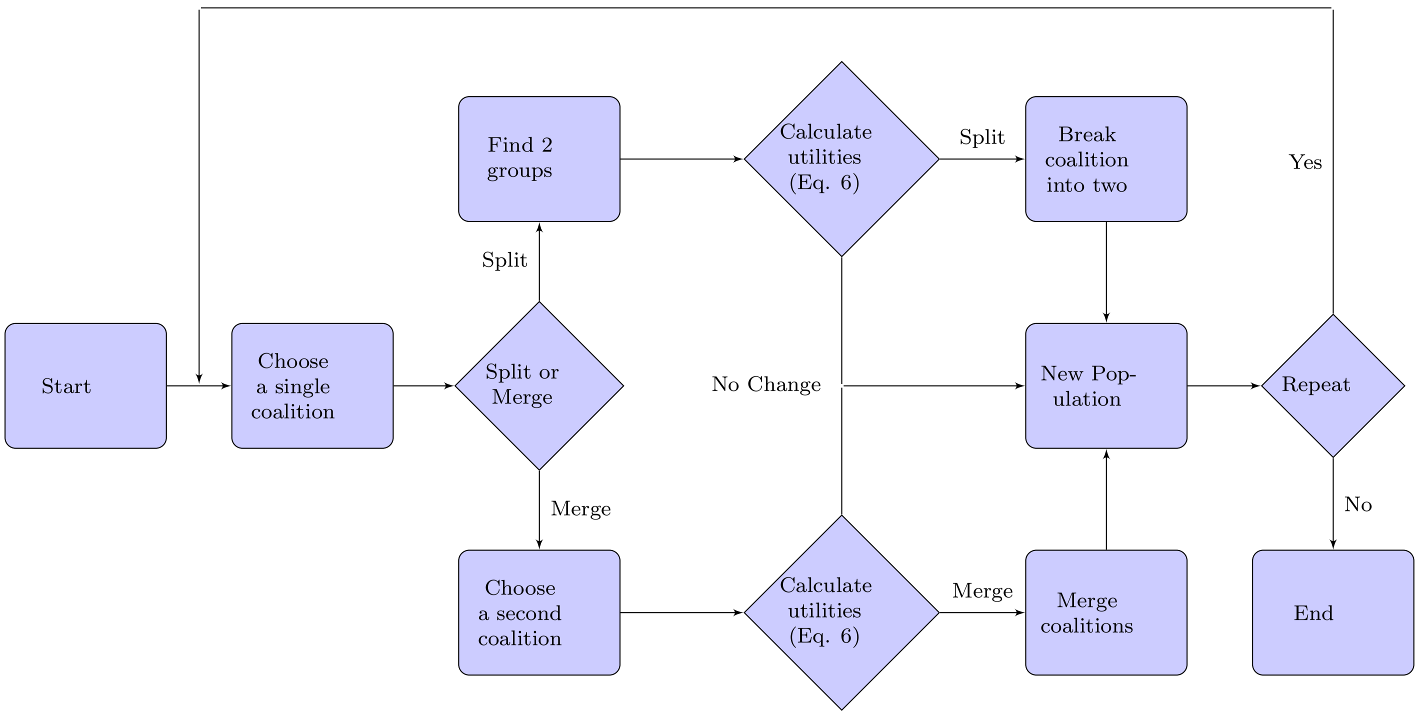


Figure 2: Flowchart detailing the order of events in the process of splitting and merging coalitions.

3.4 Single constituency results

In this section we consider the behaviour of the single constituency model set out above. This analysis is not the main focus of our paper as single constituency citizen candidate models have received much analytical attention previously, notably by BS and OC. Instead this section provides an intuition of the constituency level dynamics which will be useful in understanding the multiple constituency model.

We investigate the model for multiple policy dimensions ($k \in \{1, .., 7\}$) and for different electoral systems. We consider the election of m representatives where $m \in \{1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 24, 30, 40, 60, 120\}$ for $100m$ citizens. For these simulations only, when calculating utility the implemented policy for each citizen is that of the closest elected representative in the policy space.[20] In later sections we employ the full coalition dynamics described in Flowchart 2 for the determination of a single ideal point. Throughout the paper simulations were conducted with parameter values; $\beta = 0.99$ $\kappa = 0.1$,$\gamma = 0.2$ and all values of a_{jk} were drawn from $U(0, 1)$.[21] The model is simulated until convergence. That is, the distribution of candidates and parties has zero variance.

The key single-constituency result is that the number of candidates standing per district in equilibrium is broadly increasing in the number of policy dimensions. Table 2 reports results for 100 simulations, of the number of candidates standing for election in equilibrium as the size of the electoral district, m and the number of policy dimensions, k, are varied. The results suggest that whilst any increase in the number of candidates is smaller than the associated standard deviations for small districts, it becomes large (compared to the standard deviation) as the district size increases. Since, additional candidates are, in the language of BC 'spoiler candidates', the main result is perhaps to be expected: As the policy space becomes larger, it becomes more likely that the difference in implemented policy occasioned by an additional candidate running is sufficient to offset the cost of standing. Furthermore, it would seem that in larger constituencies either the probability of winning

[20]This assumption is solely to abstract from the coalition formation and policy game so that the constituency level dynamics are clearer.

[21]Different values of κ and γ were investigated, as long as the benefit from election is greater than the cost of standing there was little effect on the equilibrium obtained. If, however, the cost of standing were high and the benefit of winning low then under some circumstances the model could converge to a result where less than m individuals were elected in a small number of simulations.

or the change in policy position of the nearest winning candidate tends to be higher.

Figure 3.4 provides an example of these results. It can be seen that six individuals stood for election with three being elected. The three elected representatives are interspersed with the three non-elected standing representatives. Each citizen in the diagram votes for its nearest standing representative. That there is a stable pattern of individuals standing indicates that no individual can increase their utility by changing their decision to stand as such an equilibrium is found.

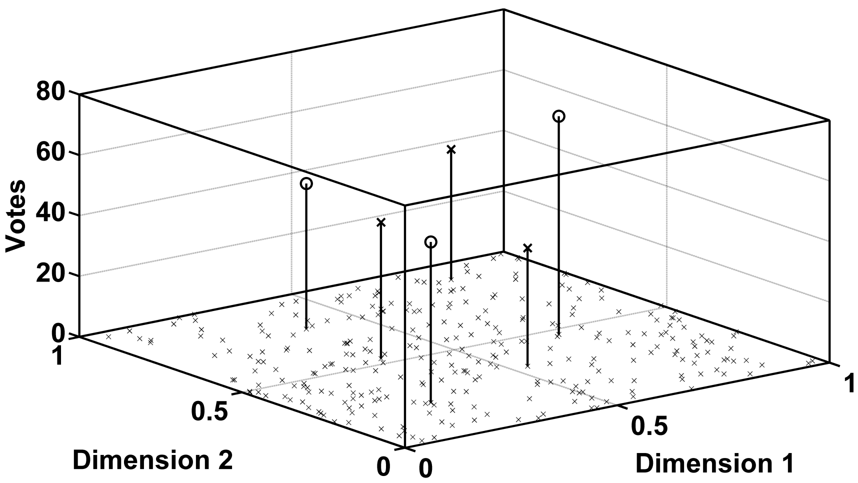


Figure 3: Results of a single constituency experiment conducted in a two dimensional policy space with 3 individual elected. Crosses represent citizens in policy space. Stalk length indicates the number of votes a citizen received. Circle markers are elected citizens.

4 Results

This section present the main results of the paper regarding the formation of coalitions when there are multiple electoral districts. Recall that, in the case of a single district, without coalitions, increasing the number of candidates elected and the number of policy dimensions both increase the number of candidates standing. But when there are potentially many districts, and coalitions are allowed to form endogenously, that an increase in the number of policy dimensions leads to fewer, larger, parties. Results are also presented describing the relative policy positions of the competing coalitions, and the number of different types of equilibria that the model gives rise to is analyzed. Again, whilst there are often very many possible equilibria in the case of a single district, without political parties, the formation of parties markedly reduces the set of possible outcomes.

We consider a democracy in which 120 candidates are elected together representing $12,000$ voters split between the C constituencies of equal size[22] each returning an equal number of representatives.[23] Larger populations may be simulated, however, this does not effect the results obtained but does dramatically increase the computational burden of the model as the running time is proportional to the squared number of individuals. As discussed in Section 3.2 the existence of coalitions isn't assumed *ex ante*, but potentially emerge endogenously. The minimal assumptions about the benefits and costs of coalition membership give rise to stable electoral coalitions - political parties. What is more, these parties seem to fulfill many of the functions of resolving ideological disagreement and, here only implicitly, providing for credible commitment to non-preferred platforms that are suggested by Levy (2004). The results suggest that an increase in the number of policy dimensions actually leads to fewer, larger, parties. That is, not only do political parties provide a way of reaching agreement in multiple dimensions, but multiple dimensions seem to provide, via parties, for more widespread agreement. We first discuss in more detail the evidence for this finding. Secondly, we provide some examples of particular outcomes of the model. Finally, the number and form of distinct types of equilibria are discussed.

Table 3 contains results describing the mean and standard deviation of the number of parties for each combination of district size and number of policy dimensions. The results show that there is a negative relationship between the number of parties and the number of dimensions. It would also seem that there is an immediate and large drop in the number of parties when the number of dimensions increases from 1 to 2. Similarly, there is a notable drop when the moving from a 2 to a 3 dimensional ideology space. But, as the number of dimensions increases from 3 to 7 there is no clear relationship. This might suggest that three dimensions provides sufficient flexibility for parties to form, that there is little benefit of further dimensions.

As Laakso and Taagepera (1979) note, many democracies are characterized by a tail of small parties. In general, these parties have little or no impact on the democratic process.

[22]Relaxing the assumption of equal sized constituencies doesn't affect the results. The choice of 120 representatives is solely because it has many factors, but again this assumption is unimportant for the results.

[23]The analogue of the previous single-constituency model, but with coalition formation, is the case where $C = 1$.

Hence, a common approach is to define the number of 'effective parties'. We employ the Laakso-Taagepera Index defined as follows:

$$N = \frac{1}{\sum_{i=1}^{n} p_i^2} \tag{8}$$

Where p_i is the vote share of party i and there are n parties.[24] A variety of alternatives to the Laakso-Taagepera (LT) measure have been proposed. One common objection to the LT measure is that it will in general suggest there are several effective parties even when one party has an overall majority and as such only that party is 'effective'. This is less problematic for the purpose here which is to use the effective number of parties as a summary statistic for the overall distribution of party sizes. A leading alternative is the Banzhaf index, which measures how often a coalition can be expected to be the 'swing voter'. Kline (2009) considers the relative empirical performance of this measure, and Gelman, Katz and Tuerlinck (2002) and Gelman, Katz and Joseph (2004) provide both a survey and a critique of this approach.[25]

Table 4 contains results in terms of the effective number of parties. These confirm the results described for the absolute number of parties, except that it is clear that additional higher dimensions do impact the size distribution of parties if not the number. That is, the continued decline in the number of effective parties as the number of dimensions increases suggests that additional dimensions provide for a greater proportion of representatives to be members of the larger parties. The size distribution of parties, and in particular the impact of district size, will be further considered in section 5.

[24] The Laakso-Taagepera index is the inverse of a standard Herfindahl-Hirschman index

[25] Another important issue is how to approach democracies employing a mixture of different electoral systems. Here, we abstract from this problem and focus simply on the largest system employed, which we argue corresponds best, although imperfectly to the simulation results. Moser and Scheiner (2004) discuss the extent to which the coexistence of multiple electoral systems can be seen to 'contaminate' outcomes within systems.

4.1 Party Size and Position

We now turn our attention to the relative 'location' of political parties. As discussed in 2 trying to predict the relative positioning of political parties has been the subject of considerable attention at least since Downs (1957). One difficulty in approaching this question in the context of multiple policy dimensions is how to display, conceptualize, and compare results. Our approach is to consider the positions of the each party relative to the largest party. To do this a Gram-Schmidt scheme (as described by Golub and Van Loan (1996)) is employed to produce an orthonormalization of the set of vectors describing party positions. We first define the location of the largest party to be the origin of a new coordinate system. From here a series of M orthogonal vectors, v^i for $i = 1 - M$ are calculated, corresponding to the axis of the new coordinate space such that for the Q^{th} largest party which has position p^Q, $v^i \dot{p}^Q = 0$ for all $i \geq Q$ and where $M \leq N$ where N was the dimensionality of the original coordinate space. We therefore produce a set of axes such that the largest party is at the origin and each additional party requires an additional dimension to represent it, up to N. That is, the second largest party falls on the x-axis, the third on the xy-plane etc. Note that a further consequence is that the second party will always have a positive x-coordinate, the third party a positive y-coordinate, but potentially negative x-coordinate,etc.

We present these results by plotting the location of parties in the first three dimensions. Each party is represented by a sphere with diameter proportional to the number of its members. The left-hand plot depicts all 3 dimensions, the right-hand side plot shows the xy-plane. A simple example is presented in Figure 4 , with just two parties competing in a 3-dimensional policy space with 3 candidates per district. It is worth noting that the ideological discrepancy between the parties is small, but non-zero. In general we find very few cases where there are large amounts of dispersion between the larger parties. It is argued that this is similar to the case of most mature democracies, where the main parties aren't normally extremist. Figure 5 considers an example with three parties. In this example, the additional party is to the 'left' of the two larger parties on the first dimension and also differentiates itself on the y dimension. Again, the ideological differences are relatively small. The final example displayed in Figure 6 involves five parties. It would seem that again we don't observe extremist behaviour, rather the parties differentiate themselves, particularly in the xy-plane, a little in several dimensions. The exception is the fifth party which appears relatively extreme, but in the first 3-dimensions at least, is about 20 percent of the total length of each dimension away from the largest party, in each dimension. Hence, whilst a smaller, more ideologically distinct, party seems to coincide with many democracies experience. That these differences would seem in some sense to be limited is considered to be both realistic and also conform to the central intuitions of the citizen candidate model. We present a more detailed analysis of the general relationships between district size, the number of policy dimensions, and equilibrium outcomes in the next section.

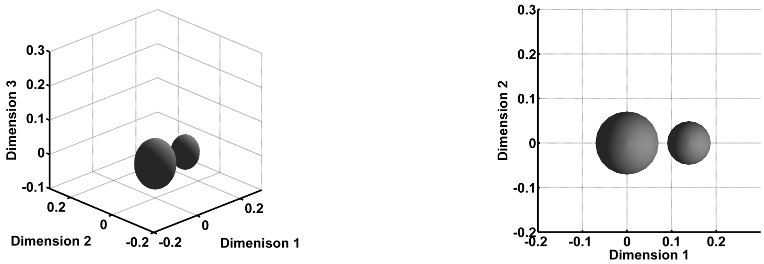


Figure 4: An example with two parties, 3 representatives per district, and 3 dimensions

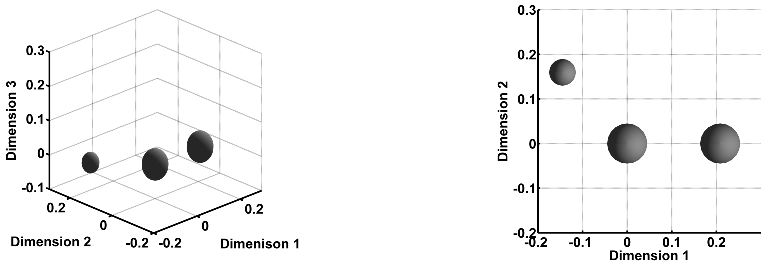


Figure 5: An example with three parties, 3 representatives per district, and 3 dimensions

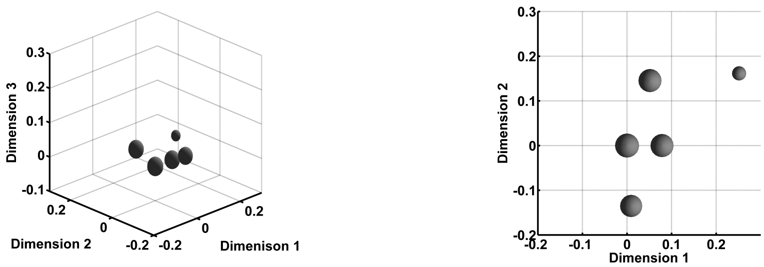


Figure 6: An example with five parties, 3 representatives per district, and 3 dimensions

4.2 Equilibria Characterization

In this section, we introduce a new computational method for the identification of simulation equilibria. Using this method, we highlight the general properties of the model rather than relying on the outcomes of a few specific runs. The approach taken is to enumerate the number of different equilibria and describe their general form. Whilst, as mentioned above the model for each combination of parameters has multiple equilibria, the equilibria have two distinct sources of variation. Firstly, different equilibria arise due to differences in the preferences of citizen-candidates as determined here by the random number seed. These differences are expected in any such model, computational or otherwise. The second source of variation is more minor differences due to path dependence. For example, if there were two citizen-candidates in a particular district with extremely similar preferences, it may be that it is in the interests of both for one but not both of them to stand. In our model, provided that they receive similar amounts of utility from standing, which of the two stands in equilibrium is potentially path-dependent. That is, there are two possible stable outcomes (in this model) one candidate stands with probability 1 and one candidates stands with probability 0. But, which of the two is which may be dependent, for example, on which is the first to randomly stand when the other doesn't. The minor differences in equilibria for reasons such as this are not as interesting as larger qualitative differences between equilibria, which result in different sized or located parties.

The distinction made above is an imperfect one, whether two equilibria count as being qualitatively different is in part subjective. As such, similarly to section 3.2, we employ a statistical approach to find the number of distinct clusters in the data. The data for each parameter combinations, are the results of 1000 repetitions of the simulation with different random seeds. As before, the results from each simulation are converted to a set of vectors in which each vector contains the policy of a party. A Gram-Schmidt process is applied to these vectors and the results combined with the relative party sizes to produce a single vector for each simulation characterizing its results. A k-means clustering algorithm is applied to the set of 1000 vectors for a range of values of k, to identify clusters of almost identical equilibria. For each value, 1000 repetitions of the algorithm are run and the minimum value of k required to explain 90% of the total variance is found.[26] Accordingly, we define the number of equilibria as, k, the number of distinct clusters identified.

It is worth noting that one consequence of the Gram-Schmidt scheme is that equilibria that are reflections or rotations of another before normalization are equivalent afterwards. We now consider an example of how the process works: The results of three simulations carried out in two dimensions are shown in figure 4.2. In each case there are three parties distributed within the policy space, however, beyond this it is not possible directly to identify any similarities between the configurations. Figures 4.2 shows the party locations after Gram-Schmidt transformation. The largest party in each case is now located at the origin with the second largest on the X-axis, all parties within each simulation maintain their relative positions, however, comparison across simulation is simplified. It can now be seen that two of the result sets are similar in the patterns of parties whilst the third differs significantly. The k-means algorithm for 2 clusters successfully identifies these (figure 4.2). The first being based on the triangle and cross results whilst the second represent solely the circle results.

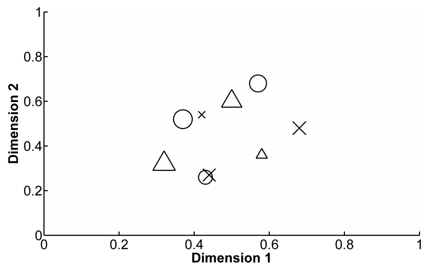


Figure 7: Results of three simulations, each marker is a party at the end of the simulation with markers of the same type coming from the same simulation. Marker size corresponds to party size

.

Figure 10 shows the results of this procedure applied to simulations of the case of 3 policy dimensions and 4 representatives per district. However, this time instead of plotting the particular outcomes of the model we plot simultaneously the different equilibria of the model, where each equilibria is represented by the mid-point of the associated cluster found by the k-means algorithm.

There is one 2-party equilibrium, two 3-party equilibria, and a 4-party equilibrium.

[26]The results are not sensitive to the choice of the percentage of variance explained.

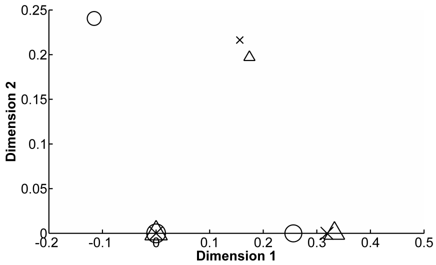


Figure 8: Results of three simulations shown in figure 1 after the application of the Gram-Schmidt scheme.

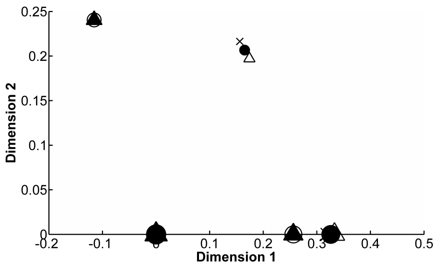


Figure 9: Results of three simulations shown in figure 1 after the application of the Gram-Schmidt scheme with two equilibria represented by filled shapes found by the k means algorithm.

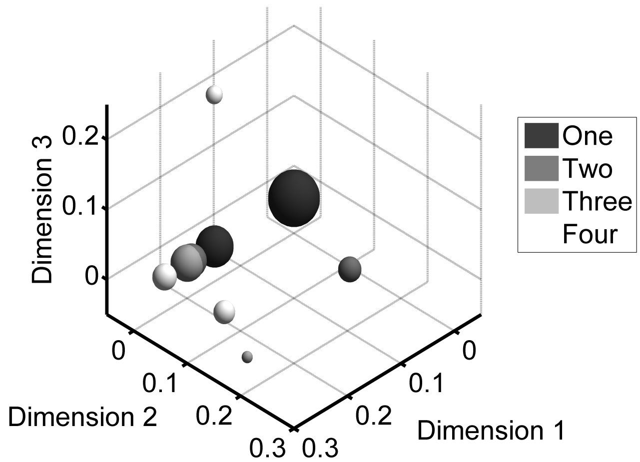


Figure 10: Four of the eight equilibria in the case of 3 policy dimensions and 4 representatives per district. In each case the largest party is at the origin.

The 2-party equilibrium, black, is straightforward to interpret and as a consequence of the Gram-Schmidt scheme the parties are only distinguished on the x-axis. The dark-grey and light-grey 3-party equilibria both contain a second large party, in virtually identical locations. The difference between the two is that the largest party is larger in the case of the light-grey equilibrium, and the third party is located further away from the origin.

In the case of the dark-grey equilibrium, the third party occupies a position between the two larger parties on the x axis, and is similarly distinct on the y axis as its equivalent in the green-equilibrium. Again, it is argued that these outcomes all are easy to interpret intuitively. The white four-party equilibrium is perhaps less intuitive. Four way competition in 3 dimensions is inevitably complex, but the outcome with four similarly sized parties, two of which are located at the same point on the x-axis but distinguished on the y-axis, with the fourth party located equidistant from the largest two parties on the x-axis but, of course, is distinguished on the z-axis. That two similarly sized parties can be similarly located and not-merge suggests that the model does not lead to parties merging too often, and similarly that there is in none of the four equilibria a tail of independent representatives suggests that similarly the parties aren't unrealistically fragmented, especially since we don't model variations in regional politics that might give rise to such parties for other reasons.[27]

Table 5 reports the number of equilibria identified for each combination of parameters. The most noticeable result is that the number of party equilibria is dramatically smaller than the number of equilibria for a given constituency. It is clear that whilst there may in some cases be many stable combinations of citizen-candidates standing for election in a given district, allowing those elected to form parties significantly alters this. One reason for this is that party membership affects the costs and benefits of standing thus reducing the number, and changing the identity, of candidates in equilibrium. Moreover, in some cases the formation of parties may lead to more extreme candidates receiving lower benefits of running since even if they reflect preferences in a given district, the requirement that all candidates belonging to a given party adopt the same platform may lead in equilibrium to those with preferences closer to the national average being elected. That is, national political concerns determine the outcome of local elections. This phenomenon is more pronounced in larger districts and for cases with more policy dimensions. The number of equilibria at the district level is increasing in the number of dimensions, but the converse is true at the national level. Conditional on the number of dimensions there isn't an obvious effect of district size at the national level, but larger districts lead to more district level equilibria. It

[27] An augmented algorithm in which small perturbations of party sizes were applied to identify seemingly different but in effect identical equilibria was applied, but the results don't change meaningfully.

would therefore seem that political parties provide for agreement in many policy dimensions, but that this is either because they alter who in equilibrium runs, or because they reduce the set of alternative compromises. In summary, whilst more dimensions lead to more distinct equilibria in any given election, once candidates are allowed to form parties, we find that more dimensions leads to fewer parties, and the number of equilibria is comparatively small.

5 Comparative Politics

This section compares the size distributions of parties conditional on average district size predicted by the model with those observed empirically. We find that the relationship between the effective number of parties and district size, is similar to the outcomes observed empirically. This empirical relationship continues to be the focus of much study, with a large, and growing literature. A classic statement is that of Lijphart (1999) whilst Gallagher and Mitchell (2005) provides an excellent recent survey. Subject to particular attention has been the empirical support and theoretical basis for Duverger's law - *"that plurality elections give rise to a two party system"* (Duverger (1951)). Dunleavy, Diwakar and Dunleavy (2008) note that there have been subsequent repeated restatements of Duverger's law in the face of growing exceptions to this rule. They further argue that previous empirical findings that lend support employ the wrong null-hypothesis. Here, the focus is solely the number of effective parties, and we don't presume to address either the theoretical logic for the Duvergian hypothesis (of which Morelli (2004) is a leading example), its empirical support, or other consequences of different electoral systems.[28] Rather, this section attempts to show consistency between the patterns in actual election data and those describing results from the model. The focus is on assessing the extent to which the model gives rise to variation across political systems that is in line with what is observed. To do so we turn first to the empirical data and identify some simple empirical regularities before arguing that the results of the model are sufficiently similar to lend it credibility.

Data for 248 post-1945 elections were collated from Modules 1 and 2 of the Comparative

[28] Other important types of variation, with more obvious normative connotations include the degree of disproportionately (the extent to which the distribution of votes is not reflected by the allocation of seats in a legislature), or the effective threshold (the extent to which an electoral system, by design or otherwise, precludes smaller parties from gaining seats).

Study of Electoral Systems (2003, 2007) project and the data contained in Caramani (2000). The combined data correspond to 38 countries, and the effective number of parties for each election is calculated as described in Equation 8. Inspection of Figure 11) suggests that the effective number of political parties(henceforth N_s) would seem to increase with district magnitude, this hypothesis is sometimes referred to as Duverger's second or other law. Moreover, the variation in N_s also seems to increase. However, Table 1 makes it clear that support for the original version of Duverger's hypothesis is scant. There are more often than not more than two effective parties in single district elections in our data. Yet, there is a difference in outcomes between single-district systems and others, as such there would seem some evidence for what might be called the weak Duvergian hypothesis. That is, there is a clear qualitative change in the distribution of effective parties between single and multi-member electoral districts.

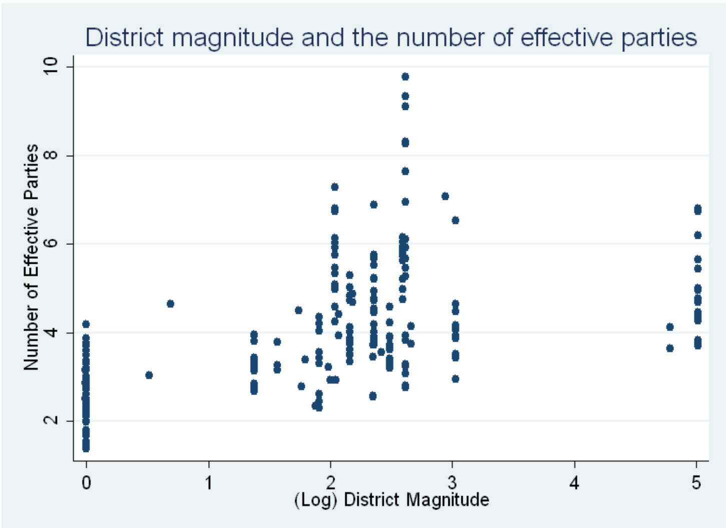


Figure 11: Scatter Plot to show the relationship between District Magnitude and the number of effective parties in 248 elections

However, it is also interesting that the number of effective parties, and the variation in that number, seems to decrease at the largest district sizes. This might represent a second discontinuity when there is just one, national, electoral district as is the case in the Netherlands.

In summary it's claimed that the data, and the previous literature, support the following three stylized facts:

1. The number of effective parties is (weakly) increasing in district size (excepting single constituency systems). (Duverger's 2^{nd} law)

2. The variation in the number of effective parties is (weakly) increasing in the size of electoral districts (excepting single constituency systems).
3. The key qualitative change, in the number and variability of the number of effective parties occurs between single and multi-member electoral districts. ('The Weak Duvergian Hypothesis')

Table 1: Summary statistics

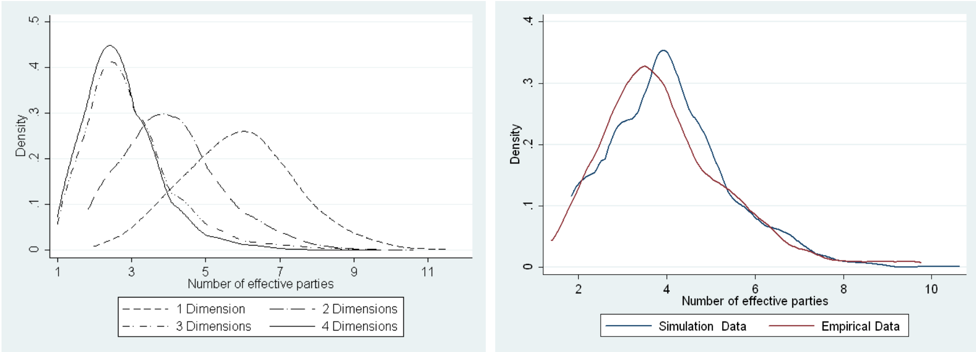
Districtsize (D)	Mean	$(Std.Dev.)$	$Min.$	$Max.$	N
$D = 1$	2.43	(0.66)	1.37	4.19	52
$1 < D \leq 5$	3.29	(0.51)	2.68	4.64	22
$5 < D \leq 10$	4.30	(1.18)	2.30	7.27	50
$10 < D \leq 20$	4.77	(1.58)	2.56	9.76	81
$20 < D$	4.36	(0.95)	2.94	6.80	42

We now consider how well the simulation of the citizen candidate model discussed above can explain these stylized facts, along with other important empirical democratic outcomes. It is argued that whilst the model indeed coincides with all three of the stylized facts, discrepancies remain with other observed outcomes. Of course, further refinement of ideological distributions and institutional features would perhaps resolve these discrepancies. However, precisely emulating reality is not this paper's purpose, per se, rather the focus remains on the remarkable success of the Citizen-Candidate model.

Figure 12(a) displays kernel density plots of the number of effective parties disaggregated by dimension for the case of 1-4 dimensions.[29] It is clear that whilst the 1-dimensional case gives rise to too many parties, and the 3 and 4 dimensional cases too few, the 2-dimensional results are extremely similar to those obtained from the empirical data.[30] This is confirmed by Figure

[29]The results for between 5 and 7 dimensions are similar to those of 3 and 4 dimensions.

[30]In all cases an Epanechnikov kernel was used, bandwidth chosen using the optimal bandwidth for the



(a) Number of Effective Parties for Different Numbers of Dimensions

(b) Comparing 2-Dimensional Simulation and Empirical Data

Figure 12: Kernel Density Plots of Number of Effective Parties for Different Numbers of Dimensions

12(b) which overlays the empirical distribution and the 2-dimensional simulation distribution. There is more mass in the right-tail than for the empirical data, but in general it is a very close match. It would be possible to find the combination of results for different numbers of dimensions which minimizes the difference between the moments of the empirical and simulation data. But, the aim is not to emulate exactly the empirical distribution but to argue that the model gives similar results with a minimum of assumptions. It would be inappropriate to argue based on these results that there are in reality often 2 policy dimensions, but the similarity for the case of 2-dimensions does demonstrate the explanatory power of the model.

Do the results of the model, and the empirical data, support the three stylized facts above? In order to answer this question regression analyses were performed on both sets of data. The use of regression analysis for simulation results is not widely employed. However, we argue that it is a useful way to understand the effects changes in an individual parameter conditional on the other parameters. We report standard errors for these results, as the use of a 100 different random number seeds for each combination of parameters represents both a sample from the underlying space of potential simulation outcomes, and something analogous to random disturbances in the conventional sense. The two sets of results are clearly not directly comparable as the sources of variation are quite different. In particular, all variations in outcomes for a given set of parameters in our simulations are due to differences in the random number seed chosen. In the empirical data, variation amongst similar electoral

empirical data. The simulation data were weighted to account for the uneven empirical distribution of district size. These weights were obtained by piecewise-linear interpolation of the empirical data.

systems can be attributed to differences in national preference distributions, differences in the detail of electoral systems, as well as other national idiosyncracies. However, much can still be gained by comparing the signs and relative magnitudes of the estimated coefficients.

We begin with the empirical data. Whilst data are available on a variety of other electoral characteristics, no attempt was made here to try to include additional variables as controls. Rather, a simple bivariate regression was estimated.[31] The results are as follows:

$$\text{parties}_i = \alpha + \widehat{\beta}\text{district\_size}_i + \epsilon_i$$

$$\widehat{\text{parties}}_i = \underset{(0.101)}{3.756} + \underset{(0.002)}{0.010}\text{district\_size}_i$$

The estimated coefficient is small, however this may be, in part, a consequence of the influence of 19 observations (of 248) for Israel, and the Netherlands which have a single national electoral district. Including a binary variable, *large* for whether district size is greater than 100 gives the following results:

$$\text{parties}_i = \alpha + \widehat{\beta}\text{district\_size}_i + \widehat{\lambda}\text{large}_i + \epsilon_i$$

$$\widehat{\text{parties}}_i = \underset{(0.127)}{2.862} + \underset{(0.017)}{0.129}\text{district\_size}_i - \underset{(2.580)}{17.062}\text{large}_i$$

These results suggest that an increase in district size by 1 member increases the number of effective parties by around 0.17. This would seem more plausible, as it is better able to explain why some single-member systems consistently give rise to less than 3.6 effective parties.It is notable that $\widehat{\text{large}}$ is negative.This implies that the effect of an increase in district size is decreasing, but remains positive.[32] This specification ignores the key importance of the difference between single-member and multi-member districts postulated in the third stylized fact. Accordingly, a specification including an additional binary variable *single* describing if a electoral system has only single-member districts, was estimated:

[31] All reported standard errors are robust.

[32] There are 17 observations for the Netherlands which has a district size of 150, and 2 for Israel which has a district size of 120. Thus, the estimated average effect of moving from district size 10 to a large district was of $(146.84 - 10) \times 0.129 - 17.062 = 0.880$.

$$\text{parties}_i = \alpha + \widehat{\beta}\text{district\_size}_i + \widehat{\lambda}\text{single}_i + \widehat{\xi}\text{large}_i + \epsilon_i$$

$$\widehat{\text{parties}}_i = \underset{(0.161)}{3.463} + \underset{(0.017)}{0.084}\text{district\_size}_i - \underset{(0.173)}{1.118}\text{single}_i - \underset{(2.456)}{11.099}\text{large}_i$$

Now, as expected, the coefficients predict the smaller number of parties observed in single-member systems. The coefficient on district size is still small, for the reasons discussed above, but overall the model would seem to conform better to the observed variation in the data even though the estimated model still can't explain the large number of effective parties observed at some elections and the R^2 only increases to 36%.

The second stylized fact, that the standard deviation of the number of effective parties is increasing in district size can be tested in a similar manner. The standard deviation of the number of effective parties for each district size is termed *sdparties*.

$$\sigma_{district\_size=j} = \alpha + \widehat{\beta}\text{district\_size}_i + \widehat{\lambda}\text{single}_i + \widehat{\xi}\text{large}_i + \epsilon_i$$

$$\sigma_{district\_size=j} = \underset{(0.054)}{0.405} + \underset{(0.049)}{0.210}\text{district\_size}_i + \underset{(0.007)}{0.040}\text{single}_i - \underset{(1.010)}{5.411}\text{large}_i$$

The results suggest that, in accordance with the second stylized fact, that the variation in the number of effective parties is increasing in district size. The positive coefficient on single implies a U-shaped relationship with single member districts and larger districts exhibiting the most variation. To what extent this is an artefact of the available sample or a more general pattern is unclear.

We now turn to the analysis of the simulation data. Such data has many advantages, large numbers of observations, well-defined variables, and no unobserved confounding factors. Given that the simulation model has two key parameters, the number of policy dimensions and district size an obvious initial specification is simply to include each variable. The results are obtained using the weights described above, and reported in Table 6. The results contained in column 1 confirm the conclusions of the discussion in Section 3.2. The number of effective parties increases with district size, and decreases with the number of policy dimensions. However, the coefficient on *PR* seems small. This may be for the same reason as in the empirical data, the influence of the observations for 120 member districts. Column 2 includes a dummy variable, for large districts, but this doesn't meaningfully increase the coefficient on *PR*. The inclusion of a binary variable for single member districts, *single*, reveals that it is instead the opposite phenomenon. The key difference is between single and multi-member districts. Indeed, the inclusion of the interaction of *single* and *dim* in column 4, dramatically increases the estimated effect of single-member districts to a reduction of over 2.5 effective parties. The positive coefficient on the interaction term, suggests that this effect is ameliorated as the number of dimensions increases. This result is important as it suggests that more than just altering the number of parties, the impact of ideology is reversed in single member districts. Perhaps most important, is that these results conform with the estimates obtained using the empirical data, as suggested by the density plots.

Overall, it would seem that the model (in general) makes predictions which conform to the three stylized facts, and the empirical size distribution of parties. Density plots suggest that the model, and in particular when 2 dimensions are assumed, gives remarkably realistic outcomes. Both the empirical and simulation results suggest, in line with the third stylized fact, that it is the single/multi-member district dichotomy that is the key source of variation. That there is greater variability in the number of effective parties is suggested by the importance of the interaction between the number of policy dimensions and the number of representatives per electoral district.

6 Conclusion

The overall conclusion of the paper is that a basic Citizen Candidate framework is remarkably well able to generate empirically and intuitively reasonable outcomes. Key to the approach of the paper are the results of Dutta, Jackson and Breton (2001) since this makes our results more readily comparable to those obtained assuming fully strategic agents. Section 4 demonstrates how equilibrium outcomes vary with the dimension of the policy space, and in particular that more dimensions lead to larger parties. The novel equilibrium identification technique employed suggests that more dimensions, are not at a national level, associated with more equilibria, and moreover that the equilibria are plausible. The endogenous emergence of plausible party systems given minimal assumptions, and the consistent variation in outcomes with different numbers of representatives and dimensions in higher dimensional party equilibria further lends credibility to the results. The previous section discussed in more detail what it can, and can not, explain in empirical terms. It is argued that overall its performance is impressive. The comparison with the empirical data highlighted several advantages of the computational approach. In particular, being able to compare the predicted results with empirical data is a useful tool in evaluating the success of the model. The similarity of results for the case of two policy dimensions suggests that comparing simulation and empirical results can be extremely fruitful.

Table 2: Single District Results

	Number of Dimensions						
PR	1	2	3	4	5	6	7
1	1.74	1.89	1.77	1.69	1.57	1.63	1.79
	0.58	0.60	0.57	0.61	0.56	0.61	0.56
2	2.98	3.21	3.25	3.37	3.13	3.36	3.33
	0.55	0.62	0.73	0.61	0.68	0.72	0.80
3	4.12	4.53	4.70	4.65	4.84	4.88	4.90
	0.71	0.72	0.76	0.85	0.88	1.01	0.97
4	5.22	5.66	5.95	6.17	6.53	6.46	6.62
	0.58	0.74	0.76	1.02	1.00	1.04	0.97
5	6.39	7.02	7.49	7.56	7.91	8.07	8.22
	0.78	0.91	0.94	1.06	1.17	1.26	1.14
6	7.62	8.22	8.83	9.32	9.50	9.49	10.10
	0.63	0.97	1.04	0.98	1.28	1.18	1.33
8	10.19	11.22	11.47	12.20	12.61	12.84	13.16
	0.90	1.11	1.26	1.26	1.28	1.61	1.55
10	12.56	13.50	14.24	14.95	15.46	15.93	16.51
	0.98	1.30	1.28	1.34	1.38	1.30	1.68
12	15.20	16.23	16.89	18.33	18.53	18.95	19.41
	1.18	1.66	1.42	1.45	1.52	1.62	1.73
15	18.80	19.90	20.98	21.81	22.75	23.65	24.49
	1.36	1.37	1.58	1.77	2.01	1.86	2.05
20	25.15	26.79	27.92	29.04	30.22	31.54	32.32
	1.51	1.75	1.92	1.80	2.20	2.02	2.36
24	29.88	31.55	33.35	34.79	36.01	37.52	38.39
	1.52	1.94	1.90	2.16	2.26	2.38	2.42
30	37.91	39.71	42.09	43.49	44.86	46.66	48.26
	2.07	2.40	2.18	2.40	2.62	2.70	3.00
40	50.30	52.69	54.62	58.27	59.90	61.98	64.07
	2.27	2.59	2.47	2.89	2.81	3.00	3.15
60	75.89	78.87	82.36	86.11	89.23	92.37	95.50
	2.72	3.14	3.10	3.64	3.49	3.71	4.06
120	151.88	157.31	163.80	169.95	176.62	182.88	189.04
	4.03	4.43	3.87	4.34	4.38	5.41	5.45

PR refers to the number of representatives elected in the district. For each pair of numbers, the upper is the mean number of candidates and the lower is the standard deviation.

Table 3: Absolute Number of Parties

	Number of Dimensions						
PR	1	2	3	4	5	6	7
1	4.36 1.09	2.3 0.48	2.25 0.44	2.13 0.34	2.19 0.42	2.13 0.34	2.11 0.31
2	5.71 0.97	2.88 0.73	2.44 0.56	2.37 0.71	2.45 1.13	2.48 1.17	2.24 0.45
3	5.71 1.35	3.73 0.95	2.66 0.65	2.53 0.81	2.9 1.48	2.96 1.99	2.9 1.9
4	5.61 1.09	4.78 0.98	2.77 0.87	2.86 1	3.2 1.69	3.33 2.21	3.47 2.44
5	6.53 1.08	4.62 0.83	2.89 0.92	3.15 1.13	3.54 1.9	3.47 2.17	3.6 2.2
6	7.27 0.94	4.61 0.99	3.26 1.17	3.18 1.1	3.6 2.36	3.47 1.82	3.72 2.46
8	8.1 1.34	4.98 1.33	3.48 1.26	3.11 0.95	3.11 1.23	3.44 1.89	3.61 2.4
10	8.14 1.38	5.31 1.46	3.72 1.61	3.23 1.17	3.14 0.97	3.17 1.63	3.75 2.49
12	7.68 1.35	5.63 1.51	3.84 1.7	3.25 1.07	3.46 0.99	3.34 1.29	3.76 2.5
15	7.72 1.55	5.59 1.75	3.62 1.59	3.16 0.93	3.53 1.61	3.3 1.31	3.55 2.11
20	7.73 1.64	5.17 1.46	3.69 1.43	3.69 1.38	3.34 1.26	3.38 1.44	3.59 2.22
24	7.32 1.43	5.03 1.63	3.65 1.53	3.41 1.17	3.47 1.61	3.42 1.64	3.64 2.51
30	7.15 1.34	4.59 1.55	3.64 1.43	3.53 1.46	3.39 1.31	3.69 1.83	3.69 2.49
40	7.24 1.51	4.7 1.4	3.5 1.24	3.42 1.24	3.46 1.54	3.76 2.14	3.15 2.11
60	6.92 1.37	4.57 1.38	3.55 1.32	3.29 1.19	3.39 1.42	3.69 2.24	3.55 2.65
120	7.21 1.38	4.16 1.36	3.57 1.59	3.29 1.05	3.19 1.68	2.8 1.53	2.59 1.83

PR refers to the number of representatives elected in the district. For each pair of numbers, the upper is the mean number of candidates and the lower is the standard deviation.

Table 4: Effective Number of Parties

	Number of Dimensions						
PR	1	2	3	4	5	6	7
1	3.46 0.77	2.18 0.35	2.11 0.24	2.06 0.22	2.09 0.29	2.08 0.26	2.04 0.2
2	4.63 0.9	2.6 0.59	2.29 0.49	2.18 0.48	1.98 0.36	1.98 0.23	2.01 0.1
3	4.2 0.86	3.42 0.81	2.51 0.61	2.25 0.58	2.11 0.48	1.93 0.48	1.88 0.39
4	4.72 0.59	4.2 0.5	2.51 0.68	2.41 0.69	2.38 0.67	2.17 0.54	2.05 0.58
5	5.65 0.52	4.07 0.64	2.64 0.79	2.61 0.78	2.43 0.6	2.31 0.59	2.17 0.74
6	6.52 0.6	4.06 0.86	2.93 1.02	2.73 0.86	2.47 0.68	2.42 0.65	2.15 0.44
8	7.23 1.16	4.39 1.2	3.09 1	2.71 0.76	2.53 0.7	2.39 0.69	2.12 0.55
10	6.95 1.24	4.69 1.28	3.35 1.38	2.79 0.92	2.62 0.65	2.43 0.63	2.36 0.62
12	6.66 1.27	4.91 1.37	3.46 1.49	2.82 0.96	2.84 0.8	2.62 0.64	2.18 0.6
15	6.75 1.37	4.83 1.52	3.38 1.44	2.72 0.76	2.79 0.8	2.61 0.76	2.24 0.64
20	6.63 1.36	4.52 1.33	3.34 1.31	3.14 1.16	2.76 0.85	2.56 0.69	2.19 0.64
24	6.31 1.3	4.47 1.44	3.33 1.4	2.96 1.02	2.73 0.76	2.55 0.7	2.2 0.67
30	6.17 1.22	4.13 1.41	3.24 1.21	3.01 1.16	2.76 0.87	2.62 0.75	2.15 0.65
40	6.28 1.3	4.26 1.27	3.14 1.15	2.9 0.86	2.78 0.84	2.49 0.66	1.97 0.51
60	6.03 1.2	4.12 1.29	3.1 1.08	2.8 0.92	2.7 0.8	2.3 0.58	1.95 0.56
120	6.28 1.12	3.79 1.22	3.15 1.39	2.81 0.88	2.18 0.54	1.97 0.3	1.95 0.33

PR refers to the number of representatives elected in the district. For each pair of numbers, the upper is the mean number of candidates and the lower is the standard deviation.

Table 5: Numbers of District and National Equilibria

	Number of Dimensions						
PR	1	2	3	4	5	6	7
PR	1	2	3	4	5	6	7
1	3 22	3 4	3 3	3 2	4 3	3 2	3 2
2	5 8	6 9	5 8	6 6	5 5	6 4	6 4
3	11 22	15 9	18 8	14 6	8 3	9 3	10 4
4	18 18	25 17	28 8	20 8	28 6	26 4	31 4
5	18 25	38 24	45 6	46 8	44 4	47 3	48 5
6	25 26	36 21	49 4	57 7	52 3	53 4	55 3
8	13 41	30 15	57 6	50 6	52 5	55 4	50 4
10	11 31	39 19	42 8	54 7	58 8	58 3	55 3
12	12 37	35 20	50 4	56 8	57 8	54 4	59 4
15	18 28	26 16	48 3	49 10	53 5	54 4	50 4
20	12 24	31 19	44 5	55 6	51 4	49 3	57 3
24	14 25	36 13	40 5	51 6	50 4	60 4	51 6
30	19 27	23 10	43 5	48 5	45 5	52 4	54 3
40	7 26	25 15	48 4	49 5	51 8	58 4	55 3
60	6 25	34 13	42 5	46 4	52 5	49 3	49 3

PR refers to the number of representatives elected in the district. For each pair of numbers, the upper is the number of district equilibria and the lower is the number of national equilibria

Table 6: Regression Analysis of Simulation Data

	(1)	(2)	(3)	(4)	(5)
pr	0.015***	0.018***	0.004***	0.004***	0.026***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
dim	−0.437***	−0.439***	−0.448***	−0.554***	−0.367***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
large		−1.857***	−0.576***	−0.642***	−0.763***
		(0.032)	(0.032)	(0.031)	(0.037)
single			−0.984***	−2.586***	−1.010***
			(0.004)	(0.008)	(0.004)
single × dim				0.394***	
				(0.001)	
pr × dim					−0.006***
					(0.000)
Constant	4.584***	4.556***	5.062***	5.512***	4.765***
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
R^2	0.383	0.388	0.453	0.509	0.469
N	452064	452064	452064	452064	452064

References

Alesina, Alberto, "Credibility and Policy Convergence in a Two-Party System with Rational Voters," *The American Economic Review*, 1988, *78* (4), 796–805.

Baylie, Dan and Guy Nason, "Cohesion of Major Political Parties," *British Politics*, 2008, *3*, 390–417.

Besley, Timothy and Stephen Coate, "An Economic Model of Representative Democracy," *The Quarterly Journal of Economics*, 1997, *112* (1), 85–114.

Caramani, Daniele, *Elections in Western Europe since 1815: Electoral Results by Constituencies*, 1st ed., Macmillian Reference, 2000.

Dhillon, Amrita, "Political Parties and Coalition Formation," in Gabrielle Demange and Myrna Wooders, eds., *Group Formation in Economics: Networks, Clubs, and Coalitions*, Cambridge University Press, 2005, chapter 9.

Downs, Anthony, *An Economic Theory of Democracy*, New York: Harper & Row, 1957.

Dunleavy, Patrick, Rekha Diwakar, and Christopher Dunleavy, "Is Duverger's Law based on a mistake?," 2008. Political Studies Association Annual Conference Paper.

Dutta, Bhaskar, Matthew O. Jackson, and Michel Le Breton, "Strategic Candidacy and Voting Procedures," *Econometrica*, 2001, *69* (4), 1013–1038.

Duverger, Maurice, *Political Parties: Their Organization and Activity in the Modern State*, Wiley, 1951.

Gallagher, Michael and Paul Mitchell, *The Politics of Electoral Systems*, 1st ed., Oxford University Press, 2005.

Gelman, Andrew, Jonathan N. Katz, and Francis Tuerlinck, "The Mathematics and Statistics of Voting Power," *Statistical Science*, 2002, *17* (4), 420–435.

_ , Jonathan N. Katz Katz, and Bafumi Joseph, "Standard Voting Power Indexes Do Not Work: An Emprical Analysis," *British Journal of Political Science*, 2004, *34*, 657–674.

Golub, Gene H. and Charles F. Van Loan, *Matrix Computations*, 3rd ed., John Hopkins University Press, 1996.

Hartigan, J. A. and M. A. Wong, "A K-means clustering algorithm," *Applied Statistics*, 1979, *28*, 100–108.

Kline, Reuben, "How we count counts: The empirical effects of using coalitional potential to measure the effective number of parties," *Electoral Studies*, 2009, *28*, 261–269.

Kollman, Ken, John H. Miller, and Scott E. Page, "Adaptive Parties in Spatial Elections," *The American Political Science Review*, 1992, *86* (4), 929–937.

_ , _ , and _ , "Political Institutions and Sorting in a Tiebout Model," *The American Economic Review*, 1997, *87* (5), 977–992.

Laakso, M. and Rein Taagepera, "Effective Number of Parties: A Measure with Application to West Europe," *Comparative Political Studies*, 1979, *12*, 3–27.

Laver, Michael, "Policy and the Dynamics of Political Competition," *The American Political Science Review*, 2005, *99* (2), 263–281.

_ **and Michel Schilperoord**, "Spatial models of political competition with endogenous political parties," *Philisophical Transactions of The Royal Society B*, 2007, *362*, 1711–1721.

Levy, Gilat, "A Model of Political Parties," *Journal of Economic Theory*, 2004, *115* (2), 250–277.

Lijphart, Arend, *Patterns of Democracy: Government Forms and Performance in Thirty-six Countries*, paperback ed., New Haven, London: Yale University Press, 1999.

Lloyd, Stuart P., "Least Squares Quantitization in PCM," *IEEE Transactions on Information Theory*, 1982, *28* (2), 129–137.

Merlo, Antonio M., "Whither Political Economy? Theories, Facts and Issues," in Richard Blundell, Whitney Newey, and Torsten Persson, eds., *Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress of the Econometric Society*, Vol. 1, Cambridge: Cambridge University Press, 2006, pp. 381–421.

Morelli, Massimo, "Party Formation and Policy Outcomes under Different Electoral Systems," *The Review of Economic Studies*, July 2004, *71* (3), 829–853.

Moser, Robert G. and Ethan Scheiner, "Mixed electoral systems and electoral system effects: controlled comparison and cross-national analysis," *Electoral Studies*, 2004, *23*, 575–599.

of Electoral Systems, University of Michigan Comparative Study, "The Comparative Study of Electoral Systems (www.cses.org). CSES Module 1," www.cses.org August 2003.

_ , "The Comparative Study of Electoral Systems (www.cses.org). CSES Module 2," www.cses.org June 2007.

Osborne, Martin J. and Al Slivinski, "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics*, Febuary 1996, *111* (1), 65–96.

_ **and Rabee Tourky**, "Party Formation in Single-Issue Politics," *Journal of the European Economic Association*, 09 2008, *6* (5), 974–1005.

Roemer, John E., "The Democratic Political Economy of Progressive Income Taxation," *Econometrica*, 1999, *67* (1), 1–19.

Rustichini, Aldo, "Optimal Properties of Stimulus-Response Learning Models.," *Games and Economic Behavior*, October 1999, *29* (1-2), 244–273.

Snyder, James M. and Michael M. Ting, "An informational rationale for political parties," *American Journal of Political Science*, 2002, *46* (1), 90–110.

Tullock, Gordon and Colin D. Campbell, "Computer Simulation of a Small Voting System," *The Economic Journal*, March 1970, *80* (317), 97–104.

Wittman, Donald, "Candidates with policy preferences: A dynamic model," *Journal of Economic Theory*, February 1977, *14* (1), 180–189.