How an Ideologically Concentrated Minority Can Trump a Dispersed Majority: Nonmedian Voter Results for Plurality, Run-Off, and Sequential Elimination Elections

Anthony J. McGann University of California, Irvine
William Koetzle U.S. House of Representatives (Staff)
Bernard Grofman University of California, Irvine

In contrast to Downs' (1957) median voter result for two-candidate elections, we should not expect multicandidate elections to produce outcomes around the median, even on average. Rather the winning candidate will tend to be between the median voter and the mode. This is true whether we consider plurality, run-off, or sequential elimination elections, and whether or not the electorate is divided into factions that control the nomination of candidates.

For an unfactionalized electorate, computer simulation with randomly positioned candidates is used to model various electoral systems. This is because no equilibrium solution exists for this case. For a factionalized electorate, a game-theoretic model of faction formation is used.

The results suggest that an ideologically cohesive minority around the mode of the population distribution may have a disproportionate influence on the outcome. These results can be applied to party leadership in the U.S. House, following Grofman, Koetzle, and McGann (forthcoming).

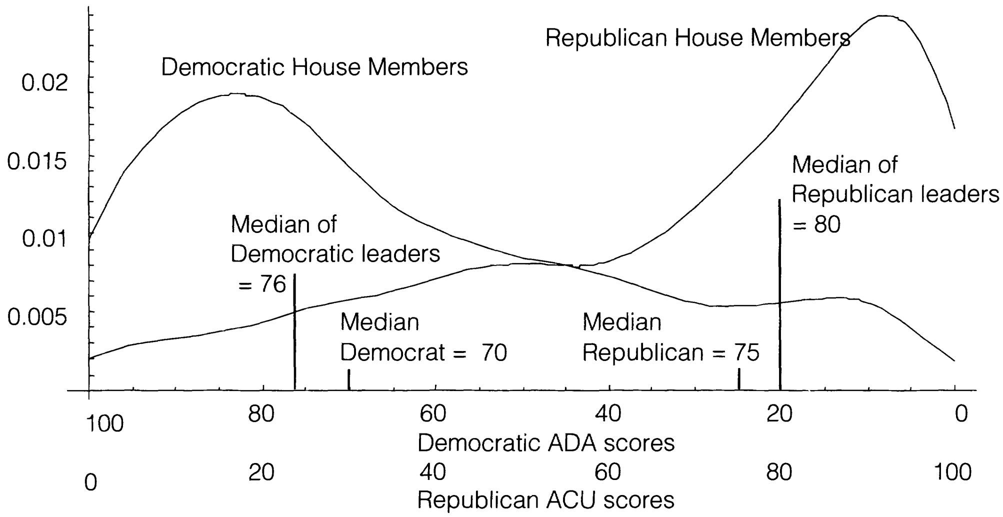
This article explains why the most common electoral systems for choosing a single position in multicandidate elections (plurality, run-off, and sequential elimination) may produce results that are biased against the median in favor of ideologically more extreme results. Some previous work (Davis, Hinich, and Ordeshook 1970; Comanor 1976; Merrill et al. 1997, Grofman, Koetzle, and McGann forthcoming; McGann 1997, 1999, forthcoming) has suggested that skewed preference distributions under a variety of multicandidate institutional settings produce results that deviate from the median voter, and instead tend towards the population mode. This contrasts with Downs' (1957) median voter result, which assumes two candidates. Grofman, Koetzle, and McGann (forthcoming), for example, argue that skewed preference distributions amongst House party Members have produced leaders who are more extreme than the median Members of their parties. Here we lay out a theory of why these kinds of effects may occur, using simulated elections with both hypothetical preference data and U.S. House data from Grofman, Koetzle, and McGann (forthcoming).

Anthony J. McGann is Assistant Professor of Political Science at the University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100 (amcgann@uci.edu). Bill Koetzle is Staff at the U.S. House of Representatives. 4 Barrington Court, Stafford, VA 22554 (bill.koetzle@mail.house.gov). Bernard Grofman is Professor of Political Science, School of Social Sciences, University of California–Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100 (bgrofman@uci.edu). The contributions of the third author were supported by National Science Foundation Grant #SBR 446740-21167, Program in Methodology, Measurement, and Statistics (to Bernard Grofman and Anthony Marley).

We would like to thank Sam Merrill, for his extremely helpful comments. We would also like to thank Emerson Niou, Hervé Moulin, John Brehm, Lin Zhou, John Aldrich, Herbert Kitschelt, Olga Shvetsova, and James Adams. All remaining errors are, of course, the responsibility of the authors.

American Journal of Political Science, Vol. 46, No. 1, January 2002, Pp. 134–148

FIGURE 1 Distribution of ADA/ACU scores for House Members, 1965–96



Consider the example provided by Grofman, Koetzle, and McGann (forthcoming). Using adjusted ADA[1] scores from 1965–96 and ACU[2] scores from 1971–96, they find that the distribution of ideological positions of Democratic House members has been strongly skewed to the right (i.e., the right tail is long and fat, so that the main concentration is on the left), while the distribution of House Republicans has been skewed to the left. This is shown in Figure 1. Furthermore Grofman, Koetzle, and McGann show that party leadership elections in the House produce results that are systematically more extreme than the median party Member. The median of the ADA scores of Democratic House leaders has been significantly to the left of the median for all Democratic Members, and the median of Republican House leaders has been significantly to the right of the median of Republican Members. Following Clausen and Wilcox (1987), this contradicts the "middleman" theory of Congressional leadership (Truman 1959) that party leaders should be selected from around the median of the party. Instead each party chooses leaders from its ideological core, which is considerably more extreme that its median Member.[3]

The reason that we should be concerned that the candidate closest to the median is often defeated is that the candidate closest to the median is the Condorcet winner and is thus preferred by the majority to any other candidate in a one-on-one contest. If, as we allege, the candidate closest to the median is systematically beaten by candidates between the median and the mode, then it would appear that an ideologically concentrated minority is having its way, despite the fact that the majority of the electorate would prefer the candidate nearest the median. Thus the will of the majority is frustrated. If the electoral system sometimes picked winners to the left of the median, and sometimes to the right, this would pose less of a problem. The majority preferred candidate would not always be chosen, but at least the median of elected candidates would be close to the median of the population as a whole. What we find, however, is that when the preference distribution is skewed, there is a systematic bias towards candidates on one side of the median—the side closest to the mode. If this is so, the key to winning may not be to make broad appeals, but to have a constituency that is ideologically concentrated.

In terms of understanding why multicandidate elections with skewed preference distributions produce outcomes that are biased towards the mode, the standard theory of spatial competition does not provide much help. When there are only two candidates, following Downs (1957) and Black ([1958] 1971), we would expect both candidates to maximize votes by situating themselves at the position of the median voter. However,

[1]Americans for Democratic Action.

[2]American Conservative Union. Grofman, Koetzle, and McGann (forthcoming) use ADA scores for Democrats and ACU scores for Republicans, because the ADA discriminates far more finely between Democrats and ACU far more finely amongst Republicans. For example, the ADA gives most Republican very low scores, and this lack of variance makes it very difficult to distinguish moderate from conservative Republicans, all of whom are very conservative in the eyes of the ADA. See Brunell et al. (1999).

[3]King and Zeckhauser (1998) propose another mechanism to explain why party leaders may be noncentrists, based on strategic behavior by party members anticipating the outcome of their leaders' negotiating behavior.

with three of more candidates, it is far from clear what the optimal strategy is in plurality, run-off, or sequential elimination elections.[4] The problem is that the optimal strategy for any candidate is highly sensitive to the strategies of the other candidates—if a particular region is crowded with candidates, it is not advantageous to take a position there.

Election systems, however, have to produce results, even if they do not produce equilibria. An election officer cannot return the result "No equilibrium." Therefore we need to consider alternative approaches to investigate the results we would expect in the absence of "preference induced equilibrium" (Shepsle 1986). This article uses two approaches, one assuming that the election is not factionalized, the second assuming that it is. Firstly, we assume that the electorate is not factionalized and use computer simulation with randomly generated lists of candidates. Secondly, we assume that the electorate is factionalized, and that the factions have a monopoly in proposing candidates. Factions are defined here as voluntary associations of voters who agree not to run independently, but instead to nominate a single candidate collectively. We model the outcomes produced by factionalization analytically, using a sorting process. In both cases we can generate predictions of what type of outcomes each electoral system will produce with specific distributions of voter preferences.

We consider three electoral systems: plurality, run-off, and sequential elimination. The relevance of plurality (the winner is simply the candidate with most votes) is obvious—it is used for elections in the United States, Canada, the United Kingdom, India, and in many other countries, including some (such as Italy) where some of the legislative seats are elected by plurality, others by proportional representation. Run-off elections (the top two vote-getters advance to the second round, which is decided by plurality) are used in France, Chile, Brazil, and the State of Louisiana. Sequential elimination elections (the bottom vote-getter is eliminated in each round, until only one candidate remains) are used in party leadership elections in Canada, the United Kingdom, and the United States.[5] If we assume that people vote sincerely and do not change their preferences between ballots, it is equivalent to single transferable vote with a single member district.

Unfactionalized Electorate

Theory

In plurality elections with two candidates, whichever candidate is closest to the median of the distribution will win. However, if there are more than two candidates, it is harder to make predictions. If the candidates are evenly spaced, then the winning candidate will come from the densest part of the distribution—that is, the region around the mode. However, if the distribution of candidates follows the same distribution as the voters, then the region that is densest in voters will also be densest in candidates. The greater density of voters increases the probability of an outcome around the mode, but the greater concentration of candidates reduces it, so there is no guarantee that a candidate from this region will actually win. In general, the outcome will be highly sensitive to the distribution of candidates.

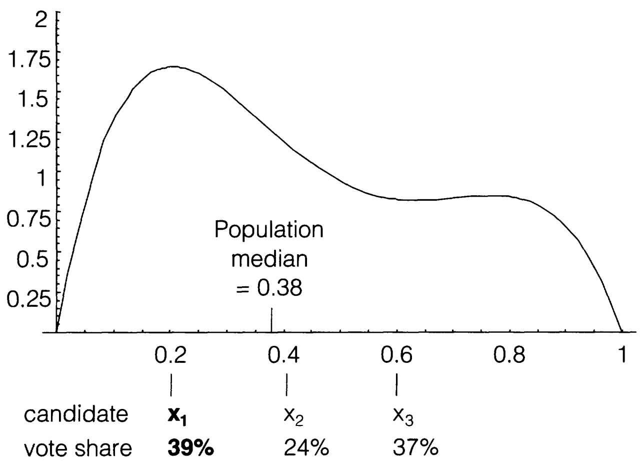
In run-off elections the two candidates with the most votes advance to a second round, which is decided by simple plurality. This to some degree addresses the shortcomings of plurality elections, in that it would seem to reduce the chance of an unpopular candidate winning because the opposition is split. Indeed, Merrill (1984, 1985) finds that with symmetric distributions of preferences both run-off and sequential elimination elections are more likely than plurality elections to choose the Condorcet winner. However, it certainly does not guarantee that the Condorcet winner will win. Consider Figure 2. In the first round, the Condorcet winner (candidate 2) is eliminated, leaving candidate 1 to defeat candidate 3 in the final round. Indeed if the area around the median is crowded with candidates, it appears unlikely that the Condorcet winner will win, because all the centrist candidates will eliminate each other in the first round.

It is hard to predict exactly what effect a skewed preference distribution will have on run-off elections. Presumably candidates in the densest part of the preference distribution will have a better chance of making it to the

[4]Eaton and Lipsey (1975) show that there is no equilibrium in multicandidate elections for vote-maximizing parties, unless the preference distribution is rectangular or the number of modes is at least half the number of candidates. However, with probabilistic voting and some nonspatial voting component, it is possible to calculate equilibrium strategies with some preference distributions (see Merrill and Adams 1999; Adams and Merrill 1999). To date, we are unaware of any equilibrium results for multicandidate run-off or elimination elections. However, it does not appear likely that there is an equilibrium set of strategies.

[5] Of course, leadership elections in the U.S. House frequently have only two candidates or are uncontested. However, the decision to run or not is influenced by consideration of what a candidate's chances would be in elimination elections. Thus much of the elimination of candidates may take place before the official election. See, for example, Brown and Peabody (1992) on the 1989 Democratic House leadership elections. In some cases, of course, multicandidate contested elections do happen (see Merrill (1988) for an account of the Democratic Majority Leader contest in 1976).

FIGURE 2 Run-off Election with Skewed Voter Distribution (Final Ballot Winner in Bold)



second round (provided that this region is not crowded with candidates). However, in the second round, the candidate who is closest to the median will win. Thus there are forces advantaging both candidates near the mode and those near the median. As with simple plurality elections, it is apparent the outcome will be highly sensitive to the distribution of candidates. However, on average, we would expect the winning candidate to be between the median and mode.

With sequential elimination elections, we would expect similar results. In sequential elimination elections, the candidate with fewest votes in each round is eliminated. If the candidate closest to the median makes it to the final round, that candidate will win. However, the candidate closest to the median may not make it to the last round, particularly if candidates on the extremes (where the voter density is lowest) are eliminated early, and their support goes to more moderate candidates on the left and right. As with run-off elections, candidates near the median have the advantage if they make it to the final round, but candidates near the mode may have a better chance of not being eliminated in the earlier rounds. Therefore, in repeated trials, we would expect that average winning candidate to lie between the median and the mode.

Methods

The electoral systems we are interested in often do not in general produce equilibria with more than two candidates, unless there are organized factions. Therefore we proceed by using computer simulation. Simulations of elections with randomly chosen slates of candidates have been performed by Fishburn and Gehrlein (1976, 1977), Merrill (1984, 1985, 1988) and Cooper and Munger (2000). We proceed in a similar manner, except that we consider skewed preference distributions as well as symmetric ones. Voter ideal-points are represented by a continuous probability density function. A slate of ten candidates[6] is generated for each iteration with random ideological positions according to some density function. The computer then calculates which candidate wins using whichever electoral system (plurality, run-off, or sequential elimination) and distribution of voters is being considered. It is assumed that each voter votes sincerely for the candidate whose ideological position is closest to their ideal-point. This process is repeated 1,000 times for each combination of electoral system and density function. By examining the distribution of winning candidates, we can gain some insight into what kind of outcome each electoral system tends to produce.

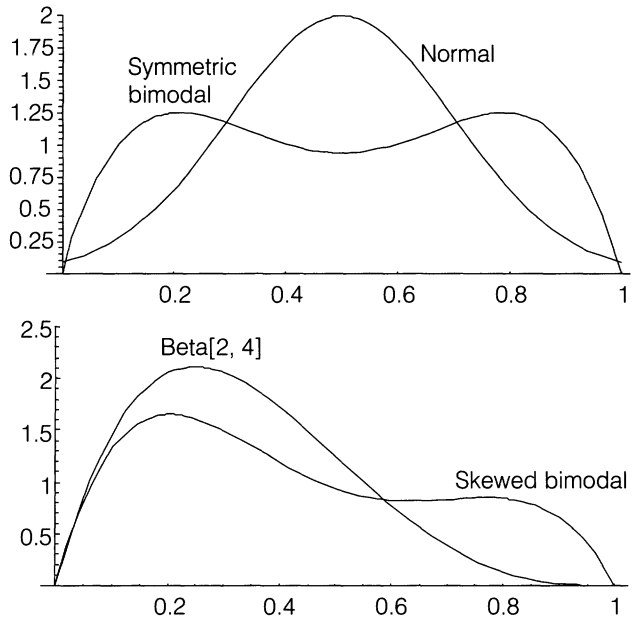
This simulation methodology is more realistic than it may at first seem. Candidates often have histories and are unable to change their position. This is especially true of legislators running for party leadership positions, in that they have years of accumulated voting records. The decision to run may be made for reasons of personal ambition and career calculations, rather than ideology, and this decision often has to be made before a candidate has full knowledge of who their competitors will be. In these circumstances, the search for an anticipatory game theoretic equilibrium is apt to be ill conceived.

Preference distributions. We consider various density functions for the electorate. For purposes of comparison we consider a normal distribution with mean 0.5 and standard deviation 0.2 and a symmetric bimodal distribution (($\beta[x, 2, 5] + \beta[x, 5, 2]/2$). What we are really concerned with, however, is how the electoral systems behave when the distribution of voters is skewed. Two skewed distributions are considered. A Beta [2, 4] distribution ($\beta[x, 2, 4]$) served as a distribution that was skewed and unimodal. We also consider a distribution that is skewed and bimodal ($\beta[x, 2, 5] + 0.5\ \beta[x, 5, 2]/1.5$). This distribution has one mode at 0.205 that is very pronounced and another smaller mode that is submerged in the right tail of the distribution. The similarity of this distribution with the actual distribution of ADA scores for Democratic House Members may be noted. These distributions are graphed in Figure 3 and summarized in Table 1.

The simulation process can also be carried out with pooled Congressional data from Grofman, Koetzle, and

[6]As a robustness test, the simulations were repeated using only five candidates per election. Results are available on request.

FIGURE 3 Voter Ideal-point Distributions Used in Simulations



McGann (forthcoming). A Gaussian kernel density function was used to smooth the data. The density function was defined: $f(x) = \Sigma\, 1/h\, k(x-w_i/h)\, W_i$, where $k(x) = 1/((2\pi)^{1/2}\exp(-x^2/2)$, h is the distance between each evenly spaced observation, w_i is the position of each observation and W_i is the relative frequency at each observation.[7] Table 1 gives the details of the distributions, which have already been graphed in Figure 1.

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 candidate distributions used were:

1. Uniform distribution in interval [0, 1]
2. The same distribution as the voters.
3. Candidates concentrated in center of distribution: Beta [2, 2] distribution.
4. Candidates concentrated at extremes of distribution: Beta [0.5, 0.5].

[7] The value of h was 5, the values of w_i were {0, 5,..., 100}, and the values of W_i were {0.005, 0.02, 0.035, 0.028, 0.028, 0.025, 0.025, 0.03, 0.038, 0.04, 0.04, 0.045, 0.05, 0.056, 0.066, 0.091, 0.096, 0.096, 0.091, 0.076, 0.061} for the Democratic ADA scores; and {0.014, 0.015, 0.016, 0.019, 0.016, 0.032, 0.026, 0.03, 0.036, 0.038, 0.043, 0.039, 0.033, 0.05, 0.05, 0.075, 0.079, 0.101, 0.129, 0.126, 0.112} for the Republican ACU scores.

Results

The simulations here address two questions. First, do skewed distributions of preferences produce results that systematically diverge from the median in the direction of the mode? Second, how do different electoral systems compare in the degree that they allow this bias, and in terms of how often they choose the Condorcet winner (in this case the candidate closest to the median)? The answer to the first question is that the median winning candidate with a skewed distribution does tend to be between the mode and median of the voter population. Regarding electoral systems, both run-off and sequential elimination have higher Condorcet efficiency (the proportion of times the method picks the Condorcet winner) than plurality election and are less biased towards the mode.

For the sake of completeness, let us briefly consider the results for the symmetric distributions of voters (normal and symmetric bimodal), which are summarized in Table 2. Our results conform to the findings of Merrill (1984, 1985). The median winning candidate (that is, the median of the candidates who win in each of the 1000 simulations) is around the median voter regardless of which election method is used. However, the election method affects the dispersion of winning candidates and the Condorcet efficiency. The standard deviation of the winning candidates is substantially less for run-off and sequential elimination elections than for plurality, and the Condorcet efficiency is substantially higher.

With skewed voter distributions we do not expect the median location of winning candidates to be at the population median, but rather to be between the median and the mode (the global mode in the case of the skewed bimodal distribution). This is what we observe, although the divergence from the median is far greater with the skewed, bimodal distribution than with the beta[2, 4] distribution. Table 3 summarizes the results with the skewed voter distributions and Figure 4 graphs the distribution of winning candidates with the skewed bimodal distribution.[8]

With plurality elections and the skewed bimodal distribution of voters, the median winning candidate is between 8.4 and 14.4 percentage points below the median—around halfway between the median of 0.38 and the mode at 0.205. With the beta[2, 4] distribution the divergence is between 0 and 7.6 percentage points, depending on the candidate distribution. (The median of the beta[2,4] distribution is 0.31 and the mode 0.25.) As

[8] The distributions are smoothed using a Gaussian kernel density function with bandwidth 0.05.

Table 1 Distributions Used in Simulations

	Normal	Symmetric Bimodal	B[2,4] Skewed	Skewed Bimodal	Democratic House ADA Scores 1965–96	Republican House ACU Scores 1971–96
Density function	N[.5, .2]	(β[2,5]+ β[5,2])/2	β[2,4]	(β[2,5]+ 0.5β[5,2])/1.5	Kernal density function	Kernal density function
Median	0.5	0.5	0.31	0.38	69.7	75.2
Mode(s)	0.5	0.21, 0.79	0.25	.205, .795	13.6, 82.8	50.4, 92
Mean	0.5	0.5	0.33	0.43	63	67.4
sd	0.2	0.27	0.18	0.26	26.1	25.8
Skewness	0	0	0.47	0.39	−0.65	−0.77
Kurtosis	3	1.75	2.63	1.99	2.36	2.53

$\beta[x, \alpha, \beta] = x^{\alpha-1} (1-x)^{\beta-1}/B[\alpha, \beta]$, $x=[0, 1]$ and $B[\alpha, \beta]$ is the integral of the beta function.

$N [x, \mu, \sigma^2] = (1 / (\sigma (2\pi)^{1/2}) \exp (- (x-\mu)^2 / 2 \sigma^2)$, $x = [0, \infty]$

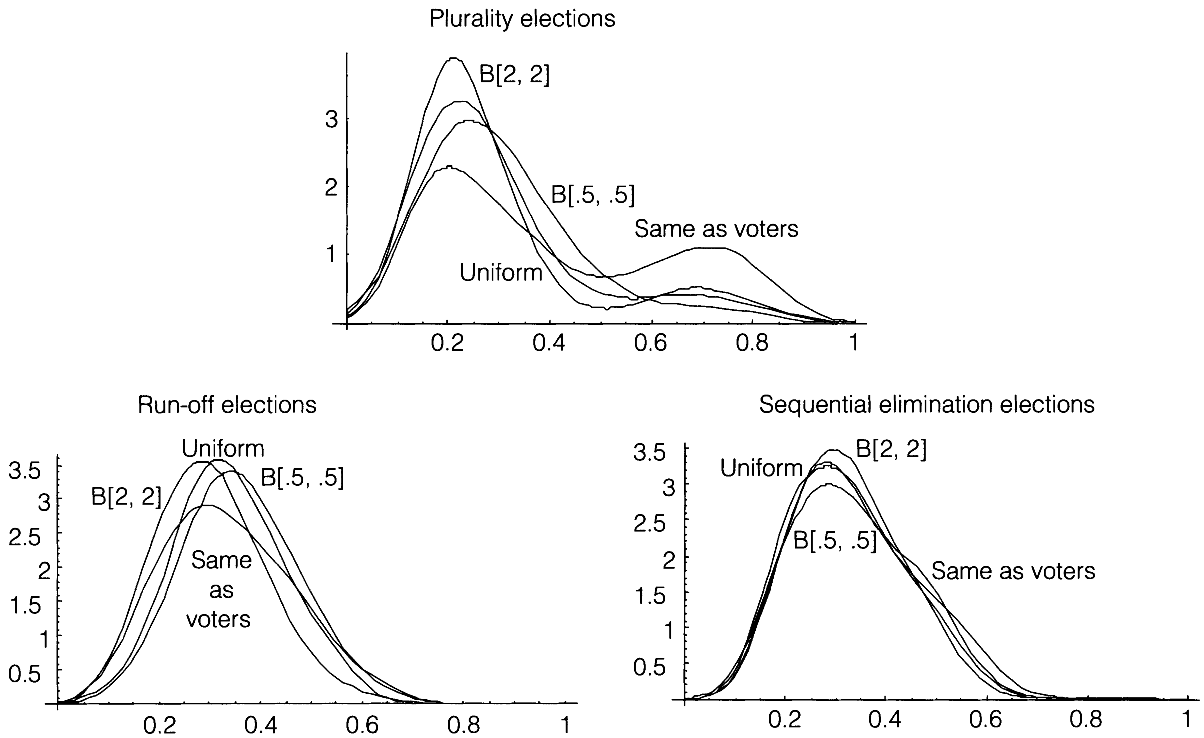
Table 2 Results of Simulated Elections with Symmetric Distribution of Voters and Candidates

Voter/Candidate Distribution	Normal			Symmetric Bimodal		
Electoral System	Plurality N=1000	Run–off N=1000	Sequential Elimination N=1000	Plurality N=1000	Run–off N=1000	Sequential Elimination N=1000
median of winning candidates	0.502	0.502	0.503	0.503	0.501	0.49
distance from median voter (0.5)	0.00234	0.00188	0.00265	0.00309	0.00117	−0.00973
mean of winning candidates	0.502	0.497	0.503	0.501	0.5	0.499
standard deviation	0.135	0.0861	0.073	0.258	0.174	0.191
standard error	0.00426	0.00272	0.00231	0.00815	0.00174	0.00603
Condorcet efficiency %	21.4	40.2	33.2	11.85	34.27	15.4

Table 3 Simulation Results with Skewed Distributions of Voters and Ten Candidates

Voter Distribution	Beta [2, 4]				Skewed bimodal			
Candidate Distribution	Uniform	Same as Voters	B[2, 2] Dense at Center	B[.5, .5] Dense at Extreme	Uniform	Same as Voters	B[2, 2] Dense at Center	B[.5, .5] Dense at Extreme
				Plurality elections				
	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000
median of winning candidates	0.273	0.31	0.238	0.29	0.253	0.296	0.236	0.285
distance from median voter	**−.0413**	**−.00402**	**−.0757**	**−.0243**	**−0.126**	**−0.0841**	**−0.144**	**−0.0944**
mean of winning candidates	0.279	0.318	0.248	0.288	0.294	0.377	0.288	0.303
standard deviation	0.0954	0.133	0.0833	0.107	0.163	0.227	0.176	0.151
standard error	0.00302	0.00421	0.00264	0.00339	0.0052	0.0072	0.0056	0.0048
Condorcet efficiency %	52.1	20.4	32.7	65.9	24.5	19.4	11.2	46.8
				Run–off elections				
	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000
median of winning candidates	0.312	0.309	0.31	0.302	0.332	0.332	0.292	0.355
distance from median voter	**−0.0016**	**−.00493**	**−.00413**	**−0.0114**	**−0.0482**	**−0.0477**	**−0.0874**	**−0.025**
mean of winning candidates	0.313	0.306	0.309	0.303	0.338	0.336	0.299	0.358
standard deviation	0.0685	0.0822	0.0628	0.09	0.0987	0.118	0.0993	0.109
standard error	0.00217	0.0026	0.0020	0.0029	0.0031	0.0038	0.0031	0.0035
Condorcet efficiency %	78.2	41.2	68.5	86.9	53.1	42.9	32.3	73.8
				Sequential elimination elections				
	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000	N=1000
median of winning candidates	0.301	0.305	0.301	0.294	0.316	0.304	0.306	0.319
distance from median voter	**−0.0128**	**−.00889**	**−0.0127**	**−0.0199**	**−0.0637**	**−0.0752**	**−0.0735**	**−0.0605**
mean of winning candidates	0.308	0.308	0.306	0.299	0.33	0.321	0.319	0.337
standard deviation	0.0783	0.0757	0.0781	0.0932	0.104	0.103	0.0964	0.118
standard error	0.00248	0.0024	0.0025	0.0030	0.0033	0.0033	0.0031	0.0037
Condorcet efficiency %	60.4	30.7	49.8	78	41.7	32.8	30.7	60.5

FIGURE 4 Distribution of Winning Candidates in Ten-Candidate Simulated Elections, with Skewed Bimodal Distribution of Voters and Various Distributions of Candidates



expected, these results are highly sensitive to the distribution of candidates. The standard deviation of the winning candidates is also considerable (between 15.1 percent and 22.7 percent in the case of the skewed bimodal distribution), indicating that although the median winning candidate is below the median of the voter population, there are winning candidates from a broad ideological range. Considering Figure 4, it is notable that when the distribution of candidates is the same as that of the voters, the distribution of winning candidates is bimodal, with a large peak near the global mode of the voter population and a smaller peak to the right of the median voter.

With run-off elections, the results are less sensitive to the distribution of candidates, as can be seen from Figure 4, which graphs the distribution of winning candidates. The median winning candidates are closer to the population median, the standard deviation of winning candidates is lower and the Condorcet efficiency is higher. With the beta distribution of voters, the median winning candidate is not significantly different from the population median. However, with the skewed bimodal distribution, the median of the winning candidates range from 2.5 to 8.7 percentage points below the population median.

Sequential elimination elections yield results similar to those of run-off elections, although the advantage for candidates closer to the mode varies less with candidate distribution. As can be seen from Figure 4, the distribution of winning candidates is similar for all the distribution of candidates considered. With a beta[2, 4] distribution, the median winning candidate is significantly below the mode, but only by about one percentage point. With the skewed, bimodal distribution, the distance between the median of the winning candidates and the median voter is between 6 and 7.5 percentage points.

It is not surprising that the bias away from the median voter in the direction of the mode is far stronger in the case of the skewed bimodal distribution than it is for the beta[2, 4] distribution. The mode and median for the beta[2,4] distribution are close and the density at the median is only slightly less than that at the mode. Therefore a candidate close to the median frequently survives to the final round of run-off or sequential elimination elections and wins. However, with a bimodal skewed distribution, a candidate at the median has a mode on either side. In this case it is far more likely that the candidate at the median will be eliminated by candidates on either side before the final round.

TABLE 4 Results of Simulated Ten-Candidate Elections with Democratic House ADA Scores 1965–96 and Republican House ACU Scores 1971–96

Candidate Distribution	Uniform	Same as Voters	B[2, 2] Dense at Center	B[.5, .5] Dense at Extreme
		Democratic ADA scores		
	N=1000	N=1000	N=1000	N=1000
median of winning candidates	76.2	76.8	75.	75.5
distance from median voter	**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
		Republican ACU scores		
	N=1000	N=1000	N=1000	N=1000
median of winning candidates	81.	83.4	79.7	82.1
distance from median voter	**5.79**	**8.18**	**4.51**	**6.96**
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 result that run-off elections produce higher Condorcet efficiency and less bias towards the mode than sequential elimination elections with the candidate distributions considered here, however, was unexpected. We can speculate why this might be the case. It is evident from the high level of Condorcet efficiency achieved by run-off elections, that a candidate close to the median is often the first or second biggest vote-getter in the first round. With run-off elections, this candidate will go through to the second round and win. With sequential elimination elections, however, candidates on the extremes may be eliminated first and their support may then go to "moderate extremists." In some cases this additional support may be sufficient for "moderate extremists" on both sides to defeat the candidate around the median before the final round.

Results with sequential elimination elections and pooled congressional data. We also simulated sequential elimination elections[9] using adjusted ADA scores for Democratic House Members 1965–96 and adjusted ACU scores for House Republicans 1971–96, using data from Grofman, Koetzle, and McGann (forthcoming). The results are given in Table 4. For the Democratic Members, the median winning candidate has a score between 5 and 7 percentage points higher (more liberal) than the median Member. (The median of the distribution of Democrats is 70 and the mode 82.) For the Republicans, the median winning candidate scores between 4.5 and 8.1 points higher (more conservative) than the median Republican, depending on the assumption about the distribution of candidates. (The median Republican has an ACU score of 75, and the mode was 92.) Grofman, Koetzle, and McGann (forthcoming) find that the median ADA score of Democratic House leaders in the period 1965–96 was 76, and the median ACU of Republican House leaders 1971–96 was 80. These are both considerably more extreme that the median party Members (70 for the Democrats and 75 for the Republicans) and are quite close to the results of our simulations.

[9]An anonymous reviewer points out that for some Democratic House positions it is not specifically forbidden for new candidates to enter between ballots of the elimination elections. We are not aware of this having happened, and are unsure whether a presiding officer would allow such entry. As a robustness test, the simulations with Congressional data were rerun with a candidate at the median being reintroduced in the final ballot. This affected the outcome in only 6 percent of the trials and did not significantly change the median of the winning candidate.

Factionalized Electorate

Theory

It is apparent that under various electoral rules it will be advantageous for a group of people in an ideological region to form a faction to coordinate who runs for office, in order to prevent similar candidates from sabotaging each other. This is certainly the case with plurality, and to a lesser degree, with run-off and sequential elimination elections. This section presents a model in which the electorate partitions itself into a number of factions, each of which may nominate one of its members as a candidate for the election. These candidates then compete in an election where all individuals vote sincerely for the candidate closest to their ideal-point. There is reason to believe that factionalization will exacerbate the tendency for outcomes to be biased away from the median and towards the mode of the distribution.

The model is based on a simple premise about the behavior of people joining factions. There is a given number of factions, and everyone joins the faction whose typical member is most similar to themselves. That is to say, everyone joins the faction whose median member is closest to their ideal-point. 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 most similar to herself—that is, the faction whose median member is closest to that member's ideal point.

Although the idea of like-minded people spontaneously grouping together into a party or faction can be traced back as far as Burke (1770/1898), the application of this type of model to political parties was first suggested by Robertson (1976) and developed formally by Aldrich (1983) and Aldrich and McGinnis (1989). McGann(1997) extended this model to more than two parties and derived results based on the shape of the preference distribution. Adams (1998) used agent-based modeling techniques on a similar model. There is also a considerable literature in economics starting with Tiebout (1956) dealing with the mathematically similar problem of how individuals choosing to move between communities can solve the problem of public goods provision. Of particular relevance is the work of Westhoff (1976, 1979), Milchtaich and Winter (1997) and Kollman, Miller, and Page (1997), in that they either use a spatial representation or explicitly consider voting mechanisms.

The model is formally defined:

Assume that the population is distributed across a one-dimensional issue space S, between points 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. The position of each faction is known by all voters with certainty. Let us make three further assumptions:

1. Each member of the population joins the faction whose typical member position is closest to their own ideal point. Each member of the population has complete knowledge of all typical member positions.
2. The typical member position of each faction is the ideal-point of the median member of the faction.
3. Factions are ordered so their positions (x_1, x_2, etc.) are so $x_1 \leq x_2$, etc. If two factions share the same typical member position, then everyone whose ideal-point is lower than this joins the lower faction, and everyone whose ideal point is higher joins the higher faction. If three or more factions share the same position, everyone with an ideal-point lower than the position of the factions joins the lowest faction, everyone with a higher ideal-point joins the highest faction, and nobody joins the faction(s) in the middle.

Given the partition of voters into factions by the above process, each faction chooses a candidate from its own membership, who then competes in an election in which every voter votes sincerely. The utility of each faction is given by a monotonically decreasing function of the distance of the winning candidate in the election from the ideal-point of the faction's median member.

We limit discussion here to the two-faction case. (See McGann [forthcoming] for a discussion of the multiple faction case.) The two-faction case is particularly significant because of the tendency of plurality elections to reduce the number of competing parties or factions to two, as observed by Duverger (1954). As illustrated in Figure 5, the model partitions the voters into two factions divided by cut-point c, with every voter with an ideal-point to the left of c joining faction 1, and every voter with an ideal-point to the right of c joining faction 2. Aldrich (1983) shows that for any continuous density function, there must be an equilibrium partition of the voters between the parties.

Given the partition of the voters between the factions, we can consider the candidates that the factions will propose and the electoral outcome. If the cut-point c

FIGURE 5 Outcome of Model with Two Factions

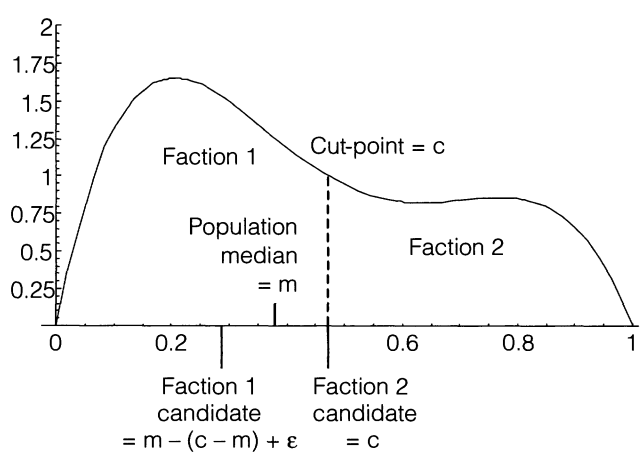


TABLE 5 Results of Two-Party Factionalized Model with Various Distributions (Winning Candidate Bold)

	Normal	Symmetric Bimodal	B[2, 4]	Skewed Bimodal
Cut-point (c)	0.5	0.5	0.34	0.47
Median voter (m)	0.5	0.5	0.315	0.38
Distance c – m	**0**	**0**	**0.025**	**0.09**
Candidates	0.5, 0.5	0.5, 0.5	**.29**, .34	**.29**, .47

is to the right of the ideal-point of the median voter, then faction 1 will be larger than faction 2. Furthermore, faction 1 will always be able to propose a candidate that can beat the candidate of faction 2. In a two-candidate election in one dimension, the candidate nearest the median voter will win. Given that the factions have to propose one of their own members as their candidate, the candidate of faction 2 cannot be closer to the median than the cut-point c. Therefore as long as faction 1 proposes a candidate closer to the median than cut-point c, its candidate will win. In equilibrium, faction 1 will propose a candidate at $m - (c - m) + e$ (the furthest point from the median voter that still guarantees it victory), or at the ideal-point of its median member, if this is closer to the median voter than $m - (c - m) + \varepsilon$. Thus if we know the distance from the cut-point to the median voter $(c - m)$, we know which faction will win, and how far its candidate can deviate from the median voter.

The numerical examples from McGann (forthcoming) suggest that if the distribution of preferences is skewed to the right (the right tail is long and fat), then the cut-point will be to the right of the median voter. If this is so, the faction to the left will propose the winning candidate, who will be to the left of the median voter. Unfortunately, we are unable to show that this result holds for all distributions of ideal-points that are skewed to the right. Therefore we will need to consider numerical results with various skewed distributions of ideal-points.

Methods

In the case of the model of the factionalized electorate, the position of the cut-point c for any continuous distribution of ideal-points can be found by numerically solving the equation: $(F^{-1}(1/2\ (F(c)) + F^{-1}(1/2\ (1-F(c))) / 2 = c$, where F is the cumulative distribution function. Graphical methods were used to check that the solutions were unique (see McGann forthcoming). Having calculated the cut-point, we can calculate the equilibrium position of each faction's candidate and the electoral outcome.

The results of the factionalized model were calculated using the same distributions of ideal-points as with the unfactionalized model. In addition, results were calculated for three families of skewed preference distributions: Beta$[\alpha, \beta]$ ($\alpha = 1\ldots6$; $\beta = (\alpha+1)\ldots6$); gamma$[\alpha, \beta]$ ($\alpha = 1\ldots6$; $\beta = 2$); and lognormal $[\mu, \sigma^2]$ ($\mu = 0$; $\sigma^2 = 0.1, 0.25, 0.5, 0.75, 1, 2$). The distributions are defined (Spanos 1999):

$\beta[x, \alpha, \beta] = x^{\alpha-1}\ (1 - x)^{\beta-1}/B[\alpha, \beta]$, $x = [0, 1]$ and $B[\alpha, \beta]$ is the integral of the beta function.

$\gamma[x, \alpha, \beta] = \beta^{-1}/\Gamma[\alpha](x/\beta)^{\alpha-1} \exp(-x/\beta)$, $x = [0, \infty]$ and $\Gamma[\alpha]$ is the integral of the gamma function

$l\ [x, \mu, \sigma^2] = 1/x\ (1/(\sigma(2\pi)^{1/2}) \exp(-(\ln(x - \mu))^2/2\sigma^2)$, $x = [0, \infty]$.

Results

We have shown in the case of the two-faction model that if the cut-point between the two factions is to the right of the median voter (that is, $c - m > 0$), then the faction to the left will win, even though it proposes a candidate whose ideal-point is $(c - m - \varepsilon)$ to the left of the median voter. Thus the winning candidate will lie between the median and the population mode. We have hypothesized that when the distribution of ideal-points is skewed to the right, the cut-point will be to the right of the median. For all the skewed distributions we consider, the results for the model are consistent with this hypothesis. Table 5 gives the results for the Beta[2, 4] distribution and the skewed bimodal distribution. Table 6 gives the results for the families of beta, gamma, and lognormal distributions considered. It is notable that the results of the two-faction model seem to be even more biased away from the median than those for the unfactionalized model with multicandidate run-off or sequential elimination elections, at least in the case of the Beta[2, 4] and skewed bimodal distributions. For example, with the skewed bimodal distribution, the distance of the winning candidate in the factionalized model from the median is 0.09, whereas in the case of the unfactionalized model with sequential elimination elections, the distance of the median winning candidate from the median ranges from 0.0605 to 0.0752 depending on the distribution of candidates.

Table 6 **Distance of Cut-point from Median Voter (c – m) in Two-party Factionalized Model with Various Distributions (Median Voter Position in Brackets)**

Gamma[α, 2] distribution

α	1	2	3	4	5	6	120
	0.405 (0.69)	0.407 (1.68)	0.407 (2.67)	0.407 (3.67)	0.407 (4.67)	0.407 (5.67)	.407 (119.7)

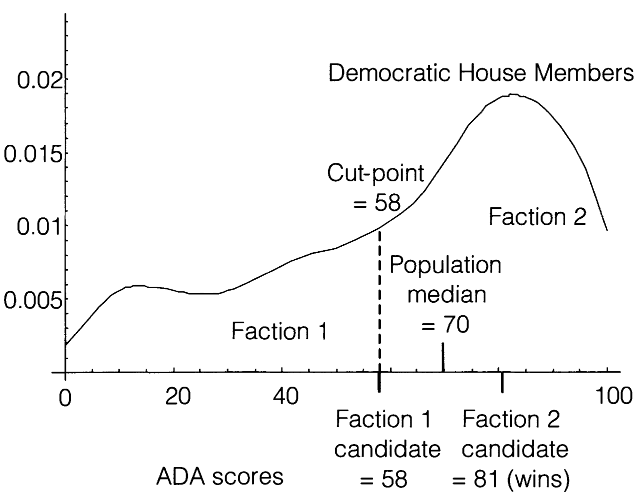
lognormal [0, σ^2] distribution

σ^2	0.1	0.25	0.5	0.75	1	2
	0.063 (1.0)	0.167 (1.0)	0.367 (1.0)	0.607 (1.0)	0.897 (1.0)	2.835 (1.0)

Beta[α, β] distribution

β	2	3	4	5	6
α					
1	0.053 (0.29)	0.057 (0.21)	0.054 (0.16)	0.049 (0.13)	0.044 (0.11)
2		0.018 (0.39)	0.024 (0.31)	0.027 (0.26)	0.027 (0.23)
3			0.009 (0.42)	0.013 (0.36)	0.016 (0.32)
4				0.005 (0.44)	0.008 (0.39)
5					0.003 (0.45)

Figure 6 **Outcome of Model with Two Factions and Democratic House ADA Scores 1965–96**



Results with pooled congressional data. We calculated the results of the two-faction model using the adjusted ADA scores for House Democrats and adjusted ACU scores for House Republicans from Grofman, Koetzle, and M^cGann (forthcoming). The outcome of the model with Democratic House ADA scores is shown in Figure 6. As expected, the outcome is between the median and mode of the distribution of ideal-points.

According to the model, the Democratic Members will partition themselves with an ADA score of 58 as the cut-point, so that everyone with an ADA score at 58 or below will join faction 1, while everyone else will join faction 2 (higher ADA scores indicate a Member is more liberal). This cut-point is lower than the median Democratic ADA score of 70. As a result, faction 2 can win if it nominates any candidate with an ADA score of 81 or lower, because the most moderate Member faction 1 can nominate has an ADA score of 58. Given that the median member of faction 2 has an ADA score of 81, we would expect faction 2 to nominate such a candidate and win, even though this candidate is 11 points more liberal than the median Democratic Member. With the Republicans we have a very similar pattern. The cut-point between the two factions of Republicans is an ACU score of 63 (higher ACU scores indicate more conservative voting). The most moderate member of faction 1 has an ACU score of 63. Given that the median of all House Republicans is 75, faction 2 can win if it proposes any candidate with an ACU score less than 87. Given that the median member of faction 2 has an ACU score of 85, we would expect faction 2 to propose a candidate far more conservative than the median Republican and still win. These results are consistent with the finding of Grofman, Koetzle, and M^cGann (forthcoming) that nonregional party leaders of House Democrats and Republicans have been significantly more extreme than the median Member from these parties.

Conclusion

The results presented here suggest that when the distribution of voter preferences is skewed, we should expect outcomes that systematically diverge from the Condorcet winner (that is, the median voter) in the direction of the population mode. This is the case whether we use plurality, run-off or sequential elimination election rules. (McGann [1997, forthcoming] shows that a similar effect may occur with proportional electoral systems.) With plurality and to a lesser degree run-off elections, the results are very sensitive to the distribution of candidates when the electorate is not factionalized. This suggests that there is a strong incentive to form factions under these election systems, because groups of voters who form factions can secure more favorable outcomes than those who do not. When we model a factionalized electorate, the results are again between the median voter and the population mode.

In terms of the performance of the different election systems, we confirm the results of Merrill (1984, 1985, 1988) that in multicandidate elections run-off and sequential elimination systems perform far better than plurality elections, in that they are more likely to pick the Condorcet winner, and have a lower variance in their outcomes. This is true even if the distribution is skewed or bimodal. However, with skewed distributions, run-off and sequential elimination elections still have a bias away from the median in the direction of the mode, although this is typically smaller than that with plurality elections.

It has been argued by some (Wright and Riker 1997, Shugart and Carey 1992) that plurality voting will produce better outcomes than run-off elections, because it leads to competition between two large coalitions rather than between many candidates. Our results cast doubt on that conclusion. Certainly plurality elections are likely to lead to factionalization. However, we find that when the distribution of voter ideal-points is skewed, multicandidate run-off and sequential elimination elections are more likely than plurality elections with two factions to produce a candidate near the median voter. The reason that two-faction competition does not necessarily produce a median voter result is that candidates need to win nomination from their faction, and both factions may be either unable or unwilling to nominate a candidate at the population median.[10] Instead the factions may prefer to nominate a candidate as close to their median member as they can get away with and still win. This can sometimes lead to significant polarization, as in the case of the United Kingdom under plurality elections in the 1980s. Of course, if plurality elections fail to reduce the number of effective candidates to two, the results will be highly unpredictable.

Grofman, Koetzle, and McGann (forthcoming) provide an empirical example of a nonmedian voter result of the type theorized in this article. Whereas much previous literature has assumed that party leaders in the U.S. House of Representatives come from the ideological median of their respective parties, Grofman, Koetzle, and McGann follow Clausen and Wilcox (1987) in finding that party leaders instead are shifted towards the mode of the distribution of Members of their party. When we simulated sequential elimination elections with the data from Grofman, Koetzle, and McGann, we find that the predicted results for the typical party leader are very similar to the observed results, lying between the median and the mode of the party Members.

The fact that a skewed preference distribution produces a nonmedian voter result means that the distribution of voters becomes an important independent variable. If the median voter result applies, all we have to know in order to predict the outcome is the position of the median voter. However, with the results here it is necessary to understand the entire voter distribution. Regions of the distribution that are dense will be advantaged over less dense regions and may prevail even if the less dense regions are broader and have a greater population. It is possible that a concentrated minority may have its way over a more dispersed majority. This puts a premium on political coordination: if one group can agree on a preferred solution, it may get its way over larger groups that cannot coordinate.

The factional model produces another interesting consequence: the same faction should always win. The Downsian median voter result is a knife-edge result, in that it predicts that the two parties will be evenly balanced, both positioned at the median with an equal number of votes. This may be seen as normatively appealing in that it forces a high degree of accountability on the parties. If the parties diverge even slightly from the median, or if they perform poorly in office, they will be replaced. The two-faction model presented here, however, gives a quite different result. Faction 1, which is ideologically more concentrated, should always win, and it does not need to take the position of the median voter. If voters consider performance in office as well as ideology, it might lose if it performs badly enough, but we would still expect a party with such an advantage in terms of ideological voting to be dominant. Mayer (1996) suggests that the internal ideological diversity of the Democratic Party has given the Republican Party this kind of advantage.

[10] See also Aranson and Ordeshook (1972) on the strategy of candidates facing both primary and general elections.

In any case, these results show that it is necessary to be extremely careful about assuming that the median voter is decisive. The results here do not contradict Downs (1957) or Black ([1958] 1971) theoretically—the assumptions are quite different. However, the median voter argument is frequently applied without considering whether all the assumption required to sustain it apply[11] (see Grofman 1993). It is apparent that if there are more than two candidates or if the electorate is factionalized (with factions controlling nomination), then the outcome may diverge systematically from the median voter. When the distribution of the electorate is skewed, outcomes will on average lie between the mode and the median. Before making an empirical argument using the median voter result, we have to be very careful that the required assumptions are met.

Manuscript submitted November 27, 2000.
Final manuscript received July 5, 2001.

References

Adams, G. 1998. "An Evolution-Based Model of Party Dynamics." Presented at the Annual Meeting of the Public Choice Society, New Orleans, LA.

Adams, J., and Merrill, S. 1999. "Party Policy Equilibrium for Alternative Spatial Voting Models: An Application to the Norwegian Storting." *European Journal of Political Research* 36:235–255.

Aldrich, J. 1983. "A Downsian Spatial Model with Party Activism." *American Political Science Review* 77:974–990.

Aldrich, J., and 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 *Probability Models of Collective Decision Making*, ed. R. Niemi and H. Weisberg. Columbus: Charles E. Merrill.

Black, D. 1958/1971. *The Theory of Committees and Elections.* Cambridge: Cambridge University Press.

Brown, L., and Peabody, R. 1992. "Foley, Gephart, and Gray, 1989." In *New Perspectives on the House of Representatives*, ed. R. Peabody and N. Polsby. Baltimore: John Hopkins 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.

Burke, E. 1770/1898. "Thoughts on the Cause of the Present Discontents." In *Burke Select Works Volume 1*, ed. E. Payne. Oxford: Clarendon Press.

[11]There may well be other assumptions (for example, the absence of voter abstention, see Davis, Hinich, and Ordeshook 1970) that are equally important in maintaining the median voter results.

Clausen, A., and Wilcox, C. 1987. "Policy Partisanship in Legislative Leadership Recruitment and Behavior." *Legislative Studies Quarterly* 122:243–263.

Comanor, W. 1976. "The Median Voter Rule and the Theory of Political Choice." *Journal of Public Economics* 5:169–177.

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.

Duverger, M. 1954. *Political Parties: Their Organization and Activity in the Modern State.* London: Methuen & Co.

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. 1993. "Towards an Institution-Rich Theory of Political Competition with a Supply Side Component." In *Information, Participation and Choice: An Economic Theory of Democracy in Perspective*, ed. B. Grofman. Ann Arbor: University of Michigan Press.

Grofman, B., Koetzle, W., and McGann, A. Forthcoming. "Congressional Leadership 1965–1996: A New Look at the Extremism versus Centrality Debate." *Legislative Studies Quarterly.*

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.

Mayer, W. 1996. *The Divided Democrats.* Boulder: Westview Press.

McGann, A. 1997. "Modeling the Effects of Intra-Party Institutions on Systems of Party Competition (Skewness, Party Dominance and The Accommodation of Shifts in Public Opinion)." Presented at the Annual Meeting of the American Political Science Association, Washington, D.C.

McGann, A. 1999. "The Modal Voter Result under Various Voting Systems." Presented at the 1999 Annual Meeting of the Public Choice Society, New Orleans, LA.

McGann, A. Forthcoming. "The Advantages of Ideological Cohesion: A Model of Constituency Representation and Electoral Competition in Multi-Party Democracies." *Journal of Theoretical Politics.*

Merrill, S. III. 1984. "A Statistical Model for Condorcet Efficiency Based on Simulations under Spatial Model Assumptions." *Public Choice* 472:389–403.

Merrill, S. III. 1985. "A Comparison of Efficiency of Multicandidate Electoral Systems." *American Journal of Political Science* 28:23–49.

Merrill, S. III. 1988. *Making Multicandidate Elections More Democratic.* Princeton: Princeton University Press.

Merrill, S., and Adams, J. 1999. "An Algorithm to Compute Nash Equilibria in Probabilistic, Multiparty Spatial Models with a Measured Nonpolicy Component." Presented at the Annual Meeting of the Public Choice Society, New Orleans, LA.

Merrill, S., Grofman, B., Brunell, T., and Koetzle, W. 1997. "The Power of Ideologically Concentrated Electorates." *Journal of Theoretical Politics* 11:57–74.

Milchtaich, I., and Winter, E. 1997. "Stability and Segregation in Group Formation." Unpublished manuscript. Washington University.

Robertson, D. 1976. *A Theory of Party Competition.* London: John Wiley and Sons.

Shepsle, K. 1986. "Institutional Equilibrium and Equilibrium Institutions." In *Political Science: The Science of Politics*, ed. H. Weisberg. New York: Agathon.

Shugart, M., and Carey, J. 1992. *Presidents and Assemblies: Constitutional Design and Electoral Dynamics.* Cambridge: Cambridge University Press 1992.

Spanos, A. 1999. *Probability Theory and Statistical Inference: Econometric Modeling with Observational Data.* Cambridge: Cambridge University Press.

Tiebout, C. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64:416–424.

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.

Wright, S., and Riker, W. 1997. "Plurality and Runoff Systems and Numbers of Candidates." *Public Choice* 60:155–175.