Public Choice **113:** 337–356, 2002.

Why party leaders are more extreme than their members: Modeling sequential elimination elections in the U.S. House of Representatives *

A.J. MCGANN[1], BERNARD GROFMAN[1] & W. KOETZLE[2]
Department of Political Science and Institute of Mathematical Behavioral Sciences, University of California, Irvine, CA 92697-5100, U.S.A.; [2]U.S. House of Representatives

Accepted 23 June 2001

Abstract. Grofman et al. (forthcoming) find that party leaders in the U.S. House of Representatives tend to be more extreme than the median member of their party, and that they tend to come from the party's ideological "heartland" between the median and the mode. This paper shows that if the distribution of preferences is skewed (as is the case with both parties in the House), then we should expect sequential elimination elections to choose on average leaders between the median and modal positions. We show that this is the case whether or not the party is factionalized.

1. Introduction

Contrary to a long line of literature going back to Truman (1959), Grofman et al. (forthcoming) find that party leaders are not in general ideological "middlemen". Rather, using adjusted ADA[1] and ACU[2] scores over a thirty year period, they conclude that party leaders in the House have been considerably more extreme than the median Member of their parties. Furthermore, they find that the distribution of scores for both parties has been highly skewed. For the Democrats, for example, there was a large concentrated mass of liberals and a long tail made up of moderates and (mostly southern) conservatives. The pattern for the Republicans was a virtual mirror image of this. Thus party leaders have tended to come not from the median of a party, but rather from its dominant wing. These results are consistent with the "policy partisanship" theory of Congressional leadership advanced by Clausen and

* An earlier version of this paper was prepared for delivery at the Annual Meeting of the American Political Science Association, Boston, Massachusetts, September 3–6, 1998. The authors wish to thank Thomas Brunell, SUNY Binghamton, and John DiNardo, University of California, Irvine, for their helpful comments, and Clover Behrend for bibliographic assistance. The contributions of the second author to the completion of this manuscript were supported by National Science Foundation Grant # SBR 446740-21167, Program in Methodology, Measurement, and Statistics (to Bernard Grofman and Anthony Marley).

Wilcox (1987), who suggest that leaders have tended to be situated in the ideological heartland of their party, somewhere between the median and the mode.

This paper shows that given a skewed distribution of preferences and sequential elimination elections, we should expect outcomes to fall between the median and the mode, as we observe with the House data from Grofman et al. (forthcoming). This is true whether or not the parties are internally factionalized. When considering leadership selection by parties there are three crucial variables:

The first is the *voting rule* used by the party. We know from the social choice literature that electoral rules can have a major impact on election outcomes (Riker, 1982; Saari, 1995). For example, the Downsian party convergence result rests on a variety of ancillary assumptions. Among the most important of these is the requirement that there is only a single election. In situations where there are party primaries, party/candidate positions are no longer expected to be identical, and the amount of divergence from the location of the overall median voter can be quite large (Aranson and Ordeshook, 1972; Coleman, 1971, 1972; Owen and Grofman, 1995). Similarly, we might not expect the same types of outcomes under simple plurality as under some form of sequential balloting involving runoffs.

There is a body of work demonstrating that outcomes under plurality voting need not be representative of group preferences. Indeed, under plurality, some simulation results are of a "just about anything can happen" nature and it is clear that the candidate closest to the median preference (the Condorcet winner) need not be chosen. Indeed, under plurality voting, there are circumstances when even a so-called "Condorcet loser," (a candidate who loses in paired competition against each and every other alternative) may be selected. However, there is reason to expect that, under majority rule runoff[3] and majority rule sequential elimination elections,[4] the candidate closest to the median preference is somewhat more likely to be chosen than would be the case under plurality (Merrill 1984, 1985). Because majority sequential runoffs are used for the selection of leaders and other representative bodies in many organizations,[5] and, in particular, because this rule is used for Congressional leadership selection in both parties in the United States, this is the voting rule we shall focus on in our subsequent discussion.[6]

The second key factor in understanding voting outcomes is the *distribution of voter preferences*. Much of the previous simulation work comparing the outcomes of different voting rules posits that the underlying distribution of voter preferences is a distribution with strong symmetry properties, such as a normal distribution, a uniform distribution, or the so-called "impartial culture" (a distribution in which all linear orderings among alternatives are

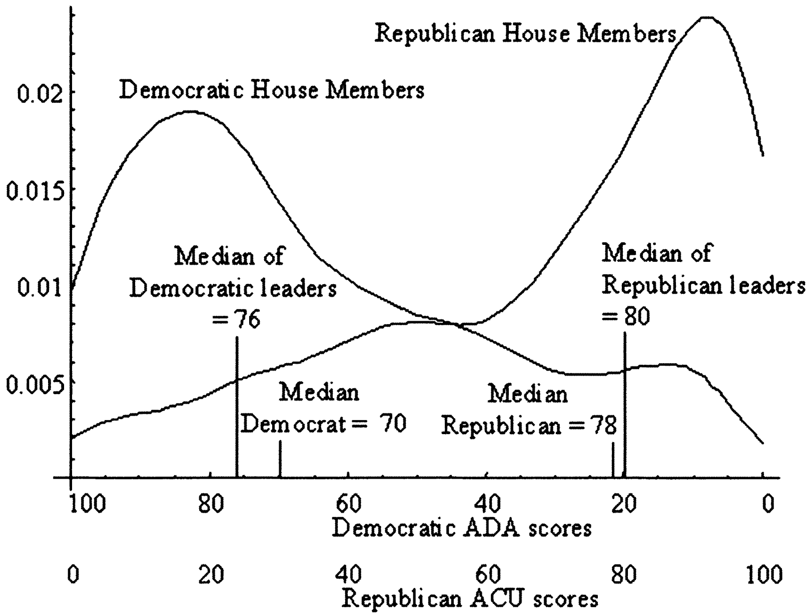

Figure 1. Kernel density plots of Democratic House adjusted ADA scores 1965–1996 and Republican House ACU scores 1971–1996.

equiprobable). We believe that results based on such distributions are likely to very unrealistic (see e.g., Regenwetter, Adams and Grofman, 2000). Real world distributions tend to be asymmetric and these asymmetries can strongly affect the political process.[7]

We base our simulations on the actual distribution of party Members in the U.S. House of Representatives. We have used ADA scores pooled over the period 1965–1996 and ACU scores pooled over the period 1971–1995 as our measures of member ideology.[8] Following the line of argument in Brunell et al. (1999),[9] we have used ADA scores to distinguish ideology among Democrats and ACU scores to distinguish ideology among Republicans.

Kernel density functions (smoothed histograms)[10] of the adjusted ADA scores[11] of House Democrats 1965–1996, and of the adjusted ACU scores of House Republicans 1971–1996, are shown in Figure 1. Note that both distributions are strongly skewed. In the case of the Democrats the mode is considerably to the left of the median (Democratic median ADA score = 70), while for the Republicans, it is considerably to the right of it (Republican median ACU score = 78).

The third key factor is the nature of the *group's factional structure* and its impact on who gets nominated for party leadership positions.[12] We shall consider what happens under two different types of assumptions about the nature of the within-party nomination process. In the first, there are no factions and k members nominate themselves at random from some underlying distribution of unidimensional ideological preferences. In the second type of nomination procedure, we assume that the nominees are the candidates of particular factions within the group. In both instances, there is no guarantee that a candidate near the actual median voter will be nominated, and even if they are, they may not win.

It is hard to tell exactly how factionalized Congressional parties are, because internal party matters are not in the public domain. Certainly it has been the case that "American party factions have characteristically been ad hoc, amorphous, and undisciplined; centered around particular personalities; and rooted in sectional divisions" (Rae 1989). However, Schousen (1994), Rohde (1991), Reiter (1981), Rae (1989), Koopman (1996) and Peabody (1967) all argue in different ways that factions have been important in coordinating leadership competition within the parties. Because of the uncertainty concerning the relevance of factions in leadership selection, we model sequential elimination elections both with and without factions.

We outline three models that explain how a skewed distribution of preferences can lead to House leaders tending to be more extreme than the median Member of their party. Firstly, we assume that candidates nominate themselves randomly. Secondly, we assume that candidates for leadership positions are nominated by factions within the party, and provide two different models of how the party divides itself into factions.

We begin with a simulation of the expected outcomes of a majority rule sequential elimination process in which there is no pre-set ideologically-based subgroup structure, but in which k nominees are randomly drawn from the actual distribution of Democratic party Members and Republican party Members in the U.S. House of Representatives over the period 1965–1996. Considering random candidates allows us to investigate the effect of the sequential elimination election rule separately from the effect of any factionalization process we impose. We show that, because the distributions of both parties from which we are drawing are strongly skewed, we tend to get outcomes between the mode and the median – outcomes in which the chosen leader will be more "extreme" than the party's median member.

Secondly, we consider two different models where the parties break into factions that nominate candidates for party leader. The first of these is a game-theoretic partitioning model, derived from McGann (1997, 2000), which is itself a multi-faction generalization of the two-party models of Robertson (1976), Aldrich (1983) and Aldrich and McGinnis (1989). It is an equilibrium model of faction formation and is similar in spirit to the models of local public goods developed in the literature following Tiebout (1956). The models of Westhoff (1977, 1979), Milchtaich and Winter (1997), Kollman et al. (1997) and Adams (1998) are particularly relevant here, in that they consider processes in which the membership of each group collectively determine the group's "position".

Using the pooled Congressional data from Grofman et al. (forthcoming), we calculate how each party would partition itself into factions according to the model. We then assume that each faction nominates its median member, and calculate the outcome of a sequential elimination election with these candidates. The nature of the predicted subgroup formation process in the partitioning model is related to the shape of party ideological distributions. In particular, if the distribution is skewed so that the density is greatest in one part of the distribution, then the faction located in the highest density portion of the ideological range is advantaged. For unimodal distributions, this tends to be the faction located closest to the mode. For the predicted Congressional party subgroupings for the actual (ideologically skewed) distributions of Democratic ADA scores and Republican ACU scores in the House pooled over the period 1965–1996, we find that the predicted leadership results will typically be closer to the mode than the median.

Finally, we use a second subgroup-based model that is cluster-theoretic in its origins – the Grofman (1982) dynamic model of proto-coalition formation.[13] Rather than positing an equilibrium partition, the model is based on specifying a process by which coalitions merge to form larger and larger coalitions. As with the game-theoretic partitioning model, we find that the predicted outcomes are considerably more extreme than the median party Member.

2. Simulation results for majority rule sequential elections with non-factionalized parties (random candidates)

As outlined in the previous section, three factors affect the outcome of leadership selection: the voting rule, the distribution of preferences, and the factional make-up of parties that determines which candidates run. Both major parties select leaders using majority rule sequential elimination elections. The distribution of ideological preferences in both parties seems to be skewed so that the modal member is substantially more extreme than the median. The degree of factionalization in Congressional parties is a matter of debate – there certainly are factions, but they are very informal by international standards. This section models the effect of skewed preference distributions on leadership selection by sequential elimination elections without factions within parties. This will demonstrate that even without factions, sequential elimination elections tend to be biased towards outcomes that vary from the median in the direction of the population mode. The following sections will show that factionalization tends to amplify this effect.

We proceed using computer simulation of sequential elimination elections with randomly generated slates of candidates and a one-dimensional distribution of preferences, using methods similar to Fishburn and Gehrlein (1976, 1977) and Merrill (1984, 1985).[14] Merrill (1984, 1985) shows that with a symmetric distribution of preferences, sequential elimination elections choose a candidate at the median on average, and have other advantageous qualities over simple plurality elections (such as having a far lower variance and being far more likely to pick the candidate closest to the population median). McGann (1999), and McGann et al. (2000), however, show that if the distribution of preferences is skewed, sequential elimination elections do not tend to pick the median position on average, but rather a position between the median and the mode. This is because the candidate closest to the median often does not make it to the final round of the election. Thus, although this candidate will beat any other candidate in a head-to-head race, he or she often will not make it to the final round. Candidates close to the mode, however, benefit from being in the area of highest preference density, and are thus more likely to make it to the final round.

Figure 2 illustrates this dynamic for five candidates, using the distribution of preferences that corresponds to the adjusted ADA scores of Democratic Representatives 1965–1996 shown in Figure 1, where higher scores are more liberal. Candidates 1, 2, 3, 4 and 5 are positioned at points 20, 45, 70, 85 and 95 respectively, where 70 is the population median. When we run sequential elimination elections using this distribution of preferences, we find that candidates 1 and 5 are eliminated in the first and second rounds respectively. Their supporters then support candidates 2 and 4. Thus in the third round, candidates 2 and 4 can eliminate candidate 3, even though candidate 3 is closest to the median. We are left with a final round in which a "champion of the left" (candidate 4) faces a "champion of the right" (candidate 2). The winner between these two candidates (candidate 4) is the one closer to the median. However, taking the median position is not advantageous, as this position is eliminated in earlier ballots (candidate 3 in our example).

In each simulation run we generated a random slate of ten candidates. Each member then voted (sincerely) for the one of the ten candidates whose position is closest to his ideal point, with voters distributed according to the distributions shown in Figure 1.[15] The candidate with the least votes was eliminated, and everyone then voted again, until only one candidate remained. We repeated the whole process 1000 times using a Mathematica routine. We can consider the median of the 1000 winning candidates, and

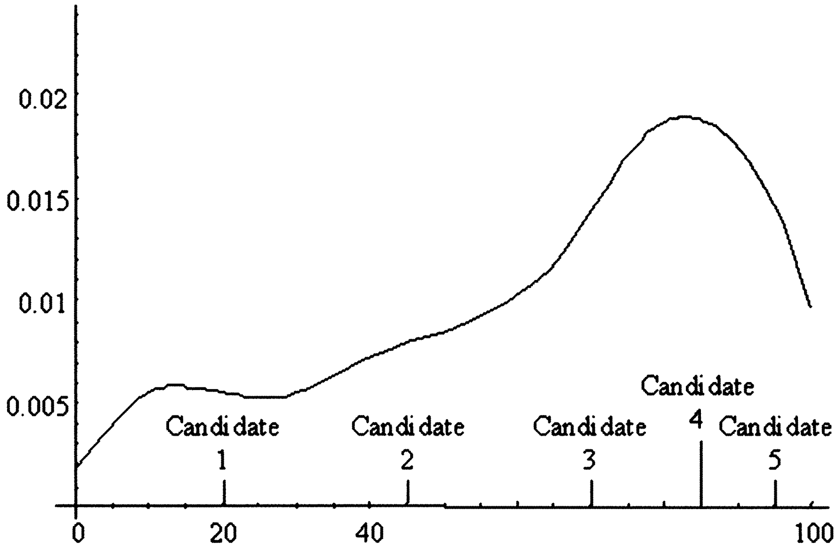

Figure 2. Hypothetical sequential elimination election with 5 candidates.

compare this to the population median. Because we have theoretical reasons to expect that the outcomes under some electoral systems will be highly sensitive to the distribution of candidates, simulations were run with several different candidate distributions. The distributions of candidates used were:

1. Uniform distribution in interval [0, 100]
2. The same distribution as the party Members.
3. Candidates concentrated in center of distribution: Beta [2, 2] distribution, scaled to the interval [0, 100]. This corresponds to the assumption that centrists are more likely to run than extremists.
4. Candidates concentrated at extremes of distribution: Beta [0.5, 0.5], scaled to the interval [0, 100]. This corresponds to the assumption that extremists are more likely to run than centrists.

We find that our simulated sequential elimination elections produce winning candidates that vary systematically and significantly from the median, and that the divergence from the median is in the direction of the mode. Furthermore, these results appear to be remarkably robust to the distribution from which the candidates are drawn.

Table 1 gives the simulation results for House Democrats. Depending on which distribution of candidates was used, the median winner was between 4

Table 1. Simulation of sequential elimination elections with candidates drawn from the democratic ADA distribution

Candidate distribution	Uniform	Same as members	Beta [2,2] (dense in center)	Beta [.5, .5] (dense at extremes)
median of winning candidates	76.2	76.8	75.	75.5
distance from median (70)	**6.52**	**7.1**	**5.27**	**5.81**
mean of winning candidates	74.5	74.8	73.6	73.6
standard deviation	9.71	9.83	8.86	10.6
standard error	0.307	0.311	0.28	0.336
Condorcet efficiency %	42.1	27.4	38.5	58.6

and 9 points (on a 100 point scale) more liberal than the median Democratic score of 70. It is notable that the median ADA score for a Democratic leader in this period was 76, very close to the results produced by the simulation (Grofman et al., forthcoming). Note that the standard deviation of the winning candidates in all cases is around 10, indicating that there are some winners from a broad ideological range either side of the median.

A similar pattern is visible for the Republicans, as can be seen in Table 2. The median winning candidate amongst the Republicans is (depending on candidate distribution) between 2 and 5 points more conservative than the median Republican score of 78. Once again this result varies little with different assumptions about the distribution of candidates. The actual median ACU score for Republican party leaders in the period studied was 80 (Grofman et al., forthcoming). The standard deviation of the winning candidates is (depending on which candidate distribution is used) between 8 and 10.

Thus, even with random candidate generation, we would expect party leaders chosen by sequential elimination elections to be more extreme than the median member of their party. This result is generated solely by the dynamics of sequential elimination elections, without any factionalization within the parties. The next section will show that, when we take factionalization into account, we should expect party leaders to be even more extreme relative to the party rank and file.

3. Simulations for majority rule sequential elimination elections with parties factionalized according to the McGann (1997) model

So far we have considered parties as atomized, with candidates for leadership randomly chosen from the distribution of Members. If, however, we consider

Table 2. Simulation of sequential elimination elections with candidates drawn from the republican ACU distribution

Candidate distribution	Uniform	Same as members	Beta [2, 2] (dense in center)	Beta [.5, .5] (dense at extremes)
median of winning candidates	81.	83.4	79.7	82.1
distance from median voter	**3.0**	**5.4**	**1.7**	**4.1**
mean of winning candidates	80.	81.5	78.4	80.1
standard deviation	8.82	8.34	8.39	9.94
standard error	0.279	0.264	0.265	0.314
Condorcet efficiency %	44.6	29.6	47.3	55.1

the effects of factionalism, sequential elimination elections produce results that are even more biased towards the mode of the distribution. We proceed by using a partitioning model of faction formation derived from Aldrich (1983) and McGann (1997, 2000). Here individuals autonomously partition themselves between factions, so that each individual joins the faction whose median member is as close as possible to that individual's ideological preference. Once we have partitioned individuals into factions using this process, we assume that each faction nominates a candidate whose position is equivalent to that of the faction's median member. These candidates then compete in a sequential elimination election where all individuals vote sincerely for the candidate closest to their position.

3.1. *A partitioning model of factionalization*

The model we use is outlined at length in McGann (1997, 2000). Here we provide a summary. Let us assume that the population (in this case the Members of Congress belonging to the party in question) is distributed across a one-dimensional issue space S, between L and U, and the distribution is defined by the cumulative distribution function F. F is continuous and monotonically increasing across the domain defined by points L and U (L < U). Assume we have n factions, which take positions $x_1 \ldots x_n$ on S, represented by a vector x in S^n. The number of factions is fixed, and there is no possibility of entry by new factions. All voters know the position of each faction with certainty.

Let us make three further assumptions:

1. Each member of the population joins the faction whose position is closest to their own ideal point. Each member of the population has complete knowledge of all faction positions.
2. Each faction adopts the position of its median supporter.
3. Factions are ordered so their positions (x_1, x_2, etc.) are so $x_1 > x_2$, etc.[16]

Equilibrium in this model is a situation in which members are partitioned between factions in such a way that every member belongs to the faction whose median member is closest to that member's ideal point. McGann (2000) shows that for any number of factions, this model must have an equilibrium. Furthermore, concavity of the distribution of voters is a sufficient condition for that equilibrium to be unique and stable.[17]

A result that is particularly significant here is that for a single peaked distribution, factions near the mode will be larger and more closely spaced than factions further from the mode (M^c^Gann 1997, 2000). This is particularly important if the distribution is skewed, as it leads to the conclusion that an ideologically cohesive minority can be more powerful than a more dispersed majority. If the distribution is skewed to the right, then the mode will be on the left side of the median. Therefore the factions on the left will tend to be closer to the mode, and will thus tend to be larger and more closely spaced. If factions nominate candidates whose positions are equivalent to the median position of the faction, this will mean that the candidates of the factions on the right will be positioned in less dense areas of the distribution. We would expect this to put these candidates at a disadvantage under sequential elimination elections.

3.2. *Application of the partitioning model to majority rule sequential elimination elections*

Let us assume that factions are formed using the model outlined in the previous section. Then let us assume that each faction "nominates" a candidate at its own median position, and that these candidates then compete in majority rule sequential elimination elections. Using the partitioning model, we can calculate the equilibrium positions of the candidates of each faction and the position of the winner for both the Democratic and Republican distributions of preferences shown in Figure 1.

Tables 3 (for Democrats and ADA scores) and Table 4 (for Republicans and ACU scores) show the positions of the candidates of each faction according to the partitioning model for the cases of two to nine factions, and the position of the winning candidate in each case.[18] Figure 3 illustrates the results for the Democrats in the two-faction case. In this case the model predicts that the median voter in faction 1 will have an ADA score of 36,

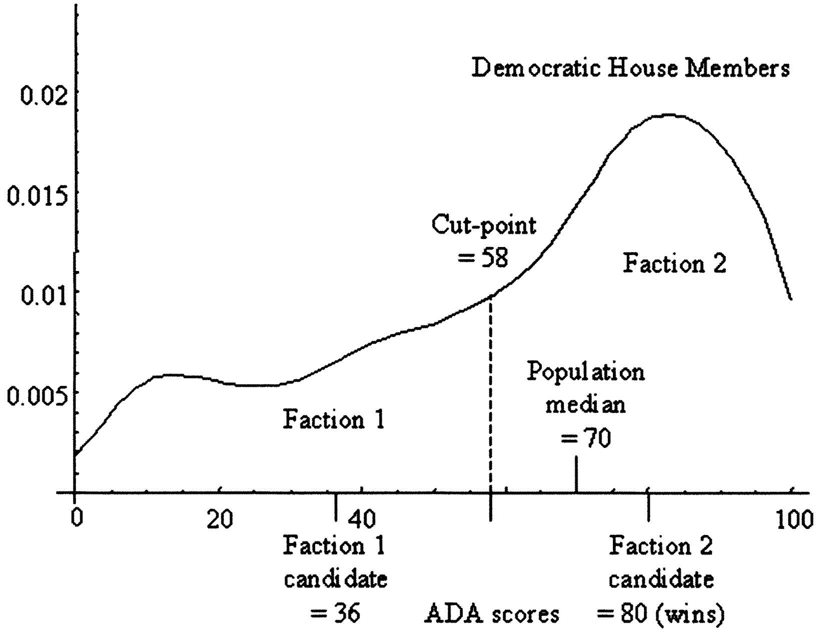

Figure 3. Result of Partitioning Model with Two Factions.

while the median voter in faction 2 will have an ADA score of 80. Thus, the boundary between the two factions would be predicted to be at an ADA score of 58 (=(36 + 80)/2). Given that the population median is at an ADA score of 70, it is apparent that faction 2 will be larger than faction 1, and that if both factions nominate their median members (positions 36 and 80 respectively), the candidate of faction 2 will win. Furthermore, even if faction 1 does not nominate its median member, but instead nominates its most moderate member (who has a score of 58), the candidate of faction 2 at position 80 would still win, being closer to the median of 70. The third column of Tables 3 and 4 gives the position (ADA for Democrats, ACU for Republicans) of the winner, when all the candidates nominated by the partitioning model were run in a sequential elimination election. While we present only the results for the *pooled* 1965–1996 adjusted ADA data and *pooled* 1971–1996 adjusted ACU data, analogous calculations can be performed for the ADA and ACU distributions in any given year.

For the Democrats, it can be seen that the winner's position is substantially to the left of the median (70) in all cases except that of three factions (see Table 3). The factions on the right of the distribution have a wider spread of ideological positions. This draws their median positions away from the

Table 3. Results of McGann factionalization model for democratic ADA scores 1965–96

Number of candidates	Candidate positions	Winner's position
2	{35.5, 80.4}	80.4
3	{22.5, 59.1, 84.9}	59.1
4	{17.5, 47.4, 71.2, 88.0}	88.0
5	{15.5, 41.2, 61.2, 77.2, 90.9}	77.2
6	{14.1, 36.4, 53.8, 68.9, 80.6, 91.9}	80.6
7	{12.5, 31.0, 46.8, 61.0, 73.1, 83.1, 93.1}	83.1
8	{11.3, 27.0, 42.0, 55.1, 66.7, 76.3, 85.0, 93.9}	85.0
9	{10.3, 24.2, 38.3, 50.1, 61.0, 70.7, 78.8, 86.4, 94.4}	78.8

Table 4. Results of McGann factionalization model for republican ACU scores 1971–96

Number of candidates	Candidate positions	Winner's position
2	{39.9, 85.3}	85.3
3	{31.0, 66.2, 89.5}	66.2
4	{22.5, 49.7, 74.8, 91.5}	91.5
5	{19.8, 44.4, 65.3, 81.0, 93.1}	81.0
6	{15.5, 36.6, 53.8, 70.9, 83.6, 93.9}	70.9
7	{13.1, 32.1, 48.2, 63.5, 76.0, 86.1, 94.7}	86.1
8	{11.4, 28.5, 43.2, 56.2, 68.9, 79.0, 87.6, 95.2}	79.0
9	{10.5, 26.6, 40.0, 51.8, 63.6, 73.6, 81.9, , 95.8}	89.1

median position of the distribution as a whole. However, the candidates on the right of the distribution still retain enough support to be able to eliminate the candidate closest to the median. As a result, in the final ballot, when a "champion of the left" competes against a "champion of the right", the "champion of the left" is closer to the median, and thus wins. We may note that these results predict that for the Democratic party a more liberal candidate is elected than we would expect from the simulations of majority rule sequential elimination elections with random candidates.

With the Republicans (using pooled adjusted ACU scores) we observe results that are similar, although somewhat less pronounced (see Table 4). With three and six factions we get results that are substantially more moderate than the Republican median (78). However, with all other number of factions, the winning candidate is more extreme than the median.

4. Simulations for majority rule sequential elimination elections with parties factionalized according to the Grofman (1982) proto-coalition formation model

4.1. *The Grofman model of factionalization*

The Grofman (1982) model treats the factional formation process as a kind of sequential dyadic marriage market, in which subgroups (proto-coalitions) play the role of potential marriage partners, with proto-coalitions growing in size with the accretion of new partners. Begin with some set of actors each characterized by a location in unidimensional space,[19] u_i, and a weight w_i. New proto-coalitions are assumed to form from two previous proto-coalitions when the two proto-coalitions agree to join each other. If two proto-coalitions join, they are posited to locate at their common center of gravity (determined by weighting each proto-coalition by the size of its membership).

Under the above assumptions, if proto-coalitions i and j join together, the location of the new combined proto-coalition will be given by:

$$(u_i w_i + u_j w_j)/(w_i + w_j)$$

and its weight will be:

$$w_i + w_j$$

Let $c(u_i, u_j)$be the change between the original position of the ith coalition, u_i, and the position of the new proto-coalition which it has joined. Under the above assumptions:

$$c(u_i, u_j) = u_i - (u_i w_i + u_j w_j)/(w_i + w_j)$$

and

$$c(u_j, u_i) = u_j - (u_i w_i + u_j w_j)/(w_i + w_j).$$

Note that change, as so defined, is asymmetric, since the smaller proto-coalition moves a greater distance from its original position than does the larger of the two proto-coalitions.

To determine how the proto-coalition process works, we posit that the ith proto-coalition examines the set of all possible "marriages" with other proto-coalitions, seeking the proto-coalition j in which $c(u_i, u_j)$is minimized.[20] Proto-coalition i proposes a partnership to this proto-coalition. A "match" between two proto-coalitions is made only when each is the other's preferred partner. The mathematician, Philip Straffin, has proven the important result that, regardless of the distributions of locations and weights, at any given round of the proto-coalition process there must be a least one "match" (Straffin and Grofman, 1984). At each new round, the new proto-coalitions formed at the previous round along with those left single, repeat the process of looking for new matches. Once a match is formed it is indissoluble. Grofman (1982) demonstrates that the set of proto-coalitions which form must be "connected" in the sense of Axelrod (1970), i.e., if j is located on the ideological continuum in between i and k, then if i and k end up in a proto-coalition together, j must also be a member of that proto-coalition.

4.2. *Application of the Grofman factionalization model to majority rule sequential elimination elections*

Given the nature of Grofman's proposed proto-coalition process, we would expect that areas of the ideological space that are densely populated will be natural breeding grounds for proto-coalitions matches, since such dense areas will give rise to coalitional opportunities with proto-coalition partners who differ little from oneself. When we apply the Grofman proto-coalition model to the kernel density distributions for major political parties in the U.S. House of Representatives we find that the center of gravity of the eventual majority coalition in a party is a location that is more extreme than the median party member.

In each simulation, we represented the Members of a party by drawing two hundred ideal-points at random from the distribution of the ADA or ACU scores for the party in question. We then ran the clustering algorithm on these party "members" until they were divided into two factions. This process was repeated 1000 times for each party, each time with a different drawing of Member ideal-points. The results can be seen in Table 5. In the case of both parties the centroid of the majority faction (u_i) tends to be considerably more extreme than the median party member. In the case of the Democratic Party, the median of the centroids of the majority faction over 1000 simulations was 78.8, compared to a population median of 70. In the case of the Republicans the median of the centroids of the majority faction was 83, against a population median of 78.

5. Discussion

There is in general no equilibrium set of strategies for plurality elections with more than two candidates,[21] and no reason to believe one exists for sequential

Table 5. Results of Grofman clustering model for democratic ADA scores 1965–96 and republican ACU scores 1971–96

	Democratic ADA	Republican ACU
Median of centroids of dominant clusters	78.8	83.0
Distance from median voter	8.77	4.97
Mean of centroids of dominant clusters	74.3	79.4
Standard deviation	13.0	11.6
Standard error	0.41	0.367

elimination elections.[22] However, using simulation techniques (either with random candidates or with particular assumptions about the factional structure that determines who is nominated), we have been able to obtain some strong results about the probable outcomes of sequential elimination elections that complement and extend those of Merrill (1984, 1985). Our results are built on three foundations. First, we made particular and realistic assumptions about the nature of the underlying preference distributions. Second, we made realistic assumptions about the mechanism used for voting. Third, we developed models of candidate nomination that were linked to the nature of party factionalization.

Thus, looking at simulations based on the actual voting mechanism used for leadership selection in the U.S. House of Representatives (majority rule sequential elimination) and the actual distribution of ideological attitudes of Democratic and Republican members of the U.S. House in recent decades, we find a theoretical explanation to support the empirical finding that, for each party, party leaders in the House tend to be more extreme than the median member of the party, in the direction of the party mode. Indeed, based on our simulation results, under any one of the several models we considered, we expect that the sequential voting process should make it likely that candidates more extreme than the median and nearer to the mode will be chosen.

This phenomenon clearly has implications for the way that two party government works as a system of representation. We have found that party leaders tend to be more extreme than their median members, who in turn will already be more extreme than the median voter in the population as a whole. As a result, the median voter may face a choice between two polarized parties, neither of which reflects that voter's position very well. Furthermore, we might expect these polarized parties to have a hard time negotiating legislative compromises. King (1998) suggests that this polarization of political parties is responsible for a decline in levels of political trust.[23]

Our results also suggest that in general ideologically concentrated groups will be disproportionately influential. We have shown that with sequential elimination elections, the outcome will tend to be biased away from the median in the direction of the population mode. McGann et al. (2000) produces similar results for plurality and run-off elections. Thus the ideologically most concentrated part of the party (the area around the mode) will have its way, even though there is a candidate that a majority of party members would prefer (the median member). As a result, winning is not simply a matter of numbers, but also of concentration and co-ordination. A concentrated minority may be able to dominate a more dispersed majority that is unable to co-ordinate. Our results complement a considerable body of work theorizing or documenting this phenomenon (Davis et al. 1970; Miller 1996; McGann 1997, 2000; Merrill et al. 1999; Grofman et al. 1999; McGann et al 2000).

It is important to consider the conditions under which elections will produce centrist outcomes. The Median Voter Result (Black 1958; Downs 1957) is normatively appealing in that it predicts a central outcome that is also a Condorcet winner. However, this result rests on the assumption that there are only two candidates and one round of elections. When we consider multi-candidate elections, it is clear that the Condorcet winner will not always win (see, for example, Merrill 1988). Our results show that the Condorcet winner need not even win on average. Of course, these results do not contradict the Median Voter Theorem – the assumptions are quite different. However, before appealing to the logic of the Median Voter Theorem, it is important to consider very carefully whether all the necessary assumption are met.

Notes

1. Americans for Democratic Action.
2. American Conservative Union.
3. In majority rule run-off elections the two candidates with the most votes proceed to a second ballot, which is by majority rule.
4. In majority rule sequential elimination elections, the candidate with least votes is eliminated in each ballot, and balloting continues until only one candidate remains.
5. For example, Banks (1999: 89) observes that majority rule sequential elimination or variants thereof is commonly used in Canada.
6. There is a considerable literature on party leadership selection mechanisms (see e.g., Carty and Blake 1999), but, as far as we are aware, this literature does not contain specific hypotheses about the effects of given voting rules on the relative extremism of leaders, although McSweeney (1999) considers the policy differences of plebiscitary vs. internal leadership selection.
7. Recent work such as Miller (1996), Merrill, Grofman, Brunell and Koetzle (1999), Grofman, Merrill, Brunell and Koetzle (1999), and McGann (1997, 2000) has demonstrated the power of ideologically concentrated minorities.

8. While using NOMINATE scores (Poole and Rosenthal, 1985, 1997) has its advantages (see e.g., Polser and Rhodes, 1997:358), we prefer to use one of the standard roll-call measure that focus on many of the most important (and usually controversial) items before Congress.
9. Brunell et al. (1999) show that ADA scores can be shown to differentiate better among liberals and ACU scores to differentiate better among conservatives: ACU scores tend to clump all strong liberals together, while ADA scores tend to clump all strong conservatives together. To assure comparability over time of ADA and ACU scores we have made use of the correction method of Groseclose, Levitt, and Snyder (1999) to adjust the ADA and ACU scores.
10. The kernel density estimate of f_h of a univariate density f based on a random sample $W_1, \ldots, W_n$ is

$$f_{h(w)} = \sum_{i=1}^{n} \frac{1}{h} K\left(\frac{w - H_i}{h}\right)$$

where h is the bandwidth and K is the kernel function. We use a Gaussian kernel and a bandwidth of 2.5. The Gaussian kernel K(z) is given by See Johnston and DiNardo (1997).
11. In order to correct for the fact that interest groups use different sets of votes each year, possibly resulting in the shifting and stretching of the underlying scale in ways that can cause problems for analysis across time, Groseclose, Levitt, and Snyder (1999) provide a way to "index" the ADA scores over time. Because of the adjustment process, some adjusted scores may fall outside the zero to one hundred range.
12. It is an obvious point, but one too easily neglected that, for the median alternative to be chosen, the median alternative must be one of the choices. Similarly, to say that the alternative closest to the party's median voter is likely be chosen is not very enlightening, unless we know how close the closest such alternative is, and whether such an alternative is apt to be consistently located on a particular side of the median. (As A. Wuffle (personal communication, April 1, 1978) once jokingly put it: "Contrary to popular belief, the race is not always to the swift – but only because the swift are not always in the race.")
13. See also Straffin and Grofman (1984); Grofman, Straffin and Noviello (1996).
14. The assumption of candidate nominations at random positions (in addition to being *widespread* in the literature: see e.g., Cooper and Munger 2000) is not completely unrealistic. Potential candidates do not have control over their positions (which may be the result of past position taking, motivated as much by the need to get re-elected to the House as by the desire to run for a leadership role) and the decision to run or not may be made for reasons of personal calculation.
15. Strictly speaking, the party Members are represented by a continuum, so each candidate wins votes according to the area of the distribution they win.
16. If two factions share the same position, the lower faction takes the territory to the left and the higher faction the territory to the right. If three factions share the same position, the lower takes the territory to the left, the higher the territory to the right, and the faction in the center wins nothing.
17. The concavity condition appears to be relatively robust. Distributions that are generally concave (concave over most of their mass) also produce unique, stable equilibria. (Many common distributions, such as the normal and lognormal distribution, fall into this category). Solving the model graphically suggests that to produce multiple equilibria requires a distribution function that is extremely convex, such as a distribution with a very steep central spike (see McGann 2000).
18. It should be noted that all these results represent unique, stable equilibria for the model.

19. The Grofman (1982) model generalizes straightforwardly to the multidimensional context, and most of its applications have been in that context. However, for present purposes, we limit ourselves to a single dimension. As Poole and Rosenthal (1985, 1997) have shown, a single dimension is quite good at capturing most of the variance in Congressional voting patterns and we would also expect that party leadership choice would reflect the most salient dimension of ongoing political conflict.
20. We shall neglect the essentially technical complication of ties.
21. With a non-uniform distribution, there is no equilibrium unless the number of modes is at least half the number of parties (Eaton and Lipsey 1975, p. 35).
22. That is, however the candidates distribute themselves, some candidates can improve their situation by moving or, if we assume that candidate positions are fixed, some faction can make itself better off by nominating a different member of that faction.
23. King and Zeckhauser (1998) propose another mechanism to explain why party leaders may be non-centrists, based on strategic behavior by party members anticipating the outcome of their leaders' negotiating behavior.

References

Adams, G. (1998). An evolution-based model of party dynamics. Mimeo. Paper presented at 1998 Public Choice Society Annual Meeting, March 13–15, 1998.

Aldrich, J. (1983). A Downsian spatial model with party activism. *American Political Science Review* 77: 974–990.

Aldrich, J. and M. McGinnis, M. (1989). A model of party constraints on optimal candidate positions. *Mathematical and Computer Modeling* 12: 437–450.

Aranson, P. and Ordeshook, P. (1972). Spatial strategies for sequential elections. In R.G. Niemi and H.F. Weisberg, (Eds.), *Probability models of collective decision making*, 298–331. Columbus, Ohio: Charles E. Merrill Publishing Company.

Axelrod, R. (1970). *Conflict of interest*. Chicago: Markham.

Banks, M. (1999). Comment on Gregg Phifer's article, "Voting Procedures". *Parliamentary Journal* 50: 89–90.

Black, D. (1958/1971). *The theory of committees and elections*. Cambridge: Cambridge University Press.

Brunell, T., Koetzle, W., DiNardo, J., Grofman, B and Feld, S. (1999). The R2 = .93, Where they do they differ? Comparing liberal and conservative interest group ratings. *Legislative Studies Quarterly* 24: 87–99.

Carty, R.K. and Blake, D. (1999). The adoption of membership votes for choosing party leaders: The experience of Canadian parties. *Party Politics* 5: 211–224.

Clausen, A. and Wilcox, C. 1987. Policy partisanship in legislative leadership recruitment and behavior. *Legislative Studies Quarterly* 122: 243–263.

Coleman, J. (1971). Foundations for a theory of collective decisions. In B. Lieberman (Ed.), *Social choice*, 27–46. New York: Gordon and Breach,

Coleman, J. (1972). The positions of political parties in elections. In R.G. Niemi and H.F. Weisberg (Eds.), *Probability models of collective decision making*, 332–357. Columbus, Ohio: Charles E. Merrill Publishing Company.

Cooper, A and Munger, M. (2000). The (un)predictability of primaries with many candidates: Simulation evidence. *Public Choice* 103: 337–355.

Davis, O., Hinich, M. and Ordeshook, P. (1970). An expository development of a mathematical model of the electoral process. *American Political Science Review* 64:426–448.

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

Eaton, B. and Lipsey, R. (1975). The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. *Review of Economic Studies* 42: 27–49.

Fishburn, P. and Gehrlein, W. (1976). An analysis of simple two stage voting systems. *Behavioral Science* 21: 1–12.

Fishburn, P. and Gehrlein, W. (1977). An analysis of voting procedures with nonranked voting. *Behavioral Science* 22: 178–185.

Grofman, B. (1982). A dynamic model of proto-coalition formation in ideological N-space. *Behavioral Science* 27: 77–90.

Grofman, B., Koetzle, W. and McGann, A. (1998). Leadership selection in the U.S. House. Mimeo. Paper presented at the Annual Meeting of the American Political Science Association.

Grofman, B., Merrill, S., Brunell, T. and Koetzle, W. (1999). The potential electoral disadvantages of a catch-all party: Ideological variance among republicans and democrats in the 50 U.S. states. *Party Politics* 5: 199–210.

Grofman, B., Straffin, P. and Noviello, N. (1996). The sequential dynamics of cabinet formation, stochastic error, and a test of competing models. In M. Schofield (Ed.), *Collective decision making: Choice and political economy*, 281–293. Boston: Kluwer-Nijhoff.

Groseclose, T., Levitt, S., and Snyder, J. Jr. (1999). Comparing interest group scores across time and chambers: Adjusted ADA scores for the U.S. Congress. *American Political Science Review* 93: 33–50.

King, D. (1998). The polarization of American political parties and mistrust of government. Kennedy School of Government Politics Research Group Working Paper.

King, D. and Zeckhauser, R. (1998). Legislators as negotiators. Kennedy School of Government Politics Research Group Working Paper.

Kollman, K., Miller, J. and Page, S. (1997) Political institutions and sorting in a Tiebout model. *American Economic Review* 87: 977–992.

Koopman, D. 1996. *Hostile takeover*. Lanham, MD: Rowman and Littlefield Publishers, Inc.

McGann, A.J. (1997). Modeling the effects of intra-party institutions on systems of party competition (Skewness, party dominance and the accommodation of shifts in public opinion). Mimeo. Paper delivered at the 1997 Annual Meeting of the American Political Science Association, Washington, DC, August 28–31, 1997.

McGann, A.J. (1999). The modal voter result under various voting systems. Mimeo. Paper presented at the 1999 Annual Meeting of the Public Choice Society, New Orleans.

McGann, A.(2000). The advantages of ideological cohesion: A model of constituency representation and electoral competition in multi-party democracies. University of California, Irvine: Institute for Mathematical Behavioral Sciences Technical Report 00-24.

McGann, A., Koetzle, W. and Grofman, B. (2000). How an ideologically concentrated minority can trump a dispersed majority: Modeling plurality, run-off and sequential elimination elections. University of California, Irvine: Institute for Mathematical Behavioral Sciences Technical Report 00-25.

McSweeney, D. (1999). Changing the rules changed the game: Selecting conservative leaders. *Party Politics* 5: 471–483.

Merrill, S. III. (1984). A comparison of the efficiency of multicandidate electoral systems. *American Journal of Political Science* 28: 23–49.

Merrill, S. III. (1985). A statistical model for Condorcet efficiency based on simulations under spatial model assumptions. *Public Choice* 47: 389–403.

Merrill, S. III. (1988). *Making multicandidate elections more democratic.* Princeton: Princeton University Press.

Merrill, S. III, Grofman, B., Brunell, T. and Koetzle, W. (1999). The power of ideologically concentrated minorities. *Journal of Theoretical Politics* 11: 57–74.

Milchtaich, I. and Winter, E. 1997. Stability and segregation in group formation. Mimeo. Hebrew University of Jerusalem and Washington University, St Louis.

Miller, Nicholas. (1996). Majority rule and minority interests. In I. Shapiro and R. Hardin (Eds.), *Political order: Nomos XXXVII*, 207–250. New York: New York University Press.

Owen, G. and Grofman, B. 1995. Two-stage electoral competition in two-party contests: Persistent divergence of party positions with and without expressive voting. Mimeo. Paper prepared for delivery at the Annual Meeting of the Public Choice Society, Long Beach, California, March 24–26, 1995. (A revised version was presented at the Conference on Strategy and Politics, Center for the Study of Collective Choice, University of Maryland, College Park, MD, April 12, 1996.)

Poole, K. and Rosenthal, H. (1985). A spatial model for legislative roll call analysis. *American Journal of Political Science* 29: 357–384.

Poole, K. and Rosenthal, H. (1997). *Congress: A political-economic history of roll call voting.* New York and London: Oxford University Press.

Polser, B. and Rhodes, C. (1997). Pre-leadership signaling in the U.S. House. *Legislative Studies Quarterly* 22: 351–368.

Rae, N. (1989). *The decline and fall of the liberal republicans.* Oxford: Oxford University Press.

Regenwetter, M., Adams, J. and Grofman, B. (2000). Don't worry about cycles, worry about representivity. Mimeo. Paper prepared for delivery at the Annual Meeting of the Public Choice Society, Charleston, South Carolina, March 10–12.

Reiter, H. (1981). Intra-party cleavages in the United States today. *Western Political Quarterly* 34: 287–300.

Riker, W. (1982). *Liberalism v. populism.* San Francisco: W.H. Freeman & Company.

Robertson, D. (1976). *A theory of party competition.* London: John Wiley & Sons.

Rohde, D. (1991). *Parties and leaders in the post-reform house.* Chicago: University of Chicago Press.

Saari, D. (1995). *The basic geometry of voting.* Heidelberg: Springer-Verlag.

Schousen, M. (1994). *Who's in charge? A study of coalitions and power in the U.S. House of Representatives.* Unpublished doctoral dissertation, Department of Political Science, Duke University.

Straffin, P. and Grofman, B. (1984). Parliamentary coalitions: A tour of models. *Mathematics Magazine* 57: 259–274.

Tiebout, C. (1956). Pure theory of local expenditures. *Journal of Political Economy* 64: □–000.

Truman, D. (1959). *The congressional party.* New York: Wiley.

Westhoff, F. (1977). Existence of equilibria with a local public good. *Journal of Economic Theory* 14: 84–112.

Westhoff, F. (1979). Policy inferences from community choice models: A caution. *Journal of Urban Economics* 6: 535–549.